

## Motion of electrons in semiconductors under inhomogeneous strain with application to laterally confined quantum wells

Yong Zhang

*Department of Physics and Astronomy, Dartmouth College, 6127 Wilder Laboratory, Hanover, New Hampshire 03755-3528*

(Received 22 June 1992; revised manuscript received 9 August 1993)

A general treatment for finding the energy levels of electrons and holes in a system with slowly varying inhomogeneous strain is given in the envelope-function approximation. An eight-band model is derived, then block diagonalized to  $2 \times 2$  and  $6 \times 6$  for the conduction and valence bands, respectively. To first order, i.e., in the so-called effective-mass approximation, the gradient of the strain tensor does not appear in the Hamiltonian; in the second-order approximation, both the strain variation of the effective masses and the gradients of the strain tensor appear. The second-order effect can be significant within the range of applicability of the theory when the strain is not sufficiently small and its variation is not sufficiently slow. The general theory is first applied to the case of homogeneous strain (for the conduction band), which gives the strain dependence of the band structure, with comparison to previous work. The theory is then applied to laterally strain-confined quantum wells, with the first-order approximation. In the strain-confined system, the valence band is treated by four simultaneous envelope-function equations. Under certain conditions, they can be reduced to a pair of independent equations. Numerical results for the energy levels in a specific quantum wire are given. In the strain-confined system, the conditions under which the second-order effect could be significant are discussed.

### I. INTRODUCTION

The effect of inhomogeneous strain in crystals was first treated by Bardeen and Shockley:<sup>1</sup> a slowly varying inhomogeneous strain introduces a deformation potential  $E_1 \Delta(\mathbf{r})$  into the effective-mass equation for the envelope function, where  $E_1$  is a deformation potential and  $\Delta(\mathbf{r}) = \varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}$  is the dilatation. They assumed that the band was nondegenerate, that strain  $\varepsilon_{ij}$  was sufficiently small, and that the variation of the strain with position was sufficiently gradual so that terms related to the position variation of the effective mass and the gradient of the strain were negligible. This can be understood as the first-order theory. A higher-order theory<sup>2,3</sup> which includes the neglected terms in Ref. 1 has been developed for the acoustic-phonon-electron interaction problem, where an inhomogeneous strain is created by the phonons. Recently, inhomogeneous strain has been used in semiconductor quantum wells to achieve lateral confinement in order to create quantum wires and dots.<sup>4-6</sup> In these structures, we have a degenerate valence band, the magnitude of the strain is on the order of 1%, and the strain may vary appreciably on the scale of the envelope function. This problem is quite distinct from the phonon-electron interaction, since the external inhomogeneous strain in general breaks the translational symmetry. To the best of my knowledge, there has not been a rigorous treatment, on the order of the approximation for the phonon-electron problem, and allowing for the degenerate band structure, of the effect of a general slowly varying inhomogeneous strain on the electronic states. Here "slowly varying" means that the variation is negligible over a unit cell, but may be significant on the scale of the envelope function. It is therefore necessary to develop a refined theory, on the basis of

Refs. 1-3, for strain-confined structures, which are currently of particular interest, and also for other possible cases in general.

Slowly varying inhomogeneous strain has been treated differently from Bardeen and Shockley by other authors,<sup>2,3</sup> who used a coordinate transformation method also used by Pikus and Bir for the homogeneous strain case.<sup>7,8</sup> The same coordinate transformation method will be used in this work.

Another important issue is to deal correctly with a degenerate band under inhomogeneous strain. This has not yet been done even in the first-order approximation. One approach<sup>4,6,9</sup> is to find position-dependent band-edge ( $k=0$ ) states under piecewise homogeneous strain. This is a zero-order theory, and is invalid since the diagonalization procedure for homogeneous strain is not applicable to the case of inhomogeneous strain, which mixes states of different  $\mathbf{k}$  with different behaviors under strain. In practice, this procedure is valid for very slowly varying strain field (on a scale of mm) as in Ref. 9, but it is not appropriate for strained-quantum-well structures<sup>4-6</sup> where quantum confinement is significant. The envelope-function method needs to be generalized to the case of degenerate bands under inhomogeneous strain, as has been done for impurities with long-range potentials.<sup>10</sup>

On the other hand, the theory for inhomogeneous strain has been well developed<sup>8</sup> since Pikus and Bir's pioneering work,<sup>7</sup> and recently an eight-band model was proposed.<sup>11</sup> Phenomenologically, one might assume that the  $\mathbf{k} \cdot \mathbf{p}$  theory for homogeneous strain<sup>11</sup> can be adapted to the inhomogeneous case by changing  $\mathbf{k}$  to  $-i\nabla$  and  $\varepsilon_{ij}$  to  $\varepsilon_{ij}(\mathbf{r})$ , and thus obtaining a set of coupled differential equations for the envelope functions. However, as we shall see, this is inadequate when the strain varies appreciably on the length scale of the envelope function. Strain gradient terms have been included<sup>12</sup> by making a

symmetric replacement to the product of the strain tensor component and the momentum operator in the theory for homogeneous strain. While this ensures the Hermiticity of the Hamiltonian, the Hamiltonian obtained is not always adequate, as will be discussed below.

Apart from the above considerations, two other effects—normal and generalized piezoelectricity,<sup>13</sup>—might also need to be taken into account. Here “generalized piezoelectricity” refers to a macroscopic polarization field proportional to the gradient of the strain. The normal piezoelectric effect appears only in crystals without inversion symmetry, such as GaAs, which in fact is weakly piezoelectric. This effect has been discussed for strain-confined quantum wires and quantum dots.<sup>14</sup> It is not present in a quantum wire structure if the orientation of the wire is along [100] ([001] being the growth direction of the quantum well). As for the generalized piezoelectricity, it has been shown<sup>15</sup> that the effect of the macroscopic field can be separated into two parts: analytic and nonanalytic. The contribution of the analytic part is equivalent to a change of deformation potential, and thus is included in the measured deformation potential, while that of the nonanalytic part is small.<sup>15</sup> Thus, if one ignores the contribution of the nonanalytic part, one need not consider the generalized piezoelectric effect separately. The normal piezoelectric effect, on the other hand, gives a potential  $V_{\text{piez}}$  that is independent of the band index,<sup>14</sup> which means that it can be added in at the places where the hydrostatic deformation potentials appear.

In Sec. II, a general treatment for slowly varying inhomogeneous strain in the envelope-function approximation is given for semiconductors with degenerate or nearly degenerate band structures. The problem is worked out in the eight-band model, and block diagonalized to  $2 \times 2$  and  $6 \times 6$  blocks by second-order perturbation theory. Applications to the limiting case of homogeneous strain are briefly discussed. In Sec. III, the theory is applied to laterally confined quantum-well structures. The nearly degenerate valence band and the nondegenerate conduction band are treated consistently to first order. Numerical results for a specific quantum wire structure are given. The second-order effect is discussed, and is shown to contribute significantly within the range of applicability of this theory, although not in the specific quantum wire structure considered in this work. Section IV is a summary of this work.

The symbolic calculation in this work was done on a Mac II computer with the help of Theorist and Mathematica.<sup>16</sup>

## II. THEORY

For an unstrained bulk crystal, the Hamiltonian for an electron is<sup>8</sup>

$$H_0 = \frac{\mathbf{p}^2}{2m} + V_0 + \frac{\hbar}{4m^2c^2} (\nabla V_0) \times \mathbf{p} \cdot \boldsymbol{\sigma} . \quad (1)$$

For a strained crystal, the Hamiltonian will be

$$H_\varepsilon = \frac{\mathbf{p}^2}{2m} + V_\varepsilon + \frac{\hbar}{2m^2c^2} (\nabla V_\varepsilon) \times \mathbf{p} \cdot \boldsymbol{\sigma} . \quad (2)$$

For homogeneous or macroscopically extended inhomogeneous strains, the difference  $V_\varepsilon - V_0$  cannot be taken directly as a perturbation, since in general it is not small.<sup>8</sup> Assuming the strain tensor is symmetric, we may make a nonlinear coordinate transformation:<sup>3,17</sup>

$$\mathbf{x}' = \mathbf{x} - \int_0^{\mathbf{x}} \boldsymbol{\varepsilon}(\mathbf{x}) d\mathbf{x} , \quad (3)$$

where the integral is the strain-induced displacement. In the coordinate system  $\mathbf{x}'$  all atoms are placed in their regular positions, as they are in an  $\mathbf{x}$  coordinate system without strain. Under this transformation, noting that  $\varepsilon_{ij}(\mathbf{x}) = \varepsilon_{ij}(\mathbf{x}')$  and  $\partial \varepsilon_{ij} / \partial x_k = \partial \varepsilon_{ij} / \partial x'_k$  to first order in  $\varepsilon$ ,

$$\mathbf{p} = [1 - \boldsymbol{\varepsilon}(\mathbf{x}')] \mathbf{p}' , \quad (4)$$

$$V_\varepsilon(\mathbf{x}) = V_0(\mathbf{x}') + \sum_{ij}^3 V_{ij}(\mathbf{x}') \varepsilon_{ij}(\mathbf{x}') . \quad (5)$$

Transformations (4) and (5) have the same forms as those for the homogeneous strain,<sup>8</sup> except that now  $\boldsymbol{\varepsilon}(\mathbf{x}')$  is position dependent. The validity of the potential expansion is based on the criteria, usually adopted in the electron-phonon interaction problem,<sup>3</sup> that the strain is small and the lattice displacements vary slowly over the scale of the unit cell, and that the nonanalytic long-range electrostatic effect<sup>15</sup> is negligible. These criteria ensure the locality of the strain potential  $V_\varepsilon(\mathbf{x})$ : that is, at a given point of space the potential depends only on the strain tensor at the same point.<sup>3,18</sup> Then the Hamiltonian  $H_\varepsilon$  in the  $\mathbf{x}'$  coordinate is

$$H_\varepsilon(\mathbf{x}') = H_0(\mathbf{x}') + D_\varepsilon(\mathbf{x}') + D_{\varepsilon\text{SO}}(\mathbf{x}') , \quad (6)$$

where

$$H_0 = \frac{\mathbf{p}'^2}{2m} + V_0 + \frac{\hbar}{4m^2c^2} (\nabla' V_0) \times \mathbf{p}' \cdot \boldsymbol{\sigma} , \quad (7)$$

$$D_\varepsilon = \sum_{ij}^3 \varepsilon_{ij} \left[ -\frac{p'_i p'_j}{m} + V_{ij} \right] - \sum_{ij}^3 (p'_i \varepsilon_{ij}) \frac{p'_j}{2m} , \quad (8)$$

$$D_{\varepsilon\text{SO}} = \frac{\hbar}{2m^2c^2} \left\{ \sum_{ij}^3 [\varepsilon_{ij} (\nabla' V_{ij}) + V_{ij} (\nabla' \varepsilon_{ij})] \right. \\ \left. \times \mathbf{p}' \cdot \boldsymbol{\sigma} - [\boldsymbol{\varepsilon} \cdot (\nabla' V_0)] \times \mathbf{p}' \cdot \boldsymbol{\sigma} \right. \\ \left. - (\nabla' V_0) \times (\boldsymbol{\varepsilon} \cdot \mathbf{p}') \cdot \boldsymbol{\sigma} \right\} . \quad (9)$$

$H_0(\mathbf{x}')$  has the same form in the  $\mathbf{x}'$  coordinate as  $H_0(\mathbf{x})$  in the  $\mathbf{x}$  coordinate. Except for strain gradient terms, all the other terms in  $D_\varepsilon$  or  $D_{\varepsilon\text{SO}}$  have the same forms as for the homogeneous strain. For a slowly varying strain,  $\varepsilon_{ij} \nabla V_{ij} \gg V_{ij} \nabla \varepsilon_{ij}$ , so the gradient term is actually negligible in  $D_{\varepsilon\text{SO}}$  and we need only consider the gradient term in  $D_\varepsilon$ . The eigenvalue problem to be solved is now

$$H_\varepsilon(\mathbf{x}') \Psi(\mathbf{x}') = E \Psi(\mathbf{x}') . \quad (10)$$

As pointed out in Refs. 2 and 3, the Hamiltonian in the  $\mathbf{x}'$  coordinate, (6), is non-Hermitian since the momentum operator  $\mathbf{p}'$  is non-Hermitian. However, the non-Hermiticity of the transformed Hamiltonian is only apparent, and the transformation will not change the eigen-

values. The Hermiticity is recovered by multiplying (10) throughout by the Jacobian of the transformation,<sup>3</sup> which is first order in the strain is  $J(\mathbf{x}') = 1 + \Delta(\mathbf{x}')$  [here  $\Delta(\mathbf{x}')$  is the dilatation]. Equation (10) then becomes

$$\mathcal{H}_\varepsilon(\mathbf{x}')\Psi(\mathbf{x}') = E[1 + \Delta(\mathbf{x}')] \Psi(\mathbf{x}') , \quad (11)$$

with

$$\mathcal{H}_\varepsilon(\mathbf{x}') = H_0(\mathbf{x}') + \Delta(\mathbf{x}')H_0(\mathbf{x}') + D_\varepsilon(\mathbf{x}') + D_{\varepsilon\text{SO}}(\mathbf{x}') . \quad (12)$$

$\mathcal{H}_\varepsilon(\mathbf{x}')$  is a Hermitian operator, and  $1 + \Delta(\mathbf{x}')$  is a positive-definite Hermitian operator, so the eigenvalues  $E$  of Eq. (11) are real.

Since the strain varies slowly on the scale of the unit cell,  $\mathcal{H}_\varepsilon(\mathbf{x}')$  can be solved by the envelope-function method.<sup>10</sup> The state  $\Psi(\mathbf{x}')$  can be expanded as

$$\Psi(\mathbf{x}') = \sum_{n\mathbf{k}} A_n(\mathbf{k}) \chi_{n\mathbf{k}}(\mathbf{x}') , \quad (13)$$

where  $\chi_{n\mathbf{k}} = \exp(i\mathbf{k} \cdot \mathbf{x}') u_n(\mathbf{x}')$ , and  $u_n(\mathbf{x}')$  are band-edge states at  $k=0$  corresponding to the Hamiltonian  $H_0(\mathbf{x}')$ , which includes spin-orbit interaction.  $\mathbf{k}$  should be understood as a wave vector in  $\mathbf{x}'$  space. Equivalent to Eq. (11), we have equations in  $\mathbf{k}$  space:

$$\left[ E_n + \frac{\hbar^2 k^2}{2m} - E \right] A_n(\mathbf{k}) + \frac{\hbar}{m} \mathbf{k} \cdot \sum_l \mathbf{P}_{nl} A_l(\mathbf{k}) + \sum_{l,\mathbf{q}} \langle n\mathbf{k} | H' | l\mathbf{q} \rangle A_l(\mathbf{q}) = 0 , \quad (14)$$

where  $E_n$  are the energies of band-edge states,  $|n\mathbf{k}\rangle \equiv |\chi_{n\mathbf{k}}\rangle$ ,  $\mathbf{P}_{nl} = \langle u_n | \mathbf{p} | u_l \rangle$ , and  $H' = \Delta(H_0 - E) + D_\varepsilon + D_{\varepsilon\text{SO}}$ . A linear  $\mathbf{k}$  term due to spin-orbit interaction in a noncentrosymmetric crystal is neglected<sup>11</sup> in (14). The situation is analogous to that for the long-range impurity problem;<sup>10</sup> however, there is a significant difference in that whereas the matrix elements  $\langle n\mathbf{k} | H' | l\mathbf{q} \rangle$  are (approximately) diagonal in the band index for the impurity potential, they are not for the strain. As has been done for the case of homogeneous strain, I will work out this problem in the eight-band manifold which includes two conduction bands and six valence bands. In the eight-band  $\mathbf{k} \cdot \mathbf{p}$  theory for unstrained crystals, it is customary to decouple the eight-band manifold from the other remote bands by the Löwdin perturbation, and then to diagonalize the “renormalized”  $8 \times 8$  matrix exactly. In the case of homogeneous strain, the strain effect on the coupling between the eight-band manifold and the remote bands is normally ignored<sup>11</sup> (except by Aspnes and Cardona<sup>19</sup>), while the strain effect within the eight-band manifold is treated exactly. This approximation will also be used here for the inhomogeneous strain case. For simplicity, effects due to the lack of inversion symmetry will be ignored.

The matrix elements  $\langle n\mathbf{k} | H' | l\mathbf{q} \rangle$  can be calculated and organized as follows:

$$\begin{aligned} \langle n\mathbf{k} | H' | l\mathbf{q} \rangle &= \langle n\mathbf{k} | \mathcal{D}_0 | l\mathbf{q} \rangle + \langle n\mathbf{k} | \mathcal{D}_\varepsilon | l\mathbf{q} \rangle \\ &+ \langle n\mathbf{k} | \mathcal{D}_{\varepsilon\text{SO}} | l\mathbf{q} \rangle + \langle n\mathbf{k} | \mathcal{D}_{\mathbf{k}\cdot\mathbf{p}} | l\mathbf{q} \rangle , \end{aligned} \quad (15)$$

where

$$\begin{aligned} \langle n\mathbf{k} | \mathcal{D}_0 | l\mathbf{q} \rangle &= \left[ E_n + \frac{\hbar^2 q^2}{2m} - E \right] \Delta(\mathbf{k} - \mathbf{q}) \delta_{nl} \\ &- \sum_{i,j} \left[ \frac{\hbar^2 q_i q_j}{m} \varepsilon_{ij}(\mathbf{k} - \mathbf{q}) \right. \\ &\left. + \frac{\hbar q_j}{2m} \varepsilon_{ij}^i(\mathbf{k} - \mathbf{q}) \right] \delta_{nl} , \end{aligned} \quad (16)$$

$$\langle n\mathbf{k} | \mathcal{D}_\varepsilon | l\mathbf{q} \rangle = \sum_{i,j} a_{nl}^{ij} \varepsilon_{ij}(\mathbf{k} - \mathbf{q}) , \quad (17)$$

$$\langle n\mathbf{k} | \mathcal{D}_{\varepsilon\text{SO}} | l\mathbf{q} \rangle = \sum_{i,j} b_{nl}^{ij} \varepsilon_{ij}(\mathbf{k} - \mathbf{q}) , \quad (18)$$

$$\begin{aligned} \langle n\mathbf{k} | \mathcal{D}_{\mathbf{k}\cdot\mathbf{p}} | l\mathbf{q} \rangle &= - \sum_{i,j} \left[ \frac{2\hbar q_i P_{nl}^j}{m} \varepsilon_{ij}(\mathbf{k} - \mathbf{q}) + \frac{P_{nl}^j}{2m} \varepsilon_{ij}^i(\mathbf{k} - \mathbf{q}) \right] \\ &+ \frac{\hbar \mathbf{q} \cdot \mathbf{P}_{nl}}{m} \Delta(\mathbf{k} - \mathbf{q}) , \end{aligned} \quad (19)$$

with

$$\begin{aligned} a_{nl}^{ij} &= \left\langle u_n \left| -\frac{P_i P_j}{m} + V_{ij} \right| u_l \right\rangle , \\ b_{nl}^{ij} &= \langle u_n | (D_{\varepsilon\text{SO}})_{ij} | u_l \rangle , \end{aligned}$$

and  $\varepsilon_{ij}(\mathbf{k})$ ,  $\varepsilon_{ij}^i(\mathbf{k})$ , and  $\Delta(\mathbf{k})$  are the Fourier transforms of the strain tensor  $\varepsilon_{ij}(\mathbf{x}')$ , its gradient  $[p_i' \varepsilon_{ij}(\mathbf{x}')$ ], and the dilatation  $\Delta(\mathbf{x}')$ , respectively.  $D_{\varepsilon\text{SO}}$  has been written in the form  $\sum_{ij} (D_{\varepsilon\text{SO}})_{ij} \varepsilon_{ij}$ , and again a linear  $\mathbf{k}$  term in  $D_{\varepsilon\text{SO}}$  has been dropped.<sup>11</sup>  $a_{nl}^{ij}$  and  $b_{nl}^{ij}$  can be calculated in the eight-band model<sup>11</sup> for the case of homogeneous strain, and they are given in (A14) and (A15) of Appendix A.

Introducing abbreviations

$$\langle n\mathbf{k} | H_0 | l\mathbf{q} \rangle = \left[ E_n + \frac{\hbar^2 k^2}{2m} \right] \delta_{nl} \delta_{\mathbf{k},\mathbf{q}} , \quad (20)$$

$$\langle n\mathbf{k} | H_{\mathbf{k}\cdot\mathbf{p}}^{\text{in}} | l\mathbf{q} \rangle = \frac{\hbar}{m} \mathbf{k} \cdot \mathbf{P}_{nl} \delta_{\mathbf{k},\mathbf{q}} \quad \text{for } n \leq 8, l \leq 8 , \quad (21)$$

$$\langle n\mathbf{k} | H_{\mathbf{k}\cdot\mathbf{p}}^{\text{out}} | l\mathbf{q} \rangle = \frac{\hbar}{m} \mathbf{k} \cdot \mathbf{P}_{nl} \delta_{\mathbf{k},\mathbf{q}} \quad \text{for } n \leq 8, l > 8 , \quad (22)$$

the total Hamiltonian corresponding to (14) is then

$$\mathcal{H}'_\varepsilon = H_0 + H_{\mathbf{k}\cdot\mathbf{p}}^{\text{in}} + H_{\mathbf{k}\cdot\mathbf{p}}^{\text{out}} + \mathcal{D} + \mathcal{D}_{\mathbf{k}\cdot\mathbf{p}} , \quad (23)$$

where  $\mathcal{D} = \mathcal{D}_0 + \mathcal{D}_\varepsilon + \mathcal{D}_{\varepsilon\text{SO}}$ , and  $E\Delta(\mathbf{x}')$  on the right-hand side of the Eq. (11) has been included in  $\mathcal{H}'_\varepsilon$  (or  $\mathcal{D}_0$ ).  $H_{\mathbf{k}\cdot\mathbf{p}}^{\text{out}}$  is the only term which couples the eight-band manifold to other remote bands, if the small coupling due to  $\mathcal{D}_{\mathbf{k}\cdot\mathbf{p}}$ ,  $\mathcal{D}_\varepsilon$ , and  $\mathcal{D}_{\varepsilon\text{SO}}$  is neglected.<sup>11</sup>

To decouple the eight-band manifold from remote bands, we make a unitary transformation  $\Psi(\mathbf{x}') = \exp(iS)\phi(\mathbf{x}')$ ,<sup>10</sup> and obtain the transformed eigenvalue equation

$$\bar{\mathcal{H}}_\varepsilon(\mathbf{x}')\phi(\mathbf{x}') = E\phi(\mathbf{x}') , \quad (24)$$

where the effective Hamiltonian is given by

$$\bar{\mathcal{H}}_\varepsilon = \exp(-iS)\mathcal{H}'_\varepsilon \exp(iS) , \quad (25)$$

and the transformed wave function is

$$\phi(\mathbf{x}') = \sum_{n=1, \mathbf{k}}^{n=8} C_n(\mathbf{k}) \chi_{n\mathbf{k}}(\mathbf{x}') \equiv \sum_{n=1}^8 \mathcal{F}_n(\mathbf{x}') u_n(\mathbf{x}'), \quad (26)$$

where  $\mathcal{F}_n(\mathbf{x}')$  is the so-called envelope function defined by

$$\mathcal{F}_n(\mathbf{x}') = \sum_{\mathbf{k}} C_n(\mathbf{k}) \exp(i\mathbf{k} \cdot \mathbf{x}'). \quad (27)$$

The eigenvalue equation in  $\mathbf{k}$  space is then

$$\sum_{l\mathbf{q}} \langle n\mathbf{k} | \bar{\mathcal{H}}_\varepsilon | l\mathbf{q} \rangle C_l(\mathbf{q}) = E C_n(\mathbf{k}). \quad (28)$$

If  $S$  is defined by

$$i[H_0, S] = -H_{\mathbf{k}\cdot\mathbf{p}}^{\text{out}}, \quad (29)$$

then it can be shown<sup>10</sup> that up to terms of second order in  $S$ ,

$$\bar{\mathcal{H}}_\varepsilon = H_0 + H_{\mathbf{k}\cdot\mathbf{p}}^{\text{in}} + \frac{1}{2}i[H_{\mathbf{k}\cdot\mathbf{p}}^{\text{out}}, S] + \mathcal{D} + \mathcal{D}_{\mathbf{k}\cdot\mathbf{p}}. \quad (30)$$

The matrix elements of  $S$  are calculated to be

$$\langle n\mathbf{k} | S | l\mathbf{q} \rangle = i \frac{(H_{\mathbf{k}\cdot\mathbf{p}}^{\text{out}})_{nl} \delta_{\mathbf{k}, \mathbf{q}}}{E_n - E_l} \quad (n \leq 8, l > 8). \quad (31)$$

With these  $\langle n\mathbf{k} | S | l\mathbf{q} \rangle$ , we can calculate the  $S$ -related term in (30):

$$\begin{aligned} & \langle n\mathbf{k} | \frac{1}{2}i[H_{\mathbf{k}\cdot\mathbf{p}}^{\text{out}}, S] | l\mathbf{q} \rangle \\ &= \frac{1}{2} \sum_{m > 8} (H_{\mathbf{k}\cdot\mathbf{p}}^{\text{out}})_{nm} (H_{\mathbf{k}\cdot\mathbf{p}}^{\text{out}})_{ml} \\ & \quad \times \left[ \frac{1}{E_n - E_m} + \frac{1}{E_l - E_m} \right] \delta_{\mathbf{k}, \mathbf{q}}. \end{aligned} \quad (32)$$

We want to convert equation (28) in  $\mathbf{k}$  space to real space in terms of the envelope functions, which can be done by multiplying (28) by  $\exp(i\mathbf{k} \cdot \mathbf{x}')$ , and summing over  $\mathbf{k}$ . We then have

$$\bar{\mathcal{H}}_\varepsilon(\mathbf{x}', \nabla') \mathcal{F}(\mathbf{x}') = E \mathcal{F}(\mathbf{x}'), \quad (33)$$

where  $\mathcal{F}$  is a  $8 \times 1$  column vector with eight envelope functions  $\mathcal{F}_l(\mathbf{x}')$ ,  $\bar{\mathcal{H}}_\varepsilon$  is an  $8 \times 8$  matrix which can be obtained from the eight-band model for the homogeneous strain (A12) by making the replacements  $\mathbf{k} \rightarrow -i\nabla'$ ,  $\varepsilon \rightarrow \varepsilon(\mathbf{x}')$ ,  $E_n \rightarrow E_n + (E_n - E)\Delta$ ,  $H_{\mathbf{k}\cdot\mathbf{p}} \rightarrow (1 + \Delta)H_{\mathbf{k}\cdot\mathbf{p}}$ ,

$$\begin{aligned} \langle n\mathbf{k} | \frac{1}{2}i[H_1, S] | l\mathbf{q} \rangle &= \frac{1}{2} \left\{ \sum_m (H_{\mathbf{k}\cdot\mathbf{p}})_{nm} (H_{\mathbf{k}\cdot\mathbf{p}})_{ml} \left[ \frac{1}{E_n - E_m} + \frac{1}{E_l - E_m} \right] \delta_{\mathbf{k}, \mathbf{q}} \right. \\ & \quad \left. + \sum_{m=c,v} [(H_{\mathbf{k}\cdot\mathbf{p}})_{nm} \langle n\mathbf{k} | \mathcal{D}_{\mathbf{k}\cdot\mathbf{p}} | l\mathbf{q} \rangle \right. \\ & \quad \left. + \langle n\mathbf{k} | \mathcal{D}_{\mathbf{k}\cdot\mathbf{p}} | m\mathbf{q} \rangle (H_{\mathbf{q}\cdot\mathbf{p}})_{ml}] \left[ \frac{1}{E_n - E_m} + \frac{1}{E_l - E_m} \right] \right\}, \end{aligned} \quad (39)$$

where only the first sum includes the remote bands in order to give the correct effective-mass tensor, while coupling to the remote bands due to the strain is ignored. We notice that this matrix divides into  $2 \times 2$  (conduction band) and  $6 \times 6$  (valence-band) blocks. Similarly,

$$\langle n\mathbf{k} | i[\mathcal{D}, S] | l\mathbf{q} \rangle = \sum_{m=c,v} \left[ \frac{(H_{\mathbf{k}\cdot\mathbf{p}})_{nm} \langle m\mathbf{k} | \mathcal{D} | l\mathbf{q} \rangle}{E_n - E_m} + \frac{\langle n\mathbf{k} | \mathcal{D} | m\mathbf{q} \rangle (H_{\mathbf{q}\cdot\mathbf{p}})_{ml}}{E_l - E_m} \right], \quad (40)$$

and  $\varepsilon_{ij}k_i$  [in  $D_{\mathbf{k}\cdot\mathbf{p}}$  given by (A18) in  $\mathbf{x}$  coordinate]  $\rightarrow -i[2\varepsilon_{ij}\nabla'_i + \frac{1}{2}(\nabla'_i\varepsilon_{ij})]$ , and adding a diagonal term  $\sum_{ij} \hbar^2/m[\nabla'_i\varepsilon_{ij}\nabla'_j - \frac{1}{2}\nabla'_j\varepsilon_{ii}\nabla'_j]\delta_{nl}$ . This effective Hamiltonian is Hermitian. Then (33) gives an eight-band model in the  $\mathbf{x}'$  coordinate system for the inhomogeneous strain. It is more convenient to work in the  $\mathbf{x}$  coordinate. Equation (33) is transformed back to the  $\mathbf{x}$  coordinate as

$$\bar{H}_\varepsilon(\mathbf{x}, \nabla) F(\mathbf{x}) = E F(\mathbf{x}), \quad (33')$$

where  $\bar{H}_\varepsilon(\mathbf{x}, \nabla)$  is the same as the effective Hamiltonian (A12) for the homogeneous strain with  $\mathbf{k} \rightarrow -i\nabla$ ,  $\varepsilon \rightarrow \varepsilon(\mathbf{x})$ , except that  $\varepsilon_{ij}k_i$  in  $D_{\mathbf{k}\cdot\mathbf{p}}$  is changed to  $-i[\varepsilon_{ij}\nabla_i + \frac{1}{2}(\nabla_i\varepsilon_{ij})]$ . This shows that strain gradient terms appear only in strain-induced interband coupling " $\mathcal{D}_{\mathbf{k}\cdot\mathbf{p}}$ ."  $F(\mathbf{x})$  is then the envelope function in the  $\mathbf{x}$  coordinate.

The effect of normal piezoelectricity can be included simply by adding the piezoelectric potential  $V_{\text{piez}}(\mathbf{x})$  (Ref. 14) to the diagonal terms of  $\bar{H}_\varepsilon(\mathbf{x}, \nabla)$  on the left-hand side of (33'); however, it will be omitted in the following for simplicity.

In fact it is possible to decouple the conduction and valence bands at the same time as the eight-band manifold is decoupled from the remote bands. Let us define

$$H_1 = H_{\mathbf{k}\cdot\mathbf{p}} + \mathcal{D}_{\mathbf{k}\cdot\mathbf{p}}, \quad (34)$$

where  $H_{\mathbf{k}\cdot\mathbf{p}} = H_{\mathbf{k}\cdot\mathbf{p}}^{\text{in}} + H_{\mathbf{k}\cdot\mathbf{p}}^{\text{out}}$ . The total Hamiltonian (23) can be rewritten as

$$\mathcal{H}'_\varepsilon = H_0 + \mathcal{D} + H_1. \quad (35)$$

Redefining  $S$  by

$$i[H_0, S] = -H_1, \quad (36)$$

then up to the terms of second order in  $S$ ,

$$\bar{\mathcal{H}}_\varepsilon = H_0 + \mathcal{D} + \frac{1}{2}i[H_1, S] + i[\mathcal{D}, S] - \frac{1}{2}([\mathcal{D}, S], S). \quad (37)$$

The matrix elements of  $S$  are calculated to be

$$\begin{aligned} \langle n\mathbf{k} | S | l\mathbf{q} \rangle &= i \frac{(H_{\mathbf{k}\cdot\mathbf{p}})_{nl} \delta_{\mathbf{k}, \mathbf{q}} + \langle n\mathbf{k} | \mathcal{D}_{\mathbf{k}\cdot\mathbf{p}} | l\mathbf{q} \rangle}{E_n - E_l} \quad (n \neq l) \\ &= 0 \quad (n = l). \end{aligned} \quad (38)$$

The last three terms in (37) can then be calculated:

$$\langle n\mathbf{k} | -\frac{1}{2}([\mathcal{D}, \mathbf{S}], \mathbf{S}) | l\mathbf{q} \rangle = \frac{1}{2} \sum_{m=c,v} \left[ \frac{(H_{\mathbf{k}\cdot\mathbf{p}})_{nm} \langle m\mathbf{k} | i[\mathcal{D}, \mathbf{S}] | l\mathbf{q} \rangle}{E_n - E_m} + \frac{\langle n\mathbf{k} | i[\mathcal{D}, \mathbf{S}] | m\mathbf{q} \rangle (H_{\mathbf{q}\cdot\mathbf{p}})_{ml}}{E_l - E_m} \right]. \quad (41)$$

Because  $\mathcal{D}$  is block diagonal in the band index and  $H_{\mathbf{k}\cdot\mathbf{p}}$  has no block-diagonal term, (40) is zero within the  $2 \times 2$  and  $6 \times 6$  blocks. Since (40) has only off-diagonal terms between the conduction and valence bands, its contribution is of the order of  $\varepsilon^2$  and is negligible. However, (41) has block-diagonal terms, and is important in describing the strain dependence of the effective masses in a homogeneous strain, which will be discussed below. Including (20), (16), (17), (18), (39), and (41), we have an accuracy that is quadratic in  $\mathbf{k}$  and first order in  $\varepsilon$ , and the  $8 \times 8$  matrix is divided into  $2 \times 2$  and  $6 \times 6$  blocks. In the  $8 \times 8$  form, the effective Hamiltonian is given by

$$\bar{H}_\varepsilon = \bar{H} + \bar{D}^{(1)} + \bar{D}^{(2)}, \quad (42)$$

where  $\bar{H}$  is the unstrained Hamiltonian in the  $\mathbf{k}\cdot\mathbf{p}$  perturbation given by (A22), with  $\mathbf{k}$  replaced by  $-i\nabla'$ ;  $\bar{D}^{(1)}$  is given by the matrix  $\mathcal{D}$  in (A12) for the homogeneous strain with replacement  $\varepsilon \rightarrow \varepsilon(\mathbf{x}')$  plus  $\Delta(\mathbf{x}')(E_n - E)$  in diagonal terms.  $\bar{D}^{(2)}$  is a higher-order effect corresponding to the strain variation of the effective masses and strain gradient terms:

$$\begin{aligned} \bar{D}_{nl}^{(2)} = & \frac{1}{2} [\tilde{\mathcal{D}}_{\mathbf{k}\cdot\mathbf{p}} H_{\mathbf{k}\cdot\mathbf{p}} - \mathcal{D}_{\mathbf{k}\cdot\mathbf{p}} \tilde{H}_{\mathbf{k}\cdot\mathbf{p}}]_{nl} + \frac{1}{2} [\tilde{\mathcal{D}}_{\mathbf{k}\cdot\mathbf{p}} H_{\mathbf{k}\cdot\mathbf{p}} - \mathcal{D}_{\mathbf{k}\cdot\mathbf{p}} \tilde{H}_{\mathbf{k}\cdot\mathbf{p}}]_{ln}^* + \frac{1}{2} [\{\tilde{H}_{\mathbf{k}\cdot\mathbf{p}} \mathcal{D}_{\mathbf{k}\cdot\mathbf{p}}\} - \{H_{\mathbf{k}\cdot\mathbf{p}} \tilde{\mathcal{D}}_{\mathbf{k}\cdot\mathbf{p}}\}]_{nl} \\ & + \frac{1}{2} [(\bar{D}^{(1)} \tilde{H}_{\mathbf{k}\cdot\mathbf{p}}^2)_{nl} + (\bar{D}^{(1)} \tilde{H}_{\mathbf{k}\cdot\mathbf{p}}^2)_{ln}^*] + [\tilde{H}_{\mathbf{k}\cdot\mathbf{p}} \{\tilde{H}_{\mathbf{k}\cdot\mathbf{p}} \bar{D}^{(1)}\}]_{nl} - (\tilde{H}_{\mathbf{k}\cdot\mathbf{p}} \bar{D}^{(1)} \tilde{H}_{\mathbf{k}\cdot\mathbf{p}})_{nl} \\ & - [\{\tilde{H}_{\mathbf{k}\cdot\mathbf{p}} \bar{D}^{(1)}\} \tilde{H}_{\mathbf{k}\cdot\mathbf{p}}]_{nl} + \sum_{i,j} \left[ \frac{\hbar^2}{m} \nabla_i \varepsilon_{ij} \nabla_j - \frac{\hbar^2}{2m} \nabla_j \varepsilon_{ii} \nabla_j \right] \delta_{nl} + \Delta(E_n - E) \delta_{nl}, \end{aligned} \quad (43)$$

where  $\mathbf{k}$  in  $H_{\mathbf{k}\cdot\mathbf{p}}$  is understood as  $-i\nabla'$ , and  $H_{\mathbf{k}\cdot\mathbf{p}}$  is given in (A12), “\*” indicates the complex conjugate, and

$$(\tilde{H}_{\mathbf{k}\cdot\mathbf{p}})_{nm} = \frac{(H_{\mathbf{k}\cdot\mathbf{p}})_{nm}}{(E_n - E_m)}, \quad (\tilde{\mathcal{D}}_{\mathbf{k}\cdot\mathbf{p}})_{nm} = \frac{(\mathcal{D}_{\mathbf{k}\cdot\mathbf{p}})_{nm}}{(E_n - E_m)},$$

with

$$\begin{aligned} (\mathcal{D}_{\mathbf{k}\cdot\mathbf{p}})_{nl} = & i \sum_{i,j} \left[ \frac{2\hbar P_{nl}^j}{m} \varepsilon_{ij} \nabla'_i + \frac{\hbar P_{nl}^j}{2m} (\nabla'_i \varepsilon_{ij}) \right] \\ & - i \frac{\hbar}{m} \Delta(\mathbf{x}') \mathbf{P}_{nl} \cdot \nabla'. \end{aligned}$$

$\mathcal{D}_{\mathbf{k}\cdot\mathbf{p}}$  can be obtained from  $D_{\mathbf{k}\cdot\mathbf{p}}$ , (A18), by replacing  $\varepsilon_{ij} k_i$  with  $-i(2\varepsilon_{ij} \nabla'_i + \frac{1}{2} \nabla'_i \varepsilon_{ij})$  plus  $\Delta(\mathbf{x}') H_{\mathbf{k}\cdot\mathbf{p}}$ . When  $H_{\mathbf{k}\cdot\mathbf{p}}$  is in front of  $\mathcal{D}_{\mathbf{k}\cdot\mathbf{p}}$  or  $\bar{D}^{(1)}$ , and when they are associated within curly brackets  $\{\}$ ,  $\nabla'$  is to be understood as acting on  $\varepsilon$ . Only first derivatives of the strain tensor have been kept in (43). Equation (42) contains only block-diagonal terms in  $2 \times 2$  and  $6 \times 6$  blocks.

Apart from the diagonal term  $\Delta(\mathbf{x}')(E_n - E) \delta_{nl}$  in  $\bar{D}^{(1)}$ ,  $\bar{H} + \bar{D}^{(1)}$  is analogous to the effective-mass approximation in the long-range impurity problem, and it is a natural extension of the theory of Ref. 1 to the case of degenerate bands.  $\bar{D}^{(2)}$  in (42) is analogous to terms which appear in the long-range impurity problem when we go beyond the

effective-mass approximation.<sup>10</sup> The general expression for  $\bar{D}^{(2)}$  is quite complicated, but it can be calculated from (43). To see how  $\bar{D}^{(2)}$  affects the electron states, we consider the conduction band only. Since the effective Hamiltonian (42) is block diagonalized, the  $2 \times 2$  block of (42) describes eigenstates associated with the conduction band. After being transformed back to the  $\mathbf{x}$  coordinate, the  $2 \times 2$  block is

$$\bar{H}_\varepsilon = \bar{H} + D^{(1)} + D^{(2)}, \quad (44)$$

where  $\bar{H}$  and  $D^{(1)}$  are diagonal, and given by  $A_0$  in (A23) with  $\mathbf{k} \rightarrow -i\nabla$ , and  $a_0$  in (A21)  $\varepsilon \rightarrow \varepsilon(\mathbf{x})$ , respectively.  $D^{(2)}$  is separated into two parts:

$$D^{(2)} = D_t^{(2)} + D_g^{(2)}, \quad (45)$$

where  $D_t^{(2)}$  describes the effect of the strain variation of the effective masses, and  $D_g^{(2)}$  describes the effect of the gradient terms. They are the following:

$$D_t^{(2)} = \begin{bmatrix} M_{11} & 0 \\ 0 & M_{11} \end{bmatrix}, \quad (46)$$

$$D_g^{(2)} = \begin{bmatrix} G_{11} & G_{12} \\ -G_{12}^* & G_{11}^* \end{bmatrix}, \quad (47)$$

where

$$\begin{aligned} M_{11} = & \frac{2P_0^2}{3} \left[ \frac{1}{E_g} + \frac{1}{E_g + E_s} \right] \sum_{i,j} \varepsilon_{ij} \nabla_i \nabla_j + \frac{P_0^2}{3} \left[ \frac{2(a_c - a')}{E_g^2} + \frac{a_c - a'}{(E_g + E_s)^2} \left( 1 + \frac{2E_s}{3a'} \right) \right] \text{tr}(\varepsilon) \nabla^2 \\ & - b' \frac{P_0^2}{3} \left[ \frac{1}{E_g^2} + \frac{2 \left( 1 + \frac{E_s}{3b'} \right)}{(E_g + E_s) E_g} \right] \left[ 3 \sum_i \varepsilon_{ii} \nabla_i^2 - \text{tr}(\varepsilon) \nabla^2 \right] - \frac{d' \sqrt{3} P_0^2}{3} \left[ \frac{1}{E_g^2} + \frac{2 \left( 1 + \frac{\sqrt{3} E_s}{3d'} \right)}{(E_g + E_s) E_g} \right] \sum_{i,(i \neq j)} \varepsilon_{ij} \nabla_i \nabla_j, \end{aligned} \quad (48)$$

$$\begin{aligned}
G_{11} = & \frac{2P_0^2}{3} \left[ \frac{2}{E_g} + \frac{1}{E_g + E_s} \right] \sum_i [\nabla_i \text{tr}(\epsilon)] \nabla_j + \frac{P_0^2}{3} \left[ \frac{2(a_c - a')}{E_g^2} + \frac{a_c - a' \left[ 1 + \frac{2E_s}{3a'} \right]}{(E_g + E_s)^2} \right] \sum_i [\nabla_i \text{tr}(\epsilon)] \nabla_i \\
& - b' \frac{P_0^2}{3} \left[ \frac{1}{E_g^2} + \frac{2 \left[ 1 + \frac{E_s}{3b'} \right]}{(E_g + E_s)E_g} \right] \left\{ 3 \sum_i (\nabla_i \epsilon_{ii}) \nabla_i - \sum_i [\nabla_i \text{tr}(\epsilon)] \nabla_i \right\} \\
& - \frac{d' \sqrt{3} P_0^2}{3} \left[ \frac{1}{E_g^2} + \frac{2 \left[ 1 + \frac{\sqrt{3}E_s}{3d'} \right]}{(E_g + E_s)E_g} \right] \sum_{i,j(i \neq j)} (\nabla_i \epsilon_{ij}) \nabla_j - ia' \frac{P_0^2}{3} \left[ \frac{1}{E_g^2} - \frac{\left[ 1 + \frac{2E_s}{3a'} \right]}{(E_g + E_s)^2} \right] \{ [\nabla_1 \text{tr}(\epsilon)] \nabla_2 - [\nabla_2 \text{tr}(\epsilon)] \nabla_1 \} \\
& - ib' \frac{P_0^2}{3} \left[ \frac{1}{E_g^2} - \frac{\left[ 1 + \frac{E_s}{3b'} \right]}{(E_g + E_s)E_g} \right] \{ [\nabla_1 (\epsilon_{11} + \epsilon_{22} - 2\epsilon_{33})] \nabla_2 - [\nabla_2 (\epsilon_{11} + \epsilon_{22} - 2\epsilon_{33})] \nabla_1 \} \\
& - id' \frac{\sqrt{3} P_0^2}{3} \left[ \frac{1}{E_g^2} - \frac{\left[ 1 + \frac{\sqrt{3}E_s}{3d'} \right]}{(E_g + E_s)E_g} \right] \{ (\nabla_1 \epsilon_{23}) \nabla_3 - (\nabla_3 \epsilon_{23}) \nabla_1 + (\nabla_3 \epsilon_{13}) \nabla_2 - (\nabla_2 \epsilon_{13}) \nabla_3 \}, \tag{49} \\
G_{12} = & a' \frac{P_0^2}{3} \left[ \frac{1}{E_g^2} - \frac{\left[ 1 + \frac{2E_s}{3a'} \right]}{(E_g + E_s)^2} \right] \{ (\nabla_1 + i \nabla_2) \text{tr}(\epsilon) \nabla_3 - [\nabla_3 \text{tr}(\epsilon)] (\nabla_1 + i \nabla_2) \} \\
& + b' \frac{P_0^2}{3} \left[ \frac{1}{E_g^2} - \frac{\left[ 1 + \frac{2E_s}{3b'} \right]}{(E_g + E_s)E_g} \right] \{ [\nabla_1 (\epsilon_{11} + \epsilon_{33} - 2\epsilon_{22})] \nabla_3 - [\nabla_3 (\epsilon_{11} + \epsilon_{33} - 2\epsilon_{22})] \nabla_1 \\
& \quad + i [\nabla_2 (\epsilon_{22} + \epsilon_{33} - 2\epsilon_{11})] \nabla_3 - i [\nabla_3 (\epsilon_{22} + \epsilon_{33} - 2\epsilon_{11})] \nabla_2 \} \\
& + d' \frac{\sqrt{3} P_0^2}{3} \left[ \frac{1}{E_g^2} - \frac{\left[ 1 + \frac{\sqrt{3}E_s}{3d'} \right]}{(E_g + E_s)E_g} \right] \{ [\nabla_1 (\epsilon_{23} - i \epsilon_{13})] \nabla_2 - [\nabla_2 (\epsilon_{23} - i \epsilon_{13})] \nabla_1 \\
& \quad + i [(\nabla_1 - i \nabla_2) \epsilon_{12}] \nabla_3 - i [\nabla_3 \epsilon_{12}] (\nabla_1 - i \nabla_2) \} \\
& - \frac{P_0^2}{3} \left[ \frac{1}{E_g} - \frac{1}{E_g + E_s} - \frac{E_c - E}{E_g^2} + \frac{E_c - E}{(E_g + E_s)^2} \right] \{ (\nabla_1 + i \nabla_2) \text{tr}(\epsilon) \nabla_3 - [\nabla_3 \text{tr}(\epsilon)] (\nabla_1 + i \nabla_2) \}. \tag{50}
\end{aligned}$$

Here  $\text{tr}(\epsilon)$  is the dilatation  $\Delta(\mathbf{x})$ ,  $E_g$  is the band gap with spin-orbit coupling,  $E_s$  is the spin-orbit splitting,  $P_0$  is a constant describing the conduction-band–valence-band coupling,  $a_c$  is the deformation potential for the conduction band, and  $a'$ ,  $b'$ , and  $d'$  are effective deformation potentials related to the normally defined deformation potentials  $a$ ,  $b$ , and  $d$  for the valence band. The constants  $P_0$ ,  $a_c$ ,  $a'$ ,  $b'$ , and  $d'$  are defined in Appendix A.  $D_i^{(2)}$  is real and diagonal,  $D_g^{(2)}$  is complex and nondiagonal.  $D_g^{(2)}$  in general will remove the spin degeneracy of the conduction band, due to coupling to the valence subbands which have different spin-orbit shifts. Without the coupling to the valence band, the spin quantum number  $m_s$  would be a good quantum number, and thus the degeneracy would remain.

Note that even when we do not explicitly include  $p'_i \epsilon_{ij}$  terms in the original Hamiltonian (8), gradient terms due to the coupling to the valence band still appear in (49) and (50). They originate from noncommutation of the strain field with the momentum operators. The gradient terms due to  $E_s \neq 0$  would have not appeared if we had made a symmetric replacement<sup>12</sup> to the homogeneous strain result following Ref. 12. In general, the symmetric replacement is adequate only for a subset of bands whose strain coupling to others bands is negligible. For instance, the eight-band model (33') can be obtained from the result for homogeneous strain, because strain-induced coupling to the remote bands is ignored. It is inadequate for the two-band model for the conduction band (44), because strain-induced coupling to the valence band is in-

cluded.

Notice that  $M_{11} + G_{11}$  and  $G_{12}$  are Hermitian operators. Although each operator in  $D^{(2)}$  is Hermitian,  $D^{(2)}$  is not Hermitian as a whole, i.e.,  $G_{21} \neq G_{12}^*$ . This is due to the hidden degeneracy of the electron bound states. In fact, the structure of  $D^{(2)}$  is similar to the matrix for the spin-orbit term in (1) in a basis without spin-orbit coupling. In spite of the removal of the spin degeneracy, the bound states are still Kramers degenerate due to time-reversal symmetry, and the Hermiticity can be recovered after the degeneracy is properly considered, as is discussed in Appendix B. If spin-orbit splitting is sufficiently small or the band gap sufficiently large, the result is much simpler, and (47) is approximately diagonal and real.

To first order, i.e., in the so-called effective-mass approximation,<sup>1,10</sup> we do not need to consider  $D^{(2)}$ . However,  $D_i^{(2)}$  may be important when the strain is so large that it significantly affects the effective mass. On the other hand, when the strain variation in position is so rapid

(though not so rapid as to invalidate the basic assumptions of this calculation) that  $D_g^{(2)}$  is comparable to  $D_i^{(2)}$ , the gradient terms in  $D^{(2)}$  are not negligible. This is the situation when  $a_B k_0 \sim 1$ , where  $a_B$  is the radius of the envelope function and  $k_0$  is a typical wave vector in the Fourier transform of the strain. Note that without  $D_g^{(2)}$ ,  $D^{(2)}$  is non-Hermitian except for the case of homogeneous strain. In narrow-band-gap semiconductors  $D^{(2)}$  is relatively more important, as in the impurity problem.<sup>20</sup>

The results above for inhomogeneous strain can be adapted to the impurity problem discussed in Ref. 20 by keeping only terms associated with the deformation potentials  $a_c$  and  $a$ . One sets  $\varepsilon_{ij} = \varepsilon_{ii} \delta_{ij}$ ,  $a_c = a$ , and  $\text{atr}(\varepsilon) = U(\mathbf{x})$ , where  $U(\mathbf{x})$  is the impurity potential. More discussion of the impurity problem is given in Appendix B.

For homogeneous strain, from (44) we have the strain dependence of the conduction band up to terms quadratic in  $\mathbf{k}$ :

$$\begin{aligned}
 E_c(\varepsilon, \mathbf{k}) = & a_c \text{tr}(\varepsilon) + \left[ \frac{\hbar^2}{2m} + A' + \frac{P_0^2}{3} \left( \frac{2}{E_g} + \frac{1}{E_g + E_s} \right) \right] k^2 - \frac{2P_0^2}{3} \left( \frac{2}{E_g} + \frac{1}{E_g + E_s} \right) \sum_{i,j} \varepsilon_{ij} k_i k_j \\
 & - \frac{P_0^2}{3} \left[ \frac{2(a_c - a')}{E_g^2} + \frac{a_c - a' \left( 1 + \frac{2E_s}{3a'} \right)}{(E_g + E_s)^2} \right] \text{tr}(\varepsilon) k^2 + b' \frac{P_0^2}{3} \left[ \frac{1}{E_g^2} + \frac{2 \left( 1 + \frac{E_s}{3b'} \right)}{(E_g + E_s) E_g} \right] \left[ 3 \sum_i \varepsilon_{ii} k_i^2 - \text{tr}(\varepsilon) k^2 \right] \\
 & + \frac{d' \sqrt{3} P_0^2}{3} \left[ \frac{1}{E_g^2} + \frac{2 \left( 1 + \frac{\sqrt{3} E_s}{3d'} \right)}{(E_g + E_s) E_g} \right] \sum_{i,j (i \neq j)} \varepsilon_{ij} k_i k_j, \quad (51)
 \end{aligned}$$

where  $\mathbf{k}$  is the wave vector in  $\mathbf{x}$  space.

For hydrostatic pressure with  $\varepsilon_{ij} = \varepsilon \delta_{ij}$ , (51) becomes

$$\begin{aligned}
 E_c(\varepsilon, \mathbf{k}) = & \left[ \frac{\hbar^2}{2m} + A' + \frac{P_0^2}{3} \left( \frac{2}{E_g} + \frac{1}{E_g + E_s} \right) \right] k^2 + 3a_c \varepsilon \\
 & - \varepsilon \left\{ \frac{2P_0^2}{3} \left( \frac{2}{E_g} + \frac{1}{E_g + E_s} \right) + P_0^2 \left[ \frac{2 \left( a_c - a + \frac{2E_s}{9} \right)}{E_g^2} + \frac{a_c - a - \frac{4E_s}{9}}{(E_g + E_s)^2} \right] \right\} k^2. \quad (52)
 \end{aligned}$$

The variation of the effective mass  $m_c$  with pressure is then given by

$$\frac{m}{m_c} = 1 + \frac{E_p}{3} \left[ \frac{2}{E_g} \left\{ 1 - \varepsilon \left[ 2 + \frac{3 \left( a_c - a + \frac{2E_s}{9} \right)}{E_g} \right] \right\} + \frac{1}{E_g + E_s} \left\{ 1 - \varepsilon \left[ 2 + \frac{3 \left( a_c - a - \frac{4E_s}{9} \right)}{E_g + E_s} \right] \right\} \right] + \frac{2m}{\hbar^2} A. \quad (53)$$

This result basically agrees with Ref. 19. The difference is due to the inclusion of the contribution of the spin-orbit interaction term  $D_{SO}$  in (53). (The absolute deformation potentials of  $\Gamma_8$  and  $\Gamma_7$  are  $a - 2E_s/9$  and  $a + 4E_s/9$ , respectively.) We notice that the pressure variation of the effective mass depends, as it must, only on the difference between the deformation potentials of the conduction band and the valence bands, contrary to Eq. (6) of Ref. 21. For GaAs at 0 K, we obtain

$m_c(\varepsilon) = m_c(0) [1 + 5.85 \times 10^{-3} (\text{kbar}^{-1}) |X|]$ , where  $X$  is the hydrostatic pressure in kbar. This result agrees well with the experimental result of  $6.15 \times 10^{-3} \text{kbar}^{-1}$  at 4.2 K.<sup>22</sup> The calculation in Ref. 19 gave  $9.39 \times 10^{-3} \text{kbar}^{-1}$  at 300 K. Parameters used are  $E_g = 1.5192$  eV,  $E_s = 0.341$  eV,  $E_p = 22.71$  eV,<sup>23</sup>  $\varepsilon = 4.29 \times 10^{-4} X$ ,<sup>24</sup> and  $a_c - (a - 2E_s/9) = -8.31$  eV corresponding to  $dE_g/dX = -10.7$  meV/kbar.<sup>25</sup> While the lack of inversion symmetry is not considered in the above calculation,

it does not affect the result for the hydrostatic pressure, as can be seen from (A14).

### III. APPLICATION TO Laterally Strain-Confined Quantum Wells

Next, the application to laterally strain-confined quantum wells, such as the quantum wire or dot structures based on GaAs/Ga<sub>1-x</sub>Al<sub>x</sub>As quantum wells,<sup>4,26</sup> will be considered. It is assumed that the quantum-well confinement is along the  $z$  direction, and that the strain modulation of the quantum-well potential is negligible. In the effective-mass approximation, i.e., neglecting  $D^{(2)}$ , and assuming large spin-orbit splitting, we have a one-band model for the conduction band and a four-band model for the valence band:

$$\left[ \frac{p_z}{2m_c} + a_c \sum_i \varepsilon_{ii}(\mathbf{x}) + V_c(z) \right] F = E_c F, \quad (54)$$

$$[-\bar{H} - D^{(1)} + V_v(z)1] \begin{pmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \end{pmatrix} = E_h \begin{pmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \end{pmatrix}, \quad (55)$$

where  $E_c$  and  $E_h$  are electron and hole levels related to the band edges of the bulk conduction and valence bands,  $V_c(z)$  and  $V_v(z)$  are the quantum-well confining potentials for the electron and holes,  $\bar{H}$  is the effective Hamiltonian for the  $4 \times 4$  block associated with the heavy and light holes in (A22), and  $D^{(1)}$  is the  $4 \times 4$  block of  $\mathcal{D}$  in (A12). In (55), the sign of the energy has been inverted.

In general,  $\varepsilon$  is a function of  $x$ ,  $y$ , and  $z$ . If in (54) and (55) we had to treat the quantum-well confinement in the  $z$  direction on an equal footing with the lateral  $(x, y)$  confinement, as would be necessary if the strain varied appreciably over the width of the well, the problem would be very difficult, even for the conduction band. However, for currently obtainable strains ( $< 1\%$ ) (Ref. 4) and narrow quantum wells, the  $z$  dependence of the strain is negligible. For the conduction band the motion along the  $z$  and  $(x, y)$  directions is then decoupled: the strain potential can be taken as  $z$  independent, with its value at the center of the well, and the calculation reduces to that of Ref. [4].

For the valence band, the problem is still complicated. Even for an unstrained quantum well, motion along the  $z$  direction is coupled to the in-plane motion ( $k_{\parallel} \neq 0$ ) because of the anisotropy of the valence bands. For the strained quantum well, the inhomogeneous strain mixes states with different  $k_{\parallel}$ , so that the motion along the  $z$  and  $(x, y)$  directions becomes coupled by off-diagonal terms of the operators  $p_x p_z$  and  $p_y p_z$  in  $\bar{H}$ . This is a common problem for any laterally confined system, not just for the strain-confined structures considered here. We notice that there will not be two independent deformation potentials for "heavy" and "light" holes, respectively. In some previous work,<sup>4,5</sup> so-called band-edge modulations for heavy and light holes were calculated by diagonalizing  $D^{(1)}$  while ignoring all terms associated with momen-

tum operators in the four simultaneous differential equations (55). In general this is not justified.

If the strain varies sufficiently slowly along the  $z$  direction, and if we neglect strain-induced coupling among the quantum-well subbands except between the first heavy-hole and first light-hole subbands [this can be shown to be similar to the effect of neglecting  $D^{(2)}$ ], (54) and (55) can be simplified to the following form:

$$\left[ \frac{(p_x^2 + p_y^2)}{2m_c} + a_c \text{tr}(\varepsilon) + E_n \right] g_c = E_c g_c, \quad (56)$$

$$(H_{\text{well}} + H_{\text{strain}}) \begin{pmatrix} g_1 \\ g_2 \\ g_3 \\ g_4 \end{pmatrix} = E_h \begin{pmatrix} g_1 \\ g_2 \\ g_3 \\ g_4 \end{pmatrix}, \quad (57)$$

where

$$H_{\text{well}} = \begin{pmatrix} H_{\text{hh}} + E_{\text{hh}} & 0 & 0 & 0 \\ 0 & H_{\text{lh}} + E_{\text{lh}} & 0 & 0 \\ 0 & 0 & H_{\text{lh}} + E_{\text{lh}} & 0 \\ 0 & 0 & 0 & H_{\text{hh}} + E_{\text{hh}} \end{pmatrix},$$

$$H_{\text{strain}} = - \begin{pmatrix} -p + q & s^+ & r^+ & 0 \\ s & -p - q & 0 & -r^+ \\ r & 0 & -p - q & s^+ \\ 0 & -r & s & -p + q \end{pmatrix}.$$

$g_i(x, y)$  are the envelope function for the in-plane motions;  $H_{\text{hh}}$  and  $H_{\text{lh}}$  are in-plane kinetic-energy operators corresponding to the in-plane dispersions of heavy and light holes in the quantum well;  $E_n$ ,  $E_{\text{hh}}$ , and  $E_{\text{lh}}$  are electron, heavy-hole, and light-hole energies, respectively, at  $k_{\parallel} = 0$ ;  $a_c \text{tr}(\varepsilon)$ ,  $p$ ,  $q$ ,  $r$ , and  $s$  are strain potentials defined in (A21) with average strain  $e_{ij}(x, y) = \langle f_n | \varepsilon_{ij}(x, y, z) | f_m \rangle$  instead of  $\varepsilon_{ij}(x, y, z)$ ;  $f_n(z)$  are envelope functions for the electron, heavy holes, or light holes. Due to the slow variation of the strain, we may set  $e_{ij}(x, y) \approx \alpha \varepsilon_{ij}(x, y, 0)$  with  $\alpha = \langle f_n | f_m \rangle$ , where the origin of the coordinates has been chosen at the center of the well. Equation (57) indicates that the in-plane motion of the heavy and light holes is coupled by the lateral confinement.

To simplify the problem further, we assume that the heavy- and light-hole splitting is much greater than the in-plane kinetic energies. Then (57) can be reduced to two independent equations by using the same method as in the impurity problem<sup>20</sup> to decouple the conduction and valence bands. For the heavy and light holes, respectively, we obtain

$$\left[ H_{\text{hh}} + p - q - \frac{|\alpha|^2(|r|^2 + |s|^2)}{\delta E + p + q - E_h} \right] g_h = E_h g_h, \quad (58)$$

$$\left[ H_{\text{lh}} + p + q + \frac{|\alpha|^2(|r|^2 + |s|^2)}{\delta E - p + q + E_l} \right] g_l = E_l g_l, \quad (59)$$

where  $\delta E$  is the splitting between heavy- and light-hole bands at  $k_{\parallel} = 0$ , and  $E_h$  and  $E_l$  are energy levels for the



heavy and light holes, referred to the band edges of the heavy- and light-hole subbands in the quantum well, respectively. Note that the eigenvalue equations (58) and (59) are nonlinear in the eigenvalues  $E_h$  or  $E_l$ . This situation is not unique, as it is found in the impurity problem and others.<sup>20</sup> In the above equations (55)–(59), we do not have a picture of a particle with a certain mass moving in a potential. However, if  $|E_h| \ll |\delta E + p + q|$  (the diagonal term for the light hole) and  $|E_l| \ll |\delta E - p + q|$  (the diagonal term for the heavy hole), we obtain two simple effective-mass equations for the heavy and light holes, respectively:

$$\left[ H_{hh} + p - q - \frac{|\alpha|^2(|r|^2 + |s|^2)(\delta E + 2p)}{(\delta E + p + q)^2} \right] g_h = E_h g_h, \quad (60)$$

$$\left[ H_{lh} + p + q + \frac{|\alpha|^2(|r|^2 + |s|^2)(\delta E - 2p)}{(\delta E - p + q)^2} \right] g_l = E_l g_l. \quad (61)$$

The specific strain-confined quantum wire structure of Ref. 4 will be taken as an example of the above theory. The wire is along the [100] direction, which is taken as the  $x$  axis. For this structure the piezoelectric effect is minimized (the piezoelectric field is zero except near the ends of the wire).<sup>14</sup> The deformation potential term  $a_c \text{tr}(\epsilon)$  in (56) and the effective potentials in (60) and (61), evaluated for this structure, are shown in Figs. 1(a) and 1(b). The band-edge modulations of the valence band calculated from equations in Ref. 4 (in its Ref. 24) are also shown. In fact, the band-edge modulations are very close to the effective potentials for this specific structure. Potentials without the off-diagonal terms  $r$  and  $s$  are also plotted to show the effect of heavy-hole–light-hole cou-

pling. With increasing  $\delta E$  the coupling should become less important. The effective potential for the heavy hole has an attractive central well, but with off-center minima. For the light hole, if the in-plane effective mass were positive, it would be barely attractive at the center, with a finite barrier. However, for the specific quantum well in Ref. 4 (with 12-nm width), the light hole has negative in-plane effective mass for small  $k_{\parallel}$ ,<sup>27</sup> so that the effective attractive potential has a double-well structure. Thus, off-center bound states are expected for both heavy and light holes.

Equations (56), (60), and (61) have been solved numerically to obtain energy levels for the quantum wire structure of Ref. 4, with the potentials shown in Fig. 1. For the conduction band, the lower-lying bound states are almost equally spaced with interval about 2.5 meV, in agreement with Ref. 4, whose authors correctly assumed that the bottom of the potential well is close to parabolic.<sup>4</sup> There are 11 bound states associated with the conduction band, as shown in Fig. 2(a). For the heavy hole, although the depth of the potential well is only about 7 meV, there are as many as 22 bound states due to the relatively large in-plane effective mass.<sup>27</sup> Due to the double-well-like potential, some lower-lying bound states are doubly degenerate, and their envelope functions are concentrated near two “wells” which are close to the edge or even outside the wire. However, the envelope functions of some shallow bound states have significant amplitude near the center of the wire. For the light hole with negative in-plane effective mass,<sup>27</sup> the effective potential has a double-well structure, and there exist some eigenstates with negative binding energies. Except for their negative binding energies, these states are similar to the heavy-hole bound states, for instance, their envelope functions have finite extents. The depth of the effective potential is about the same as that for the heavy hole, but

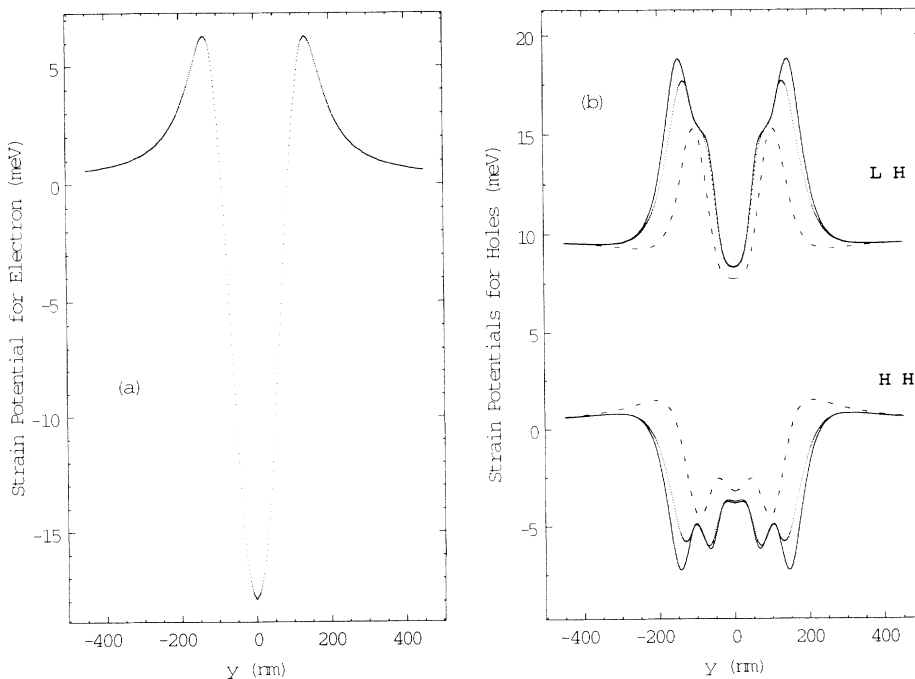


FIG. 1. Strain potentials for the quantum wire structure in Ref. 4 (a) for the conduction band and (b) for the heavy and light holes. The potential for the light hole is displaced by  $\Delta E = 10$  meV. Solid line: effective potentials including heavy-hole–light-hole coupling. Dotted line: band-edge modulations including heavy-hole–light-hole coupling by the theory of Ref. 4. Dashed line: potentials given by the diagonal terms. The sign of the energy for the valence band has been inverted from the conduction band.

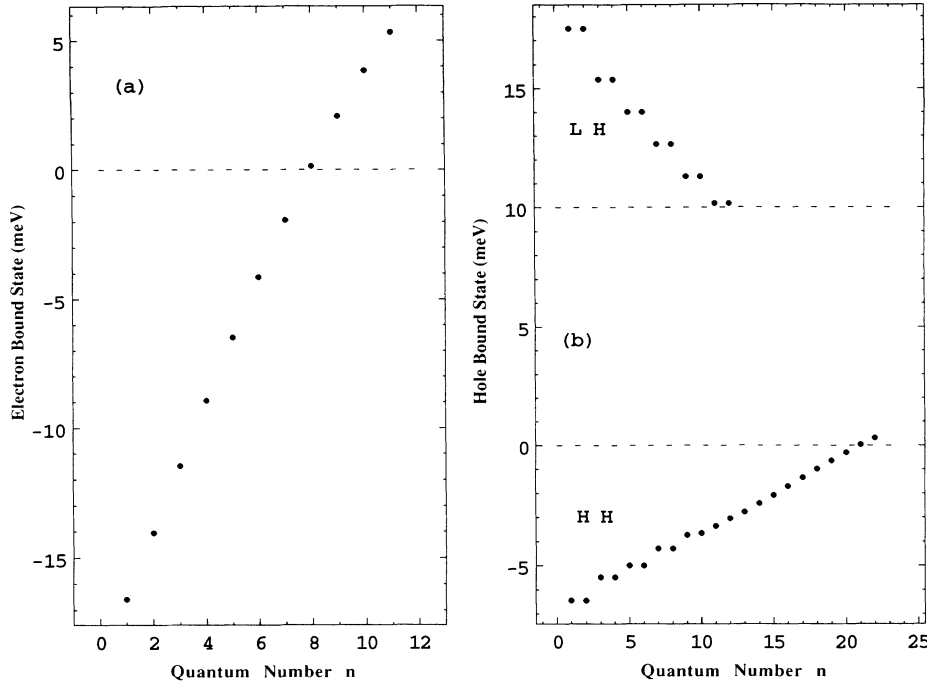


FIG. 2. Energy levels of the bound states in the quantum wire (a) for the conduction band, and (b) for the heavy and light holes. The energy references are the bottom of the conduction band and the top of the valence band in the quantum well, respectively. Odd  $n$  for even-parity envelope functions, even  $n$  for odd parity.

there are only 12 “bound states” for the light hole because of its smaller in-plane effective mass. Due to the inversion symmetry of the strain potentials, all the envelope functions have definite parity.

Because the in-plane dispersion of the light hole has a “camel’s back” (Ref. 28) structure, some light-hole bound states are resonant with band states with relative large  $k_{\parallel}$ . However, their envelope functions are mostly made up of band states with small  $k_y$  (the wire being along  $x$ ). On the other hand, the envelope function of the light-hole exciton will include contributions from states with quite large  $k_x$ , outside the negative mass region. As the light-hole effective mass changes greatly with the width of the well,<sup>27</sup> it will be interesting to make quantum wires or dots on the quantum wells with different width.

One of the major requirements for the validity of the theory is that the strain must be slowly varying over the unit cell. The rapidity of the strain variation is determined mainly by the size of the stressor. For the 180-nm-wide wire in Ref. 4, the variation of the strain over one lattice constant is smaller than  $3 \times 10^{-4}$  (about 1% of the strain).

In all the above calculations,  $D^{(2)}$  has not been considered. In general, when  $D^{(2)}$  is non-negligible, the strain effect on effective masses of the quantum well should be taken into account at the same time. This is a very complicated problem. However, without going to the detailed calculation, it can be shown that  $D^{(2)}$  can indeed make a significant contribution within the range of applicability of the theory, but it is actually negligible in the above example.

Before going to the case of inhomogeneous strain, let us consider the contribution of  $D^{(2)}$  to the hydrostatic term. For the conduction band of GaAs, assuming  $\epsilon = 1\%$  (corresponding to about 23-kbar hydrostatic pressure), (53) gives a change in the effective mass of 12%,

which indicates that the contribution of  $D^{(2)}$  can be significant within the range of applicability of the theory. Although this result is for homogeneous strain, it gives us a rough idea when  $D^{(2)}$  may become important for the more general case of inhomogeneous strain. For inhomogeneous strain, the effect of the strain gradient may also contribute when  $a_B k_0 \sim 1$ . The contribution of  $D^{(2)}$  depends on the specific strain configuration. For the wire structure of Ref. 4, the strain is smaller than 0.2%, so neglect of the  $D^{(2)}$  term is valid in this case. However, stronger confinement has been achieved in some strain-confined systems; for instance, about a 60-meV redshift was achieved in a quantum dot structure,<sup>26</sup> compared to about 20 meV in the wire.<sup>4</sup> When the strain approaches 1%, the  $D^{(2)}$  term has to be considered.

To examine  $D^{(2)}$  more closely, let us consider the conduction band of the GaAs. Ignoring terms related to the spin-orbit splitting of the valence band, the  $D^{(2)}$  term in the effective Hamiltonian is then

$$\delta \bar{H}_c = \frac{\hbar^2}{2m_c} \left\{ 5.32 \sum_i \nabla_i \epsilon_{ii} \nabla_i - 6.06 \sum_i \nabla_i \text{tr}(\epsilon) \nabla_i + 6.84 \sum_{i \neq j} \nabla_i \epsilon_{ij} \nabla_j \right\}. \quad (62)$$

Including  $D^{(2)}$ , the motion along the  $z$  and  $(x, y)$  directions becomes coupled in general. For the above-mentioned [100] wire, the only nonzero off-diagonal component of the strain tensor is  $\epsilon_{yz}$ , thus motion along the  $y$  and  $z$  directions is coupled. In this case, the correction of  $\delta \bar{H}_c$  to the kinetic energy in the strain-confined direction  $y$  is only about 1–2%. Roughly speaking, the correction in a dot is twice of that in a wire with the same magnitude of the strain. Thus, for a dot with strain of 1%, the correction may be necessary. In (62), the terms related to

variation of the effective mass with strain are combined with the terms related to a strain gradient to give a Hermitian operator. Their contributions are comparable for structures with a stressor size of order of 100 nm, in which the extension is similar to that of the envelope function.

#### IV. SUMMARY

The envelope-function method has been applied to slowly varying inhomogeneous strain in a semiconductor with a degenerate band structure. The effects of strain variation of the kinetic energy and the gradient of the strain have been considered in general in the eight-band manifold, and given explicitly for the conduction bands. Applications to homogeneous strains and hydrostatic pressure were discussed.

The theory has been applied to laterally confined quantum wells, quantum wires, and quantum dots. Under the effective-mass approximation, the motion of the holes in degenerate valence bands under the inhomogeneous strain should be described by multiband envelope-function equations (55). For the case of the strain varying negligibly on the scale of the width of the quantum well, they can be simplified to (57). If the strain confinement is weak relative to the heavy- and light-hole splitting in the wells, they can be further simplified to two decoupled equations (58) and (59) or (60) and (61).

To first order, the strain variation of the effective mass and the gradient terms does not affect the electronic states. When the strain is not sufficiently small and does not vary sufficiently slowly, we must go beyond the effective-mass approximation. This second-order effect can be significant within the range of applicability of the theory. It can alter not only the energy levels of the electronic states but also their degeneracy. When applied to strain-confined quantum wires or quantum dots, it is necessary to consider the second-order effect for structures with relative large strain.

Numerical results for the energy levels in a specific quantum wire structure were also given.

#### ACKNOWLEDGMENTS

I am very grateful to Professor C. Herring, Professor E. I. Rashba, Professor Weikun Ge, and Professor W. E. Lawrence, and to Dr. T. B. Bahder, Dr. C. Van de Walle, and Dr. K. Kash for helpful discussions and comments on the manuscript, to Dr. D. Gershoni for a copy of unpublished material, and especially to Professor M. D. Sturge for his supervision and critical reading and discussion throughout the course of this work. This work was supported by the Office of Naval Research under Grant No. N001491J1429.

#### APPENDIX A

In  $\mathbf{k}\cdot\mathbf{p}$  perturbation theory, the strain Hamiltonian is given by<sup>8,11</sup>

$$H_\varepsilon = H_0 + H_k + H_{\mathbf{k}\cdot\mathbf{p}} + H'_{\text{SO}} + D_0 + D_k + D_{\mathbf{k}\cdot\mathbf{p}} + D_{\text{SO}} + D'_{\text{SO}}, \quad (\text{A1})$$

where

$$H_0 = \frac{\mathbf{p}^2}{2m} + V_0 + \frac{\hbar}{4m^2c^2}(\nabla V_0) \times \mathbf{p} \cdot \boldsymbol{\sigma}, \quad (\text{A2})$$

$$H_k = \frac{\hbar^2 k^2}{2m}, \quad (\text{A3})$$

$$H_{\mathbf{k}\cdot\mathbf{p}} = \frac{\hbar}{m} \mathbf{k} \cdot \mathbf{p}, \quad (\text{A4})$$

$$H'_{\text{SO}} = \frac{\hbar^2}{4m^2c^2}(\nabla V_0) \times \mathbf{k} \cdot \boldsymbol{\sigma}, \quad (\text{A5})$$

$$D_k = -\sum_{ij}^3 \varepsilon_{ij} \frac{\hbar^2 k_i k_j}{m}, \quad (\text{A6})$$

$$D_0 = \sum_{ij}^3 \varepsilon_{ij} \left[ -\frac{p_i p_j}{m} + V_{ij} \right], \quad (\text{A7})$$

$$D_{\mathbf{k}\cdot\mathbf{p}} = \frac{2\hbar}{m} \mathbf{k} \cdot \boldsymbol{\varepsilon} \cdot \mathbf{p}, \quad (\text{A8})$$

$$D_{\text{SO}} = \frac{\hbar}{4m^2c^2} \left\{ \sum_{ij}^3 \varepsilon_{ij} (\nabla V_{ij}) \times \mathbf{p} \cdot \boldsymbol{\sigma} - [\boldsymbol{\varepsilon} \cdot (\nabla V_0)] \times \mathbf{p} \cdot \boldsymbol{\sigma} - (\nabla V_0) \times (\boldsymbol{\varepsilon} \cdot \mathbf{p}) \cdot \boldsymbol{\sigma} \right\}, \quad (\text{A9})$$

$$D'_{\text{SO}} = \frac{\hbar^2}{4m^2c^2} \left\{ \sum_{ij}^3 \varepsilon_{ij} (\nabla V_{ij}) \times \mathbf{k} \cdot \boldsymbol{\sigma} - [\boldsymbol{\varepsilon} \cdot (\nabla V_0)] \times \mathbf{k} \cdot \boldsymbol{\sigma} \right\}, \quad (\text{A10})$$

where  $H_\varepsilon$  is in the  $\mathbf{x}'$  space, and  $\mathbf{k}$  is the wave vector in  $\mathbf{x}'$  space.  $H_\varepsilon$  has been calculated<sup>11</sup> in an eight-band model by using Löwdin perturbation. To derive Eq. (33) for the inhomogeneous strain, we need the result for the homogeneous strain. The eight band-edge states given by Kane,<sup>29</sup> which are different from those in Ref. 11, are chosen in our calculation:

$$\begin{aligned} u_1 &= |\Gamma_6, -\frac{1}{2}\rangle = |S\rangle\beta, \\ u_2 &= |\Gamma_6, \frac{1}{2}\rangle = |S\rangle\alpha, \\ u_3 &= |\Gamma_8, -\frac{3}{2}\rangle = 1/\sqrt{2}(|X\rangle - i|Y\rangle)\beta, \\ u_4 &= |\Gamma_8, -\frac{1}{2}\rangle = -1/\sqrt{6}[(|X\rangle - i|Y\rangle)\alpha + 2|Z\rangle\beta], \\ u_5 &= |\Gamma_8, \frac{1}{2}\rangle = 1/\sqrt{6}[(|X\rangle + i|Y\rangle)\beta - 2|Z\rangle\alpha], \\ u_6 &= |\Gamma_8, \frac{3}{2}\rangle = 1/\sqrt{2}(|X\rangle + i|Y\rangle)\alpha, \\ u_7 &= |\Gamma_7, -\frac{1}{2}\rangle = 1/\sqrt{3}[-(|X\rangle - i|Y\rangle)\alpha + |Z\rangle\beta], \\ u_8 &= |\Gamma_7, \frac{1}{2}\rangle = 1/\sqrt{3}[(|X\rangle + i|Y\rangle)\beta + |Z\rangle\alpha], \end{aligned} \quad (\text{A11})$$

where  $|S\rangle, |X\rangle, |Y\rangle$ , and  $|Z\rangle$  are band-edge states without spin-orbit interaction, and  $\alpha$  and  $\beta$  are spinors for spin-up and -down.  $H_\varepsilon$  is calculated in the above basis by Löwdin perturbation as in Ref. 11, neglecting  $H'_{\text{SO}}$ ,  $D'_{\text{SO}}$  and the first term in  $D_{\text{SO}}$ .<sup>11</sup> Keep in mind that (A1)

is presented in  $\mathbf{x}'$  space. For homogeneous strain, the matrix representation of  $H_\varepsilon$  is easy to change back to  $\mathbf{x}$  space, simply by deleting  $D_k$  and the factor of 2 in  $D_{k,p}$ . Thus, with the wave vector in  $\mathbf{x}$  space,  $H_\varepsilon$  is given as the following:

$$H_\varepsilon = \mathcal{H} + H_{k,p} + \mathcal{D} + D_{k,p}, \quad (\text{A12})$$

where  $\mathcal{H}$  includes  $H_0$ ,  $H_k$ , and the coupling to remote bands by  $H_{k,p}$ ,  $\mathcal{D} = D_0 + D_{SO} \equiv \mathcal{D}_1 + \mathcal{D}_2$ , and  $H_{k,p}$  and  $D_{k,p}$  contain only coupling between the conduction and valence bands.  $\mathcal{H}$  contains only quadratic terms in  $\mathbf{k}$ , but  $H_{k,p}$  and  $D_{k,p}$  will give corrections to  $\mathcal{H}$  and  $\mathcal{D}$  to any order of  $\mathbf{k}$  within the eight-band manifold. These matrices are

$$\mathcal{H} = \begin{pmatrix} A_0 & 0 & \sqrt{3}T & -\sqrt{2}W & T^+ & 0 & W & \sqrt{2}T^+ \\ 0 & A_0 & 0 & -T & -\sqrt{2}W & \sqrt{3}T^+ & -\sqrt{2}T & W \\ \sqrt{3}T^+ & 0 & -P+Q & S^+ & R^+ & 0 & -\frac{1}{\sqrt{2}}S^+ & \sqrt{2}R^+ \\ -\sqrt{2}W & -T^+ & S & -P-Q & 0 & -R^+ & \sqrt{2}Q & \sqrt{\frac{3}{2}}S^+ \\ T & -\sqrt{2}W & R & 0 & -P-Q & S^+ & -\sqrt{\frac{3}{2}}S & \sqrt{2}Q \\ 0 & \sqrt{3}T & 0 & -R & S & -P+Q & -\sqrt{2}R & -\frac{1}{\sqrt{2}}S \\ W & -\sqrt{2}T^+ & -\frac{1}{\sqrt{2}}S & \sqrt{2}Q & -\sqrt{\frac{3}{2}}S^+ & -\sqrt{2}R^+ & Z & 0 \\ \sqrt{2}T & W & \sqrt{2}R & \sqrt{\frac{3}{2}}S & \sqrt{2}Q & -\frac{1}{\sqrt{2}}S^+ & 0 & Z \end{pmatrix}, \quad (\text{A13})$$

$$\mathcal{D}_1 = \begin{pmatrix} a_0 & 0 & \sqrt{3}t & -\sqrt{2}w & t^+ & 0 & w & \sqrt{2}t^+ \\ 0 & a_0 & 0 & -t & -\sqrt{2}w & \sqrt{3}t^+ & -\sqrt{2}t & w \\ \sqrt{3}t^+ & 0 & -p+q & s^+ & r^+ & 0 & -\frac{1}{\sqrt{2}}s^+ & \sqrt{2}r^+ \\ -\sqrt{2}w & -t^+ & s & -p-q & 0 & -r^+ & \sqrt{2}q & \sqrt{\frac{3}{2}}s^+ \\ t & -\sqrt{2}w & r & 0 & -p-q & s^+ & -\sqrt{\frac{3}{2}}s & \sqrt{2}q \\ 0 & \sqrt{3}t & 0 & -r & s & -p+q & -\sqrt{2}r & -\frac{1}{\sqrt{2}}s \\ w & -\sqrt{2}t^+ & -\frac{1}{\sqrt{2}}s & \sqrt{2}q & -\sqrt{\frac{3}{2}}s^+ & -\sqrt{2}r^+ & z & 0 \\ \sqrt{2}t & w & \sqrt{2}r & \sqrt{\frac{3}{2}}s & \sqrt{2}q & -\frac{1}{\sqrt{2}}s^+ & 0 & z \end{pmatrix}, \quad (\text{A14})$$

$$\mathcal{D}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}}s'^+ & \sqrt{2}r'^+ \\ 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2}q' & \sqrt{\frac{3}{2}}s'^+ \\ 0 & 0 & 0 & 0 & 0 & 0 & -\sqrt{\frac{3}{2}}s' & \sqrt{2}q' \\ 0 & 0 & 0 & 0 & 0 & 0 & -\sqrt{2}r' & -\frac{1}{\sqrt{2}}s' \\ 0 & 0 & -\frac{1}{\sqrt{2}}s' & \sqrt{2}q' & -\sqrt{\frac{3}{2}}s'^+ & -\sqrt{2}r'^+ & z' & 0 \\ 0 & 0 & \sqrt{2}r' & \sqrt{\frac{3}{2}}s' & \sqrt{2}q' & -\frac{1}{\sqrt{2}}s'^+ & 0 & z' \end{pmatrix}, \quad (\text{A15})$$

$$H_{k,p} = \begin{pmatrix} 0 & KP \\ KP^+ & 0 \end{pmatrix}, \quad (\text{A16})$$

with

$$KP = \begin{pmatrix} \sqrt{3}V^+ & \sqrt{2}U^+ & -V & 0 & U & -\sqrt{2}V \\ 0 & -V^+ & \sqrt{2}U^+ & -\sqrt{3}V & -\sqrt{2}V^+ & U \end{pmatrix}, \quad (\text{A17})$$

$$D_{k,p} = \begin{pmatrix} 0 & kp \\ kp^+ & 0 \end{pmatrix}, \quad (\text{A18})$$

with

$$kp = \begin{pmatrix} \sqrt{3}v^+ & \sqrt{2}u^+ & -v & 0 & u & -\sqrt{2}v \\ 0 & -v^+ & \sqrt{2}u^+ & -\sqrt{3}v & -\sqrt{2}v^+ & u \end{pmatrix}, \quad (\text{A19})$$

where

$$\begin{aligned} A_0 &= E_1 + (A' + \hbar^2/2m)(k_1^2 + k_2^2 + k_3^2), \\ P &= -E_2 + \gamma_1 \hbar^2/2m(k_1^2 + k_2^2 + k_3^2), \\ Q &= -\gamma_2 \hbar^2/2m(k_1^2 + k_2^2 - 2k_3^2), \\ R &= -\sqrt{3} \hbar^2/2m[\gamma_2(k_1^2 - k_2^2) - 2i\gamma_3 k_1 k_2], \\ S &= 2\sqrt{3} \gamma_3 \hbar^2/2m k_3(k_1 - ik_2), \\ Z &= E_2 - E_s - \gamma_1 \hbar^2/2m(k_1^2 + k_2^2 + k_3^2), \\ V &= P_0/\sqrt{6}(k_2 - ik_1), \\ U &= iP_0/\sqrt{3}k_3, \\ T &= B/\sqrt{6}(k_2 - ik_1)k_3, \\ W &= B/\sqrt{3}k_1 k_2, \end{aligned} \quad (\text{A20})$$

$$\begin{aligned} a_0 &= a_c(\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}), \\ p &= -a'(\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}), \\ q &= b'/2(\varepsilon_{11} + \varepsilon_{22} - 2\varepsilon_{33}), \\ r &= \sqrt{3}/2b'(\varepsilon_{11} - \varepsilon_{22}) - id'\varepsilon_{12}, \\ s &= -d'(\varepsilon_{13} - i\varepsilon_{23}), \\ z &= a'(\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}), \\ v &= P_0/\sqrt{6}[(i\varepsilon_{11} - \varepsilon_{21})k_1 + (i\varepsilon_{12} - \varepsilon_{22})k_2 \\ &\quad + (i\varepsilon_{13} - \varepsilon_{23})k_3], \\ u &= -iP_0/\sqrt{3}(\varepsilon_{31}k_1 + \varepsilon_{32}k_2 + \varepsilon_{33}k_3), \\ t &= B'/\sqrt{6}(\varepsilon_{23} - i\varepsilon_{13}), \\ w &= B'/\sqrt{3}\varepsilon_{12}, \\ q' &= \delta b/2(\varepsilon_{11} + \varepsilon_{22} - 2\varepsilon_{33}), \\ r' &= \sqrt{3}/2\delta b(\varepsilon_{11} - \varepsilon_{22}) - i\delta d\varepsilon_{12}, \\ s' &= -\delta d(\varepsilon_{13} - i\varepsilon_{23}), \\ z' &= \delta a(\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}). \end{aligned} \quad (\text{A21})$$

where  $E_1$ ,  $E_2$ , and  $E_2 - E_s$  are band edges of  $\Gamma_6$ ,  $\Gamma_8$ , and  $\Gamma_7$ , respectively;  $E_g = E_1 - E_2$  is then the band gap, and  $E_s$  is the spin-orbit splitting.  $A'$  is due to the coupling between the conduction band and the remote bands,<sup>11</sup>  $\gamma_1$ ,  $\gamma_2$ , and  $\gamma_3$  are so-called modified Luttinger parameters due to the coupling between the valence band and the remote bands;<sup>11</sup>  $B$  and  $B'$  (for  $b'$  in Ref. 11) are parameters for the lack of inversion symmetry;<sup>11</sup>  $P_0 = -i\hbar/m \langle S | p_x | X \rangle$ ;  $a' = a - 2E_s/9$ ,  $b' = b - 2E_s/9$ , and  $d' = d - 2\sqrt{3}E_s/9$ , where  $a_c$ ,  $a$ ,  $b$ , and  $d$  are deformation potentials defined in Ref. 8,  $a_c = \langle S | -p_x p_x + V_{xx} | S \rangle$ ,  $a = (l + 2m)/3$ ,  $b = (l - m)/3$ , and  $d = n/\sqrt{3}$ , where  $l = \langle X | -p_x p_x + V_{xx} | X \rangle$ ,  $m = \langle Y | -p_x p_x + V_{xx} | Y \rangle$ , and  $n = 2 \langle X | -p_x p_z + V_{xz} | Z \rangle$ ; and  $\delta a = 2E_s/3$ ,  $\delta b = E_s/3$ , and  $\delta d = \sqrt{3}E_s/3$ . These three new deformation potentials appear because  $D_0 + D_{SO}$  is reorganized as  $\mathcal{D}_1 + \mathcal{D}_2$ . + has been used for the Hermitian conjugate in this Appendix.

It should be pointed out that due to the difference in the basis used here and Ref. 11, the dispersion and strain behavior of heavy-hole  $|\Gamma_8, \pm \frac{3}{2}\rangle$  and light-hole states  $|\Gamma_8, \pm \frac{1}{2}\rangle$  are switched, for example, along the [001] direction or under [001] uniaxial stress.

Another perturbation procedure ("transformation method") different from Löwdin's was proposed,<sup>30</sup> which is in fact based on the same idea used by Luttinger and Kohn<sup>10</sup> in the impurity problem, and is employed in the main text of this paper. The advantage of this "transformation method" is that the perturbation can be worked out order by order. Up to terms quadratic in  $\mathbf{k}$ , the lack of inversion symmetry has no effect in this order for the unstrained crystal, and the effective Hamiltonian for the unstrained crystal will be block diagonal:

$$\bar{H} = \begin{pmatrix} A_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & A_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -P+Q & S^+ & R^+ & 0 & -\frac{1}{\sqrt{2}}S'^+ & \sqrt{2}R'^+ \\ 0 & 0 & S & -P-Q & 0 & -R^+ & \sqrt{2}Q' & \sqrt{\frac{3}{2}}S'^+ \\ 0 & 0 & R & 0 & -P-Q & S^+ & -\sqrt{\frac{3}{2}}S' & \sqrt{2}Q' \\ 0 & 0 & 0 & -R & S & -P+Q & -\sqrt{2}R' & -\frac{1}{\sqrt{2}}S' \\ 0 & 0 & -\frac{1}{\sqrt{2}}S' & \sqrt{2}Q' & -\sqrt{\frac{3}{2}}S'^+ & -\sqrt{2}R'^+ & Z & 0 \\ 0 & 0 & \sqrt{2}R' & \sqrt{\frac{3}{2}}S' & \sqrt{2}Q' & -\frac{1}{\sqrt{2}}S'^+ & 0 & Z \end{pmatrix}, \quad (\text{A22})$$

where

$$\begin{aligned}
A_0 &= E_1 + \{ A' + \hbar^2/2m + P_0^2/3[2/E_g + 1/(E_g + \Delta)] \} \\
&\quad \times (k_1^2 + k_2^2 + k_3^2), \\
P &= -E_2 + (\gamma_1 + \delta\gamma_1)\hbar^2/2m (k_1^2 + k_2^2 + k_3^2), \\
Q &= -(\gamma_2 + \delta\gamma_2)\hbar^2/2m (k_1^2 + k_2^2 - 2k_3^2), \\
R &= -\sqrt{3}\hbar^2/2m [(\gamma_2 + \delta\gamma_2)(k_1^2 - k_2^2) \\
&\quad - 2i(\gamma_3 + \delta\gamma_3)k_1k_2], \\
S &= 2\sqrt{3}(\gamma_3 + \delta\gamma_3)\hbar^2/2mk_3(k_1 - ik_2), \\
Z &= E_2 - E_s - (\gamma_1 + \delta\gamma_1)\hbar^2/2m (k_1^2 + k_2^2 + k_3^2), \\
Q' &= -(\gamma_2 + \delta\gamma_2)\hbar^2/2m (k_1^2 + k_2^2 - 2k_3^2),
\end{aligned} \tag{A23}$$

$$\begin{aligned}
R' &= -\sqrt{3}\hbar^2/2m [(\gamma_2 + \delta\gamma_2)(k_1^2 - k_2^2) \\
&\quad - 2i(\gamma_3 + \delta\gamma_3)k_1k_2], \\
S' &= 2\sqrt{3}(\gamma_3 + \delta\gamma_3)\hbar^2/2mk_3(k_1 - ik_2),
\end{aligned}$$

where  $\delta\gamma_1 = E_p/(3E_g)$ ,  $\delta\gamma_2 = \delta\gamma_3 = E_p/(6E_g)$ ,  $\delta\gamma_1' = E_p/[3(E_g + E_s)]$ , and  $\delta\gamma_2' = \delta\gamma_3' = E_p/12[1/E_g + 1/(E_g + E_s)]$ . We notice that in  $\mathcal{H}$  unique values of  $Q$ ,  $R$ , and  $S$  appear in the  $6 \times 6$  block, but not in  $\bar{H}$ .

## APPENDIX B

From results in Sec. II, the envelope-function equations for the bound states associated with the conduction band under an inhomogeneous strain can be written as

$$\begin{pmatrix} \frac{p^2}{2m_c} + a_c \text{tr}(\epsilon) + M_{11} + G_{11} & G_{12} \\ -G_{12}^* & \frac{p^2}{2m_c} + a_c \text{tr}(\epsilon) + M_{11} + G_{11}^* \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = E \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}. \tag{B1}$$

In this appendix, we want to demonstrate that even when the coupling to the valence band removes the spin degeneracy of the conduction band, the bound states are still Kramer degenerate, and the Hermiticity of the effective Hamiltonian can be recovered after considering the degeneracy. Also, the extension to the impurity problem will be discussed.

We notice that the operator  $D_g^{(2)}$  is complex and can be written as  $G_{11} = G_{11}^R + iG_{11}^I$  and  $G_{12} = G_{12}^R + iG_{12}^I$ , where  $G_{11}^R$  and  $G_{12}^R$  stand for the real parts, and  $G_{11}^I$  and  $G_{12}^I$  for the imaginary parts. We also try to write the envelope functions as  $F_1 = F_{11} + iF_{12}$  and  $F_2 = F_{21} + iF_{22}$ . Substituting these definitions in (B1), we then have

$$\begin{pmatrix} H_d & -G_{11}^I & G_{12}^R & -G_{12}^I \\ G_{11}^I & H_d & G_{12}^I & G_{12}^R \\ -G_{12}^R & -G_{12}^I & H_d & G_{11}^I \\ G_{12}^I & -G_{12}^R & -G_{11}^I & H_d \end{pmatrix} \begin{pmatrix} F_{11} \\ F_{12} \\ F_{21} \\ F_{22} \end{pmatrix} = E \begin{pmatrix} F_{11} \\ F_{12} \\ F_{21} \\ F_{22} \end{pmatrix}, \tag{B2}$$

where

$$H_d = \frac{p^2}{2m_c} + a_c \text{tr}(\epsilon) + M_{11} + G_{11}^R. \tag{B3}$$

This is a set of equations for four eigenstates, and the effective Hamiltonian is Hermitian. Introducing a unitary transformation

$$U = \begin{pmatrix} \frac{i}{\sqrt{2}} 0 & 0 \frac{1}{\sqrt{2}} \\ 0 \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} 0 \\ 0 \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} 0 \\ \frac{1}{\sqrt{2}} 0 & 0 \frac{i}{\sqrt{2}} \end{pmatrix}, \tag{B4}$$

and applying it to Eq. (B2), we arrive at

$$\begin{pmatrix} H_d + iG_{12}^I & -G_{11}^I - iG_{12}^R & 0 & 0 \\ G_{11}^I - iG_{12}^R & H_d - iG_{12}^I & 0 & 0 \\ 0 & 0 & H_d + iG_{12}^I & G_{11}^I + iG_{12}^R \\ 0 & 0 & -G_{11}^I + iG_{12}^R & H_d - iG_{12}^I \end{pmatrix} \begin{pmatrix} F'_{11} \\ F'_{12} \\ F'_{21} \\ F'_{22} \end{pmatrix} = E \begin{pmatrix} F'_{11} \\ F'_{12} \\ F'_{21} \\ F'_{22} \end{pmatrix}, \tag{B5}$$

where  $F'_{ij}$  are linear combinations of  $F_{ij}$ . The solutions then give twofold degenerate levels, with  $F'_{11} \sim F'_{22}^*$  and  $F'_{12} \sim F'_{21}^*$ .

The theory for the inhomogeneous strain can be used to derive the equations for the long-range impurity problem as discussed in Sec. II of the main text. For donor states associated with the conduction band, we have

$$\begin{pmatrix} \frac{p^2}{2m_c} + U(\mathbf{x}) + G_{11} & G_{12} \\ -G_{12}^* & \frac{p^2}{2m_c} + U(\mathbf{x}) + G_{11}^* \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = E \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}, \quad (\text{B6})$$

where  $U(\mathbf{x})$  is the impurity potential, and

$$G_{11} = -\frac{P_0^2}{3} \left[ \frac{2}{E_g^2} + \frac{1}{(E_g + E_s)^2} \right] \sum_i (\nabla_i U) \nabla_i - i \frac{P_0^2}{3} \left[ \frac{1}{E_g^2} - \frac{1}{(E_g + E_s)^2} \right] [(\nabla_1 U) \nabla_2 - (\nabla_2 U) \nabla_1], \quad (\text{B7})$$

$$G_{12} = \frac{P_0^2}{3} \left[ \frac{1}{E_g^2} - \frac{1}{(E_g + E_s)^2} \right] \{ [(\nabla_1 + i\nabla_2)U] \nabla_3 - (\nabla_3 U)(\nabla_1 + i\nabla_2) \}. \quad (\text{B8})$$

Compared with the results given in Ref. 20, it is found that  $G_{22}(=G_{11}^*)$  and  $G_{21}(=-G_{12}^*)$  in fact agree with (B27) and (B28) of Ref. 20 [the order of the basis functions is changed from  $(S\alpha, S\beta)$  there to  $(S\beta, S\alpha)$  here, and (B27) there can be simplified to (B7)]. While the diagonal terms were given to be equal in Ref. 20, they are actually not.

<sup>1</sup>J. Bardeen and W. Shockely, Phys. Rev. **80**, 72 (1950).

<sup>2</sup>G. D. Whitfield, Phys. Rev. **121**, 720 (1961).

<sup>3</sup>L. J. Sham and J. M. Zimann, *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1963), Vol. 15, p. 221.

<sup>4</sup>K. Kash, B. P. Van der Gaag, Derek D. Mahoney, A. S. Gozdz, L. T. Florez, J. P. Harbison, and M. D. Sturge, Phys. Rev. Lett. **67**, 1326 (1991).

<sup>5</sup>D. Gershoni, J. S. Weiner, S. N. G. Chu, G. A. Baraff, J. M. Vandenberg, L. N. Pfeiffer, K. West, R. A. Logan, and T. Tanbun-Ek, Phys. Rev. Lett. **65**, 1631 (1990).

<sup>6</sup>Z. Xu and P. M. Petroff, J. Appl. Phys. **69**, 6564 (1991).

<sup>7</sup>G. E. Pikus and G. L. Bir, Fiz. Tverd. Tela. (Leningrad) **1**, 1642 (1959) [Sov. Phys. Solid State **1**, 1502 (1960)].

<sup>8</sup>G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (Wiley, New York, 1974), p. 295.

<sup>9</sup>R. S. Markiewicz, J. P. Wolfe, and C. D. Jeffries, Phys. Rev. B **15**, 1988 (1977).

<sup>10</sup>J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955); J. Callaway, *Quantum Theory of the Solid State*, 2nd ed. (Academic, New York, 1991), p. 412.

<sup>11</sup>T. B. Bahder, Phys. Rev. B **41**, 11 992 (1989); **46**, 9913 (1992).

<sup>12</sup>G. A. Baraff and D. Gershoni, Phys. Rev. B **43**, 4011 (1991); D. Gershoni, C. H. Henry, and G. A. Baraff, IEEE J. Quantum Electron **QE-29**, 2433 (1993).

<sup>13</sup>K. B. Tolpygo, Fiz. Tverd. Tela. (Leningrad) **4**, 1765 (1962) [Sov. Phys. Solid State **4**, 1297 (1963)]; Sh. M. Kogan, *ibid.* **5**, 2829 (1963) [**5**, 2069 (1964)].

<sup>14</sup>K. Kash and Derek D. Mahoney, Bull. Am. Phys. Soc. **37**, 187 (1992); K. Kash (private communication). The piezoelectric potential is given by

$$V_{\text{piez}}(\mathbf{x}) = \frac{3}{4\pi\epsilon_0\epsilon} \int_0^x \frac{\nabla \cdot \mathbf{P}}{|\mathbf{x} - \mathbf{r}|} d\mathbf{r},$$

where  $\mathbf{P} = e_{14}(\epsilon_{yz}, \epsilon_{zx}, \epsilon_{xy})$  is the strain-induced polarization,  $e_{14}$  is the piezoelectric constant, and  $\epsilon$  is the dielectric con-

stant of the semiconductor.

<sup>15</sup>P. Lawaetz, Phys. Rev. **183**, 730 (1969); C. G. Van de Walle and R. M. Martin, Phys. Rev. Lett. **62**, 2028 (1989).

<sup>16</sup>Theorist, Allan Bonadio Associates, Prescience Corporation, 1989; S. Wolfram, *Mathematica-A System for Doing Mathematics by Computer* (Addison-Wesley, New York, 1991).

<sup>17</sup>A. I. Lur'e, *Three-Dimensional Problems of the Theory of Elasticity* (Interscience, New York, 1964), p. 14.

<sup>18</sup>C. Herring (private communication) points out that, in general, the strain potential is not local but will have an integral form. I have ignored this complication.

<sup>19</sup>D. E. Aspnes and M. Cardona, Phys. Rev. B **17**, 726 (1978); **17**, 741 (1978).

<sup>20</sup>G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Les Éditions de Physique, Les Ulis Cedex, 1988), p. 59.

<sup>21</sup>T. B. Bahder, Phys. Rev. B **45**, 1629 (1992).

<sup>22</sup>Z. Wasilewski and R. A. Stradling, Semicond. Sci. Technol. **1**, 264 (1986).

<sup>23</sup>G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Ref. 20), p. 49.

<sup>24</sup>S. Adachi, J. Appl. Phys. **58**, R1 (1985).

<sup>25</sup>D. J. Wolford and J. A. Bradely, Solid State Commun. **53**, 1069 (1985).

<sup>26</sup>K. Kash, Derek D. Mahoney, B. P. Van der Gaag, A. S. Gozdz, J. P. Harbison, and L. T. Florez, J. Vac. Sci. Technol. B **10**, 2030 (1992).

<sup>27</sup>G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Ref. 20), p. 110.

<sup>28</sup>P. Lawaetz, Solid State Commun. **16**, 65 (1975).

<sup>29</sup>E. O. Kane, *Handbook of Semiconductors*, edited by T. S. Moss (North-Holland, Amsterdam, 1982), Vol. 1, p. 212.

<sup>30</sup>W. Fawcett and T. P. McLean, Proc. Phys. Soc. **85**, 1315 (1965).