PHYSICAL REVIEW B VOLUME 49, NUMBER 11 15 MARCH 1994-I

Effective band Hamiltonian in semiconductor quantum wells

Ahmet Elçi
Phillips Laboratory-LIDA, Kirtland Air Force Base, Albuquerque, New Mexico 87117

(Received 19 January 1993; revised manuscript received 19 July 1993)

The Luttinger-Kohn canonical transformations are used to derive the effective band Hamiltonian for a quantum well. The transformations produce δ-function potentials in the effective-mass approximation. The boundary conditions on envelope functions, obtained from this effective Hamiltonian, differ from the BenDaniel-Duke ansatz. The effective Hamiltonian depends on the electronic coherence length. It also has a velocity-dependent term that is proportional to the sum of band lineup energies.

I. INTRODUCTION

In this paper, I use the Luttinger-Kohn representation[1] and a canonical transformation to obtain an effective band Hamiltonian for an electron in a semiconductor quantum well. The method gives a systematic procedure for various approximations involved in confining the electronic motion to one band. The resulting Hamiltonian requires a set of boundary conditions different from those used in the current literature.

Contemporary investigations of electronic states in quantum wells are usually based on a phenomenological model of a heterostructure that uses the effective-mass approximation or the configurational interaction of a small number of bands.[2-4] The basic assumption of this model is that actual electronic states in the vicinity of a crystalline interface may be approximated by products of some envelope functions with the Bloch functions of the relevant homogeneous crystals. For an example, let us consider a conduction electron in the vicinity of a perfect interface between two lattice-matched crystals shown in Fig. 1. Using the effective-mass approximation, the model postulates that the electronic wave function is given by

$$\psi(\mathbf{x})=\begin{cases}F_c(z)\psi_{c0}^{A}(\mathbf{x}) & \text{for } z<0\\ F_c(z)\psi_{c0}^{B}(\mathbf{x}) & \text{for } z>0\ ,\end{cases} \tag{1.1}$$

where $\psi_{c0}^{A,B}(\mathbf{x})$ are the Bloch wave functions of the homogeneous A and B crystals at $\mathbf{k}=0$. The model further assumes that, for the purpose of finding energy eigenvalues and eigenfunctions in the effective-mass approximation, the difference between ψ_{c0}^{A} and ψ_{c0}^{B} can be neglected.[2-6] The envelope function F_c is then continuous at $z=0$. Its derivative, however, is allowed to change in a stepwise fashion at $z=0$. If m_c^A and m_c^B are the effective masses in the crystal segments A and B, then it is postulated that[2,7]

$$\frac{1}{m_c^A}\frac{dF_c}{dz}(z=0^-)=\frac{1}{m_c^B}\frac{dF_c}{dz}(z=0^+)\ . \tag{1.2}$$

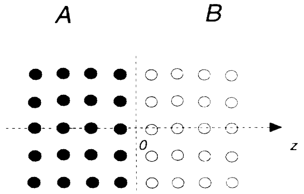

FIG. 1. A heterostructure junction.

This boundary condition may be referred to as the BenDaniel-Duke ansatz. In a structure in which the interface of Fig. 1 is repeated periodically as in superlattices, the BenDaniel-Duke ansatz is made at each interface and is supposed to conserve the probability current of the electron. However, its function is not as clear in a quantum well, where the interest is generally in bound states that do not carry current. The theoretical framework based on the assumptions outlined above may be referred to as the conventional method. Not all calculations concerning semiconductor microstructures fall within this framework. Nevertheless, one is justified in calling it conventional, since it is the most widely used method in the literature.

The conventional method suffers from several defects. Some of these have been discussed in the literature. Burt has pointed out that neglecting the difference between ψ_{n0}^{A} and ψ_{n0}^{B} may cause substantial errors, on the order of 1.0 eV, in the computation of energy eigenvalues.[8] Note also that effective masses are determined from band gaps and $\mathbf{k}\cdot\mathbf{p}$ coupling terms that are evaluated in the representations of ψ_{n0}^{A} and ψ_{n0}^{B}. These parameters take the same values on both sides of the interface, when $\psi_{n0}^{A}=\psi_{n0}^{B}$. Self-consistency requires one to neglect the difference between m_n^A and m_n^B, if the difference between ψ_{n0}^{A} and ψ_{n0}^{B} is neglected. The effective masses then drop out of ansatz (1.2).

If the difference between ψ_{n0}^{A} and ψ_{n0}^{B} is not neglected, $F_n(z)$ cannot be continuous across an interface. This follows from the fact that the Hamiltonian underlying the postulate (1.1) is given by

$$H=\frac{\mathbf{p}^2}{2m}+\theta(-z)v_A(\mathbf{x})+\theta(z)v_B(\mathbf{x})\ , \tag{1.3}$$

where m is the ordinary electron mass, $v_A(\mathbf{x})$ and $v_B(\mathbf{x})$ are the crystal potentials for homogeneous crystals of A and B, and $\theta(z)$ is the step function. The stationary Schrödinger equation corresponding to H requires ψ to be continuous at $z=0$:

$$F_c(0^-)\psi_{c0}^A(x,y,z=0^-)=F_c(0^+)\psi_{c0}^B(x,y,z=0^+)\ . \tag{1.4}$$

When ψ_{c0}^A changes into ψ_{c0}^B in a stepwise fashion, F_c must also change in a stepwise fashion. Note also that the Hamiltonian (1.3) requires $\nabla\psi$, as well as ψ, to be continuous at $z=0$. The continuity of the former leads to

$$\frac{dF_c}{dz}(0^-)\psi_{c0}^A(x,y,0^-)+F_c(0^-)\frac{d\psi_{c0}^A}{dz}(x,y,0^-) = \frac{dF_c}{dz}(0^+)\psi_{c0}^B(x,y,0^+)+F_c(0^+)\frac{d\psi_{c0}^B}{dz}(x,y,0^+)\ . \tag{1.5}$$

The boundary conditions (1.5) cannot take the simple form (1.2) for two distinct functions $\psi_{c0}^A(\mathbf{x})$ and $\psi_{c0}^B(\mathbf{x})$.

A second defect of the conventional method has been pointed out by Nag and Mukhopadhyay.[9] It is associated with the ambiguity of the effective-mass concept itself. There is more than one way to define an effective mass. If the homogeneous crystal bands are parabolic, then these various definitions coincide and one has a pair of unambiguously determined effective masses to use in (1.2). However, if one or both bands of the heterostructure are nonparabolic, then different definitions of effective masses depend on energy differently. Without an additional assumption, one does not know which effective mass should be used in (1.2).

Another difficulty with the conventional method is indicated indirectly from the comparison of its predictions with the exact solutions for a model superlattice.[10] In the model, barrier potentials are δ functions. This permits one to solve the Schrödinger equation exactly. One finds that if the effective masses are omitted from (1.2), there is significantly better agreement between the results of the exact calculation and the results of the envelope-function approximation as in (1.1).

The origin of the problems referred to above appears to be the incorrect application of the perturbation theory in the conventional method. The effective-mass approximation corresponds to a perturbative treatment of the electronic motion. When one is developing a perturbative solution, one must choose a basis. This could be the set of Bloch functions for A, or for B, but not both together. The fact that one must use the same Bloch functions on both sides of an interface has been particularly emphasized by Burt.[8] If the set of Bloch functions for A is chosen as the basis, one may expand the complete wave function as

$$\psi(\mathbf{x})=\sum_{n\mathbf{k}} C_{n\mathbf{k}}\psi_{n\mathbf{k}}^A(\mathbf{x})\ . \tag{1.6}$$

To determine $C_{n\mathbf{k}}$, one treats the difference between the crystal potentials of A and B as a perturbation. This is equivalent to rewriting the Hamiltonian of (1.3) in the form

$$H=\frac{\mathbf{p}^2}{2m}+v_A(\mathbf{x})+U_{\text{pert}}(\mathbf{x})\ , \tag{1.7}$$

where the perturbation potential is given by

$$U_{\text{pert}}(\mathbf{x})=\theta(z)[v_B(\mathbf{x})-v_A(\mathbf{x})]\ . \tag{1.8}$$

Thus one has a set of basis states $\{\psi_{n\mathbf{k}}{}^A(\mathbf{x})\}$ which is complete and well-defined throughout the entire space. Using U_{pert}, one calculates the corrections to these basis states. Of course, one does not need to choose the set $\{\psi_{n\mathbf{k}}{}^A(\mathbf{x})\}$ or the set $\{\psi_{n\mathbf{k}}{}^B(\mathbf{x})\}$ as the basis. Another complete set might be more convenient. Smith and Mailhiot[11] choose the set of the Bloch functions corresponding to a reference crystal with the potential $[v_A(\mathbf{x})+v_B(\mathbf{x})]/2$. In their theory, the differences between this reference potential and v_A and v_B are treated as perturbations.

The basic concern of the present work is the elimination of another inconsistency in the conventional method, as well as in others that use the effective-mass approximation. The effective-mass approximation corresponds to the elimination of the interband couplings represented by the $\mathbf{k}\cdot\mathbf{p}$ term in the homogeneous crystal Hamiltonian to the lowest order. If one considers the example of the coupling between one conduction band and one valence band, the relevant coupling energy is given by $\Delta E=m^{-1}\hbar k P_{cv}$, where P_{cv} is the interband matrix element of the momentum operator. Typically, $\hbar^{-1}P_{cv}\sim 10^8$ cm^{-1}. Assuming a large k value, say $k\sim 10^7$ cm^{-1}, one finds $\Delta E\sim 500$ meV. In a quantum-well geometry, the difference between the potentials of the well and the barrier crystals also produces interband coupling. For typical quantum-well structures, the corresponding energy is comparable to ΔE. Therefore, the interband coupling induced by the quantum-well potential must be eliminated simultaneously with the $\mathbf{k}\cdot\mathbf{p}$ coupling. To eliminate one without the other is not a self-consistent application of the effective-mass approximation.

To treat the interband couplings on an equal basis, one can use the method of approximate canonical transformations in a Luttinger-Kohn representation.[1] This method is used in Sec. II to derive an effective band Hamiltonian and a set of boundary conditions on envelope functions at the interfaces of the well. One welcome by-product of the method is the transformation of the stationary Schrödinger equation to a Sturm-Liouville problem. The boundary conditions are thus determined by the effective Hamiltonian itself. The main results of the analysis may be summarized as follows.

(a) For a given band, there is a position-independent effective mass in and around a quantum well. This effective mass depends on the band lineup energies of the crystals of A and B, as well as on the coherence length within which a single electron wave function may be defined.

(b) The effective Hamiltonian has δ-function potentials. These require discontinuities in the derivative of an envelope function across the quantum-well interfaces. The stepwise jumps of the derivative are proportional to a component of a complex vector quantity designated by $\mathbf{d}_{nn}$. $\mathbf{d}_{nn}$ depends on the interband matrix elements of the momentum operator, as well as on the ratios of the band lineup energies to the band gaps. The effective Hamiltonian has, of course, the usual step function potential that depends on the band lineup energies.

(c) The effective Hamiltonian has a velocity-dependent term whose coefficient is determined by the real part of

$\mathbf{d}_{nn}$. This velocity-dependant term is present only inside the well and modifies energy dispersions in the transverse directions.

(d) Unlike the solutions of the conventional method, the eigenstates of the present effective Hamiltonian do not split into even and odd functions. There is only one eigenvalue equation.

(e) Bound-state eigenfunctions have certain propagating phase factors within the well which can have an observable effect on electronic transition matrix elements.

These results, as well as the range of validity for the effective Hamiltonian given in (2.66), are discussed in greater detail in Sec. III.

II. CANONICAL TRANSFORMATION

In this section, I use the original Luttinger-Kohn method to derive the effective band Hamiltonian for an electron in and around a quantum well. The method is based on canonical transformations in the Luttinger-Kohn representation for eliminating the coupling between the chosen band and the other bands to some approximate order. Because of the quantum well, the electron acquires a two-dimensional character that is manifested through the well-localized states, as well as through the dependence of the effective Hamiltonian on the ratio $L/\mathcal{L}$, where L is the well width and $\mathcal{L}$ is the coherence length.

For the basis states, the derivation uses the Luttinger-Kohn states corresponding to the center of the Brillouin zone (BZ) of the homogeneous crystal of A. These states are given by[1]

$$\chi_{n\mathbf{k}}(\mathbf{x})=\frac{1}{(\mathcal{L}_1\mathcal{L}_2\mathcal{L}_3)^{1/2}}e^{i\mathbf{k}\cdot\mathbf{x}}u^A_{n0}(\mathbf{x})$$

$$=\frac{1}{(\mathcal{L}_1\mathcal{L}_2\mathcal{L}_3)^{1/2}}e^{i\mathbf{k}\cdot\mathbf{x}}\sum_{\mathbf{G}}e^{i\mathbf{G}\cdot\mathbf{x}}\phi^A_n(\mathbf{G}) \ . \tag{2.1}$$

Here $\mathcal{L}_1\mathcal{L}_2\mathcal{L}_3$ is the quantization volume, $\{\mathbf{G}\}$ is the set of reciprocal lattice vectors, and ϕ^A_n are the momentum Bloch functions. $\chi_{n\mathbf{k}}$ are orthornormal:

$$\int d\mathbf{x}\,\chi^*_{n\mathbf{k}}(\mathbf{x})\chi_{n'\mathbf{k}'}(\mathbf{x})=\delta_{nn'}\delta(\mathbf{k},\mathbf{k}') \ , \tag{2.2}$$

where $\delta(\mathbf{k},\mathbf{k}')$ is related to the Dirac δ function by

$$\delta(\mathbf{k},\mathbf{k}')=\frac{(2\pi)^3}{\mathcal{L}_1\mathcal{L}_2\mathcal{L}_3}\delta(\mathbf{k}-\mathbf{k}') \ . \tag{2.3}$$

The set $\{\chi_{n\mathbf{k}}\}$ is complete:

$$\sum_{n\mathbf{k}}\chi_{n\mathbf{k}}(\mathbf{x})\chi^*_{n\mathbf{k}}(\mathbf{x}')=\delta(\mathbf{x}-\mathbf{x}') \ . \tag{2.4}$$

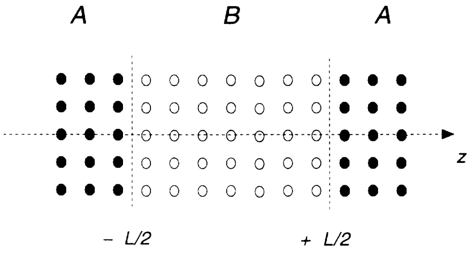

FIG. 2. A semiconductor quantum well.

Let us consider the Hamiltonian for the quantum well illustrated in Fig. 2:

$$H=H_A+W \ , \tag{2.5a}$$

where

$$H_A=\frac{\mathbf{p}^2}{2m}+v_A(\mathbf{x}) \ , \tag{2.5b}$$

$$W=\epsilon(z)[v_B(\mathbf{x})-v_A(\mathbf{x})] \ , \tag{2.5c}$$

$$\epsilon(z)=\theta\left[z+\frac{L}{2}\right]-\theta\left[z-\frac{L}{2}\right] \ . \tag{2.5d}$$

Equation (2.5c) assumes that the interfaces are perfect planes and that the potential is determined completely by the atomic layers. Under this assumption, the ratio of L to a, the lattice constant of the homogeneous crystal, is an odd or even integer. In Fig. 2, it is assumed to be odd:

$$L=(2l_0+1)a \ , \tag{2.6}$$

where l_0 is a positive integer. W given by (2.5c) may be referred to as a structural well potential.

Actual quantum-well potentials do not need to be structural. They may depend on the coupling among electrons and on the distribution of electrons over various states. An individual electron may therefore move in a potential that is produced dynamically in the vicinity of interfaces. Tersoff theory of heterostructure junction,[12] for example, attributes a dynamical origin to a barrier potential. Nevertheless, (2.5c) provides a good starting point, since the states produced by a structural W yield reasonable transition spectra. One may use these states as the basis to formulate electronic many-body interactions.

Using (2.1), one finds that the matrix elements of H_A are given by

$$\langle n\mathbf{k}|H_A|n'\mathbf{k}'\rangle=\left[\frac{\hbar^2k^2}{2m}+E^A_n(0)\right]\delta_{nn'}\delta(\mathbf{k},\mathbf{k}')$$

$$+\frac{\hbar\mathbf{k}\cdot\mathbf{P}^A_{nn'}}{m}\delta(\mathbf{k},\mathbf{k}') \ , \tag{2.7a}$$

where

$$\mathbf{P}^A_{nn'}=\hbar\sum_{\mathbf{G}}\mathbf{G}\phi^{A*}_n(\mathbf{G})\phi^A_{n'}(\mathbf{G}) \ . \tag{2.7b}$$

Similarly, the matrix elements of W are given by

$$\langle n\mathbf{k}|W|n'\mathbf{k}'\rangle=\frac{2}{\mathcal{L}_3}\delta_\perp(\mathbf{k},\mathbf{k}')\sum_{\mathbf{G}\mathbf{G}'G_3''}\phi_n^{A*}(\mathbf{G})\phi_{n'}^{A}(\mathbf{G}')[V_B(\mathbf{G}_\perp-\mathbf{G}_\perp'+\hat{\mathbf{z}}G_3'')-V_A(\mathbf{G}_\perp-\mathbf{G}_\perp'+\hat{\mathbf{z}}G_3'')]$$

$$\times\frac{\sin\left[\frac{L}{2}(k_3-k_3'+G_3-G_3'-G_3'')\right]}{(k_3-k_3+G_3-G_3'-G_3'')-i\delta}\;. \tag{2.8}$$

Here

$$\delta_\perp(\mathbf{k},\mathbf{k}')=\delta(k_1,k_1')\delta(k_2,k_2')\;, \tag{2.9}$$

and $V_{A,B}(\mathbf{G})$ are the Fourier coefficients of the periodic crystal potentials:

$$v_{A,B}(\mathbf{x})=\sum_{\mathbf{G}}V_{A,B}(\mathbf{G})e^{i\mathbf{G}\cdot\mathbf{x}}\;. \tag{2.10}$$

The matrix elements of the full Hamiltonian are given by the sum of (2.7a) and (2.8):

$$\langle n\mathbf{k}|H|n'\mathbf{k}'\rangle=\langle n\mathbf{k}|H_A|n'\mathbf{k}'\rangle+\langle n\mathbf{k}|W|n'\mathbf{k}'\rangle\;. \tag{2.11}$$

Comparing (2.8) with (2.7a), one sees that different bands are coupled not only by the $\mathbf{k}\cdot\mathbf{p}$ term of (2.7a), but also by W. The effective-mass approximation for an isolated band corresponds to the removal of these interband couplings from the Hamiltonian in the lowest order. Before proceeding with the discussion of the canonical transformation that removes the interband couplings to first order, let us discuss some features of $\langle n\mathbf{k}|W|n'\mathbf{k}'\rangle$.

Let us set $k_3=k_3'$ in (2.8). This means that $\mathbf{k}=\mathbf{k}'$, since the transverse wave-vector components are required to be equal by the transverse δ function. The sine function approaches zero as $\mathbf{k}\to\mathbf{k}'$:

$$\sin\left[\frac{L}{2}(G_3-G_3'-G_3'')\right]=\sin[\pi(2l_0+1)\iota]=0\;, \tag{2.12}$$

where ι is some integer. Clearly, the limit $\mathbf{k}\to\mathbf{k}'$ yields a finite result only if

$$G_3-G_3'=G_3''\;. \tag{2.13}$$

Then,

$$\lim_{\substack{k_3\to k_3'\\ \delta\to 0}}\frac{\sin\left[\frac{L}{2}(k_3-k_3')\right]}{k_3-k_3'-i\delta}=\frac{L}{2} \tag{2.14}$$

and

$$\langle n\mathbf{k}|W|n'\mathbf{k}\rangle=\frac{L}{\mathcal{L}_3}\sum_{\mathbf{G}\mathbf{G}'}\phi_n^{A*}(\mathbf{G})\phi_{n'}^{A}(\mathbf{G}')\times[V_B(\mathbf{G}-\mathbf{G}')-V_A(\mathbf{G}-\mathbf{G}')]$$

$$=\langle n0|W|n'0\rangle\equiv W_{nn'}\;. \tag{2.15}$$

Thus $\langle n\mathbf{k}|W|n'\mathbf{k}\rangle$ is independent of $\mathbf{k}$. It does, however, depend on the ratio $L/\mathcal{L}_3$. This forces us to reconsider the meaning of the quantization volume.

At first glance, the size of $\mathcal{L}_3$ in (2.15) appears to be arbitrary. If one lets the crystal sample, as well as the quantization volume, be infinite, then $\mathcal{L}_3\to\infty$ and $W_{nn'}\to 0$ (as long as L is finite). This is to be expected, since in an infinitely large three-dimensional crystal, the contribution of a thin sheet of material to the total energy of the sample will be infinitesimally small. However, the choice of quantization volume is not arbitrary. A single-particle Hamiltonian such as (2.5a) can describe the dynamics of the electron with reasonable accuracy only over a finite range. Over this range, the electron has a wave function whose phase and amplitude are well-defined functions of position. Beyond this range, the electron suffers many elastic and inelastic scatterings, and the phase, as well as the amplitude, of its wave function becomes indefinite. Therefore, there exists a coherence length $\mathcal{L}$ for the electron such that the electron has a well-defined wave function within the volume $\mathcal{L}^3$. This volume may be called the coherence volume. In the interior of the coherence volume, the evolution of the electron is determined by H in (2.5a). At the boundaries of the coherence volume and beyond, the evolution of the electron is not determined by H. The wave function must be specified independently of H at the boundaries of the coherence volume. Thus the quantization volume must be identified with a volume which falls within the coherence volume. The coherence length $\mathcal{L}$ determines the maximum size for the quantization volume. For the remainder of the discussion, I will assume that the three directions are equivalent and set

$$\mathcal{L}_1=\mathcal{L}_2=\mathcal{L}_3=\mathcal{L}\;. \tag{2.16}$$

For a true quantum well, the ratio of L to $\mathcal{L}$ must be less than one:

$$0<\frac{L}{\mathcal{L}}<1\;. \tag{2.17}$$

This can be seen from the following consideration. Suppose that the ratio is larger than unity and that the electron is in the middle of the well. Because its coherence length is less than the width of the well, the electron sees only the potential of the B-type crystal. The electron's evolution is thus determined by

$$H_B=\frac{\mathbf{p}^2}{2m}+v_B(\mathbf{x})\;, \tag{2.18}$$

not by H in (2.5a). The electron does not see the boundaries of the well and therefore one does not have a quantum well. The electronic states correspond to those in the homogeneous B crystal.

The absorption and emission spectrum of a particular sample may gain considerable complexity if $L/\mathcal{L}$ approaches 1. Indeed, if $L/\mathcal{L}$ is near 1, one may observe three types of spectral lines in electronic transitions. If the electron is far to the left of the well so that its Hamiltonian is H_A, then its spectrum corresponds to the band spectrum of the homogeneous crystal of A. If the electron is close to the left A/B interface but cannot discern the interface on the right, then its Hamiltonian is the single-interface Hamiltonian (1.3). Its spectrum will contain transitions which depend on the interface states. Finally, if the electron is in the quantum well, then its Hamiltonian is given by (2.5a). Its spectrum will have lines that correspond to transitions between quantum-well localized states.

The electronic coherence length in typical semiconductors is in the range 500–1000 Å. The upper part of this range is achieved in pure crystals at low temperatures. The coherence length may vary not only from one material to another, but also from one device geometry to another. In the rest of this paper, I will treat $\mathcal{L}$ as a parameter which is determined individually for each sample.

The quantity $(\mathcal{L} W_{nn}/L)$ is related to the lineup energy between the bands of A and B. One may see this as follows. $V(\mathbf{G})$ decreases rapidly for finite $\mathbf{G}$.[13] Hence the sum over the reciprocal lattice vectors in W_{nn} may be restricted to the zero and to the three basis vectors of the reciprocal lattice (designated by $\mathbf{g}$):

$$\begin{aligned}\frac{\mathcal{L}}{L} W_{nn} \simeq {} & [V_B(0)-V_A(0)] \sum_{\mathbf{G}} \phi_n^{A*}(\mathbf{G})\phi_n^A(\mathbf{G}) \\ & + \sum_{\mathbf{g}} \{\phi_n^{A*}(\mathbf{g})\phi_n^A(0)[V_B(\mathbf{g})-V_A(\mathbf{g})] \\ & + \phi_n^{A*}(0)\phi_n^A(\mathbf{g})[V_B(-\mathbf{g})-V_A(-\mathbf{g})]\} \; . \end{aligned} \tag{2.19}$$

The sum in the first line of (2.19) equals one, since the momentum Bloch functions are orthonormal:[14]

$$\sum_{\mathbf{G}} \phi_n^*(\mathbf{G})\phi_{n'}(\mathbf{G}) = \delta_{nn'} \; . \tag{2.20}$$

The coefficients $V(\mathbf{G})$ are related to the band energies at the center of the BZ:[14]

$$V(\mathbf{G}) = \sum_n E_n(0)\phi_n^*(-\mathbf{G})\phi_n(0) \; . \tag{2.21}$$

If this relation is inverted using (2.20), one finds

$$\sum_{\mathbf{G}} \phi_n(\mathbf{G})V(-\mathbf{G}) = E_n(0)\phi_n(0) \; . \tag{2.22}$$

Restricting the sum in (2.22) to $\{0,\mathbf{g}\}$, one has

$$[E_n(0)-V(0)]\phi_n(0) \simeq \sum_{\mathbf{g}} \phi_n(\mathbf{g})V(-\mathbf{g}) \; . \tag{2.23}$$

Let us assume that ϕ_n^A and ϕ_n^B are sufficiently similar so that

$$\frac{\phi_n^B(\mathbf{g})}{\phi_n^A(\mathbf{g})} \simeq \frac{\phi_n^B(0)}{\phi_n^A(0)} \; . \tag{2.24}$$

With these assumptions, (2.19) becomes

$$\begin{aligned}\frac{\mathcal{L}}{L} W_{nn} \simeq {} & [E_n^B(0)-E_n^A(0)]|\phi_n^A(0)|^2 \\ & + [V_B(0)-V_A(0)][1-|\phi_n^A(0)|^{2]} \; . \end{aligned} \tag{2.25}$$

Finally, it is reasonable to expect ϕ_n to decrease rapidly for finite $\mathbf{G}$ just like V. This implies that the normalization condition on the momentum Bloch functions is essentially exhausted by $\phi_n(0)$, that is,

$$|\phi_n^A(0)|^2 \approx 1 \; . \tag{2.26}$$

Thus

$$W_{nn} \approx \frac{L}{\mathcal{L}}[E_n^B(0)-E_n^A(0)] \; . \tag{2.27}$$

Let us consider an $\mathrm{Al}_{1-x}\mathrm{Ga}_x\mathrm{As}/\mathrm{GaAs}$ structure and let the band lineup energies be on the order of 0.2–0.4 eV. If $L \sim 100$ Å and $\mathcal{L} \sim 400$ Å, then $W_{nn} \sim 0.05$–0.1 eV.

It is necessary to separate the interband coupling terms from the intraband terms in (2.5a) for the canonical transformation. Define the Hamiltonians H_0, H_1, and H_2, such that

$$H = H_0 + H_1 + H_2 \; , \tag{2.28a}$$

$$\langle n\mathbf{k}|H_2|n'\mathbf{k}'\rangle = \delta_{nn'}[1-\delta(k_3,k_3')]\langle n\mathbf{k}|W|n\mathbf{k}'\rangle \; . \tag{2.28b}$$

$$\begin{aligned}\langle n\mathbf{k}|H_1|n'\mathbf{k}'\rangle = (1-\delta_{nn'}) \Bigg[& \langle n\mathbf{k}|W|n'\mathbf{k}'\rangle \\ & + \delta(\mathbf{k},\mathbf{k}')\frac{\hbar\mathbf{k}\cdot\mathbf{P}_{nn'}^A}{m} \Bigg] \; , \end{aligned} \tag{2.28c}$$

$$\langle n\mathbf{k}|H_2|n'\mathbf{k}'\rangle = \delta_{nn'}[1-\delta(k_3,k_3')]\langle n\mathbf{k}|W|n\mathbf{k}'\rangle \; . \tag{2.28d}$$

H_0 contains the diagonal terms. H_1 contains the interband coupling terms. H_2 has intraband terms but is off-diagonal in wave vectors. The effective-mass approximation corresponds to the elimination of H_1 from H by means of a canonical transformation to first order in H_1. Let us define

$$H' = e^{-iS}He^{iS} \tag{2.29a}$$

and choose S such that

$$i[H_0,S] = -H_1 \; . \tag{2.29b}$$

This eliminates H_1 from H to first order in H_1:

$$H' = H_0 + H_2 + \frac{i}{2}[H_1,S] + i[H_2,S] + O(W^3) \; . \tag{2.30}$$

The effective Hamiltonian for band n is given by

$$\begin{aligned}\langle n\mathbf{k}|H_{\mathrm{eff}}|n\mathbf{k}'\rangle = \langle n\mathbf{k}| \Bigg\{ & H_0 + H_2 + \frac{i}{2}[H_1,S] \\ & + i[H_2,S] \Bigg\} |n\mathbf{k}'\rangle \; . \end{aligned} \tag{2.31}$$

From (2.28b) and (2.28c), one readily finds that

$$\langle n\mathbf{k}|S|n'\mathbf{k}'\rangle = \frac{i\langle n\mathbf{k}|H_1|n'\mathbf{k}'\rangle}{[E_n^0(\mathbf{k})-E_{n'}^0(\mathbf{k})]} \; , \tag{2.32a}$$

where

$$E_n^0(\mathbf{k}) \equiv E_n^A(0) + \frac{\hbar^2 \mathbf{k}^2}{2m} + W_{nn} \; . \tag{2.32b}$$

The third and fourth terms in H' are therefore given by

$$\begin{aligned} &\frac{i}{2}\langle n\mathbf{k}|[H_1,S]|n'\mathbf{k}'\rangle \\ &\quad = -\tfrac{1}{2}\sum_{n''\mathbf{k}''} \langle n\mathbf{k}|H_1|n''\mathbf{k}''\rangle\langle n''\mathbf{k}''|H_1|n'\mathbf{k}'\rangle \\ &\qquad \times \left\{ \frac{1}{E_{n''}^0(\mathbf{k}'') - E_{n'}^0(\mathbf{k}')} + \frac{1}{E_{n''}^0(\mathbf{k}'') - E_n^0(\mathbf{k})} \right\} \end{aligned} \tag{2.33}$$

and

$$\begin{aligned} &i\langle n\mathbf{k}|[H_2,S]|n'\mathbf{k}'\rangle \\ &\quad = -\sum_{n''\mathbf{k}''} \left[\frac{\langle n\mathbf{k}|H_2|n''\mathbf{k}''\rangle\langle n''\mathbf{k}''|H_1|n'\mathbf{k}'\rangle}{E_{n''}^0(\mathbf{k}'') - E_{n'}^0(\mathbf{k}')} + \frac{\langle n\mathbf{k}|H_1|n''\mathbf{k}''\rangle\langle n''\mathbf{k}''|H_2|n'\mathbf{k}'\rangle}{E_{n''}^0(\mathbf{k}'') - E_n^0(\mathbf{k})} \right] . \end{aligned} \tag{2.34}$$

Let us consider (2.33). When (2.28c) is substituted into (2.33), one finds three groups of terms:

$$V^{(1)}_{n\mathbf{k};n'k'} = -\frac{\hbar^2}{2m^2}\delta(\mathbf{k},\mathbf{k}') \sum_{n''(n''\neq n,\, n''\neq n')} \left\{ \frac{\mathbf{k}\cdot\mathbf{P}^A_{nn''}\mathbf{k}\cdot\mathbf{P}^A_{n''n'}[2E^0_{n''}(0) - E^0_n(0) - E^0_{n'}(0)]}{[E^0_{n''}(0) - E^0_n(0)][E^0_{n''}(0) - E^0_{n'}(0)]} \right\} , \tag{2.35}$$

$$\begin{aligned} V^{(2)}_{n\mathbf{k};n'\mathbf{k}'} = -\frac{\hbar}{2m} \sum_{n''(n''\neq n,\, n''\neq n')} &\left\{ \frac{\mathbf{k}\cdot\mathbf{P}^A_{nn''}\langle n''\mathbf{k}|W|n'\mathbf{k}'\rangle[E^0_{n''}(\mathbf{k}) + E^0_{n''}(0) - E^0_n(0) - E^0_{n'}(\mathbf{k}')]}{[E^0_{n''}(0) - E^0_n(0)][E^0_{n''}(\mathbf{k}) - E^0_{n'}(\mathbf{k}')]} \right. \\ &\left. + \frac{\mathbf{k}'\cdot\mathbf{P}^A_{n''n'}\langle n\mathbf{k}|W|n''\mathbf{k}'\rangle[E^0_{n''}(\mathbf{k}') + E^0_{n''}(0) - E^0_n(\mathbf{k}) - E^0_{n'}(0)]}{[E^0_{n''}(0) - E^0_{n'}(0)][E^0_{n''}(\mathbf{k}') - E^0_n(\mathbf{k})]} \right\} , \end{aligned} \tag{2.36}$$

$$V^{(3)}_{n\mathbf{k};n'\mathbf{k}'} = -\tfrac{1}{2} \sum_{n''\mathbf{k}''(n''\neq n,\, n''\neq n')} \langle n\mathbf{k}|W|n''\mathbf{k}''\rangle\langle n''\mathbf{k}''|W|n'\mathbf{k}'\rangle \left[\frac{1}{E^0_{n''}(\mathbf{k}'') - E^0_n(\mathbf{k})} + \frac{1}{E^0_{n''}(\mathbf{k}'') - E^0_{n'}(\mathbf{k}')} \right] . \tag{2.37}$$

Note that $V^{(1)}$ is diagonal in the momentum space. The intraband element $V^{(1)}_{n\mathbf{k};n\mathbf{k}}$ adds to the k^2 term in H_0 to give the inverse effective-mass tensor of the band n in and around the quantum well:

$$\left[\frac{1}{m_n}\right]_{ij} = \frac{1}{m}\delta_{ij} - \frac{2}{m^2} \sum_{\substack{n'' \\ (n''\neq n)}} \frac{P^{Ai}_{nn''}P^{Aj}_{n''n}}{E^A_{n''}(0) + W_{n''n''} - E^A_n(0) - W_{nn}} \; . \tag{2.38}$$

If W_{nn} and $W_{n''n''}$ are omitted, (2.38) is the effective-mass equation for the homogeneous crystal of A:

$$\left[\frac{1}{m^A_n}\right]_{ij} = \frac{1}{m}\delta_{ij} - \frac{2}{m^2} \sum_{\substack{n'' \\ (n''\neq n)}} \frac{P^{Ai}_{nn''}P^{Aj}_{n''n}}{E^A_{n''}(0) - E^A_n(0)} \; . \tag{2.39}$$

In contrast to $V^{(1)}$, $V^{(2)}$ and $V^{(3)}$ are not diagonal in the momentum space. The intraband matrix elements $V^{(2)}_{n\mathbf{k};n\mathbf{k}'}$ and $V^{(3)}_{n\mathbf{k};n\mathbf{k}'}$ yield the potential that may confine the electron to the well. To determine the form of this potential in ordinary space, let us consider the effect of $V^{(2)}_{n\mathbf{k};n\mathbf{k}'}$ and $V^{(3)}_{n\mathbf{k};n\mathbf{k}'}$ on an arbitrary function $f(\mathbf{x})$ in the Luttinger-Kohn representation.

Let us first expand $f(\mathbf{x})$ in the Luttinger-Kohn basis:

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{\mathcal{L}^{3/2}} \sum_{n\mathbf{k}} F_n(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} u_{n0}(\mathbf{x}) \\ &= \frac{1}{\mathcal{L}^{3/2}} \sum_{n\mathbf{k}\mathbf{G}} F_n(\mathbf{k}) e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{x}} \phi^A_n(\mathbf{G}) \; . \end{aligned} \tag{2.40}$$

The expansion coefficient $F_n(\mathbf{k})$ is related to the Fourier transform of $f(\mathbf{x})$. This Fourier transform is given by

$$F(\mathbf{p}) = \frac{1}{\mathcal{L}^{3/2}} \int d\mathbf{x}\, e^{-i\mathbf{p}\cdot\mathbf{x}} f(\mathbf{x}) \; . \tag{2.41}$$

Substituting (2.40) into (2.41) and performing the spatial integral, one finds

$$F(\mathbf{p}) = \left[\frac{2\pi}{\mathcal{L}}\right]^3 \sum_{n\mathbf{k}'\mathbf{G}'} \delta(\mathbf{p} - \mathbf{k}' - \mathbf{G}') F_n(\mathbf{k}') \phi^A_n(\mathbf{G}') \; . \tag{2.42}$$

Let

$$\mathbf{p} = \mathbf{k} + G, \quad \mathbf{k} \in \mathrm{BZ} \; . \tag{2.43}$$

Since $\mathbf{k}'$, as well as $\mathbf{k}$, is confined to the BZ,

$$\delta(\mathbf{p} - \mathbf{k}' - \mathbf{G}') = \delta_{\mathbf{G}\mathbf{G}'}\delta(\mathbf{k} - \mathbf{k}') \; . \tag{2.44}$$

Therefore,

$$F(\mathbf{k} + \mathbf{G}) = \sum_n F_n(\mathbf{k}) \phi^A_n(\mathbf{G}) \; . \tag{2.45}$$

Using the completeness of the momentum Bloch functions,[14]

$$\sum_{n} \phi_n^*(\mathbf{G})\phi_n(\mathbf{G}')=\delta_{\mathbf{G}\mathbf{G}'}\ , \tag{2.46}$$

one may invert (2.45) to find

$$F_n(\mathbf{k})=\sum_{\mathbf{G}} \phi_n^{A*}(\mathbf{G})F(\mathbf{k}+\mathbf{G})\ . \tag{2.47}$$

Next, let us consider the operation of $V^{(2)}$ on f in the Luttinger-Kohn representation:

$$I_n^{(2)}(\mathbf{k})=\sum_{n'\mathbf{k}'} V^{(2)}_{n\mathbf{k};n'\mathbf{k}'}F_{n'}(\mathbf{k}')\ . \tag{2.48}$$

The matrix elements $\langle n''\mathbf{k}|W|n'\mathbf{k}'\rangle$ and $\langle n\mathbf{k}|W|n''\mathbf{k}'\rangle$ in $V^{(2)}_{n\mathbf{k};n'\mathbf{k}'}$ have the sine factor of the form

$$\sin\left[\frac{L}{2}(k_3-k_3'+G_3-G_3'-G_3'')\right]=(-1)^l\sin\left[\frac{L}{2}(k_3-k_3')\right]\ . \tag{2.49}$$

This function is significantly different from zero only within the range $\Delta k\sim\pi/L$. After setting the transverse wave-vector components equal, the denominators in (2.36) have the form

$$E^0_{n''}(\mathbf{k}')-E^0_n(\mathbf{k})=E^0_{n''}(0)-E^0_n(0)+\frac{\hbar^2}{2m}(k_3'^2-k_3^2)\ . \tag{2.50a}$$

The kinetic part of this expression has the range

$$\frac{\hbar^2}{2m}(k_3'^2-k_3^2)\simeq\frac{\hbar^2}{m}k_3(k_3'-k_3)<\frac{\hbar^2\pi^2}{maL}\equiv\delta E\ . \tag{2.50b}$$

$\delta E\sim0.1$ eV for $a\sim5$ Å and $L\sim100$ Å. Let us assume that the narrowest band gap is on the order of 1 eV, hence substantial compared to δE. The wavelength dependence of the denominators in (2.36) may then be neglected. One can also approximate the sums over the reciprocal lattice vectors in $\langle n\mathbf{k}|W|n'\mathbf{k}'\rangle$ as follows:

$$\sum_{G_3} Z(G_3)\frac{2\sin\left[\frac{L}{2}(k_3-G_3)\right]}{\mathcal{L}(k_3-G_3-i\delta)}\simeq Z(0)\frac{2\sin\left[\frac{L}{2}k_3\right]}{\mathcal{L}(k_3-i\delta)}\ . \tag{2.51}$$

The right-hand side follows from the fact that, for finite G_3, the amplitude of the sine factor obeys the inequality

$$\frac{2}{\mathcal{L}G_3}<\frac{a}{\mathcal{L}\pi}\sim1.6\times10^{-3}\ \text{ for } \mathcal{L}\sim10^3\ \text{Å}\ . \tag{2.52}$$

Using (2.46) and (2.51), one then finds

$$I_n^{(2)}(\mathbf{k})\simeq\sum_{n'\mathbf{k}'}\delta_\perp(\mathbf{k},\mathbf{k}')\left\{\frac{2\sin\left[\frac{L}{2}(k_3-k_3')\right]}{\mathcal{L}(k_3-k_3'-i\delta)}\right\}\times[\mathbf{d}_{nn'}\cdot\mathbf{k}+d^*_{nn'}\cdot\mathbf{k}']F_{n'}(\mathbf{k}')\ , \tag{2.53}$$

where

$$\mathbf{d}_{nn'}=-\frac{\hbar}{2m}\sum_{n''(n''\neq n,n''\neq n')\mathbf{G}\mathbf{G}'}\mathbf{P}^A_{nn''}\phi^{A*}_{n''}(\mathbf{G})\phi^A_{n'}(\mathbf{G}')[V_B(\mathbf{G}-\mathbf{G}')-V_A(\mathbf{G}-\mathbf{G}')]\times\left[\frac{1}{E^0_{n''}(0)-E^0_n(0)}+\frac{1}{E^0_{n''}(0)-E^0_{n'}(0)}\right]\ . \tag{2.54}$$

If $f(\mathbf{x})$ is slowly varying over many lattice sites, then its spectral components for reciprocal lattice vectors are small compared to the components for the BZ:

$$|F(\mathbf{k}+\mathbf{G})|\ll|F(\mathbf{k})|\quad\text{for finite } \mathbf{G}\ . \tag{2.55}$$

Indeed, (2.55) may be taken as the definition of a slowly varying function. Taking only $\mathbf{G}=0$ term in the sum in (2.47), one has

$$F_n(\mathbf{k})\simeq\phi_n^{A*}(0)F(\mathbf{k})\ . \tag{2.56}$$

If (2.56) is substituted into (2.53), $I_n^{(2)}(\mathbf{k})$ takes the form of a convolution integral:

$$I_n^{(2)}(\mathbf{k})\simeq\sum_{\mathbf{k}'}\delta_\perp(\mathbf{k},\mathbf{k}')\left\{\frac{2\sin\left[\frac{L}{2}(k_3-k_3')\right]}{\mathcal{L}(k_3-k_3'-i\delta)}\right\}\times[\mathbf{d}_n\cdot\mathbf{k}+d_n'\cdot\mathbf{k}']F(\mathbf{k}')\ , \tag{2.57}$$

where

$$\mathbf{d}_n=\sum_{n'}\mathbf{d}_{nn'}\phi^{A*}_{n'}(0),\quad \mathbf{d}_n'=\sum_{n'}\mathbf{d}^*_{nn'}\phi^{A*}_{n'}(0)\ . \tag{2.58}$$

To treat (2.57) as a convolution, the sum over $\mathbf{k}'$ must be extended outside the BZ. This is permissible, since the contribution of the higher zones to the integral is negligibly small. Note that the approximations that led to (2.53) simplify $\langle n\mathbf{k}|W|n\mathbf{k}'\rangle$ to

$$\langle n\mathbf{k}|W|n\mathbf{k}'\rangle=W_{nn}\delta_\perp(\mathbf{k},\mathbf{k}')\left\{\frac{2\sin\left[\frac{L}{2}(k_3-k_3')\right]}{\mathcal{L}(k_3-k_3'-i\delta)}\right\}\ . \tag{2.59}$$

One can deduce the effect of $V^{(3)}$ on f in a similar manner:

$$I_n^{(3)}(\mathbf{k})=\sum_{n'\mathbf{k}'}V^{(3)}_{n\mathbf{k};n'\mathbf{k}'}F_{n'}(\mathbf{k}')\ . \tag{2.60}$$

Using

$$\frac{\mathcal{L}}{2\pi}\int_{-\infty}^{+\infty}dk''\frac{4\sin\left[\frac{L}{2}(k-k'')\right]\sin\left[\frac{L}{2}(k''-k')\right]}{\mathcal{L}^2(k-k''-i\delta)(k''-k'-i\delta)}=\frac{2\sin\left[\frac{L}{2}(k-k')\right]}{\mathcal{L}(k-k'-2i\delta)}\ , \qquad (2.61)$$

and making the same approximations that led to (2.53), one finds

$$I_n^{(3)}(\mathbf{k})\simeq\sum_{n'\mathbf{k}'}\delta_\perp(\mathbf{k},\mathbf{k}')\times\left\{\frac{2\sin\left[\frac{L}{2}(k_3-k_3')\right]}{\mathcal{L}(k_3-k_3'-2i\delta)}\right\}U_{nn'}F_{n'}(\mathbf{k}')\ , \qquad (2.62)$$

where

$$U_{nn'}=-\tfrac{1}{2}\sum_{n''(n''\neq n,n''\neq n')\mathbf{G}_1\mathbf{G}_1'\mathbf{G}_2\mathbf{G}_2'}\phi_n^{A*}(\mathbf{G}_1)\phi_{n''}^{A}(\mathbf{G}_1')\phi_{n''}^{A*}(\mathbf{G}_2)\phi_{n'}^{A}(\mathbf{G}_2')$$
$$\times[V_B(\mathbf{G}_1-\mathbf{G}_1')-V_A(\mathbf{G}_1-\mathbf{G}_1')][V_B(\mathbf{G}_2-\mathbf{G}_2')-V_A(\mathbf{G}_2-\mathbf{G}_2')]$$
$$\times\left[\frac{1}{E_{n''}^0(0)-E_n^0(0)}+\frac{1}{E_{n''}^0(0)-E_{n'}^0(0)}\right]. \qquad (2.63)$$

Finally, let us consider the commutator of H_2 with S. It is clear from (2.28c) and (2.28d) that the expression (2.34) yields only an interband correction term:

$$\langle n\mathbf{k}|[H_2,S]|n'\mathbf{k}'\rangle=i(1-\delta_{nn'})\sum_{n''\mathbf{k}''}\left\{\frac{[1-\delta(k_3,k_3'')]\langle n\mathbf{k}|W|n\mathbf{k}''\rangle}{E_n^0(\mathbf{k}'')-E_{n'}^0(\mathbf{k}')}\left[\langle n\mathbf{k}''|W|n'\mathbf{k}'\rangle+\delta(\mathbf{k}'',\mathbf{k}')\frac{\hbar\mathbf{k}'\cdot\mathbf{P}_{\mathrm{nn}'}^A}{m}\right]\right.$$
$$\left.+\frac{[1-\delta(k_3',k_3'')]\langle n'\mathbf{k}''|W|n'\mathbf{k}'\rangle}{E_{n'}^0(\mathbf{k}'')-E_n^0(\mathbf{k})}\left[\langle n\mathbf{k}|W|n'\mathbf{k}''\rangle+\delta(\mathbf{k},\mathbf{k}'')\frac{\hbar\mathbf{k}\cdot\mathbf{P}_{\mathrm{nn}'}^A}{m}\right]\right\}. \qquad (2.64)$$

Thus this commutator does not contribute to the effective-mass Hamiltonian defined by (2.31).

Using (2.38), (2.53), (2.59), and (2.62), the effective band Hamiltonian (2.31) can be written in the explicit form:

$$\langle n\mathbf{k}|H_{\mathrm{eff}}|n\mathbf{k}'\rangle=\left[E_n^A(0)+\frac{\hbar^2\mathbf{k}^2}{2m_n}\right]\delta(\mathbf{k},\mathbf{k}')$$
$$+\delta_\perp(\mathbf{k},\mathbf{k}')\frac{2\sin\left[\frac{L}{2}(k_3-k_3')\right]}{\mathcal{L}(k_3-k_3'-i\delta)}\times\left[\frac{\mathcal{L}}{L}W_{nn}+U_{nn}+\mathbf{d}_{\mathrm{nn}}\cdot\mathbf{k}+\mathbf{d}_{\mathrm{nn}}^*\cdot\mathbf{k}'\right]. \qquad (2.65)$$

If (2.65) is transformed to the ordinary space, one finds

$$H_n^{\mathrm{eff}}(\mathbf{x})=E_n^A(0)-\frac{\hbar^2\nabla^2}{2m_n}+\left[\frac{\mathcal{L}}{L}W_{nn}+U_{nn}\right]\epsilon(z)$$
$$-i\mathbf{d}_{nn}\cdot\nabla\epsilon(z)-i\epsilon(z)d_{nn}^*\cdot\nabla\ , \qquad (2.66)$$

where $\epsilon(z)$ is given by (2.5d). To simplify the notation, let

$$\Delta(z)=\delta\left[z+\frac{L}{2}\right]-\delta\left[z-\frac{L}{2}\right],$$
$$B_n=\frac{\mathcal{L}}{L}W_{nn}+U_{nn}\ ,$$
$$\mathbf{D}_n=\mathbf{d}_{nn}+\mathbf{d}_{nn}^*\ ,$$
$$d_{n3}=\hat{e}_3\cdot\mathbf{d}_{nn}\ . \qquad (2.67)$$

The effective Schrödinger equation becomes

$$i\hbar\frac{\partial\psi}{\partial t}=\left[-\frac{\hbar^2\nabla^2}{2m_n}-i\epsilon(z)\mathbf{D}_n\cdot\nabla\right.$$
$$\left.-id_{n3}\Delta(z)+B_n\epsilon(z)+E_n^A(0)\right]\psi\ . \qquad (2.68)$$

If $\psi(\mathbf{x})$ is an eigenfunction of H^{eff}, then it must be continuous across the quantum-well interfaces, because the right-hand side of (2.68) is second order in space derivatives. The derivative of $\psi(\mathbf{x})$ with respect to z, however, may be stepwise discontinuous across the interfaces. The values of the steps are determined by d_{n3}. Consider the stationary Schrödinger equation corresponding to (2.68):

$$\left[-\frac{\hbar^2\nabla^2}{2m_n}+E_n^A(0)-E\right]\psi=0 \quad \text{for } z<-\frac{L}{2} \text{ or } z>\frac{L}{2}\,,$$

$$\left[-\frac{\hbar^2\nabla^2}{2m_n}-i\mathbf{D}_n\cdot\nabla+B_n+E_n^A(0)-E\right]\psi=0 \quad \text{for } -\frac{L}{2}<z<\frac{L}{2}\,. \tag{2.69}$$

There are two boundary conditions from the continuity of ψ across the interfaces ($\delta\rightarrow0$):

$$\begin{aligned}\psi\left[x,y,-\frac{L}{2}-\delta\right]&=\psi\left[x,y,-\frac{L}{2}+\delta\right],\\ \psi\left[x,y,\frac{L}{2}-\delta\right]&=\psi\left[x,y,\frac{L}{2}+\delta\right].\end{aligned} \tag{2.70}$$

Integrating the right-hand side of (2.68) across the interfaces, one finds two other boundary conditions:

$$\frac{d\psi}{dz}\left[x,y,-\frac{L}{2}+\delta\right]-\frac{d\psi}{dz}\left[x,y,-\frac{L}{2}-\delta\right]=-\frac{2im_n}{\hbar^2}d_{n3}\psi\left[x,y,-\frac{L}{2}\right], \tag{2.71a}$$

$$\frac{d\psi}{dz}\left[x,y,\frac{L}{2}+\delta\right]-\frac{d\psi}{dz}\left[x,y,\frac{L}{2}-\delta\right]=\frac{2im_n}{\hbar^2}d_{n3}\psi\left[x,y,\frac{L}{2}\right]. \tag{2.71b}$$

Equations (2.71a) and (2.71b) take the place of the BenDaniel-Duke ansatz.

III. DISCUSSION

The inspection of the effective Schrödinger equation in (2.68) shows that one has a position-independent effective mass m_n in a quantum-well structure. m_n depends on the matrix elements $\{W_{n'n'}\}$. $\{W_{n'n'}\}$ in turn depend on the band lineup energies and on the ratio $(L/\mathcal{L})$. In order to utilize the effective-mass concept in a given problem, one must be able to make the effective-mass approximation. A self-consistent application of the effective-mass approximation to a quantum well requires simultaneous elimination of the interband couplings arising from the homogeneous crystal $\mathbf{k}\cdot\mathbf{p}$ term and the well potential, since they are usually comparable. As illustrated by the derivation of (2.68), this leads to a position-independent effective mass in and around a quantum well.

Conversely, the use of a position-dependent effective mass as in the BenDaniel-Duke ansatz must postulate that the interband coupling induced by the well potential must be negligibly small compared to the homogeneous crystal $\mathbf{k}\cdot\mathbf{p}$ coupling. Clearly, such a postulate can hold only if $\mathbf{k}$ is not near the center of the BZ and the well potential is extremely shallow. In fact, the ansatz has usually been applied in cases where these conditions do not hold. Crystals used in quantum-well fabrication have their band edges mostly at the center of the BZ. Well depths are usually significant fractions of band gaps.

The preceding argument also rules out uses of energy-dependent effective masses taken from homogeneous crystals.[4,8,9] Such masses have been used for quantum wells whose depths are larger fractions of band gaps than GaAs/Al$_x$Ga$_{1-x}$As wells. Clearly, this is not a consistent procedure. An energy-dependent effective mass arises from elimination of the $\mathbf{k}\cdot\mathbf{p}$ coupling between conduction and valence bands. Since a typical quantum-well potential induces comparable coupling between the same bands, the well-potential-induced coupling must be eliminated simultaneously with the $\mathbf{k}\cdot\mathbf{p}$ coupling. Effective mass of a band would then depend on $W_{nn'}$. Its energy dependence would differ from those of homogeneous crystals.

In (2.68), the energy parameter B_n acts as the effective band lineup energy. B_n is given by

$$B_n=\sum_{\mathbf{G}\mathbf{G}'}\phi_n^{A*}(\mathbf{G})\phi_n^A(\mathbf{G}')[V_B(\mathbf{G}-\mathbf{G}')-V_A(\mathbf{G}-\mathbf{G}')]+U_{nn}\,. \tag{3.1}$$

U_{nn} is the perturbative correction given by (2.63), which is independent of the ratio $(\mathcal{L}/L)$ or its inverse. In contrast to m_n, B_n is not related to the coherence length and depends purely on the homogeneous crystal potentials.

The effective band Hamiltonian in (2.68) has δ-function singularities that yield the boundary conditions on the wave function. The origin of these singularities is the well-potential-induced coupling, as well as the $\mathbf{k}\cdot\mathbf{p}$ coupling. The possibility of a δ-function potential in a heterojunction was discussed earlier by Zhu and Kroemer.[15] They arrived at a δ-function potential from the extrapolation of the wave function across the interface in the Wannier representation and from the renormalization of the wave function in terms of the effective masses on the two sides of the interface. In contrast, the $\Delta(z)$ term in (2.68) is obtained from the canonical transformation that eliminates all interband couplings simultaneously to first order.

In (2.68), there is a velocity-dependent term proportional to $\mathbf{D}_n$. This term is nonzero only inside the well. The origin of the $\mathbf{D}_n\cdot\nabla$ term is the same as the origin of the δ-function potentials. We shall see that its size is of the same order as the $\mathbf{k}\cdot\mathbf{p}$ term for the homogeneous crystal. Thus the $\mathbf{D}_n\cdot\nabla$ term can significantly modify energy dispersion. Note that $\mathbf{D}_n\cdot\nabla$ is three dimensional. Energy dispersion is modified not only along the z direction, but also in the transverse directions.

Both the $\mathbf{D}_n\cdot\nabla$ term and the $\Delta(z)$ term depend on the vector $\mathbf{d}_{nn}$. One can estimate the size of $\mathbf{d}_{nn}$ by using the approximations (2.23) and (2.24) in (2.54):

$$\mathbf{d}_{nn}\simeq\frac{\hbar}{m}\sum_{n'(n'\neq n)}\phi_n^A(0)\mathbf{P}_{nn'}^A\phi_{n'}^{A*}(0)\left[\frac{E_n^B(0)-E_n^A(0)+E_{n'}^B(0)-E_{n'}^A(0)-2V_B(0)+2V_A(0)}{E_n^A(0)-E_{n'}^A(0)+W_{nn}-W_{n'n'}}\right]. \tag{3.2}$$

The quantity $V_A(0)-V_B(0)$ in the bracketed term measures the difference between the spatially averaged potentials of the homogeneous crystals of A and B. It may be identified with the difference between the electronic affinities of the two materials. The remaining terms in the numerator represent the sum of the band lineup energies for bands n and n'. The denominator represents the gap between n and n'. Since ϕ_n^A are rapidly decreasing functions of the reciprocal lattice vectors, the normalization condition on these functions is essentially exhausted by $\phi_n^A(0)$ and

$$|\phi_n^{A*}(0)\phi_n^A(0)|\sim 1 \ . \tag{3.3}$$

If the sum of the band lineup energies is comparable to the gap energy between two bands for at least one pair, then the bracketed term is on the order of 1, and $m\mathbf{d}_{nn}/\hbar$ is on the order of $\mathbf{P}_{nn'}^A$ corresponding to this pair. For example, for the two-band model,

$$\frac{m}{\hbar}|d_{cc}|=\frac{m}{\hbar}|\mathbf{d}_{vv}|\approx|P_{cv}^A| \ . \tag{3.4}$$

One may therefore conclude that the strength of the δ-function potentials, as well as the size of the $\mathbf{D}_n\cdot\nabla$ terms, in (2.68) is comparable to the $\mathbf{k}\cdot\mathbf{p}$ term in the ordinary Bloch Hamiltonian when it is written in the Luttinger-Kohn representation.

The z component of the $\mathbf{D}_n\cdot\nabla$ term contributes a z-dependent phase factor to bound-state wave functions. One sees from (2.68) that, inside the well, a bound-state wave function has the form

$$\begin{aligned} F_{n\lambda}(z)&=e^{-i\eta_n z}[A_{n\lambda}\cos k_{n\lambda}z+B_{n\lambda}\sin k_{n\lambda}z] \\ &=A'_{n\lambda}e^{i(k_{n\lambda}-\eta_n)z}+B'_{n\lambda}e^{-i(k_{n\lambda}+\eta_{n\lambda})z} \ . \end{aligned} \tag{3.5}$$

Here $A_{n\lambda},\ldots,B'_{n\lambda}$ are constants. n and λ designate the band and the bound states, respectively. η_n is given by

$$\eta_n=\frac{m_n}{2\hbar^2}\hat{\imath}_3\cdot\mathbf{D}_n \ . \tag{3.6}$$

η_n depends on the band lineup energies, as well as on the coherence length through the effective mass m_n. It is, however, independent of λ. According to (3.5), the wave number $k_{n\lambda}$ is shifted by $\pm\eta_n$. This may affect transition matrix elements. If $O(z)$ is an operator, its matrix element between $F_{n\lambda}$ and $F_{n'\lambda'}$ is given by

$$O_{n\lambda;n'\lambda'}=\int_{-L/2}^{L/2}dz\,F_{n\lambda}^{*}OF_{n'\lambda'}+\text{other terms} \ . \tag{3.7}$$

For intraband transitions, $n=n'$ and the η_n-phase factors drop out of (3.7). But for interband transitions, $n\neq n'$ and $O_{n\lambda;n'\lambda'}$ depends on the difference $\eta_n-\eta_{n'}$. For typical semiconductors, $\eta_n\sim 10^6-10^7$ cm^{-1}, and hence may cause substantial shifts in $k_{n\lambda}$. The $\eta_n-\eta_{n'}$ dependence may therefore be observable.

While on the subject of transition matrix elements, let me note that the canonical transformation of Sec. II provides a systematic procedure for calculating radiative and nonradiative transition matrix elements between quantum well states in terms of the parameters for homogeneous crystals. The accuracy of a calculation may be improved by performing a series of canonical transformations to eliminate higher-order coupling terms in the overall Hamiltonian. For radiative transitions, the procedure yields a hierarchy of selection rules.[16] This topic will be discussed elsewhere.

In the conventional method, the eigenfunctions for quantum-well-localized states separate into even and odd functions. One has two eigenvalue equations, one for odd and one for even eigenfunctions. In contrast, the solutions of (2.69), with the boundary conditions (2.70), (2.71a), and (2.71b), do not separate into even and odd functions. There is only one eigenvalue equation. This difference arises entirely from the interband coupling induced by the well potential. When one lets $\mathbf{d}_{nn}\to 0$, one recovers even and odd solutions.

The breakdown of the parity with respect to z is to be expected. When the interband coupling due to the well potential is taken into account, there are two special directions of the quantum-well structure for the band n. One is, of course, the z direction along which the crystal layers are grown. The other is the direction of $\mathbf{d}_{nn}$. It is clear from (3.2) that the direction of $\mathbf{d}_{nn}$ is determined by the interband momentum matrix element between n and the nearest band to n. As long as $\mathbf{d}_{nn}$ is not perpendicular to the z axis, the plus z direction differs from the minus z direction. The z parity is broken. If $\hat{\imath}_3\cdot\mathbf{d}_{nn}=0$, the plus and minus directions along the z axis are not distinguishable. The z parity is restored. Indeed, the solutions of (2.69) separate into even and odd functions if $\hat{\imath}_3\cdot\mathbf{d}_{nn}=0$, as well as if $\mathbf{d}_{nn}=0$.

Despite the z-parity breakdown, interband transition energies calculated with (2.69) agree with those obtained from the conventional method to within 10–20 meV.[17] There are significant differences between the results of the two methods, however, when one looks at the number of localized states in the valence and conduction bands and the separations of bound-state energies within a given band. The solutions of (2.69) are far more sensitive to band lineup energies than the solutions obtained from the conventional method. The breakdown of the z parity also affects the strength of some radiative transitions.

It is clear from the derivation that H^{eff} given by (2.68) is valid only if the effective approximation holds, that is,

$$\left|\frac{\hbar\mathbf{k}\cdot\mathbf{P}_{nn'}^{\mathbf{A}}}{m\,[E_n^A(\mathbf{k})-E_{n'}^A(\mathbf{k})]}\right|<1 \ . \tag{3.8}$$

The matrix elements of $(\hbar\mathbf{k}\cdot\mathbf{p}/m)$ between bands n and n' must be smaller than the gap separating them. Furthermore, the interband coupling Hamiltonian given by (2.63) must be negligible relative to $E_n^A(0)+B_n$:

$$\begin{aligned} &\frac{\langle n\mathbf{k}|[H_2,S]|n'\mathbf{k}'\rangle}{E_n^A(0)+B_n} \\ &\qquad\simeq\frac{W_{nn'}}{[E_n^A(0)+B_n][E_n^A(0)-E_{n'}^A(0)]} \\ &\qquad\quad\times\left[W_{nn}-W_{n'n'}+\frac{\hbar(\mathbf{k}-\mathbf{k}')\cdot\mathbf{P}_{nn'}^A}{m}\right]<1 \ . \end{aligned} \tag{3.9}$$

Clearly, $W_{nn'}$ must be small relative to $E_n^A(0)+B_n$, as well as relative to $E_n^A(0)-E_{n'}^A(0)$. This would also guarantee that the higher commutators in (2.30) are negligible. It is interesting that the matrix elements $W_{nn'}$ appear in (3.9). The inequality therefore depends on the coherence length. Everything else being equal, H^{eff} applies better to a purer sample.

The inequality (3.8) implies that the band n for which H^{eff} is derived is sufficiently isolated in energy. If this were not true and n had several nearby bands that were coupled to it through $\mathbf{P}_{nn'}^A$, as well as through the well potential, then this entire set of bands, including n, would have to be decoupled from the remaining bands.[18] The configurational decoupling could still be effected by a canonical transformation. This will be discussed elsewhere.

When there are perfect interfaces in a heterostructure crystal, as in Figs. 1 and 2, the potential that the electron "sees" varies rapidly in the vicinity of an interface. This creates the impression that the electronic wave function cannot be determined perturbatively, since single-band effective Hamiltonians are usually used in conjection with slowly varying perturbations. The analysis of Sec. II and (3.9) shows that such a view is not correct. As long as $W_{nn'}$ is small relative to the original band gaps, as well as to those gaps that are shifted by B_n, a perturbative determination of the wave function is valid.

ACKNOWLEDGMENTS

I am grateful to P. Sharma, A. Ongstad, and G. Dente for many useful discussions.

---

[1] J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955).

[2] G. Bastard, *Wavemechanics Applied to Semiconductor Heterostructures* (Halsted, New York, 1988).

[3] R. Dingle, in *Feskörperprobleme XV (Advances in Solid Physics)*, edited by H. J. Queisser (Pergamon, Vieweg, 1975), p. 21.

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

[5] N. F. Johnson, H. Ehrenreich, K. C. Hass, and T. C. McGill, Phys. Rev. Lett. **59**, 2352 (1987).

[6] N. F. Johnson, H. Ehrenreich, P. M. Hui, and P. M. Young, Phys. Rev. B **41**, 3655 (1990).

[7] D. J. BenDaniel and C. B. Duke, Phys. Rev. **152**, 683 (1966).

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

[9] B. R. Nag and S. Mukhopadhyay, Appl. Phys. Lett. **58**, 1056 (1991).

[10] T. Cunningham, R. C. Barker, and L. C. Chiu, J. Appl. Phys. **63**, 5393 (1988).

[11] D. L. Smith and C. Mailhiot, Rev. Mod. Phys. **62**, 173 (1990).

[12] J. Tersoff, Phys. Rev. B **30**, 4874 (1984).

[13] W. Jones and N. H. March, *Theoretical Solid State Physics, Vol. 1* (Dover, New York, 1985).

[14] A. Elçi and E. D. Jones, Phys. Rev. B **34**, 8611 (1986).

[15] Q. Zhu and H. Kroemer, Phys. Rev. B **27**, 3519 (1983).

[16] A similar calculation was performed for impurity superlattices. See A. Elçi, Phys. Rev. B **46**, 2208 (1992).

[17] A. Elçi (unpublished).

[18] E. O. Kane, in *Semiconductors and Semimetals I: Physics of III-V Compounds*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1966).