Auger recombination in semiconductor quantum wells

Anatoli S. Polkovnikov and Georgy G. Zegrya
Ioffe Physico-Technical Institute, Politechnicheskaya st. 26, 194021, St. Petersburg, Russia
(Received 27 August 1997; revised manuscript received 27 February 1998)

The principal mechanisms of Auger recombination of nonequilibrium carriers in semiconductor heterostructures with quantum wells (QW's) are investigated. It is shown that there exist three fundamentally different Auger recombination mechanisms of (i) thresholdless, (ii) quasithreshold, and (iii) threshold types. The rate of the thresholdless Auger process depends on temperature only slightly. The threshold energy of the quasithreshold process essentially varies with QW width and is close to zero for narrow QW's. It is shown that the thresholdless and the quasithreshold Auger processes dominate in narrow QW's, while the threshold and the quasithreshold processes prevail in wide QW's. The limiting case of a three-dimensional (3D) Auger process is reached for infinitely wide QW's. The critical QW width is found at which the quasithreshold and threshold Auger processes merge into a single 3D Auger process. Also studied is phonon-assisted Auger recombination in QW's. It is shown that for narrow QW's the act of phonon emission becomes resonant, which in turn increases substantially the coefficient of phonon-assisted Auger recombination. [S0163-1829(98)00628-6]

I. INTRODUCTION

Two recombination processes are predominant in semiconductors at high excitation levels: radiative and nonradiative Auger recombination (AR). For homogeneous semiconductors, mechanisms of AR have been extensively studied.[1–4] In narrow gap semiconductors there occur AR processes involving two electrons and a heavy hole (CHCC Auger process) or an electron and two heavy holes, with transition of one of the holes to the spin-orbit split-off (SO) band (CHHS Auger process).[2,4,5] Both of these processes are of threshold nature, and the rate of Auger recombination changes with temperature exponentially.[1,2] The only exceptions are semiconductors in which the spin-orbit splitting is close to the energy gap (GaSb and InAs). Under certain conditions the rate of the CHHS process in these semiconductors depends on temperature only slightly.[6] It is commonly believed that the phonon-assisted AR process dominates the direct one at low temperature.[3,5] Because of the large momentum transferred to a phonon, the threshold is eliminated and the rate of phonon-assisted Auger process becomes a power-law function of temperature. However, the carrier-carrier scattering can also cause large momentum transfer to the electron-hole subsystem and thus eliminate the threshold conditions. The problem of the predominant AR mechanism in bulk semiconductors needs further consideration.

Single semiconductor heterostructures, quantum wells (QW's), quantum wires, and quantum dots are spatially inhomogeneous owing to the existence of barriers. The presence of a heteroboundary affects not only the energy and wave functions of carriers, but also the macroscopic properties of heterostructures.[7] It is commonly believed that the AR mechanism in QW's is similar to that in homogeneous semiconductors.[5,8–11] Nevertheless, the heteroboundary lifts restrictions imposed on the electron-electron interaction process by the energy and momentum conservation laws. Namely, the conservation of quasimomentum perpendicular to the heteroboundary breaks down. In turn, this leads to the appearance in heterostructures of new thresholdless channels of Auger recombination.[7] The rate of the thresholdless AR process is a power function of temperature. A direct experiment aimed at observing the thresholdless AR channel at $T=77$ K was reported in Ref. 12.

A detailed analysis of the threshold and thresholdless AR mechanisms has been performed for a single heterobarrier.[7] Conditions were studied under which the thresholdless channel dominates the threshold one. For QW's no such detailed analysis has been done. Taylor *et al.*[13] considered the possibility of threshold for AR in QW's being eliminated upon transition of excited carriers to the continuous part of the spectrum. However, no microscopic theory of the thresholdless process was given in this work, and no theoretical analysis performed of the competition between different AR mechanisms at various temperatures and QW widths. Only the thresholdless AR channel, corresponding to small momenta transferred in the Coulomb interaction of particles, for the CHCC process with spin-orbit interaction neglected, was considered in Refs. 14 and 15.

The aim of the present work is to investigate theoretically the principal mechanisms of AR for nonequilibrium carriers in semiconductor QW's. It will be shown that there exist three fundamentally different AR mechanisms: (i) threshold mechanism similar to the Auger process in a homogeneous semiconductor, (ii) quasithreshold mechanism with a threshold energy strongly depending on the QW width, and (iii) thresholdless mechanism inoperative in a homogeneous semiconductor. All three processes will be shown to have different dependences on temperature and QW parameters. Rates of the processes corresponding to these AR mechanisms at different temperatures and QW parameters will be compared. Conditions will be found under which the quasithreshold and threshold Auger processes merge to form a single three-dimensional (3D) AR process. Also, the phonon-assisted AR in QW's will be studied. The act of phonon emission will be shown to become resonant for sufficiently narrow quantum wells.

The paper is organized as follows. In Sec. II wave functions and the energy spectrum of electrons and holes in a QW are investigated, using Kane's model. In Sec. III the matrix element of Auger transition is calculated. Section IV is concerned with investigation of AR coefficients for three different mechanisms existing in a QW. In Sec. V the phonon-assisted AR mechanism in a QW is considered. Section VI summarizes the results of the paper.

II. PRINCIPAL EQUATIONS

To analyze the AR mechanisms and find the rate of the Auger process, wave functions of charge carriers are to be known. As already established for bulk Auger processes, the wave functions of electrons and holes must be calculated using the multiband approximation.[2] We will use four-band Kane's model, the most adequate describing the wave functions and energy spectrum of carriers in narrow-gap $A_{\mathrm{III}}B_{\mathrm{V}}$ semiconductors.[16]

A. Wave functions in a homogeneous semiconductor

For most $A_{\mathrm{III}}B_{\mathrm{V}}$ semiconductors, wave functions of electrons and holes in the center of the Brillouin zone are described by the Γ_6^+ representation for the conduction band and by the Γ_7^+ and Γ_8^+ representations for the valence band. Of these the first two and the last are doubly and fourfold degenerate, respectively. The corresponding equations for wave functions may be written in differential form. Commonly, the basis wave functions of the conduction and valence bands are taken in form of eigenfunctions of the angular momentum.[16,17] However, another representation of the basis functions is more appropriate for our purposes:

$$|s\uparrow\rangle,\quad |s\downarrow\rangle,\quad |x\uparrow\rangle,\quad |x\downarrow\rangle,\quad |y\uparrow\rangle,\quad |y\downarrow\rangle,\quad |z\uparrow\rangle,\quad |z\downarrow\rangle, \tag{1}$$

where $|s\rangle$ and $|x\rangle$, $|y\rangle$, $|z\rangle$ are the Bloch functions of s and $\mathbf{p}$ type with angular momenta of 0 and 1, respectively. The former describe the state of the conduction band and the latter the state of the valence band at the Γ point. Arrows denote the direction of spin. The wave function of carriers ψ can be written in the form

$$\psi=\Psi_s|s\rangle+\boldsymbol{\Psi}|\mathbf{p}\rangle,$$

where Ψ_s and $\boldsymbol{\Psi}$ are spinors. In the vicinity of the Γ point the equations for Ψ_s and $\boldsymbol{\Psi}$ envelopes written in the spherical approximation are as follows:

$$\begin{gathered}(E_c-E)\Psi_s-i\hbar\,\gamma\nabla\boldsymbol{\Psi}=0,\\ (E_v-\delta-E)\boldsymbol{\Psi}-i\hbar\,\gamma\nabla\Psi_s+\frac{\hbar^2}{2m}(\tilde{\gamma}_1+4\tilde{\gamma}_2)\nabla(\nabla\boldsymbol{\Psi})\\ -\frac{\hbar^2}{2m}(\tilde{\gamma}_1-2\tilde{\gamma}_2)\nabla\times[\nabla\times\boldsymbol{\Psi}]+i\,\delta[\boldsymbol{\sigma}\times\boldsymbol{\Psi}]=0.\end{gathered} \tag{2}$$

Here, γ is Kane's matrix element having dimension of velocity, $\tilde{\gamma}_1$ and $\tilde{\gamma}_2=\tilde{\gamma}_3$ are the generalized Luttinger parameters,[17] $\delta=\Delta_{so}/3$, Δ_{so} is the spin-orbit splitting, E_v and E_c are the energies of the lower edge of the conduction band and the upper edge of the valence band, m is the free-electron mass, and $\boldsymbol{\sigma}=(\sigma_x,\sigma_y,\sigma_z)$ are the Pauli spin matrices. If, instead of using the Luttinger parameters, the heavy-hole mass describing the interaction with higher bands is introduced phenomenologically, then Eq. (2) is transformed into equations derived by Suris.[18] It is easy to verify that Eq. (2) is identical to those commonly used.[17,19–21] In the first equation in the system (2) we neglect the term with the free-electron mass.

1. Hole states

The expression for Ψ_s can be found from the first equation of the system (2). Substitution of Ψ_s into the second equation gives

$$-E\boldsymbol{\Psi}+\frac{\hbar^2}{2m_l}\nabla(\nabla\boldsymbol{\Psi})-\frac{\hbar^2}{2m_h}\nabla\times[\nabla\times\boldsymbol{\Psi}]+i\,\delta[\boldsymbol{\sigma}\times\boldsymbol{\Psi}]=0, \tag{3}$$

where

$$m_l^{-1}=\frac{2\gamma^2}{E_g+\delta-E}+m^{-1}(\tilde{\gamma}_1+4\tilde{\gamma}_2),$$

$$m_h^{-1}=m^{-1}(\tilde{\gamma}_1-2\tilde{\gamma}_2).$$

Here m_h coincides with the heavy-hole mass, and m_l with the light-hole mass in the case of zero constant of spin-orbit interaction; $E_g=E_c-E_v$ is the semiconductor band gap. For the sake of convenience, it is assumed that $E_v=\delta$. This choice is due to an increase in the heavy-hole and light-hole energies at the Γ point by δ and a decrease in the SO hole energy by $2\,\delta$ under the action of spin-orbital interaction [Eq. (8)]. Equation (3) can be simplified by introducing new functions

$$\phi=\mathrm{div}\ \boldsymbol{\Psi}\quad\text{and}\quad\eta=\boldsymbol{\sigma}\ \mathrm{rot}\ \boldsymbol{\Psi}. \tag{4}$$

After taking the divergence and rotor of Eq. (3) multiplied by $\boldsymbol{\sigma}$, it is transformed into a system of two differential equations

$$\begin{gathered}-E\phi+\frac{\hbar^2}{2m_l}\Delta\phi+i\,\delta\eta=0,\\ -(E+\delta)\,\eta+\frac{\hbar^2}{2m_h}\Delta\,\eta-2i\,\delta\phi=0.\end{gathered} \tag{5}$$

Fourier transform of these equations gives hole spectra for a homogeneous semiconductor

$$\begin{bmatrix} E+\dfrac{\hbar^2}{2m_l}k^2 & i\,\delta \\ -2i\,\delta & E+\dfrac{\hbar^2}{2m_h}k^2+\delta \end{bmatrix}\begin{pmatrix}\phi\\ \eta\end{pmatrix}=0. \tag{6}$$

The characteristic equation has two roots

$$E_{1,2}=-\frac{\delta}{2}-\frac{\hbar^2k^2}{4}(m_l^{-1}+m_h^{-1})\pm\sqrt{2\,\delta^2+\left(\frac{\delta}{2}-\frac{\hbar^2k^2}{4}(m_l^{-1}-m_h^{-1})\right)^2}. \tag{7}$$

It should be noted that m_l depends on energy [see Eq. (3)]. At the Γ point ($k=0$) we have the roots $E_1=\delta$ and $E_2=-2\delta$. The positive solution corresponds to light holes, and that with a negative sign, to SO holes.

In the vicinity of the Γ point the energies $E_{1,2}$ can be expanded into a series in terms of a wave vector to relate the effective masses of light and SO holes m_{hl}, m_{so}, and the Luttinger parameters:

$$E_1 \approx \delta - \frac{\hbar^2 k^2}{2m_{hl}}, \quad E_2 \approx -2\delta - \frac{\hbar^2 k^2}{2m_{so}}, \tag{8}$$

where

$$m_{hl}^{-1} = \frac{4\gamma^2}{3E_g} + \frac{(\tilde{\gamma}_1 + 2\tilde{\gamma}_2)}{m}, \quad m_{so}^{-1} = \frac{2\gamma^2}{3(E_g+3\delta)} + \frac{\tilde{\gamma}_1}{m}.$$

An approximate light-hole spectrum can be obtained by means of a widely used 4×4 Hamiltonian.[20] However, the range of its applicability is rather narrow, since commonly $m_l \sim 0.1\, m_h$ and the expansion (8) is only valid when $E \ll (m_l/m_h)\,\Delta_{so}$. Moreover, such a model cannot describe Auger transitions at all, since the basis states of carriers in different bands are orthogonal. The same applies to the spectrum of the SO band.

The Fourier amplitudes of the wave functions of both light and SO holes can be presented in the form

$$\boldsymbol{\Psi} = \mathbf{k}f + \frac{i\delta}{E+\delta+\dfrac{\hbar^2 k^2(E)}{2m_h}}[\mathbf{k}\times\boldsymbol{\sigma}f], \qquad \Psi_s = -\frac{\hbar\gamma k^2(E)}{E_g+\delta-E}f, \tag{9}$$

where f is an arbitrary spinor related to the previously introduced function ϕ by $\phi = k^2(E)f$.

The third solution of Eq. (3) pertaining to heavy holes satisfies the relations div $\boldsymbol{\Psi}=0$ (as a consequence $\Psi_s=0$) and $\boldsymbol{\sigma}$ rot $\boldsymbol{\Psi}=0$. This follows from Eq. (5), since, if $\phi=0$, then $\eta=0$ and vice versa. It can be readily seen that $[\boldsymbol{\sigma}\times\boldsymbol{\Psi}_h] = -i\boldsymbol{\Psi}_h$. Thus, the dispersion law describing the heavy-hole spectrum looks like

$$E_h = \delta - \frac{\hbar^2 k_h^2}{2m_h}. \tag{10}$$

The components of the heavy hole wave function must satisfy the equations

$$\left.\begin{aligned}\Psi_{z\downarrow} &= (\Psi_{x\uparrow} + i\Psi_{y\uparrow})\\ \Psi_{z\uparrow} &= (-\Psi_{x\downarrow} + i\Psi_{y\downarrow})\end{aligned}\right\} \Leftrightarrow [\boldsymbol{\sigma}\times\boldsymbol{\Psi}] = -i\boldsymbol{\Psi}, \tag{11}$$

$$\left.\begin{aligned}k_z\Psi_{z\uparrow} + k_x\Psi_{x\uparrow} + k_y\Psi_{y\uparrow} &= 0\\ k_z\Psi_{z\downarrow} + k_x\Psi_{x\downarrow} + k_y\Psi_{y\downarrow} &= 0\end{aligned}\right\} \Leftrightarrow \text{div } \boldsymbol{\Psi} = 0. \tag{12}$$

By solving these equations, one may obtain explicit expressions for the wave functions. For a QW they are given in Appendix A.

2. Electron states

In principle, the conventional equations for electrons have the same form as those for holes. Since the Γ point in the conduction band is only doubly degenerate and the crystal field causes no additional splitting, there is no need to retain terms with the parameters $\tilde{\gamma}_i$. Moreover, the presence of these terms in the equations for electrons gives a far too exact model. Thus, a simplified set of equations will be used for electrons

$$\begin{aligned}(E_c - E)\Psi_s - i\hbar\gamma\nabla\boldsymbol{\Psi} &= 0,\\ (E_v - \delta - E)\boldsymbol{\Psi} - i\hbar\gamma\nabla\Psi_s + i\delta[\boldsymbol{\sigma}\times\boldsymbol{\Psi}] &= 0.\end{aligned} \tag{13}$$

The electron energies can be conveniently reckoned from the lower edge of the conduction band ($E_c=0$). This energy will be denoted by $\mathcal{E}$, so that it would not be confused with the full energy of electron E, reckoned from the same level as the hole energy. Introducing into Eq. (13) the functions ϕ and η in the same form as in Eq. (4), we have

$$\begin{aligned}-(E_g + \delta + \mathcal{E})\phi + \frac{\hbar^2\gamma^2}{\mathcal{E}}\Delta\phi + i\delta\eta &= 0,\\ -(E_g + \mathcal{E} + 2\delta)\eta - 2i\delta\phi &= 0.\end{aligned} \tag{14}$$

Passing to the Fourier transform, we find the electron dispersion law

$$k^2 = \frac{\mathcal{E}}{\hbar^2\gamma^2}\,\frac{\mathcal{E}^2 + \mathcal{E}(2E_g+3\delta) + (E_g+3\delta)E_g}{E_g+\mathcal{E}+2\delta}. \tag{15}$$

If $\mathcal{E} \ll E_g$ then the energy is quadratic in wave vector,

$$\mathcal{E} = \frac{\hbar^2 k^2}{2m_c} \quad \text{where} \quad m_c^{-1} = 2\gamma^2\,\frac{E_g+2\delta}{(E_g+3\delta)E_g}. \tag{16}$$

The Fourier amplitude of the wave function is given by

$$\Psi_s = f, \quad \boldsymbol{\Psi} = \frac{\mathcal{E}}{\hbar\gamma k^2(\mathcal{E})}\left[\mathbf{k}f + \frac{i\delta}{\mathcal{E}+E_g+2\delta}[\mathbf{k}\times(\boldsymbol{\sigma}f\ )]\right], \tag{17}$$

where f is an arbitrary spinor [see Eq. (9)].

3. Probability flux and the equations near the heteroboundary

An expression for the probability flux density can be derived from Eq. (2) by substituting $E \to -i\hbar\,\partial/\partial t$ and then using a procedure similar to that employed in quantum mechanics.[22] It can also be derived by the $\mathbf{k}\cdot\mathbf{p}$ method in the second-order perturbation theory. As a result, the following expression is obtained in the case of holes:

$$\mathbf{j}_h = \frac{E_g+\delta-E}{2m_l\gamma}[\Psi_s\boldsymbol{\Psi}^* + \Psi_s^*\boldsymbol{\Psi}] - \frac{i\hbar}{2m_h}(\boldsymbol{\Psi}\times\text{rot }\boldsymbol{\Psi}^* - \boldsymbol{\Psi}^*\times\text{rot }\boldsymbol{\Psi}). \tag{18}$$

For electrons in the conduction band this expression takes a simpler form,

$$\mathbf{j}_e = \gamma[\Psi_s\boldsymbol{\Psi}^* + \Psi_s^*\boldsymbol{\Psi}]. \tag{19}$$

The exact procedure for deriving the boundary conditions for wave functions still remains to be devised. However, some approximate methods for solving this problem have been developed in recent years. Usually in a $A_{III}B_V$ semiconductor heterostructure Kane's parameter γ varies only slightly; hence continuity of γ is often supposed (see, for example Ref. 17). The discrepancy of the parameter γ in a QW and barrier region results in a small change of the Auger coefficient (see Sec. IV). Following the method elaborated by Burt[19] and assuming the contunuity of Kane's parameter, we derive from the system (2) equations that can be integrated across the heterobarrier:

$$\begin{aligned}&(E_g+\delta-E)\Psi_s-i\hbar\gamma\nabla\boldsymbol{\Psi}=0,\\ &-E\boldsymbol{\Psi}-i\hbar\gamma\nabla\Psi_s+\frac{\hbar^2}{2m}\nabla[6\tilde{\gamma}_2\nabla\boldsymbol{\Psi}]\\ &\quad+\frac{\hbar^2}{2m}\frac{\partial}{\partial x_k}(\tilde{\gamma}_1-2\tilde{\gamma}_2)\frac{\partial}{\partial x_k}\boldsymbol{\Psi}+i\delta[\boldsymbol{\sigma}\times\boldsymbol{\Psi}]=0.\end{aligned}\tag{20}$$

Using these equations and the probability flux density conservation law the boundary conditions for the wave-function envelopes will be derived.

B. Carrier states in a quantum well

The wave functions of carriers in a QW can be found using the symmetry properties of the Hamiltonian. The spinless Hamiltonian $\mathcal{H}_0$ is invariant with respect to the substitution $x\to -x$. Consider an operator $\mathcal{R}$ such that

$$\begin{aligned}&\mathcal{R}:(x,y,z)\to(-x,y,z),\quad \mathcal{R}=\mathcal{I}\mathcal{C}_{\pi x},\\ &\mathcal{H}_0\mathcal{R}=\mathcal{R}\mathcal{H}_0,\end{aligned}\tag{21}$$

where $\mathcal{I}$ is the inversion operator, and $\mathcal{C}_{\pi x}$ is the operator of rotation by an angle π around the x axis perpendicular to the plane of the QW.

With the account of the spin-orbit interaction the Hamiltonian can be written in the form

$$\mathcal{H}=\mathcal{H}_0+\frac{\hbar}{4m^2c^2}[\nabla V\times\mathbf{p}]\boldsymbol{\sigma},\tag{22}$$

where $\mathbf{p}$ is the momentum operator and V is the potential energy of an electron in the crystal. The last term does not commute with R. Therefore, the symmetry operator $\mathcal{D}$ may be sought as a product of the operator $\mathcal{R}$ and some spin matrix S to be found: $\mathcal{D}=\mathcal{R}\otimes S$. Since inversion leaves unchanged the sign of the vector product, the matrix S must satisfy the relations

$$\begin{aligned}&S\sigma_x=\sigma_x S,\\ &S\sigma_y=-\sigma_y S,\\ &S\sigma_z=-\sigma_z S.\end{aligned}\tag{23}$$

Obviously, a Pauli spin matrix σ_x may be taken for the matrix S: $S=\sigma_x$.

The functions $\Psi(x,y,z)$ and $\mathcal{D}\Psi(-x,y,z)$ satisfy the same equation. For this reason the eigenfunctions of the Hamiltonian may be sought as eigenfunctions of the operator $\mathcal{D}$.

$$\Psi(x,y,z)=\nu\mathcal{D}\Psi(-x,y,z)\quad\text{where}\quad\nu=\pm1.\tag{24}$$

The values $\nu=\pm1$ correspond to carrier states with different symmetry. With the wave functions chosen in such a way, the boundary conditions can be satisfied at one heteroboundary only, since at the other they will be fulfilled automatically. Solving Eq. (24) we find the necessary conditions for various components of the symmetrized wave function.

$$\begin{aligned}&\Psi_{s\uparrow}(x,y,z)=\pm\Psi_{s\downarrow}(-x,y,z),\\ &\Psi_{x\uparrow}(x,y,z)=\mp\Psi_{x\downarrow}(-x,y,z),\\ &\Psi_{y\uparrow}(x,y,z)=\pm\Psi_{y\downarrow}(-x,y,z),\\ &\Psi_{z\uparrow}(x,y,z)=\pm\Psi_{z\downarrow}(-x,y,z),\end{aligned}$$

where the sign ''+'' corresponds to $\nu=1$, and ''−'' to $\nu=-1$ for the s, y, z components and vice versa for the x component. The corresponding expressions for the components of electron and hole-wave functions are given in Appendix A.

III. MATRIX ELEMENT OF AUGER RECOMBINATION

The differential rate of AR can be calculated in terms of the first-order perturbation theory in electron-electron interaction:

$$W_{i\to f}=\frac{2\pi}{\hbar}|M_{fi}|^2\delta(\varepsilon_f-\varepsilon_i),\tag{25}$$

where

$$M_{fi}=\left\langle\Psi_f(\mathbf{r}_1,\mathbf{r}_2,\nu_1,\nu_2)\left|\frac{e^2}{\kappa_0|\mathbf{r}_1-\mathbf{r}_2|}+\tilde{\Phi}(\mathbf{r}_1,\mathbf{r}_2)\right|\Psi_i(\mathbf{r}_1,\mathbf{r}_2,\nu_1,\nu_2)\right\rangle\tag{26}$$

is the matrix element of electron-electron interaction, $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are the carrier coordinates, ν_1 and ν_2 are spin variables [see Eq. (24)], e is the electron charge, κ_0 is the dielectric constant of the intrinsic semiconductor, and $\tilde{\Phi}(\mathbf{r}_1,\mathbf{r}_2)$ is the additional potential arising because of the difference between the QW and barrier region dielectric constants. An explicit expression for $\tilde{\Phi}(\mathbf{r}_1,\mathbf{r}_2)$ is given in Appendix B.

With an account of the antisymmetrized form of the wave functions, the matrix element of the Auger transition is the following:

$$M_{fi}=M_{\mathrm{I}}-M_{\mathrm{II}},\tag{27}$$

$$M_{\mathrm{I}}=\left\langle \Psi_3(\boldsymbol{r}_1,\nu_1)\Psi_4(\boldsymbol{r}_2,\nu_2)\left|\frac{e^2}{\kappa_0|\boldsymbol{r}_1-\boldsymbol{r}_2|}\right.\right.$$
$$\left.\left.+\tilde{\Phi}(\mathbf{r}_1,\mathbf{r}_2)\right|\Psi_1(\boldsymbol{r}_1,\nu_1)\Psi_2(\boldsymbol{r}_2,\nu_2)\right\rangle, \qquad (28)$$

an expression for M_{II} can be derived from Eq. (28) by interchanging indices 1 and 2 in the wave functions Ψ_1 and Ψ_2. Hereafter the indices I and II in the expressions for the matrix elements will be omitted.

We shall consider two AR processes, CHCC and CHHS, since in fact only these two determine the rate of Auger recombination. Strictly speaking, such a terminology is inapplicable to carriers in a QW, since there exists mixing between heavy-hole, light-hole, and SO-hole subbands. However, the mixing of SO holes with heavy and light holes is negligible at $\Delta_{so}\gg T$. The last condition is fulfilled for the majority of $A_{\mathrm{III}}B_{\mathrm{V}}$ semiconductors. For this reason we may rely on the above terminology.

Procedures for evaluating matrix elements for the CHCC and CHHS Auger processes are similar. For the sake of simplicity, mainly the matrix element of the CHCC Auger transition will be discussed further in this section. In the following section approximate expressions for the Auger coefficient will be given for both the CHCC and the CHHS processes.

The matrix element of an electron-electron Coulomb interaction can be most conveniently calculated using a Fourier transform. We take into account that the wave functions of carriers in a QW are plane waves propagating along the lateral direction:

$$\Psi_i(\mathbf{r})=\psi_i(x,\mathbf{q}_i)\exp(i\mathbf{q}_i\boldsymbol{\rho}),$$

where $\mathbf{q}$ and $\boldsymbol{\rho}$ are, respectively, the lateral wave vector and coordinate of carriers. Then

$$M=\frac{4\pi e^2}{\kappa_0}\frac{1}{2q}\int_{-\infty}^{\infty}\int_{-a/2}^{a/2}\psi_4^*(x_1)\psi_3^*(x_2)[e^{-q|x_1-x_2|}$$
$$+\tilde{\phi}(x_1,x_2,q)]\psi_1(x_1)\psi_2(x_2)dx_1dx_2, \qquad (29)$$

$q=|\mathbf{q}_1-\mathbf{q}_4|=|\mathbf{q}_3-\mathbf{q}_2|$ is the momentum transferred in the plane of the QW in the Coulomb interaction, a is the QW width, and $\tilde{\phi}$, for which an expression is given in Appendix B, corresponds to the potential $\tilde{\Phi}$. Integration over x_2 is limited to within the QW owing to the fact that heavy holes are usually strongly localized inside the well because of their relatively large mass. Hereafter, x and $\boldsymbol{\rho}$ denote the coordinates orthogonal and parallel to the QW plane respectively, $\mathbf{q}$ and k are the lateral and x quasimomentum components of particles.

As seen from Eq. (29), Auger scattering occurs on a one-dimensional exponentially decaying potential that depends on the transferred lateral momentum. The state of an excited particle may lie in either continuous or discrete spectrum, with the latter situation occurring when the longitudinal momentum of the particle much exceeds the transverse momentum. [We assume that, as is commonly the case, $(V_c,\ V_v)<E_g$.] In determining the rate of AR, both localized and delocalized states must be considered as final states of the excited particle. The possibility of an electron (hole) transition into a localized or a free state leads to the existence of different AR mechanisms in QW's.

A. Calculation of the matrix element for a transition of an excited particle into the continuous spectrum

For evaluating the matrix element we use the approximation $V_c, V_v \ll E_g$, where V_c and V_v are the barrier heights for electrons and holes, respectively. Obviously, this approximation also implies that $k_4^2+q^2\gg k_1^2$, i.e., the total momentum of an excited electron, is much larger than that of a localized one. The integral over the x_1 coordinate can be found by integrating by parts. The nth antiderivative of the function $\psi_4 e^{-qx}$ is

$$F_4^n(q,x)=(-1)^n\frac{[e^{qx}\psi_4(x)]^{(n)}}{(k_4^2+q^2)^n}e^{-2qx}.$$

Then an approximate expression for the matrix element M_I can be obtained:

$$M\approx M^{(1)}+M^{(2)}, \qquad (30)$$

where

$$M^{(1)}=-\frac{4\pi e^2}{\kappa_0(q^2+k_4^2)}\left(\mathcal{F}(a/2)\int_{-a/2}^{a/2}e^{qx_2}\psi_3^*(x_2)\psi_2(x_2)dx_2\right.$$
$$\left.-\mathcal{F}(-a/2)\int_{-a/2}^{a/2}e^{-qx_2}\psi_3^*(x_2)\psi_2(x_2)dx_2\right). \qquad (31)$$

Here

$$\mathcal{F}(a/2)=e^{-qa/2}\psi_{4s}^*(a/2)\psi_{1s}(a/2)\left(\frac{3V_c+V_v}{4E_g}-\frac{\kappa_0-\tilde{\kappa}_0}{\kappa_0+\tilde{\kappa}_0}\right). \qquad (32)$$

The index s in ψ_{4s} and ψ_{1s} implies that only the s components of the wave functions are taken, $\tilde{\kappa}_0$ is the dielectric constant in the barrier region. For $M^{(2)}$ we have,

$$M^{(2)}=\frac{4\pi e^2}{\kappa_0(q^2+k_4^2)}\int_{-a/2}^{a/2}\psi_4^*(x)\psi_3^*(x)\psi_2(x)\psi_1(x)dx. \qquad (33)$$

Note that since the wave functions are spinors, the components of ψ_4^* should be multiplied by components of ψ_1, and, vice versa, the components of ψ_3^* should be multiplied by those of ψ_2.

In this way it appears that the matrix element of Auger transition splits into two parts. The first is related to the presence of heteroboundaries and the second corresponds to the short-range Coulomb scattering. The latter can be easily understood since during Auger transition a large energy is transmitted and this is possible only if the scattering particles find themselves very close to each other. Note that both $M^{(1)}$ and $M^{(2)}$ are in fact thresholdless matrix elements. Indeed, they are not subject to any restrictions imposed on the initial momenta of carriers: k_1, k_2, and k_3. However, the mechanisms responsible for the momentum nonconservation $(k_1+k_2\neq k_3+k_4)$ in $M^{(1)}$ and $M^{(2)}$ are different. In $M^{(1)}$ the latter is related to carrier scattering at the heteroboundary,

and the same mechanism gives rise to a thresholdless Auger process in scattering on a single heterobarrier.[7] The reason why the conservation law breaks down for $M^{(2)}$ is that the volume of integration with respect to x is restricted to the QW region, which results in the appearance of a function of the type $\sin(ka/2)k$ instead of $\delta(k)$. Substituting Eq. (32) into Eq. (31) and integrating over x_2, we obtain

$$M^{(1)} \approx \frac{8\pi e^2}{\kappa_0(q^2+k_4^2)(q^2+k_3^2)} \left(\frac{3V_c+V_v}{4E_g} - \frac{\kappa_0-\tilde{\kappa}_0}{\kappa_0+\tilde{\kappa}_0} \right) \times [\psi_4^*(a/2)\psi_1(a/2)][\psi_3^*(a/2)\psi_2(a/2)]'(1\pm e^{-qa}). \tag{34}$$

The sign in the last pair of parentheses is chosen according to the parity of the product $\psi_3^*(x)\psi_2(x)$: ''+'' corresponds to an even product and ''−'' to an odd one. In the case $qa \gg 1$ the exponent may be omitted and the matrix element $M^{(1)}$ corresponds to independent scattering at two heteroboundaries. When evaluating Eq. (34), we took into account that $k_3^2+q^2 \gg k_2^2$, which reflects the fact that the average momentum of holes is larger than that of electrons. The term $(\kappa_0-\tilde{\kappa}_0)/(\kappa_0+\tilde{\kappa}_0)$ in Eq. (34) arises from taking into consideration the additional potential $\tilde{\Phi}(\mathbf{r}_1,\mathbf{r}_2)$ [see Eq. (28)]. Note that the matrix element M is zero if the parities of the products $\psi_3^*(x)\psi_2(x)$ and $\psi_4^*(x)\psi_1(x)$ are different.

Let us now analyze $M^{(2)}$. The integral in $M^{(2)}$ is proportional to the sum

$$\int_0^a \psi_4^*(x)\psi_3^*(x)\psi_1(x)\psi_2(x)dx \propto \sum \pm \frac{\sin(k_4-k)a/2}{k_4-k}, \tag{35}$$

where k runs through eight different values $k=\pm k_1 \pm k_2 \pm k_3$. Of all terms in the sum from Eq. (35), the largest is that for which $k=k_1+k_2+k_3$. (The reason is that this term has the lowest threshold energy. By the threshold energy we understand the mean energy of a heavy hole taking part in an Auger transition.) The contributions to the sum from other terms are less significant and will be neglected for the sake of simplicity. Then the expression for the matrix element $M^{(2)}$ takes the following form:

$$M^{(2)} \approx \frac{\pi e^2}{\kappa_0(q^2+k_4^2)} e^{i\delta} \frac{\hbar\gamma}{E_g} \frac{1+2/3\alpha}{1+\alpha} A_c^2 A_f A_h \times \frac{\sin(k_f-k_{c1}-k_{c2}-k_h)a/2}{k_f-k_{c1}-k_{c2}-k_h} \times \begin{cases} q_h k_c e^{i\phi_{2,3}} + q_c k_h, & \nu_c = \pm\nu_h \\ q_c q_h \sin\phi_{2,3}, & \nu_c = \mp\nu_h. \end{cases} \tag{36}$$

Here δ is an insignificant phase coefficient, A_i denotes the normalizing constant, ν_c and ν_h are the spin indexes introduced according to Eq. (24), and $\phi_{2,3}$ is the angle between the lateral momenta of the electron and the hole. As follows from Eq. (36), in the limit $a\to\infty$ the matrix element $M^{(2)}$ becomes proportional to $\delta(k_f-k_1-k_2-k_h)$ and hence tends to the bulk matrix element. On the contrary, $M^{(1)}$ exhibits no extremum in the same limit. Therefore, Auger processes corresponding to $M^{(1)}$ and $M^{(2)}$ will be referred to as thresholdless and quasi-threshold processes, respectively. Note that in this work the matrix element for transition into continuous spectrum (M) is split into $M^{(1)}$ and $M^{(2)}$ in a different way than in our previous work,[23] to make the corresponding expressions more clear.

B. Calculation of the matrix element of Auger recombination for a transition of an excited particle into the discrete spectrum

We now turn our attention to analyzing the matrix element of an Auger transition in which the high-energy particle ψ_4 remains in the bound state. This case corresponds to the condition $q_4 \gg k_4$. The matrix element can be calculated similarly to $M^{(2)}$ above:

$$M^{(3)} \approx \frac{4\pi e^2}{\kappa_0(q^2+k_4^2)} \int_{-a/2}^{a/2} (\psi_4^*\psi_1)(\psi_3^*\psi_2)dx. \tag{37}$$

This integral can be readily calculated; however, the general formula is rather cumbersome and will not be presented here. We shall only make an estimate of M_3, valid in the case when bound carriers are in the ground quantum state. Then we have

$$M^{(3)} \approx \frac{1}{q^2+k_4^2} e^{i\delta} \frac{\hbar\gamma a}{2Z} A_c^2 A_f A_h \alpha q_c q_h \sin\phi_{2,3}, \tag{38}$$

where $\phi_{2,3}$ is the angle between the lateral quasimomentum components of an electron and a heavy hole, and α is a coefficient of order unity, resulting from integrating the product of the envelopes of the carrier wave functions over the quantum well:

$$\int_0^{a/2} f_1 f_2 f_3 f_4 dx = \alpha a/2, \tag{39}$$

where $f_i = \cos k_i x$, and i numerates the initial and final states of particles taking part in the AR process.

IV. AUGER RECOMBINATION COEFFICIENT

To calculate the rate of AR, the probabilities of Auger transition (25) should be summed over all initial and final states of carriers with appropriate weights-occupation numbers.

$$G = \frac{2\pi}{\hbar} \sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4} \langle M^2\rangle f_1 f_2 (1-f_3)(1-f_4) \times \delta(E_3+E_4-E_1-E_2). \tag{40}$$

Here f_1, f_2 are the occupancies of the initial states and f_3, f_4 are those of the final states,

$$\langle M^2\rangle = \sum_{\nu_1,\nu_2,\nu_3,\nu_4} |M_{fi}|^2$$

is a sum of squared Auger matrix elements over spins of the initial and final states. It is more convenient to choose electrons and holes as carriers for the CHCC and CHHS processes, respectively. For high-excited states the distribution function f_4 may be set to zero. Note that instead of $1-f_3$,

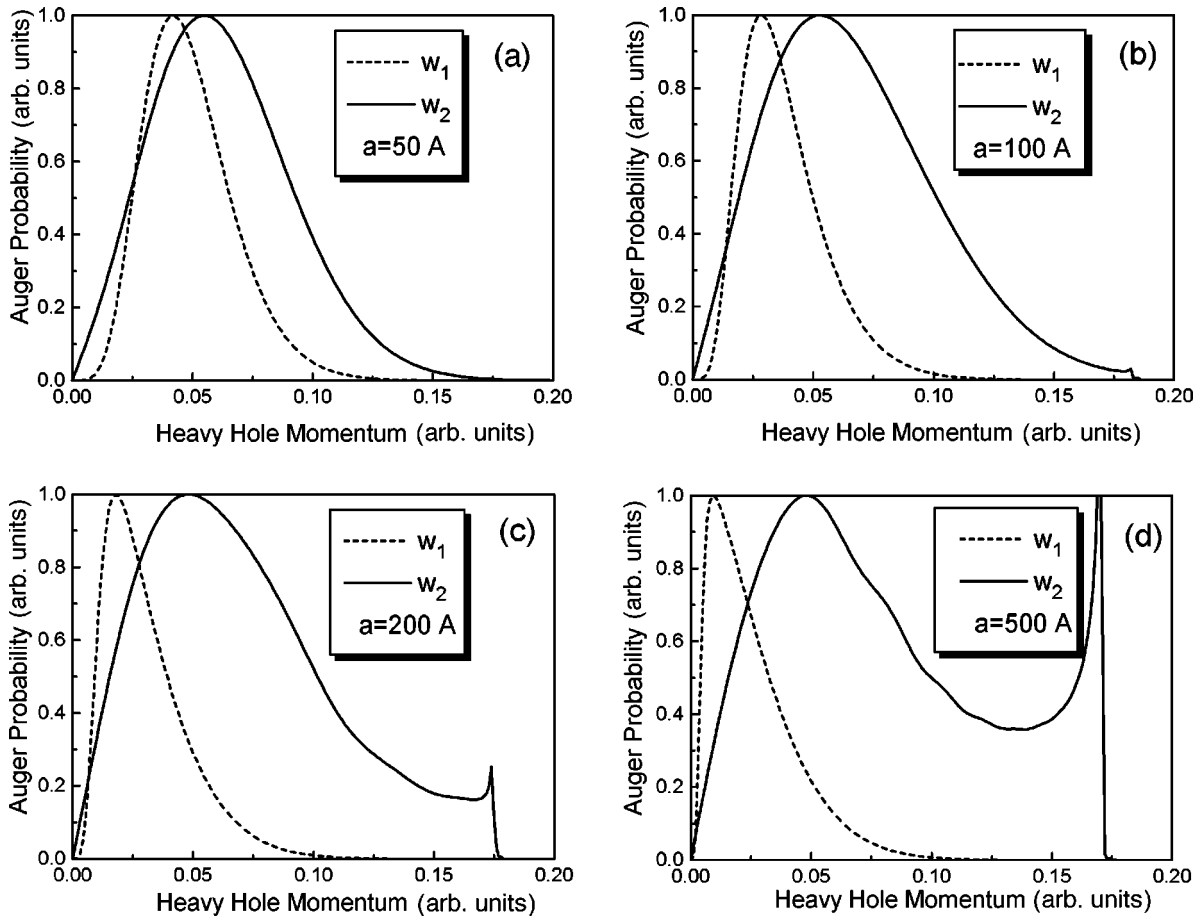

FIG. 1. Auger transition probabilities $w_1(q)$ and $w_2(q)$, corresponding to the thresholdless and quasithreshold matrix elements M_1 and M_2 as functions of the longitudinal momentum of heavy holes at $T=300$ K for different QW widths [(a) $a=50$ Å, (b) $a=100$ Å, (c) $a=200$ Å, and (d) $a=500$ Å].

we may write $\tilde{f}_3$, where $\tilde{f}_3$ is the distribution function for carriers of the opposite sign: holes for the CHCC process and electrons for the CHHS process.

The matrix elements $M^{(1)}$ and $M^{(2)}$, on the one hand, and $M^{(3)}$, on the other, describe transitions in which the excited particle occupies a state of the continuous or discrete spectrum, respectively. Therefore, the contributions to the Auger rate from the first two matrix elements and the last one can be separated. It is more difficult to separate the contributions from $M^{(1)}$ and $M^{(2)}$. Even though the physical difference between these terms does exist, there is a term of interference between them, essential at small QW widths, since both processes are in fact thresholdless. However, even when the interference is neglected we still obtain a result of correct order, reflecting all the main specific features of the AR coefficient as a function of temperature and parameters of a structure with a QW. In sufficiently wide QW's the modulus of $M^{(2)}$ exhibits a maximum at the point $k_4(q)+k_3=k_1+k_2$, whereas $M^{(1)}$ as a function of quasimomentum shows no extremum. When the QW width tends to infinity, $M^{(2)}$ at this point has the form of a δ function. In accordance with the aforesaid, the AR probability for wide QW's, corresponding to the matrix element $M^{(2)}$, has a maximum (as a function of the longitudinal momentum of the heavy hole q_h) at higher q_h values than the probability associated with $M^{(1)}$. With decreasing QW width the maxima of these probabilities approach each other, and the region of overlapping between these matrix elements becomes larger (see Fig. 1). It should be noted that the AR probabilities presented in Fig. 1 are rather smooth functions of the longitudinal momentum of heavy hole (q_h), since, in calculating them, the summation is done over discrete quantum states of carriers. At q_h close to the maximum value determined by the conservation of longitudinal momentum and energy, the AR probability shows a square-root divergence eliminated upon integration with respect to q_h, i.e., in calculating the rate of AR. The probability of Auger transition for the CHHS process has a form similar to that for the CHCC process.

In line with the aforesaid, let us present the rate of AR as follows:

$$G=G_1+G_2+G_3, \tag{41}$$

where the rate G_1 corresponds to a thresholdless Auger process with the matrix element M_1, rate G_2 to a quasithreshold Auger process with the matrix element M_2, and rate G_3 to a threshold Auger process with the matrix element M_3.

The expressions for the rates G_1 and G_2 can be derived from Eq. (40) by passing from summation over k_4 to integration and from δ function with respect to energy to δ function with respect to momentum. In what follows we shall study the AR coefficients C related to the rate G by

$$G=Cn^2p \quad \text{and} \quad G=Cp^2n$$

for the CHCC and CHHS Auger processes, respectively. Here n and p are the 2D densities of electrons and holes.

For the CHCC process we have

$$C_1\approx\frac{32\pi^2e^4}{\kappa_0^2}\frac{\hbar\gamma^2}{E_g^3}\frac{F(\Delta_{so}/E_g)}{a(a+2/\kappa_c)^2}\frac{k_c^2\kappa_c^2}{(k_c^2+\kappa_c^2)^2}\frac{V_c}{E_g}\times\left(\frac{3V_c+V_v}{4E_g}-\frac{\kappa_0-\tilde{\kappa}_0}{\kappa_0+\tilde{\kappa}_0}\right)^2\left\langle\frac{q_h^2k_h^2}{(q_h^2+k_h^2)^3}\frac{1}{k_4(q_h)}\right\rangle, \tag{42}$$

where

$$F(x)=\left(\frac{1+2x/3}{1+x}\right)^2\frac{1+7x/9+x^2/6}{(1+x/2)(1+4x/9)}.$$

Note that, if Kane's parameter γ is discontinuous then the term

$$\frac{E_{0c}}{2E_g}\left(\frac{\delta\gamma}{\gamma}\right)^2$$

should be added to

$$\left(\frac{3V_c+V_v}{4E_g}-\frac{\kappa_0-\tilde{\kappa}_0}{\kappa_0+\tilde{\kappa}_0}\right)^2,$$

where E_{0c} is the electron energy in the ground level and $\delta\gamma=\gamma-\tilde{\gamma}$ is the difference between Kane's parameters in the QW and the barrier region. However, this addition is usually negligible. Similarly, the term $(\kappa_0-\tilde{\kappa}_0)/(\kappa_0+\tilde{\kappa}_0)$ is usually small compared to $(3V_c+V_v)/4E_g$ and can be omitted. It can be shown that the influence exerted by the difference between the dielectric constants in the quantum well and barrier region is even less for the case of the CHHS AR process. For this reason the corresponding terms for this process will be omitted.

The angular brackets in Eq. (42) and below denote averaging over the heavy-hole distribution function. In the case of a Boltzmann distribution, which is commonly the case for holes, this averaging looks like

$$\langle f(q_h,k_h)\rangle=\frac{1}{Z}\sum_n\int_0^\infty q_hf(q_h,k_{hn})e^{-(k_{hn}^2+q_h^2)/q_T^2}dq_h,$$

where

$$Z=\frac{2}{q_T^2}\sum_n e^{-(k_{hn}^2)/q_T^2},$$

$q_T=\sqrt{2m_hT}/\hbar$ is the thermal wave vector of heavy holes, and k_{hn} is the wave vector corresponding to the nth level of heavy holes.

For the CHHS process the following expression for C_1 can be derived:

$$C_1\approx\frac{2\pi^2e^4}{\kappa_0^2\hbar}\frac{V_c}{E_g}\frac{k_c^2\kappa_c^2}{(k_c^2+\kappa_c^2)^2}\frac{\tilde{F}(\Delta_{so}/E_g)}{a^2(a+2/\kappa_c)}\frac{\hbar^3}{m_{so}^3(E_g-\Delta_{so})^3}\times\left\langle\frac{k_{h1}^2k_{h2}^2q_{h1}^2(q_{h1}^2+q_{h2}^2)}{(q_{h1}^2+k_{h1}^2)^3(q_{h2}^2+k_{h2}^2)}\right\rangle, \tag{43}$$

where

$$\tilde{F}(x)=\frac{[2x+3(1-x)(1-m_{so}/m_h)]^2}{2x^2+[x+3(1-x)(1-m_{so}/m_h)]^2}\frac{1+2x/3}{1+x}.$$

In the last case averaging is performed over distribution functions of two holes. In deriving Eq. (43) we assumed that $E_g-\Delta_{so}\gg T(m_h/m_{so})$.

Let us turn to the quasithreshold Auger process. For the CHCC process we have

$$C_2\approx\frac{\pi^2e^4}{\kappa_0^2}\frac{\hbar^3\gamma^4}{E_g^5}\frac{F(\Delta_{so}/E_g)}{a(a+2/\kappa_c)^2}\times\left\langle\frac{q_c^2k_h^2+q_h^2\left(k_c^2+\frac{1}{2}q_c^2\right)}{(q_h^2+k_h^2)k_f}\frac{1-\cos(k_f-k_h-2k_c)a}{2(k_f-k_h-2k_c)^2}\right\rangle. \tag{44}$$

Direct calculation of the Auger coefficient C_2 for the CHHS process gives a cumbersome result. We present here a simplified expression valid for sufficiently narrow QW's at $k_c\gg q_c$:

$$C_2\approx\frac{\pi^2e^4}{4\kappa_0^2}\frac{E_c}{E_g}\frac{\hbar^3}{m_{so}^2(E_g-\Delta)^3}\frac{\tilde{F}(\Delta_{so}/E_g)}{a^2(a+2/\kappa_c)}\left\langle\frac{1-\cos(k_{so}-k_{h1}-k_{h2}-k_c)a}{2(k_{so}-k_{h1}-k_{h2}-k_c)^2}\right.$$

$$\left.\times\frac{q_{h2}^2[(k_{so}^2+k_{h1}^2)q_{h1}^2+q_{h2}^2k_{h1}^2+2k_{h1}^2(\mathbf{q}_{h1}\mathbf{q}_{h2})+(\mathbf{q}_{h1}\times\mathbf{q}_{h2})^2]}{(q_{h1}^2+k_{h1}^2)(q_{h2}^2+k_{h2}^2)k_{so}}\right\rangle. \tag{45}$$

And, finally, we have C_3 for the CHCC process:

$$C_3 \approx \frac{32\pi^2 e^4}{\kappa_0^2 \hbar E_g} \frac{a}{(a+1/\kappa_c)^3} \frac{1+\frac{7}{9}x+\frac{1}{6}x^2}{(1+x/3)^2} \frac{1+\frac{2}{3}x}{1+x} \times \left\langle \frac{q_{th}^2}{q_T^2} \frac{q_c^2}{(q_{th}^2+k_h^2)^3} e^{-q_{th}^2/q_T^2 \alpha^2} \right\rangle_n . \quad (46)$$

Here $x=\Delta_{so}/E_g$, α is a multiplier introduced in Eq. (38). In the last case, averaging only over discrete quantum states of heavy holes is performed. The threshold momentum q_{th} is found from the conservation law for energy and longitudinal component of momentum:

$$E_f(\sqrt{k_f^2+q_{th}^2}) = E_g + \frac{\hbar^2(q_{th}^2+k_h^2)}{2m_h} + \frac{\hbar^2(k_{c1}^2+k_{c2}^2)}{2m_c}.$$

For simplicity we neglected here the small longitudinal momenta of electrons. Expanding the energy of an excited electron E_f into a series in terms of momenta in the vicinity of $q_{th}=Q$, where Q is the electron momentum corresponding to an energy equal to E_g ($Q \approx \sqrt{4m_c E_g/\hbar^2}$), we get the following estimate for the threshold momentum:

$$q_{th} \approx \sqrt{\frac{4m_c E_g}{\hbar^2} + \frac{3}{2}\left(k_c^2 + \frac{m_c}{m_h}k_h^2\right)}. \quad (47)$$

If the QW width tends to infinity, the threshold momentum approaches its bulk value.[2] Account must be taken of the fact that for wide quantum wells with a large number of levels the introduced multiplier α [see Eq. (38)] tends to a δ function expressing the conservation law for the transverse quasimomentum component:

$$\alpha^2 \to \frac{\pi}{128} a \sum \delta(k_h \pm k_{c1} \pm k_{c2} \pm k_{c4}).$$

For large QW widths and $V_c \ll E_g$, the inequality $C_3 \ll C_2$ is valid, since the ratio $C_3/C_2 \approx \sqrt{V_c/E_g}$. Hence, for wide QW's C_3 may be neglected as compared with C_2. For narrow QW's the threshold energy of the CHCC process increases [see Eq. (47)] and the AR coefficient (46) decreases relative to the bulk value by a factor

$$e^{3k_c^2/2q_T^2} \approx e^{(3m_c/2m_h)(E_{0c}/T)}.$$

The characteristic width of a QW for which this phenomenon becomes essential can be readily evaluated from the condition that the exponent is unity:

$$E_{0c} \approx T\frac{2m_h}{3m_c} \Leftrightarrow a \approx \frac{\pi}{q_T}. \quad (48)$$

Thus, at QW widths a less than several reciprocal thermal momenta $a \lesssim \pi/q_T$ the threshold energy E_{th}^3 becomes considerably higher than the bulk value E_{th}^{bulk} (Fig. 2). For semiconductor compounds $A_{\text{III}}B_{\text{V}}$ the equality (48) is fulfilled at room temperature at a quantum-well width of order of 100 Å.

For the threshold CHHS process the heavy-hole momenta are not specified by the threshold conditions and, therefore, integration with respect to them has to be performed. It seems impossible to derive analytically the exact Auger coefficient C_3 for the CHHS process in view of the fact that the matrix element M_3 is rather cumbersome. However, an approximate expression can readily be obtained by factoring out the averaged squared matrix element from the integrand sign:

$$C_3 \approx \frac{2\pi}{\hbar}\langle M_3^2\rangle \frac{1}{2\pi^2 q_T^4} \times \int d^2q_{h1} d^2q_{h2} e^{-(q_{h1}^2+q_{h2}^2)/q_T^2} \times \delta\left(\tilde{E}_g - \Delta - \frac{\hbar^2(\mathbf{q_{h1}}+\mathbf{q_{h2}})^2}{2m_{so}} + \frac{\hbar^2 q_{h1}^2}{2m_h} + \frac{\hbar^2 q_{h2}^2}{2m_h}\right), \quad (49)$$

Here $\tilde{E}_g = E_g + E_{0c} + 2E_{0h} - E_{0\,so}$, where E_{0h} $E_{0\,so}$ are the ground-state energies of heavy and SO holes, respectively. Let us introduce a threshold momentum, setting it equal to

$$Q_{th}^2 = \frac{2(\tilde{E}_g - \Delta)m_{so}}{\hbar^2(2-\mu_{so})},$$

where $\mu_{so}=m_{so}/m_h$. Then the expression for C_3 takes the form

$$C_3 \approx \frac{2m_{so}}{\hbar^3 Q_{th}^2} e^{-Q_{th}^2/q_T^2}\langle M_3^2\rangle. \quad (50)$$

Taking into account the fact that $Q_{th} \gg k_{so}$, we obtain

$$C_3 \approx \frac{256 m_{so}\pi^2 e^4}{\hbar^3 \kappa_0^2} \frac{Q_{th}^2}{(Q_{th}^2+2k_h^2)^4} \times \frac{V_c}{E_g} \frac{k_c^2}{k_c^2+\kappa_c^2} \frac{(1-\lambda_{so})^2}{1+2\lambda_{so}^2} \tilde{\alpha}^2 e^{-Q_{th}^2/q_T^2}, \quad (51)$$

where $\tilde{\alpha}$ is a multiplier defined as in the case of the CHCC process (38), λ_{so} is derived from λ_l [see Eq. (A3)] by substituting k_{so} instead of k_l.

Let us consider in more detail the Auger recombination coefficient C_2 for the quasi-threshold CHCC process. When the QW width approaches infinity, a substitution can be made in the function being averaged in Eq. (44):

$$\frac{1-\cos(k_f-k_h-2k_c)a}{2(k_f-k_h-2k_c)^2} \to \frac{\pi a}{2}\delta(k_f-k_h-2k_c) \quad \text{at } a \to \infty. \quad (52)$$

This formula clearly shows the occurrence of a threshold in this limit, and the coefficient C_2 transforms into a 3D expression on being multiplied by a^2. For comparison, we present both the result of Gelmont for C_{3D} (Ref. 2) and our limiting expression.

$$C_{3D} = 6\sqrt{2\pi^5}\frac{e^4\hbar^3}{\kappa_0^2} \frac{1}{E_g^{5/2}T^{1/2}m_c^{1/2}m_h^{3/2}} e^{-E_{th}/T}, \quad (53)$$

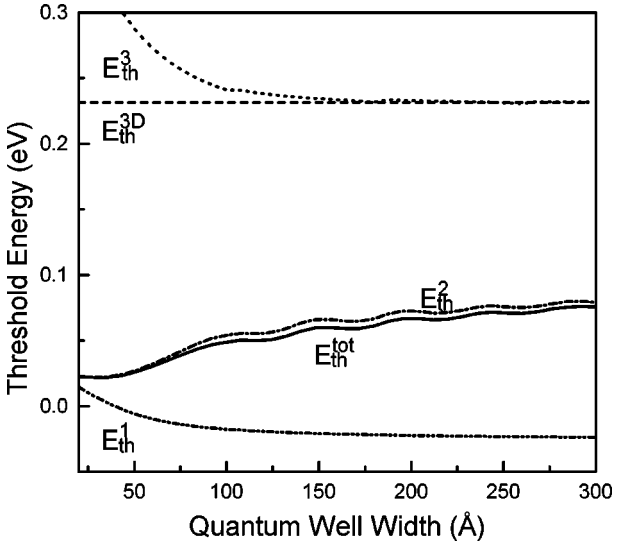

FIG. 2. Threshold energy of the CHCC process as a function of QW width for three mechanisms of Auger recombination: thresholdless (E_{th}^1), quasithreshold (E_{th}^2), and threshold (E_{th}^3) at $T=300$ K. The solid curve corresponds to the threshold energy (E_{th}^{tot}) of the total Auger coefficient ($C=C_1+C_2+C_3$). The horizontal dashed line corresponds to the threshold energy E_{th}^{3D} for the bulk Auger process.

$$C_2a^2=4\frac{8\sqrt{2\pi^5}}{9}\frac{e^4\hbar^3}{\kappa_0^2}\frac{1}{E_g^{5/2}T^{1/2}m_c^{1/2}m_h^{3/2}}e^{-E_{th}/T}. \quad (54)$$

The factor 4 in Eq. (54) results from the necessity to take into account, in calculating M_2 in accordance with Eq. (36), not only the terms with $k=k_{c1}+k_{c_2}+k_h$, but also those with $k=k_{c1}-k_{c_2}+k_h$, $k=-k_{c1}+k_{c_2}+k_h$, and $k=-k_{c1}-k_{c_2}+k_h$. When the QW width tends to infinity, all the four terms make equal contributions to C_2. As can be seen, the only difference between expressions (53) and (54) is in a numerical coefficient. The small discrepancy of about 2/3 times is due to the necessity to distinguish between the momenta of the two localized electrons: $k_{c1}\neq k_{c2}$. Furthermore, expression (53) was derived for the case when Δ_{so} tends to infinity, while in obtaining expression (54) it was assumed that $\Delta_{so}\lesssim E_g$. In calculating Eq. (54) we neglected the quantity V_c as compared with E_g. In the general case, $(C_2+C_3)a^2$ should be written instead of C_2a^2 to make expression (54) valid. However, the limiting transition from the quasithreshold to the threshold Auger process [see Eq. (52)] can be realized only for very large QW's. When analyzing the probability of Auger transition as a function of the heavy-hole momentum, one can obtain a qualitative criterion for this transition. The probability of the quasithreshold Auger process has two characteristic extrema [see Eq. (44)]. The first of them corresponds to the maximum of squared Auger transition matrix element in the vicinity of the threshold momentum. The second one lies in the vicinity of the thermal momentum of heavy holes. The Auger coefficient C_2 can be estimated as a sum of contributions from these extrema. Then

$$C_2\approx C_2^{th}(Q_h\approx q_{th})+C_2^T(Q_h\approx q_T), \quad (55)$$

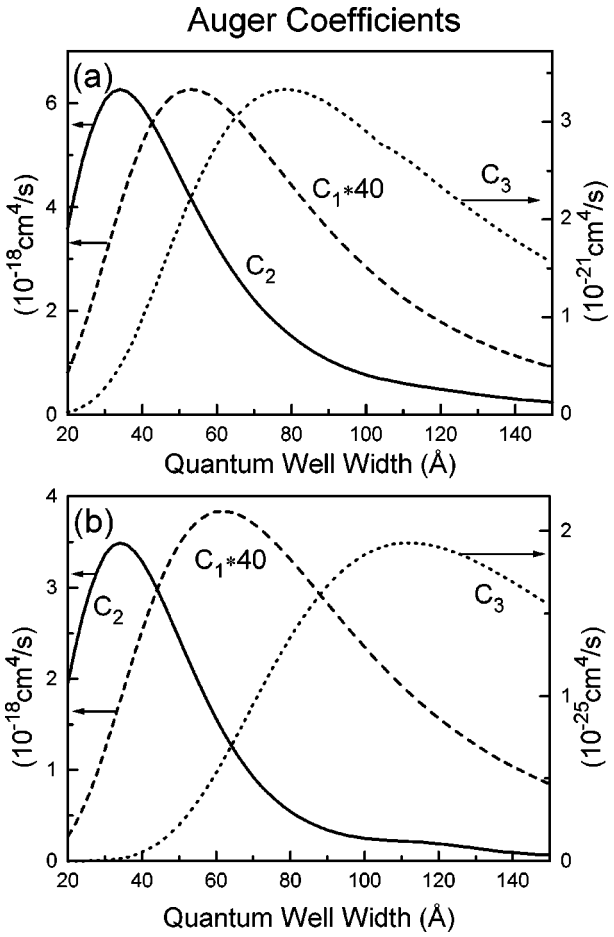

FIG. 3. Auger coefficients C_1, C_2, and C_3 for the thresholdless, quasithreshold, and threshold processes as functions of QW width at different temperatures [(a) $T=150$ K and (b) $T=300$ K].

where Q_h is the value of the heavy-hole momentum: $Q_h^2=k_h^2+q_h^2$;

$$\frac{C_2^T}{C_2^{th}}\approx\frac{\lambda_{E_g}}{a}\left(\frac{T}{E_{th}}\right)^{3/2}e^{E_{th}/T}. \quad (56)$$

Here $\lambda_{E_g}\approx 2\pi/q_{th}$ is a characteristic wavelength of an electron having energy close to E_g. Comparing the terms C_2^{th} and C_2^T, one can obtain a criterion for the transition from the quasithreshold to the threshold Auger process:

$$a\gg a_c \quad \text{where } a_c=\lambda_{E_g}\left(\frac{T}{E_{th}}\right)^{3/2}e^{E_{th}/T}. \quad (57)$$

For semiconductors with an energy gap of the order of 1 eV the critical width (a_c) may be as large as several thousand angstroms. However, the value of a_c is considerably larger than the mean free path of carriers in semiconductors. This obviously shows that correct derivation of the Auger rate in homogeneous semiconductors must involve the momentum scattering process if the critical width a_c exceeds the mean-free-path.[24]

With decreasing QW width, the average momentum of holes participating in the Auger transition shifts to lower

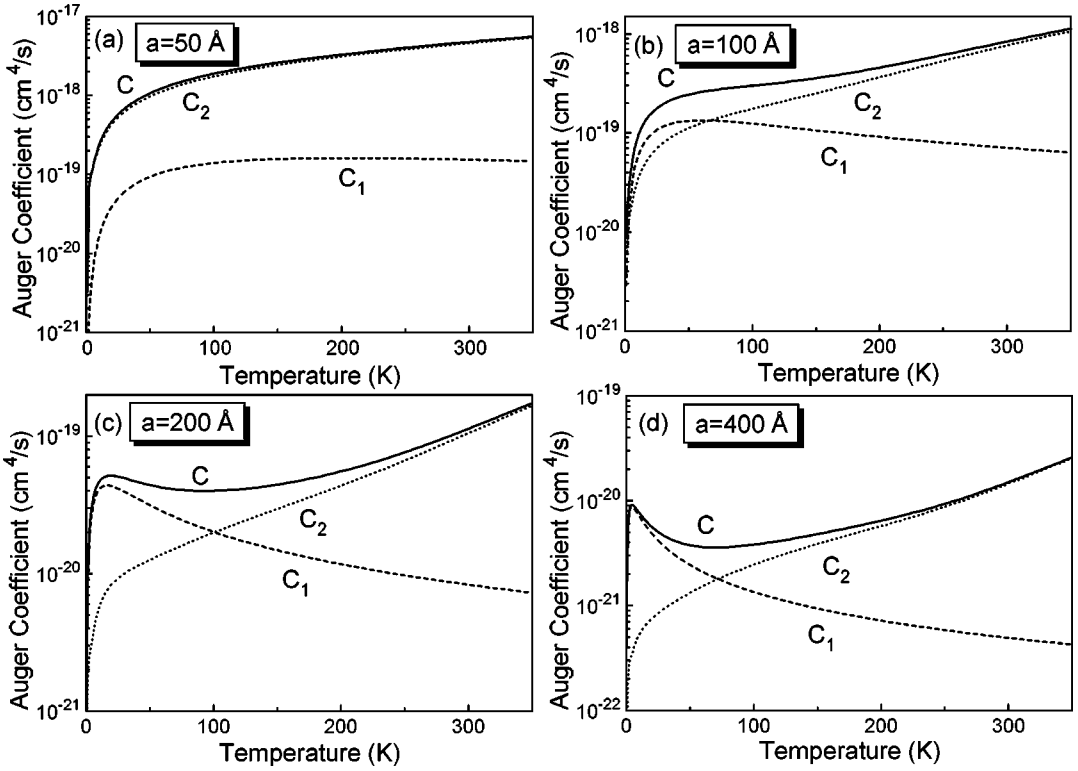

FIG. 4. Temperature dependence of the total Auger coefficient and the partial contributions of the thresholdless and quasithreshold mechanisms at different QW widths.

values (see Fig. 1). This reduces the threshold energy of the process, and makes weaker the temperature dependence of the AR coefficient.

Figure 2 shows the threshold energy of the CHCC process as a function of QW width for all the three mechanisms of Auger recombination C_1, C_2, and C_3 separately and for the overall process $C=C_1+C_2+C_3$, found from the formula

$$E_{th}^{(i)}(T)=T^2\frac{d\ \ln C_i}{dT}, \quad i=1,2,3. \tag{58}$$

The threshold energy for the quasithreshold Auger process is less than its bulk value. The reason is that the critical QW width $a_c\approx 1000$ Å is greater than the maximum width shown in the figure. The value of E_{th} for the thresholdless Auger process decreases with QW width and becomes negative. This is due to the fact that the Auger coefficient C_1 decreases with increasing temperature for wide enough QW's (see Fig. 5, below). With increasing quantum-well width, the total threshold energy tends to its limiting value E_{th}^{3D} denoted in the figure.

We now turn our attention to the thresholdless Auger process. As already noted, the probability of a thresholdless Auger transition shows no extrema as a function of the heavy hole momentum. Therefore, the coefficient C_1 has a weak nonexponential temperature dependence. This phenomenon was first studied by Zegrya and Kharchenko.[7] In addition, the function $C_1(T)$ is nonmonotonic and has a maximum. The presence of this maximum can readily be explained. At low temperature and, correspondingly, small longitudinal momenta of carriers, their wave functions are nearly orthogonal and the C_1 value is small. With increasing temperature, the characteristic momentum transferred in the Coulomb interaction (approximately equal to the thermal momentum of heavy hole) grows. This is the reason why at low temperature the Auger coefficient increases with temperature. As the temperature is elevated further, the AR coefficient C_1 passes through a maximum and starts to decrease, since Coulomb interaction responsible for the Auger process is low for large transferred momenta. The temperature at which the maximum occurs can be evaluated by equating the ground-state energy holes to temperature: $T\approx\hbar^2\pi^2/2m_h a^2$. Note that there would be no such maximum if the overlap integral between the electron and the heavy hole were taken to be proportional to the momentum transferred.[25] Such an assumption, having in our opinion no justification for the majority of structures investigated, is frequently used and gives incorrect expressions for the rate of AR and incorrect dependence of this quantity on temperature and QW parameters.

The AR coefficient C_1 depends rather strongly on the QW width a. For wide QW's, even being multiplied by a^2, C_1 remains a decreasing function of a. At $a\approx 1/\kappa_c$ the coefficient C_1 exhibits a maximum related to the weak overlapping of electron and hole wave-functions. With the quantum-well width decreasing further, the rate of the thresholdless Auger process gradually falls. A similar expression for C_1 in the CHCC process was obtained by Dyakonov and Kachorovskii, and Zegrya *et al.*[14,15]

Figure 3 shows AR coefficients C_1, C_2, and C_3 for the CHCC process as functions of QW width at different temperatures for a model structure based on InGaAsP. It can be seen that all curves show a sharply pronounced maximum. The positions of the maxima for C_1 and C_2 are practically temperature independent. The maximum for the threshold

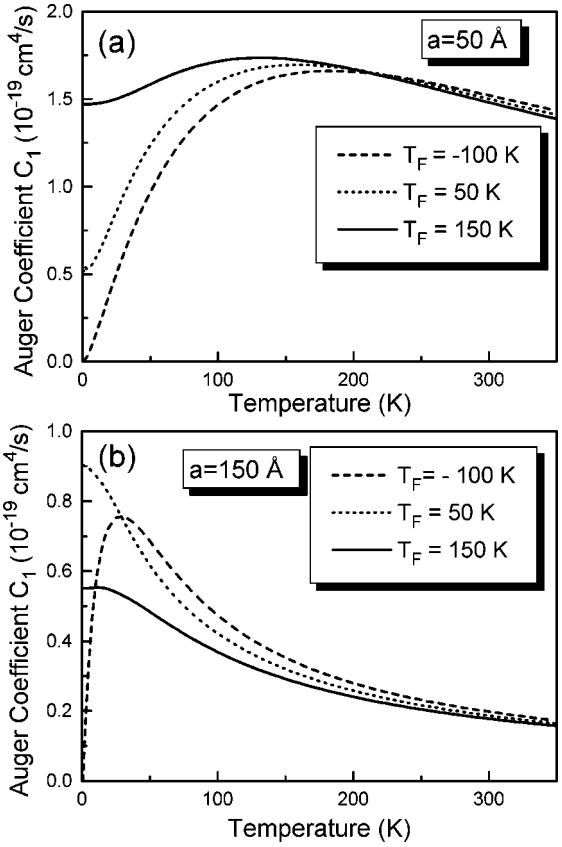

FIG. 5. Comparison of the thresholdless Auger coefficient (C_1) as a function of temperature at various Fermi energies of holes for two different QW widths [(a) $a=50$ Å and (b) $a=200$ Å]. T_F denotes the Fermi energy expressed in Kelvins. The curve with $T_F=-100$ K approximately corresponds to the Boltzmann statistics.

process (C_3) is achieved at a wider QW than for the quasithreshold or thresholdless processes and its position shifts with temperature. This is in the first place due to the reduction of the threshold energy of the threshold process with increasing quantum-well width (Fig. 2), rather than to the overlapping of the wave functions.

Figure 4 shows the temperature dependence of the overall AR coefficient for the CHCC process and the partial contributions from the thresholdless and quasithreshold mechanisms at different QW widths. It can be seen that at low temperature and sufficiently wide QW's the thresholdless Auger process predominates ($C_1>C_2$), and at high temperature, conversely, the quasithreshold process becomes more important ($C_2>C_1$). Therefore, the curve describing the temperature dependence of the overall AR coefficient has a characteristic shape with a maximum and a minimum. With increasing QW width, both the maximum and the minimum of the AR coefficient shift to lower temperature and, in the limit of infinitely wide QW, disappear. Thus, in the case of a homogeneous semiconductor the AR coefficient is a monotonic function of temperature. Note that the Boltzmann distribution of carriers was used in calculating the Auger coefficients as functions of temperature. At low temperature both electrons and holes are, as a rule, described by the Fermi-Dirac distribution function. Thus, the average momenta of particles participating in the Auger transition depends on temperature only slightly. As a result, at low temperatures the Auger coefficient is a smoother function of temperature than in the case of the Boltzmann distribution and it does not tend to zero at $T\rightarrow 0$. Figure 5 shows the thresholdless Auger coefficient C_1 versus temperature at various Fermi energies for QW's with different widths. Essential discrepancies between the values of the Auger coefficients for the Fermi-Dirac and Boltzmann distributions take place only in the case $T\ll E_F$, where E_F is the Fermi energy for holes. This condition can be realized only at very low temperatures when the Auger process ceases to be an important mechanism of recombination.

V. PHONON-ASSISTED AUGER RECOMBINATION IN QUANTUM WELLS

At low temperatures a threshold Auger process becomes exponentially weak ($C_{3D}\propto e^{-E_{th}/T}$). In this case mechanisms leading to threshold elimination are to be taken into account. It is commonly believed that the primary mechanism of this kind is emission or absorption of a virtual optical phonon. At the expense of a large momentum transferred to the phonon, the AR threshold is eliminated, and the rate of such an Auger process is a power-law function of temperature.[3,5,27] The rate of the phonon-assisted AR is calculated in terms of the second-order perturbation theory in electron-electron and electron-phonon interaction.[26] However, at high carrier densities, carrier-carrier scattering may become a more effective mechanism of threshold elimination. This is why the problem of the AR mechanism in homogeneous semiconductors at low temperatures still remains open. We shall discuss this problem in more detail elsewhere. In the present work we follow the commonly accepted viewpoint that there exists a competition between the phonon-assisted and direct AR processes. For quantum wells the situation differs strongly from the 3D case, owing to the presence of a direct thresholdless process. Therefore, it is *a priori* evident that the conditions under which the phonon-assisted AR process dominates the direct one strongly depend on the QW width.

As already noted, in the 2D case there exist three AR processes: threshold, quasi-threshold, and thresholdless. The AR coefficient for a phonon-assisted CHCC process with the threshold matrix element of electron-electron interaction is comparatively easily calculated when[27]

$$E_g\gg 2\mu E_g\gg\hbar\omega_{LO},T, \tag{59}$$

where ω_{LO} is the optical phonon frequency, $\mu=m_c/m_h$. It can be shown that the coefficient of phonon-assisted AR is related to the previously calculated one, Eq. (46), by

$$C_{ph}^3\approx C_3\frac{e^2\hbar\omega_{LO}}{2\bar{\kappa}\alpha}\frac{T}{E_{th}^{2D}}g(a,k_{th})\frac{1}{e^{\hbar\omega_{LO}/T}-1}\times\left[\frac{e^{\hbar\omega_{LO}/T}}{(E_{th}^{2D}-\hbar\omega_{LO})^2}+\frac{1}{(E_{th}^{2D}+\hbar\omega_{LO})^2}\right]e^{E_{th}^{2D}/T}, \tag{60}$$

where $\bar{\kappa}=\kappa_0\kappa_\infty/\kappa_0-\kappa_\infty$, κ_∞ is the high-frequency dielectric constant of the medium, $g(a,k_{th})$ is a factor reflecting the 2D character of holes[28]

$$g(a,k_{th})=\left[\frac{1}{2}+\frac{k_{th}^2a^2}{4(4\pi^2+k_{th}^2a^2)}\right]\times\left[1-\frac{1-e^{-k_{th}a}}{k_{th}a}\times\frac{32\pi^4}{(4\pi^2+k_{th}^2a^2)(8\pi^2+3k_{th}^2a^2)}\right]. \quad (61)$$

For comparison, we present a 3D phonon-assisted AR coefficient:

$$C_{ph}^{3D}\approx C^{3D}\frac{e^2\hbar\omega_{LO}}{2\sqrt{\pi}\overline{\kappa}}\left(\frac{T}{E_{th}^{3D}}\right)^{3/2}\frac{k_{th}}{e^{\hbar\omega_{LO}/T}-1}\times\left[\frac{e^{\hbar\omega_{LO}/T}}{(E_{th}^{3D}-\hbar\omega_{LO})^2}+\frac{1}{(E_{th}^{3D}+\hbar\omega_{LO})^2}\right]e^{E_{th}^{3D}/T}. \quad (62)$$

It can be seen that the results for the 2D and 3D processes with threshold matrix elements of electron-electron interaction are closely allied. A noticeable difference for the case of narrow quantum wells is that the threshold energy E_{th}^{2D} increases [see Eq. (47)]. Correspondingly, the criterion for predominance of the phonon-assisted AR process (C_3^{ph}) over the direct one (C_3) in QW's is met at somewhat higher temperature than in the 3D case. As already noted (see Sec. III), the rate of the threshold Auger process in narrow QW's is in itself by several orders of magnitude lower than those of the thresholdless and quasithreshold processes $[C_3 \ll (C_1, C_2)]$. Hence, the phonon-assisted AR process with a threshold matrix element of electron-electron interaction cannot compete with these processes either $(C_{ph}^3 \ll C_1, C_2)$.

Let us now consider a phonon-assisted CHCC Auger process with the other two matrix elements ($M_{ee}=M^{(1)}+M^{(2)}$). For the sake of simplicity we will use the momentum-conservation approximation for the hole–optical phonon scattering.[28] In this case the virtual hole state is fixed and we obtain

$$w_{i\to f}=\pm\frac{2\pi}{\hbar}\frac{|M_{ee}|^2|M_{hp}|^2}{(E_s\mp\hbar\omega_{LO}-E_h)^2}\frac{e^{\pm\hbar\omega_{LO}/T}}{e^{\pm\hbar\omega_{LO}/T}-1}\times\delta(E_i-E_f)d\nu_f, \quad (63)$$

where E_s is the energy of the virtual hole, and M_{hp} is the matrix element of scattering of the virtual hole by an optical phonon, with the signs plus and minus corresponding to phonon emission and absorption, respectively. It is seen that there is a singularity in the expression (63), when the denominator is equal to zero. To eliminate this divergence, account must be taken of transitions into quasistationary states, i.e., states with complex energy. In this case the pole (63) will transform into a region of complex energies:

$$w_{i\to f}\propto\frac{1}{(E_s\mp\hbar\omega_{LO}-E_h)^2+\Gamma^2},$$

where $\Gamma=\hbar/\tau$. The characteristic lifetimes τ corresponding to these states may vary over a wide range, depending on temperature, free-carrier density, etc. It only makes sense to consider a resonant phonon-assisted process in terms of the second-order perturbation theory when the halfwidths of the quasistationary hole and the phonon states are less than the energy of the optical phonon ($\hbar\omega_{LO}$). Otherwise, the Auger coefficient must be calculated in the first order of the perturbation theory, using the Lorentz function

$$f(\Delta E)=\frac{1}{\pi}\frac{\Gamma}{\Delta E^2+\Gamma^2}$$

instead of the δ function expressing the energy conservation law. For a phonon-assisted AR process with a quasithreshold matrix element of electron-electron interaction, both the resonant and virtual Auger processes are possible, with the former predominant in narrow QW's and the latter in sufficiently wide QW's.

In the general case the Auger coefficient for a phonon-assisted process may be written as

$$C_{ph}=C_{ph}^1+C_{ph}^2, \quad (64)$$

where

$$C_{ph}^{1,2}=\pm\frac{\pi\omega e^2}{\overline{\kappa}Z}\frac{e^{\pm\hbar\omega_{LO}/T}}{e^{\pm\hbar\omega_{LO}/T}-1}\sum_{m,n,\nu_n}\int\frac{d^2Q}{(2\pi)^2}\frac{d^2q_h}{(2\pi)^2}\left(\frac{\partial E_4}{\partial k_4}\right)^{-1}\times\frac{|M_{ee}(n,\nu_n,\mathbf{q_h}+\mathbf{Q}|^2}{\left(\frac{\hbar^2(m^2-n^2)\pi^2}{2a^2m_h^2}-\frac{\hbar^2(\mathbf{q_h}+\mathbf{Q})^2}{2m_h}+\frac{\hbar^2q_h^2}{2m_h}\pm\hbar\omega_{LO}\right)^2+\Gamma^2}J_{n,m}(Q)f_h(m,q_h). \quad (65)$$

Here

$$Z=\sum_m\int\frac{d^2q_h}{(2\pi)^2}f_h(m,\mathbf{q_h}),\quad J_{n,m}(Q)\approx\frac{a}{2}\frac{(1+\delta_{m,n})[(m+n)^2\pi^2+Q^2a^2]+(m-n)^2\pi^2+Q^2a^2}{[(m+n)^2\pi^2+Q^2a^2][(m-n)^2\pi^2+Q^2a^2]},$$

$f(m,q_h)$ is the hole-distribution function on the mth quantum level. The function $J_{n,m}(Q)$ has been calculated for a nondegenerate band.[28] In the case of phonon scattering by heavy holes its value will be somewhat lower. However, this fact is insignificant for our purposes. For momenta of bound electrons in the matrix element of electron-electron interaction in Eq. (65) their mean thermal values should be substituted. The Auger coefficients C_{ph}^1 (sign ''+'') and C_{ph}^2 (sign ''−'') correspond, respectively, to phonon emission and absorption. Irrespective of the type of the matrix element of Coulomb interaction, the phonon-assisted Auger process is thresholdless. This corresponds to the main contribution to the Auger coefficient C_{ph} coming from the hole momenta of the same order of magnitude as the thermal momentum. Therefore, in calculating C_{ph}, we may substitute for the longitudinal hole momentum q_h its mean thermal value. Expression (65) can be analyzed easily when the temperature is much lower than the optical phonon energy. In this case the thermal momenta of holes q_h may be neglected as compared with the phonon momentum Q, approximately equal to the threshold one. It is readily seen that the probability of Auger transition as a function of Q for phonon emission has two extrema. The first of these corresponds to the minimum of the denominator in Eq. (65) and is related to a resonant Auger transition. Note that for an Auger transition with phonon absorption no extremum of this kind is observed and there is no resonant process. The second extremum corresponds to the maximum of the squared matrix element and, as a rule, is related to a virtual Auger transition. For sufficiently wide QW's the matrix element of the electron-electron interaction as a function of the heavy-hole momentum has a form close to the δ function. In this case the second extremum predominates, and the process of scattering by phonons is virtual. With decreasing QW width, the δ function broadens for the quasithreshold matrix element, and, in addition, the role of the thresholdless matrix element, only slightly depending on Q, becomes more significant. This enhances the resonant Auger transition and weakens the virtual process. For narrow QW's the matrix element of Coulomb electron-electron interaction depends on Q only slightly, and, therefore, the resonant process is predominant. It can be shown that in this case the following estimation is valid for the AR coefficient of the phonon-assisted Auger process:

$$C_{ph} \approx \frac{\omega_{LO} e^2 m_h a}{8 \tilde{\kappa} \hbar \Gamma} J_{1,1}(Q_0) \frac{2\pi}{\hbar} \frac{3k(E_g)}{4E_g} |M_{ee}(Q_0)|^2, \tag{66}$$

where $Q_0 = \sqrt{2 m_h \omega_{LO}/\hbar}$. Hence, it immediately follows that the phonon-assisted to direct AR coefficient ratio has the form

$$\frac{C_{ph}}{C} \approx \frac{\Gamma_{hp}}{\Gamma} \frac{[M_{ee}(Q_O)]^2}{[M_{ee}(q_T)]^2}, \tag{67}$$

where $C = C_1 + C_2$ is the Auger coefficient for the direct process, $\Gamma_{hp} = \hbar/\tau_{hp}$, τ_{hp} is the time of hole scattering by an optical phonon, and q_T is the thermal momentum of holes. It can be seen that the phonon-assisted Auger process may dominate the direct one only in the case when the values of Γ_{ph} and Γ are close to each other or at an extremely low temperature when the ratio of the matrix elements taken at the momenta Q_O and q_T is large. Note that at high nonequilibrium carrier densities, when Auger recombination becomes at all significant, the hole-hole scattering turns out to be generally a much more effective mechanism of relaxation than the hole-phonon scattering. This results in a small Γ_{ph}/Γ ratio. Therefore, the direct AR mechanism dominates the phonon-assisted one down to a very low temperature. Figure 6 shows the coefficients of the phonon-assisted Auger transition (C_{ph} and direct one as a function of temperature for different quantum-well widths). The parameter Γ is taken to be equal to a characteristic value of 20 meV.

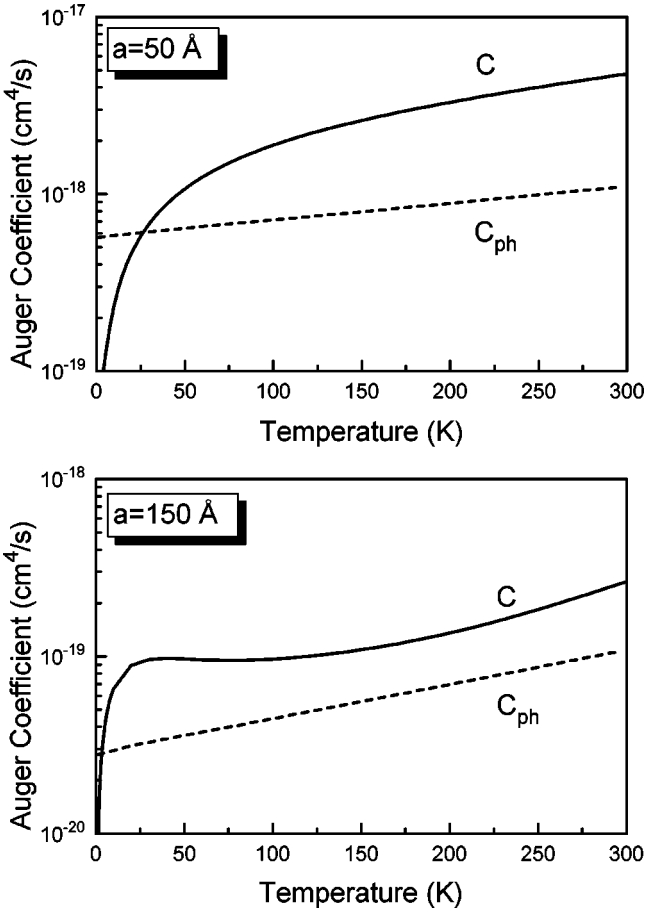

FIG. 6. Phonon-assisted and direct Auger coefficients as functions of temperature at different QW widths.

VI. SUMMARY

Our analysis has shown that for the CHCC and CHHS processes in semiconductor structures with QW's there exist three AR mechanisms: thresholdless, quasithreshold, and threshold. The first one depends on temperature only slightly. The effective threshold energy for the second process substantially varies with the QW width (Fig. 2). Namely, it is close to zero for sufficiently narrow QW's and approaches the bulk value in the limit $a \to \infty$. For this reason there is no clear distinction between the thresholdless and

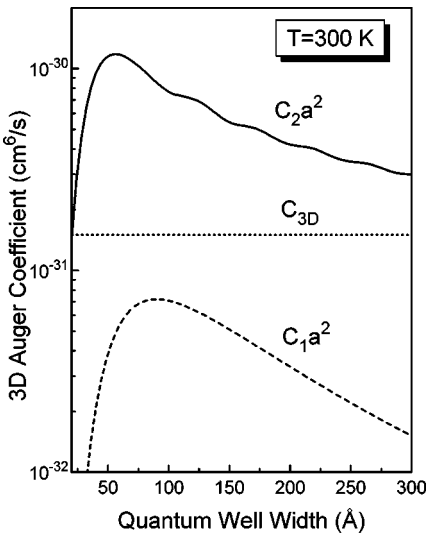

FIG. 7. Three-dimensional Auger coefficients $C_1\ a^2$ and $C_2\ a^2$ as functions of QW width at $T=300$ K. The horizontal dotted line corresponds to the bulk Auger coefficient (C_{3D}).

quasithreshold mechanisms in narrow quantum wells, and they may be considered as a single thresholdless process. The third, threshold process, is similar to that in bulk semiconductors. The only distinction is that its threshold energy is somewhat higher than in the bulk semiconductor because of an increase in the effective band gap [Eq. (47)]. The rate of this process is small compared with the rates of the first two processes in narrow QW's. In the limit $a \gg a_c$ the quasithreshold and threshold Auger processes merge and form a bulk AR process ($C_2a^2+C_3a^2 \rightarrow C^{3D}$). The critical QW width (a_c) is a strong (exponential) function of temperature and may be up to several thousand angstroms at room temperature for semiconductors with $E_g \approx 1$ eV. The thresholdless Auger process ceases to be operative on passing to the homogeneous semiconductor. For narrow QW's the 2D Auger coefficient multiplied by a^2 exceeds the 3D value owing to predominance of the thresholdless and quasithreshold processes (Fig. 7). Thus, the Auger recombination in QW's is enhanced as compared with that in a homogeneous semiconductor. This enhancement is more pronounced at a low temperature. Note that the entire analysis of the AR coefficients (C_1, C_2, C_3) as functions of temperature and QW parameters is qualitatively applicable to the same extent to both the CHCC and CHHS Auger processes. However, since no model structures with QW's have been specified, we illustrated these relations by the example of the CHCC process.

Note that the AR in QW's may be suppressed substantially if the following conditions are met: $(V_c,V_v)>E_g$ and $E_2-E_1>E_g$ (E_1 and E_2 are the energies of the first and second levels in the QW),[29] i.e., in the case when the energy of an excited particle is insufficient for a transition into the continuous spectrum or to a next energy level. For these conditions to be fulfilled, a structure is to be created with deep and narrow QW's for both electrons and holes. The structures of this kind can be fabricated on the base of InAs/AlSb (Ref. 30) or InAs/GaSb/AlSb (Ref. 31) compounds. In these deep QW's only the threshold AR mechanism, corresponding to the coefficient C_3, is operative. This coefficient may be smaller by several orders of magnitude than the Auger coefficients for the thresholdless and quasithreshold processes (C_1, C_2) in shallow QW's [$(V_c,V_v)<E_g$].

It is also shown that the phonon-assisted AR process in QW's undergoes significant changes. Similar to the direct AR there exist three different phonon mechanisms (C_{ph}^3, C_{ph}^2, C_{ph}^1) corresponding to the threshold, quasithreshold, and thresholdless matrix elements of electron-electron interaction. The first process is quite similar to its 3D counterpart. However, for narrow QW's this process is much weaker than the thresholdless and threshold Auger processes. It is this process with the participation of phonons that is often considered to be the principal AR process in QW's.[32,33] AR phonon-assisted processes with quasithreshold and thresholdless matrix elements of electron-electron interaction can be resonant. At low temperatures they can compete with direct AR processes. However, owing to the lack of any strong temperature dependence in the latter, such a competition is possible at lower temperatures than in the 3D case (Fig. 6). With increasing QW width, the resonant scattering by phonons becomes weaker, and we pass to the conventional 3D phonon-assisted AR process.

It should be emphasized once again that at high densities of nonequilibrium carriers in a homogeneous semiconductor the phonon-assisted AR process may be less intensive than the Auger recombination, with the subsequent carrier-carrier scattering eliminating the threshold.[24]

ACKNOWLEDGMENTS

The authors wish to thank R. A. Suris and V. I. Perel for initiating this work and discussing some results. This work was supported in part by the Russian Foundation for Basic Research (Grants Nos. 96-02-17952 and 97-02-18151) and the Russian State Program ''Physics of Solid State Nanostructures'' (Grants Nos. 97-0003 and 97-1035).

APPENDIX A: WAVE FUNCTIONS OF CARRIERS IN A RECTANGULAR QUANTUM WELL

1. Holes

Selecting a coordinate system so that the longitudinal components of the wave vector coincide with the y axis, and performing a Fourier transform in this plane we obtain the following expressions for the wave functions of carriers.

For heavy holes:

$$\psi_h(q,x)=H_1\begin{pmatrix} q\ \cos k_h x\xi \\ -ik_h\ \sin k_h x\xi \\ -k_h\ \sin k_h x\xi+q\ \cos k_h x\eta \end{pmatrix} + H_2\begin{pmatrix} q\ \sin k_h x\eta \\ ik_h\ \cos k_h x\eta \\ -q\ \sin k_h x\xi-k_h\ \cos k_h x\eta \end{pmatrix}, \quad \text{(A1)}$$

where q and k_h are the y and x components of the heavy-hole momentum,

$$\xi=\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \eta=\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

and H_1 and H_2 are the normalizing constants.

For light holes:

$$\boldsymbol{\psi}_l(q,x)=L_1\begin{pmatrix} k_l \sin k_l x\eta-\lambda_l q\cos k_l x\xi \\ -iq\cos k_l x\eta+i\lambda_l k_l \sin k_l x\xi \\ -\lambda_l k_l \sin k_l x\xi+\lambda_l q\cos k_l x\eta \end{pmatrix} + L_2\begin{pmatrix} -k_l\cos k_l x\xi-\lambda_l q\sin k_l x\eta \\ -i\lambda_l k_l\cos k_l x\eta-iq\sin k_l x\xi \\ -\lambda_l q\sin k_l x\xi-\lambda_l k_l\cos k_l x\eta \end{pmatrix}, \tag{A2}$$

$$\psi_{sl}=\frac{i\hbar\gamma(k_l^2+q^2)}{E_g+\delta-E}[L_1\cos k_l x\eta+L_2\sin k_l x\xi]. \tag{A3}$$

$$\lambda_l=\frac{\delta}{E+2\delta-\hbar^2k_l^2/2m_h}.$$

The wave functions of SO holes are similar to those of light holes.

The functions of another symmetry can be found from the above expressions by formal substitution $\xi\leftrightarrow\eta$ for the s,x,y components and $\xi\leftrightarrow-\eta$ for the z component. The wave functions of carriers in the barrier region may be obtained similarly to Eqs. (A1)–(A3).

If wave functions of two or more particles are considered together, then in general their z components of quasimomentum cannot become zero simultaneously. Wave functions with arbitrary direction of quasi momentum can be obtained using the rotation matrix:

$$D_\varphi=R_\varphi\otimes S_\varphi, \tag{A4}$$

where R_φ acts on the coordinate components of the wave function, and S_φ on the spinor components. The Euler angles for a rotation in the yz plane by an angle φ are

$$\Phi=-\pi/2,\quad \Theta=\varphi,\quad \Psi=\pi/2.$$

Thus

$$R_\phi=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\varphi & \sin\varphi \\ 0 & 0 & -\sin\varphi & \cos\varphi \end{bmatrix}, \tag{A5}$$

$$S_\varphi=\begin{bmatrix} \cos\varphi/2 & -i\sin\varphi/2 \\ i\sin\varphi/2 & \cos\varphi/2 \end{bmatrix}.$$

If the vector $\mathbf{q}$ has components $q(0,\cos\varphi,\sin\varphi)$ in the coordinate system x,y,z, then

$$\psi_{\mathbf{q}}\equiv\psi_\varphi=D_{-\varphi}\psi_0. \tag{A6}$$

The previously found wave function has a zero subscript. The wave function of heavy holes, found using Eq. (A6), is written below, as it will be used later,

$$\boldsymbol{\psi}_h(q,x,\phi)=H_1\begin{bmatrix} q\cos k_h x e^{-i\phi}\xi \\ -ik_h\sin k_h x\xi-q\cos k_h x\sin\phi\eta \\ -k_h\sin k_h x\xi+q\cos k_h x\cos\phi\eta \end{bmatrix} + H_2\begin{bmatrix} q\sin k_h x e^{i\phi}\eta \\ ik_h\cos k_h x\eta+q\sin k_h x\sin\phi\xi \\ -k_h\cos k_h x\eta-q\sin k_h x\cos\phi\xi \end{bmatrix}. \tag{A7}$$

The boundary conditions for hole wave functions can be derived by integrating Kane's equations across the heteroboundary (see Sec. II C). We also consider the generalized Luttinger parameters $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ continuous for the sake of simplicity. Taking into account the fact that $m_l^{-1}\approx 2\gamma^2/(E_g+\delta-E)\gg m_h^{-1}$ we obtain continuity conditions at the heteroboundary for the following quantities:

$$(1)\quad \boldsymbol{\psi},$$

$$(2)\quad \frac{\partial}{\partial x}\boldsymbol{\psi}_\perp, \tag{A8}$$

$$(3)\quad \frac{1}{E_g+\delta-E}\,\mathrm{div}\,\boldsymbol{\psi}.$$

Generally speaking, the wave functions of holes in a QW are a superposition of three subbands of the valence band: of heavy, light, and SO holes. However, the last subband strongly (exponentially) decays away from the heteroboundary with an exponent $\kappa_{so}\approx\sqrt{(m_h\Delta/\hbar^2)^{\frac{4}{3}}}$. As a consequence, this branch mainly affects the derivative of the wave function near the heteroboundary, and its influence on the wave function itself is negligible. It should be emphasized that such an approximation is not equivalent to using a 4×4 Hamiltonian from the very beginning. We shall seek the wave function as a superposition of the heavy- and light-hole subbands. Near the upper edge of the valence band $|\lambda_{so}|\approx m_h/m_l\gg1$. This means that only the first and the third of the boundary conditions (A8) are applicable. In this approximation, light and heavy holes do not mix with each other and have different spectra. The heavy-hole spectrum coincides with the quantum-mechanical spectrum of a particle in a rectangular QW. The dispersion equations are the following:

$$\tan k_h a/2=\frac{\kappa_h}{k_h}\quad \text{for even states},$$
$$\cot k_h a/2=-\frac{k_h}{\kappa_h}\quad \text{for odd states}. \tag{A9}$$

For light holes the states with different parities cannot be separated, and the dispersion equation becomes more cumbersome:

$$\left[\frac{E_g+\delta+V_c-E}{E_g+\delta-E}\,\frac{k_l^2+q^2}{\kappa_l^2-q^2}\,\kappa_l\,\cot\,k_l a/2+k_l\frac{2\lambda_l-1}{2\tilde{\lambda}_l-1}\right]$$
$$\times\left[\frac{E_g+\tilde{\delta}+V_c-E}{E_g+\delta-E}\,\frac{k_l^2+q^2}{\kappa_l^2-q^2}\,\kappa_l\,\tan\,k_l a/2-k_l\frac{2\lambda_l-1}{2\tilde{\lambda}_l-1}\right]$$
$$=q^2\left[\frac{2\lambda_l-1}{2\tilde{\lambda}_l-1}+\frac{E_g+\tilde{\delta}+V_c-E}{E_g+\delta-E}\,\frac{k_l^2+q^2}{\kappa_l^2-q^2}\right]^2. \tag{A10}$$

Here κ_l and κ_h denote the moduli of x quasimomentum components of light and heavy holes in the barrier region, respectively,

$$\tilde{\lambda}_i=\frac{\tilde{\delta}}{U_v+E+2\tilde{\delta}+\hbar^2\kappa_l^2/2m_h},\quad \tilde{\delta}=\frac{\tilde{\Delta}_{so}}{3}.$$

Note that at $q=0$ the light hole states also split into states with different parities. The constants H_i and L_i in Eqs. (A1), and (A2) are determined by normalization conditions. In particular:

$$H_i=\frac{1}{\sqrt{q^2+k_h^2}}\,\frac{1}{\sqrt{a+\dfrac{2}{\kappa_h}\,\dfrac{q^2}{q^2+k_h^2}}}.$$

The opposite is the case for SO holes. The components of the wave functions of light and heavy holes oscillate rapidly, and the contribution from them to the overlap integrals is negligibly small. Similarly, it is easy to verify that ψ_x and $(E_g+\delta-E)^{-1}\,\mathrm{div}\,\boldsymbol{\psi}$ are to be considered continuous in this case. The type of wave functions of SO holes is similar to that given above for light holes. Strictly speaking, with the condition $E_g-\Delta>U_v$ fulfilled, the spectrum of spin-split holes is continuous. However, when the rapidly oscillating contributions from the subbands of light and heavy holes are neglected, the spectrum may be both continuous and discrete. The dispersion equation of localized SO holes is similar to that of light holes.

2. Electrons

Electrons obey the same symmetry rules as holes. Their wave functions are similar to those of light holes, and inside the QW can be written as

$$\psi_{sc}=A_1\,\cos\,k_c x\,\eta+A_2\,\sin\,k_c x\xi, \tag{A11}$$

$$\boldsymbol{\psi_c}=\frac{i\hbar\gamma}{Z}A_1\begin{pmatrix} k_c\,\sin\,k_c x\;\eta-\lambda_c q\,\cos\,k_c x\xi\\ -iq\,\cos\,k_c x\,\eta+i\lambda_c k_c\,\sin\,k_c x\xi\\ -\lambda_c k_c\,\sin\,k_c x\xi+\lambda_c q\,\cos\,k_c x\,\eta\end{pmatrix}$$
$$+\frac{i\hbar\gamma}{Z}A_2\begin{pmatrix} -k_c\,\cos\,k_c x\xi-\lambda_c q\,\sin\,k_c x\,\eta\\ -i\lambda_c k_c\,\cos\,k_c x\,\eta-iq\,\sin\,k_c x\xi\\ -\lambda_c q\,\sin\,k_c x\xi-\lambda_c k_c\,\cos\,k_c x\,\eta\end{pmatrix}, \tag{A12}$$

where

$$Z=\frac{\mathcal{E}^2+\mathcal{E}(2E_g+2\delta)+(E_g+3\delta)E_g}{\mathcal{E}+E_g+2\delta},\quad \lambda_c=\frac{\delta}{\mathcal{E}+E_g+2\delta}. \tag{A13}$$

Here q and k_c denote the y and x component of the quasimomentum of the electron. Functions with another symmetry can be derived by the same procedure as that used for holes. From the boundary condition follows that ψ_s and ψ_x must be continuous. This yields the following dispersion equation:

$$\left(k_c\,\tan\,k_c a/2-\frac{Z}{\tilde{Z}}\kappa_c\right)\left(k_c\,\cot\,k_c a/2+\frac{Z}{\tilde{Z}}\kappa_c\right)$$
$$=-q^2\left(\lambda_c-\tilde{\lambda}_c\,\frac{Z}{\tilde{Z}}\right)^2, \tag{A14}$$

where κ_c is the modulus of the x quasimomentum component of electrons in the barrier region,

$$\tilde{Z}=\frac{\mathcal{E}^2+\mathcal{E}(2E_g+2U_v+2\tilde{\delta})+(E_g+U_v+3\tilde{\delta})(E_g+U_v)}{\mathcal{E}+E_g+U_v+2\tilde{\delta}},$$

$$\tilde{\lambda}_c=\frac{\tilde{\delta}}{\mathcal{E}+E_g+U_v+2\tilde{\delta}}.$$

The spectrum splits into even and odd states if the longitudinal wave vector (q) is small or the expression in parentheses in the right-hand side of the equation is close to zero. The last condition is commonly fulfilled, since, as a rule, $U_v\ll E_g$, which corresponds to semiconductors with about the same band structure. Note that in the case of discontinuous Kane's parameter the $\gamma\neq\mathrm{const}$, the continuity of $\gamma\psi_x$ and ψ_s should be used.[19]

APPENDIX B: COULOMB POTENTIAL IN THE PRESENCE OF HETEROBOUNDARIES

Near the interface of two media with different dielectric constants, the potential of a point charge differs from that in a homogeneous medium.[34] A similar situation takes place in the presence of two interfaces; therefore, in a QW the electron potential takes the form

$$\Phi(\mathbf{r}_0,\mathbf{r})=\frac{e}{\kappa_0|\mathbf{r}-\mathbf{r}_0|}+\tilde{\Phi}(\mathbf{r}_0,\mathbf{r}), \tag{B1}$$

where $\mathbf{r}_0$ is the coordinate of the particle and $\mathbf{r}$ is the coordinate of the point where the potential is observed. We consider only the case when the particle is inside the QW $(|x_0|<a/2)$. Using the reflection method[34] we obtain

$$\tilde{\Phi}=\sum_{n\geqslant 1}\frac{e}{\kappa_0}\left(\frac{\kappa_0-\tilde{\kappa}_0}{\kappa_0+\tilde{\kappa}_0}\right)^{2n-1}\left(\frac{1}{\sqrt{[x+x_0-(2n-1)a]^2+\rho^2}}+\frac{1}{\sqrt{[x+x_0+(2n-1)a]^2+\rho^2}}\right)$$
$$+\sum_{n\geqslant 1}\frac{e}{\kappa_0}\left(\frac{\kappa_0-\tilde{\kappa}_0}{\kappa_0+\tilde{\kappa}_0}\right)^{2n}\left(\frac{1}{\sqrt{(x-x_0-2na)^2+\rho^2}}+\frac{1}{\sqrt{(x-x_0+2na)^2+\rho^2}}\right)\quad \text{at } |x|<a/2$$
$$\tilde{\Phi}=\frac{e}{\kappa_0\sqrt{(x-x_0)^2+\rho^2}}\frac{\kappa_0-\tilde{\kappa}_0}{\kappa_0+\tilde{\kappa}_0}+\frac{2e}{\kappa_0+\tilde{\kappa}_0}\sum_{n\geqslant 1}\left(\frac{\kappa_0-\tilde{\kappa}_0}{\kappa_0+\tilde{\kappa}_0}\right)^{2n}\frac{1}{\sqrt{(x-x_0+2na)^2+\rho^2}}$$
$$+\frac{2e}{\kappa_0+\tilde{\kappa}_0}\sum_{n\geqslant 1}\left(\frac{\kappa_0-\tilde{\kappa}_0}{\kappa_0+\tilde{\kappa}_0}\right)^{2n-1}\frac{1}{\sqrt{[x+x_0+(2n-1)a]^2+\rho^2}}\quad \text{at } x>a/2. \tag{B2}$$

Here $\rho^2=(y-y_0)^2+(z-z_0)^2$. These potentials are rather cumbersome. However, they can be simplified if the dielectric constants κ_0 and $\tilde{\kappa}_0$ are supposed to be close to each other. After taking Fourier transform along lateral coordinates (y and z) we get

$$\phi(x,x_0,q)\approx\frac{e}{2q\kappa_0}\left(e^{-q|x-x_0|}+2\frac{\kappa_0-\tilde{\kappa}_0}{\kappa_0+\tilde{\kappa}_0}\cosh[q(x+x_0)]e^{-qa}\right)\quad \text{at } |x|<a/2,$$
$$\phi(x,x_0,q)\approx\frac{e}{q(\kappa_0+\tilde{\kappa}_0)}\left(e^{-q(x-x_0)}+\frac{\kappa_0-\tilde{\kappa}_0}{\kappa_0+\tilde{\kappa}_0}e^{-q(x+x_0+a)}\right)\quad \text{at } x>a/2. \tag{B3}$$

It can be seen that while the potential itself is a continuous function across the interface and the difference between its left and right derivatives is proportional to $(\kappa_0-\tilde{\kappa}_0)/(\kappa_0+\tilde{\kappa}_0)$.

[1] A. R. Beattie and P. T. Landsberg, Proc. R. Soc. London, Ser. A **249**, 16 (1959).
[2] B. L. Gel'mont, Zh. Eksp. Teor. Fiz. **75**, 536 (1978) [Sov. Phys. JETP **48**, 268 (1978)].
[3] A. Haug, J. Phys. C **16**, 4159 (1983).
[4] M. Takeshima, Phys. Rev. B **26**, 917 (1982); **28**, 2039 (1983).
[5] G. P. Agrawal and N. K. Dutta, *Long-Wavelength Semiconductor Lasers* (Van Nostrand Reinhold, New York, 1993).
[6] M. Aidaraliev *et al.*, Fiz. Tekh. Poluprovodn. **26**, 249 (1992) [Sov. Phys. Semicond. **26**, 138 (1992)].
[7] G. G. Zegrya and V. A. Kharchenko, Zh. Eksp. Teor. Fiz. **101**, 32 (1992) [Sov. Phys. JETP **74**, 173 (1992)].
[8] *Quantum Well Lasers*, edited by Peter S. Zory, Jr. (Academic Press, San Diego, 1993).
[9] R. I. Taylor *et al.*, IEE Proc. **132**, 364 (1985).
[10] W. W. Lui *et al.*, Phys. Rev. B **48**, 8814 (1993).
[11] N. K. Dutta, J. Appl. Phys. **54**, 1236 (1983).
[12] G. G. Zegrya *et al.*, in 23rd International Symposium on Compound Semiconductors (ISCS-23). (St. Petersburg, Russia, 1996), [Inst. Phys. Conf. Ser. **155**, 795 (1996)].
[13] R. I. Taylor *et al.*, Semicond. Sci. Technol. **5**, 90 (1990).
[14] M. I. Dyakonov and V. Yu. Kachorovskii, Phys. Rev. B **49**, 17 130 (1994).
[15] G. G. Zegrya *et al.*, Proc. SPIE **2399**, 307 (1995).
[16] E. O. Kane, J. Phys. Chem. Solids **1**, 249 (1957).
[17] P. C. Sercel and K. J. Vahala, Phys. Rev. B **42**, 3690 (1990).
[18] R. A. Suris, Fiz. Tekh. Poluprovodn. **20**, 2008 (1986) [Sov. Phys. Semicond. **20**, 1258 (1986)].
[19] M. G. Burt, J. Phys.: Condens. Matter **4**, 6651 (1992).
[20] B. A. Foreman, Phys. Rev. B **49**, 1757 (1994).
[21] G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (Wiley, New York, 1974).
[22] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, New York, 1977).
[23] G. G. Zegrya and A. S. Polkovnikov, Zh. Eksp. Teor. Fiz. **113**, 1491 (1998) [JETP (to be published)].
[24] G. G. Zegrya and A. S. Polkovnikov (unpublished).
[25] J. Wang, P. von Allmen, J.-P. Leburton, and K. J. Linden, IEEE J. Quantum Electron. **QE-31**, 864 (1995).
[26] A. Haug, J. Phys. Chem. Solids **49**, 599 (1988).
[27] A. Haug, Appl. Phys. A: Solids Surf. **51**, 354 (1990)
[28] B. K. Ridley, J. Phys. C **15**, 5899 (1982).
[29] G. G. Zegrya, in *Antimonide Related Strained Layer Heterostructures*, edited by M. O. Manasreh (Gordon and Breach Science Publishers, Amsterdam, 1997).
[30] S. Ideshita *et al.*, Appl. Phys. Lett. **60**, 2594 (1992).
[31] M. Sweeny and J. Xu, Appl. Phys. Lett. **54**, 546 (1989).
[32] E. P. O'Reilly and M. Silver, Appl. Phys. Lett. **63**, 3318 (1993).
[33] E. P. O'Reilly and A. R. Adams, IEEE J. Quantum Electron. **30**, 366 (1994).
[34] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon, New York, 1977).