# New mechanism of Auger recombination of nonequilibrium current carriers in semiconductor heterostructures

G. G. Zegrya and V. A. Kharchenko

*A. F. Ioffe Physicotechnical Institute of the Academy of Sciences of the USSR*
(Submitted 23 July 1991)
Zh. Eksp. Teor. Fiz. **101**, 327–343 (January 1992)

The effect of a heteroboundary on nonradiative recombination of nonequilibrium carriers in semiconductor structures was investigated. It is shown that the presence of the heteroboundary results in the appearance of a new no-threshold mechanism of Auger recombination: an electron-hole pair is annihilated in the volume of the narrow-band-gap semiconductor and a fast Auger particle is ejected from the subbarrier-motion region located next to the heteroboundary. It was established that the new recombination mechanism has a weak power-law temperature dependence and, owing to the efficient long-range action, it is the dominant process of nonradiative recombination in micron-size heterostructures at sufficiently low temperatures. The rates of nonradiative recombination were calculated for heterostructures with different values of the parameters.

The principal objects of investigation in the physics of semiconductors are at present heterostructures, such as quantum drops, quantum wells, superlattices, etc. The distinguishing feature of these structures is their strong spatial nonuniformity, resulting from the existence of the heteroboundaries. It is obvious that the heteroboundaries have a fundamental effect on the behavior of the energies and wave functions of carriers in quantum-size systems, while their effect on the properties of macroscopic objects was believed to be negligibly small to the extent that the ratio of the number of surface (boundary) states and volume states is small. In particular, the efficiency of nonradiative Auger-recombination processes was assumed to correspond to the standard volume characteristics.[1,2] The presence of a heteroboundary, as such, nonetheless substantially affects the electron-electron interaction in semiconductor structures, and this effect is fundamental. The heteroboundary removes the restrictions which the energy and momentum conservation laws impose on the interelectron collision processes, and this results in the appearance of no-threshold, weakly temperature-dependent channels for Auger recombination.

In this work we observed and investigated a new mechanism of Auger recombination of volume states—so-called combined Auger decay. The rate of this no-threshold process exhibits a weak power-law temperature dependence, and the process itself is the dominant mechanism of nonradiative recombination of nonequilibrium carriers in micron-size and smaller heterostructures at sufficiently low temperatures. For example, at room temperature the combined Auger process competes with the standard volume mechanism of Auger decay, but in contradistinction to the latter it does not have a threshold dependence. We give a qualitative interpretation of the processes occurring with the combined type of Auger recombination.

**1.** The restrictions imposed on the Auger decay process by the momentum conservation law are lifted if at least one of the quasiparticles (say, the electron) is located in the region of subbarrier motion near the heteroboundary. Actually, the wave function of the volume motion of quasiparticles in this region is a wave packet consisting of quasimomenta, including momenta which correspond to the final momentum of the fast Auger electron.

**2.** In order for a combined Auger transition to be realized an energy equal to the energy released when an electron-hole pair is annihilated must be imparted to the particle in the subbarrier region. In contrast to the standard volume process the quasimomentum transferred in a Coulomb collision of the particles need not be large, and therefore there is no need for close particle collisions. The processes in which energy is transferred to a subbarrier electron occur in a large volume region of free motion of the collapsing electron-hole pair. The size of this region is determined by the long-range character of the Coulomb interaction and is limited by the screening radius or by the thermal wavelength of the recombining particles.

**3.** The process of energy transfer to the fast Auger electron is resonant (to a large extent analogous to resonance photoexcitation) and does not have the standard exponential temperature dependence due to the existence of a kinematic threshold of the reaction for free quasiparticles. The absence of a threshold is a consequence of the strong spatial nonuniformity of the heteromedium.

**4.** The fast Auger electrons are ejected predominantly in the direction perpendicular to the heteroboundary, i.e., the direction of maximum spatial nonuniformity in the system. In what follows we shall study heterostructures in which an Auger electron arises as a result of recombination ($n$-type structure). The basic results can nonetheless be extended to processes in which Auger holes are formed ($p$-type structures).

The aim of our theoretical investigation is to clarify the role of the heteroboundary in the formation of no-threshold mechanisms of Auger recombination. We intentionally avoid systems with strong size quantization,[1)] in order to study the exclusive role of the heteroboundary as a source of spatial inhomogeneity. We shall calculate below the rate of Auger recombination at different temperatures and for different values of the physical parameters of a semiconductor structure with a flat heterojunction.

 0038-5646/92/010173-09$05.00  

## 1. WAVE FUNCTIONS OF CARRIERS IN THE HETEROSTRUCTURE AND THE AUGER RECOMBINATION RATE

To find the Auger recombination rage $G$ it is first necessary to calculate the wave functions of the carriers in the presence of the heteroboundary. Figure 1 shows schematically the energy-band diagram of an ideal flat heterojunction of a narrow-gap semiconductor ($x<0$) and a wide-gap semiconductor ($x>0$) with band gaps $E_g^-$ and $E_g^+$, respectively. The nonequilibrium carriers are generated in the narrow-gap semiconductor crystal under conditions of stationary excitation. As established for volume Auger processes,[5] the wave functions of the carriers must be calculated in the multiband approximation, since the probability of an Auger process is proportional to the small overlap of the electron and hole wave functions. We employ below the simplest multiband approximation—Kane's model[6]—with vanishingly small spin-orbit interaction. In this model the basis wave functions at the bottom of the conduction band and at the top of the valence band are chosen to be of the form of $|S\rangle$ and $|X\rangle$, $|Y\rangle$, and $|Z\rangle$ states (the $x$ axis is perpendicular to the plane of the heterojunction).

The wave functions of the quasiparticles are sought in the form of a superposition of band states

$$\psi = u(\mathbf{r})|S\rangle + v_x(\mathbf{r})|X\rangle + v_y(\mathbf{r})|Y\rangle + v_z(\mathbf{r})|Z\rangle,$$

where $u(\mathbf{r})$ and $\mathbf{v}(\mathbf{r}) = \{v_x(\mathbf{r}); v_y(\mathbf{r}); v_z(\mathbf{r})\}$ are smooth Bloch envelope functions. The system of equations for the envelope functions is obtained by the standard method:[7]

$$[\Delta(x) - E]u + \gamma \hat{\mathbf{k}}\mathbf{v} = 0, \quad \gamma \hat{\mathbf{k}} u + [-\Delta(x) - E]\mathbf{v} = 0,$$

where $\hat{\mathbf{k}} = -i\hat{\nabla}$,

$$\gamma = -\frac{i\hbar^2}{m}\left\langle S\left|\frac{\partial}{\partial x}\right|X\right\rangle$$

is the Kane matrix element, which is the same for the wide- and narrow-gap semiconductors, and $\Delta(x) = E_g/2$ is the only $x$-dependent structural parameter in Eq. (1). Everywhere below we shall be concerned with a structure having a vertical barrier

$$\Delta(x) = \begin{cases} \Delta_-, & x<0 \\ \Delta_+, & x>0. \end{cases}$$

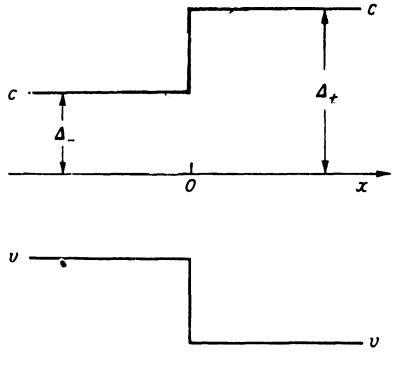



FIG. 1. Schematic band diagram of a heterojunction. The point $x=0$ corresponds to the heteroboundary; $E_g^{\pm} = 2\Delta_{\pm}$.

The spectrum of energies $E$ in the volume of the conductor is divided, according to Eq. (1), into an electron branch $E_e$ and a light-hole branch $E_{hl}$ with Dirac dispersion relation $E_{e,hl} \pm (\Delta^2 + \gamma^2\mathbf{k}^2)^{1/2}$ and a heavy-hole branch $E_h = \hbar^2\mathbf{k}^2/2m_h$ that does not interact with the other two branches. The effective mass $m_h$ of a heavy hole in the Kane model, which we have chosen, is an independent parameter and is determined by the interaction with other bands not included in the basis.

To be specific, we study the wave functions of the electrons and light holes. The boundary conditions at the heterojunction have the form[7]

$$u^<(x=0) = u^>(x=0), \quad \left(\frac{1}{E+\Delta_-}\frac{\partial u^<}{\partial x}\right)_{x=0} = \left(\frac{1}{E+\Delta_+}\frac{\partial u^>}{\partial x}\right)_{x=0}. \tag{2}$$

The wave function components $v_y$ and $v_z$ parallel to the interface are discontinuous. The thermal kinetic energy of the carriers is, as a rule, less than the barrier height of the heterostructure, i.e., $T \ll \Delta_+ - \Delta_-$, so that for the initial states of the quasiparticles the heterojunction is a perfectly reflecting boundary. Using Eqs. (1) and (2), we express the amplitude of the reflected wave in terms of the amplitude of the wave incident on the heteroboundary with quasimomentum $\mathbf{k}$. The electron or light-hole waves formed on reflection have the same functional form and differ only by the sign of $E$ ($E>0$ for electrons and $E<0$ for holes): for $x<0$ (the region of free motion)

$$\Psi_{\mathbf{k}}(\mathbf{r}) = A_{\mathbf{k}} e^{i\mathbf{q}\rho} \begin{vmatrix} \exp(ik_x x) + \exp(-ik_x x + 2i\delta) \\ \frac{\gamma k_x}{E+\Delta_-}[\exp(ik_x x) - \exp(-ik_x x + 2i\delta)] \\ \frac{\gamma q_y}{E+\Delta_-}[\exp(ik_x x) + \exp(-ik_x x + 2i\delta)] \\ \frac{\gamma q_z}{E+\Delta_-}[\exp(ik_x x) + \exp(-ik_x x + 2i\delta)] \end{vmatrix} \tag{3}$$

and for $x \geqslant 0$ (subbarrier part)

$$\Psi_{\mathbf{k}}(\mathbf{r}) = 2A_{\mathbf{k}} \cos\delta e^{i\delta} e^{i\mathbf{q}\rho - \varkappa x} \begin{vmatrix} 1 \\ \frac{i\gamma\varkappa}{E+\Delta_+} \\ \frac{\gamma q_y}{E+\Delta_+} \\ \frac{\gamma q_z}{E+\Delta_+} \end{vmatrix}. \tag{4}$$

In the formulas (3) and (4) $2\delta$ is the phase shift arising when the wave is reflected; $\mathbf{k} \equiv (K_x, \mathbf{q})$ is the wave vector of the incident particle ($k_x > 0$); $\varkappa$ is the subbarrier decay constant of the wave; $A_{\mathbf{k}}$ is a normalization factor; and,

$$\begin{aligned} \operatorname{tg}\delta &= -\frac{\varkappa}{k_x}\frac{E+\Delta_-}{E+\Delta_+}, \\ k_x^2 &= \frac{1}{\gamma^2}(E^2 - \Delta_-^2 - \gamma^2 q^2), \\ \varkappa^2 &= \frac{1}{\gamma^2}(\Delta_+^2 - \Delta_-^2 - \gamma^2 k_x^2), \\ A_{\mathbf{k}}^2 &= \frac{1}{2}\left(1 + \frac{\Delta_-}{E}\right). \end{aligned} \tag{5}$$

The normalization corresponds to an incident wave of unit amplitude. From the formulas (3)–(5) presented above it follows that the electron and hole waves are reflected from the heteroboundary differently for different components of the envelopes. Thus, in contrast to all components in Eq. (3), the $|X\rangle$ component of the reflected wave acquires, apart from the standard phase $2\delta$ due to reflection, an additional minus sign owing to the change in the sign of the **x**-component of the quasimomentum on reflection: $k_x \to -k_x$. The sign of the components of the valence band depends on the direction of **k** because of the vector nature of the Bloch amplitudes of the states at the top of the valence band. In the limit of small quasimomenta the behavior of the wave function should actually correspond to the single-band approximation, i.e., it should correspond to the standard reflection from a potential barrier. Indeed, in the limit $k \to 0$ it follows from Eq. (5) that for electrons ($E > 0$)

$$\operatorname{tg}\delta_e \approx -\frac{2\Delta_-}{\gamma k_x}\left(\frac{\Delta_+-\Delta_-}{\Delta_++\Delta_-}\right)^{1/2}, \quad 2\delta_e \approx \pi + \frac{\gamma k_x}{\Delta_-}\left(\frac{\Delta_++\Delta_-}{\Delta_+-\Delta_-}\right)^{1/2}. \tag{6}$$

while for holes ($E = -(\Delta^2+\gamma^2\mathbf{k}^2)^{1/2} < 0$) incident at an angle $\theta$ with respect to the interface

$$\operatorname{tg}\delta_h = \frac{\gamma k_x}{2\Delta_-}\left(\frac{\Delta_++\Delta_-}{\Delta_+-\Delta_-}\right)^{1/2}\left(1+\frac{\mathbf{q}^2}{k_x^2}\right),$$
$$2\delta_h = \frac{\gamma k_x}{\Delta_-}\left(\frac{\Delta_++\Delta_-}{\Delta_+-\Delta_-}\right)^{1/2} \Big/ \cos^2\theta. \tag{7}$$

The phases of the reflections for electrons and holes in the case of normal incidence on the interface ($\theta = 0$) differ only by $\pi$; this is taken into account by the change in the sign of the hole function on reflection. The dependence of $\delta_h$ on the angle of incidence on the heteroboundary is related with the different behavior of the parallel **q** and normal $k_x$ components on reflection: **q** is conserved while $k_x$ changes sign. The values of the scattering phases are important for calculating the rate of the combined Auger process, since the phases, essentially, fix the value of the wave functions at the heteroboundary and in the subbarrier region [see the formula (4) for $x \geqslant 0$].

The wave function of fast Auger electrons is a superposition of the incident and reflected waves in the narrow-gap part ($x \leqslant 0$) and the transmitted electron wave in the wide-gap part of the heterostructure:

$$\Psi^{<}_{\mathbf{k}_f}(\mathbf{r}) = A_{\mathbf{k}_f}\exp(i\mathbf{q}_f\boldsymbol{\rho}) \times \begin{vmatrix} \exp(ik_{fx}^{-}x) + f\exp(-ik_{fx}^{-}x) \\ \dfrac{\gamma k_{fx}^{-}}{E_f+\Delta_-}\left[\exp(ik_{fx}^{-}x) - f\exp(-ik_{fx}^{-}x)\right] \\ \dfrac{\gamma q_{fy}}{E_f+\Delta_-}\left[\exp(ik_{fx}^{-}x) + f\exp(-ik_{fx}^{-}x)\right] \\ \dfrac{\gamma q_{fz}}{E_f+\Delta_-}\left[\exp(ik_{fx}^{-}x) + f\exp(-ik_{fx}^{-}x)\right] \end{vmatrix}, \quad x < 0,$$

$$\Psi^{>}_{\mathbf{k}_f}(\mathbf{r}) = A_{\mathbf{k}_f} d_f \begin{vmatrix} 1 \\ \dfrac{\gamma k_{fx}^{+}}{E_f+\Delta_+} \\ \dfrac{\gamma q_{fy}}{E_f+\Delta_+} \\ \dfrac{\gamma q_{fz}}{E_f+\Delta_+} \end{vmatrix} \exp\left[i\left(k_{fx}^{+}x + \mathbf{q}_f\boldsymbol{\rho}\right)\right], \quad x > 0, \tag{8}$$

where $f$ and $d_f$ are the reflection and transmission amplitudes; $\mathbf{k}_f^{+} \equiv (k_{fx}^{+}, \mathbf{q}_f)$ and $\mathbf{k}_f^{-} \equiv (k_{fx}^{-}, \mathbf{q}_f)$ are the wave vectors of the Auger electron in the wide- and narrow-gap materials, respectively; $A_{kf}$ is a normalization factor; and,

$$f = \frac{k_{fx}^{-} - k_{fx}^{+}[(E_f+\Delta_-)/(E_f+\Delta_+)]}{k_{fx}^{-} + k_{fx}^{+}[(E_f+\Delta_-)/(E_f+\Delta_+)]},$$
$$d_f = 2k_{fx}^{-}\Big/\left[k_{fx}^{-} + k_{fx}^{+}\left(\frac{E_f+\Delta_-}{E_f+\Delta_+}\right)\right],$$
$$A_{\mathbf{k}_f} = \left\{\frac{(E_f+\Delta_-)(E_f+\Delta_+)}{2E_f[2E_f+f^2(\Delta_+-\Delta_-)+\Delta_++\Delta_-]}\right\}^{1/2}. \tag{9}$$

We note that the wave functions (8) correspond to situations when the narrow-gap material is the source of Auger electrons, while in the wide-gap material there are no free carriers, with the exception of electrons which have passed through the heterobarrier from the narrow-gap part. Using Eq. (8), we calculate the probability that electrons are transferred from the narrow-gap into the wide-gap region for the most important physical process.

According to the standard rules of the theory of Auger processes,[5] the rate of nonradiative Auger recombination is calculated to first-order perturbation theory in the electron-electron interaction as follows:

$$G = \frac{2\pi}{\hbar}\int\ldots\int_{(k_{fx}\geqslant 0)} \langle|M(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_h,\mathbf{k}_f)|^2\rangle g(\mathbf{k}_1) \times g(\mathbf{k}_2)\,g_h(\mathbf{k}_h)\,[1-g(\mathbf{k}_f)] \times \delta(E_1+E_2-E_h-E_f)\,\frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3}\frac{d^3k_h}{(2\pi)^3}\frac{d^3k_f}{(2\pi)^3}, \tag{10}$$

where $g(\mathbf{k})$ is the momentum distribution function of nonequilibrium electrons and $g_h(\mathbf{k}) = 1 - g(\mathbf{k})$ is the momentum distribution function of holes; $E_1$ and $E_2$ are the initial energy states and $E_f$ and $E_h$ are the final energy states of the electrons (we regard the hole state as the final state for one of the electrons participating in the Auger process); $M$ is the Auger transition matrix element, calculated with antisymmetrized electronic wave functions of the initial and final states; after taking the statistical average over the spin states of the system, the squared modulus of the matrix element has the form

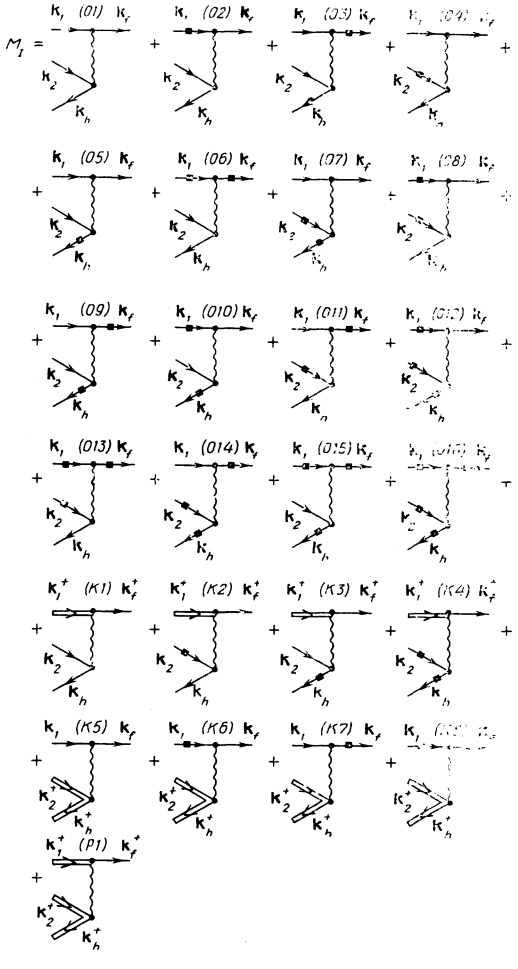
$$\langle|M(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_h,\mathbf{k}_f)|^2\rangle = M_{\mathrm{I}}^2 + M_{\mathrm{II}}^2 - M_{\mathrm{I}}M_{\mathrm{II}},$$
$$M_{\mathrm{I}} = \int \Psi_{\mathbf{k}_f}(\mathbf{r})\,\Psi_{\mathbf{k}_1}^{*}(\mathbf{r})\,d^3r\,\frac{e^2}{\varkappa_0|\mathbf{r}-\mathbf{r}'|}\,\Psi_{\mathbf{k}_h}(\mathbf{r}')\,\Psi_{\mathbf{k}_2}^{*}(\mathbf{r}')\,d^3r',$$
$$M_{\mathrm{II}} = \int \Psi_{\mathbf{k}_h}(\mathbf{r})\,\Psi_{\mathbf{k}_1}^{*}(\mathbf{r})\,d^3r\,\frac{e^2}{\varkappa_0|\mathbf{r}-\mathbf{r}'|}\,\Psi_{\mathbf{k}_f}(\mathbf{r}')\,\Psi_{\mathbf{k}_2}^{*}(\mathbf{r}')\,d^3r', \tag{11}$$

where $\varkappa_0$ is the permittivity of the medium. We recall that the wave functions in Eq. (11) are column vectors whose components are the corresponding components of the envelope functions, while $M_{\mathrm{I,II}}$ are called the direct and exchange matrix elements.[8] For the processes which we are studying $M_{\mathrm{I}}$ and $M_{\mathrm{II}}$ are obtained by interchanging the indices of the quasimomenta of the initial states $\mathbf{k}_1$ and $\mathbf{k}_2$ (or the final states $\mathbf{k}_f$ and $\mathbf{k}_h$). We shall analyze in detail the Auger decay matrix elements. This will allow us to determine the most important recombination processes.

## 2. CLASSIFICATION OF AUGER DECAY PROCESSES IN THE PRESENCE OF A HETEROBOUNDARY

We now study the matrix element $M_{\mathrm{I}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_h,\mathbf{k}_f)$ for Auger decay in the presence of an ideal flat heteroboundary. In this case, according to Eq. (11), the integrals over the coordinates of the interacting particles separate, in a natural manner, into the wide- and narrow-gap regions: $x, x' \leqslant 0$ for processes occurring in the narrow-gap region, which we shall call below volume processes ($O$ processes); $x, x' \geqslant 0$ for processes occurring in the wide-gap region, where all wave functions (except that of the final electron) are subbarrier wave functions ($P$ processes); $x<0, x'\geqslant 0$ and $x>0, x'\leqslant 0$ for combined processes, when one of the electrons is located in the subbarrier region and the other is located in the volume part of the narrow-gap semiconductor ($K$ processes).

It is convenient to set in correspondence to all types of terms in a matrix element diagrammatic representations of the amplitudes in which both the waves incident on and reflected from the heteroboundaries as well as the subbarrier terms are taken into account:




(12)

where the thin lines designate electron and hole waves with momenta $\mathbf{k}_1$ and $\mathbf{k}_h$, respectively, while the same lines with rectangles designate waves arising on reflection from the heteroboundary; a wavy line designates the effective Coulomb interaction between particles while double lines designate subbarrier electron and hole states (we recall that $\mathbf{k}_f^+$ is a superbarrier state of the Auger electron in the wide-gap region). To the diagrams (12) indicated above it is easy to set in correspondence analytical expressions for the probability amplitudes; in so doing, as follows from Eq. (12), $M_{\mathrm{I}}$ separates into three types of processes:

$$M_{\mathrm{I}}=M_{\mathrm{I}}^{(O)}+M_{\mathrm{I}}^{(K)}+M_{\mathrm{I}}^{(P)}, \tag{13}$$

where $M_{\mathrm{I}}^{(O)}$ are volume processes ($O1$–$O16$); $M_{\mathrm{I}}^{(K)}$ are combined processes ($K1$–$K8$); and, $M_{\mathrm{I}}^{(P)}$ are subbarrier processes ($P1$). In deriving the analytical form of $M_{\mathrm{I}}$ it is necessary to take into account the fact (see, for example, $O1$) that the standard Coulomb interaction of the particles, $4\pi e^2|\mathbf{k}_f-\mathbf{k}_1|^{-2}/\varkappa_0$, is multiplied by the product of overlap integrals of the Bloch wave functions of the initial and final states of the electrons, $I_{cc}(\mathbf{k}_1,\mathbf{k}_f)\cdot I_{cv}(\mathbf{k}_2,\mathbf{k}_h)$. We write out the explicit form of the matrix element for the volume Auger processes ($O1$–$O16$), where to each diagram there is associated the law of conservation of quasimomentum of the recombining particles:

$$M_{\mathrm{I}}^{(O)}=2\pi^3\delta(\mathbf{q}_1+\mathbf{q}_2-\mathbf{q}_h-\mathbf{q}_f)\frac{4\pi e^2}{\varkappa_0}\Bigg\{\frac{I_{cc}(\mathbf{k}_1,\mathbf{k}_f)I_{cv}(\mathbf{k}_2,\mathbf{k}_h)}{(\mathbf{q}_f-\mathbf{q}_1)^2+(k_{xf}-k_{x1})^2}$$
$$\times\{1+f\exp[2i(\delta_h-\delta_1-\delta_2)]\}\delta(k_{x1}+k_{x2}-k_{xh}-k_{xf})$$
$$+\frac{I_{cc}(\mathbf{k}_{1r},\mathbf{k}_f)I_{cv}(\mathbf{k}_2,\mathbf{k}_h)}{(\mathbf{q}_f-\mathbf{q}_1)^2+(k_{xf}+k_{x1})^2}$$
$$\times\{\exp(-2i\delta_1)+f\exp[2i(\delta_h-\delta_2)]\}\delta(k_{x2}-k_{x1}-k_{xf}-k_{xh})$$
$$+\frac{I_{cc}(\mathbf{k}_1,\mathbf{k}_{fr})I_{cv}(\mathbf{k}_2,\mathbf{k}_h)}{(\mathbf{q}_f-\mathbf{q}_1)^2+(k_{xf}+k_{x1})^2}$$
$$\times\{\exp[2i(\delta_h-\delta_1-\delta_2)]+f\}\delta(k_{x1}+k_{x2}+k_{xf}-k_{xh})$$
$$+\frac{I_{cc}(\mathbf{k}_1,\mathbf{k}_f)I_{cv}(\mathbf{k}_{2r},\mathbf{k}_h)}{(\mathbf{q}_f-\mathbf{q}_1)^2+(k_{xf}-k_{x1})^2}$$
$$\times\{\exp(-2i\delta_2)+f\exp[2i(\delta_h-\delta_1)]\}\delta(k_{x1}-k_{x2}-k_{xf}-k_{xh})$$
$$+\frac{I_{cc}(\mathbf{k}_1,\mathbf{k}_f)I_{cv}(\mathbf{k}_2,\mathbf{k}_{hr})}{(\mathbf{q}_f-\mathbf{q}_1)^2+(k_{xf}-k_{x1})^2}$$
$$\times\{\exp(2i\delta_h)+f\exp[-2i(\delta_1+\delta_2)]\}\delta(k_{x1}+k_{x2}+k_{xh}-k_{xf})$$
$$+\frac{I_{cc}(\mathbf{k}_1,\mathbf{k}_f)I_{cv}(\mathbf{k}_2,\mathbf{k}_h)}{(\mathbf{q}_f-\mathbf{q}_1)^2+(k_{xf}-k_{x1})^2}$$
$$\times\{\exp[2i(\delta_h-\delta_2)]+f\exp(-2i\delta_1)\}\delta(k_{x2}+k_{xf}-k_{x1}-\mathbf{k}_{xh})$$
$$+\frac{I_{cc}(\mathbf{k}_{1r},\mathbf{k}_f)I_{cv}(\mathbf{k}_{2r},\mathbf{k}_h)}{(\mathbf{q}_f-\mathbf{q}_1)^2+(k_{xf}+k_{x1})^2}$$
$$\times\{\exp[-2i(\delta_1+\delta_2)]+f\exp(2i\delta_h)\}\delta(k_{x1}+k_{x2}+k_{xf}+k_{xh})$$

$$+\frac{I_{cc}(\mathbf{k}_{1r},\mathbf{k}_f)I_{cv}(\mathbf{k}_2,\mathbf{k}_{hr})}{(\mathbf{q}_f-\mathbf{q}_1)^2+(k_{xf}+k_{x1})^2}$$

$$\times\{\exp[2i(\delta_h-\delta_1)]+f\exp(-2i\delta_2)\}\delta(k_{x2}+k_{xh}-k_{xf}-k_{x1})\Big\}. \tag{14}$$

The overlap integrals in the narrow-gap region have the form

$$I_{cc}(\mathbf{k}_1,\mathbf{k}_f)=I_{cc}(\mathbf{k}_{1r},\mathbf{k}_{fr})=A_{\mathbf{k}_1}^{*}A_{\mathbf{k}_f}\left[1+\frac{\gamma^2\mathbf{k}_1\mathbf{k}_f}{(E_1+\Delta_-)(E_f+\Delta_-)}\right],$$

$$I_{cc}(\mathbf{k}_1,\mathbf{k}_{fr})=I_{cc}(\mathbf{k}_{1r},\mathbf{k}_f)=A_{\mathbf{k}_1}^{*}A_{\mathbf{k}_f}\left[1+\frac{\gamma^2\mathbf{k}_1\mathbf{k}_{fr}}{(E_1+\Delta_-)(E_f+\Delta_-)}\right],$$

$$I_{cv}(\mathbf{k}_2,\mathbf{k}_h)=I_{cv}(\mathbf{k}_{2r},\mathbf{k}_{hr})=A_{\mathbf{k}_2}^{*}A_{\mathbf{k}_h}\left[1+\frac{\gamma^2\mathbf{k}_2\mathbf{k}_h}{(E_2+\Delta_-)(E_h+\Delta_-)}\right],$$

$$I_{cv}(\mathbf{k}_2,\mathbf{k}_{hr})=I_{cv}(\mathbf{k}_{2r},\mathbf{k}_h)=A_{\mathbf{k}_2}^{*}A_{\mathbf{k}_h}\left[1+\frac{\gamma^2\mathbf{k}_2\mathbf{k}_{hr}}{(E_2+\Delta_-)(E_h+\Delta_-)}\right], \tag{15}$$

where $\mathbf{k}_{ir}=(-k_{ix};\mathbf{q})$ is the wave vector of the reflected wave. We note that in deriving Eq. (14) from the diagrammatic representation (12) the reflected states of the incoming lines $\mathbf{k}_{1,2}$ correspond to multiplication of the amplitudes by $\exp(-2i\delta_{1,2})$, while the outgoing lines $\mathbf{k}_f$ and $\mathbf{k}_h$ correspond to multiplication by $f$ and $\exp(2i\delta_h)$, respectively. The appearance in Eq. (14) of delta functions corresponding to momentum conservation is a characteristic of volume processes, while the role of the heteroboundary in Eq. (14) reduces to the appearance of interference terms which depend on the phases $\delta_i$ of the reflected waves and on $f$. Thus the processes (14) satisfy a more complicated form of the law of conservation of energy-momentum and, as will become evident in what follows, they have a kinematic threshold. This means that the rate of the volume ($O$) processes depends exponentially on the temperature of the nonequilibrium carriers.

We write out the explicit expressions for the amplitudes of the combined Auger processes ($K$ 1–$K$ 8), for which the law of conservation of "momentum" along the $x$ axis is not satisfied because at least one of the quasiparticles is located in the region of subbarrier motion $x\geqslant 0$:

$$M_1^{(K)}=2\pi^3\delta(\mathbf{q}_1+\mathbf{q}_2-\mathbf{q}_h-\mathbf{q}_f)\frac{4\pi e^2}{\varkappa_0}$$

$$\times\Big\{\frac{2\cos\delta_1\cdot d_f\exp(-i\delta_1)}{(\mathbf{q}_2-\mathbf{q}_h)^2+(k_{xh}-k_{x2})^2}I_{cc}(\mathbf{k}_1^{+},\mathbf{k}_f)$$

$$\times I_{cv}(\mathbf{k}_2,\mathbf{k}_h)\left[\frac{1}{\varkappa_1-i(k_{xf}+k_{xh}-k_{x2})}+\frac{\exp[2i(\delta_h-\delta_2)]}{\varkappa_1-i(k_{xf}-k_{xh}+k_{x2})}\right]\frac{1}{\pi}$$

$$+\frac{2\cos\delta_1\cdot d_j\exp(-i\delta_1)}{(\mathbf{q}_2-\mathbf{q}_h)^2+(k_{xh}+k_{x2})^2}I_{cc}(\mathbf{k}_1^{+},\mathbf{k}_f)I_{cv}(\mathbf{k}_{2r},\mathbf{k}_h)$$

$$\times\left[\frac{\exp(-2i\delta_2)}{\varkappa_1-i(k_{xf}+k_{xh}+k_{x2})}+\frac{\exp(2i\delta_h)}{\varkappa_1-i(k_{xf}-k_{xh}-k_{x2})}\right]\frac{1}{\pi}$$

$$+\frac{2\cos\delta_2\cdot 2\cos\delta_h\exp[i(\delta_h-\delta_2)]}{(\mathbf{q}_f-\mathbf{q}_1)^2+(k_{xf}-k_{x1})^2}I_{cc}(\mathbf{k}_1,\mathbf{k}_f)$$

$$\times I_{cv}(\mathbf{k}_2^{+},\mathbf{k}_h^{+})\left[\frac{1}{\varkappa_2+\varkappa_h-i(k_{xf}-k_{x1})}+\frac{f\exp(-2i\delta_1)}{\varkappa_2+\varkappa_h+i(k_{xf}-k_{x1})}\right]\frac{1}{\pi}$$

$$+\frac{2\cos\delta_2\cdot 2\cos\delta_h\exp[i(\delta_h-\delta_2)]}{(\mathbf{q}_f-\mathbf{q}_1)^2+(k_{x1}+k_{xf})^2}I_{cc}(\mathbf{k}_{1r},\mathbf{k}_f)I_{cv}(\mathbf{k}_2^{+},\mathbf{k}_h^{+})$$

$$\times\left[\frac{\exp(-2i\delta_1)}{\varkappa_2+\varkappa_h-i(k_{xf}+k_{x1})}+\frac{f}{\varkappa_2+\varkappa_h+i(k_{x1}+k_{xf})}\right]\frac{1}{\pi}\Big\}, \tag{16}$$

where $\mathbf{k}_i^{+}=(\mathbf{q}_i;i\varkappa_i)$ is the quasimomentum in the subbarrier region for the states $i=1$, 2, $h$ of the particles, and $I_{cc}(\mathbf{k}_1^{+},\mathbf{k}_f)$ and $I_{cv}(\mathbf{k}_2^{+},\mathbf{k}_h^{+})$ are the overlap integrals of the Bloch functions in the subbarrier region and are obtained from the corresponding integrals (15) by making the substitution $\mathbf{k}_i\to\mathbf{k}_i^{+}$ and then forming the complex conjugate of the wave functions of the initial states. For example,

$$I_{cc}(\mathbf{k}_1^{+},\mathbf{k}_f)=A_{\mathbf{k}_1}^{*}A_{\mathbf{k}_f}\left[1+\frac{\gamma^2\mathbf{k}_f(\mathbf{k}_1^{+})^{*}}{(E_1+\Delta_+)(E_f+\Delta_+)}\right]$$

$$\equiv A_{\mathbf{k}_1}^{*}A_{\mathbf{k}_f}\left[1+\frac{\gamma^2(\mathbf{q}_1\mathbf{q}_f-i\varkappa_1k_{xf})}{(E_1+\Delta_+)(E_f+\Delta_+)}\right]. \tag{17}$$

In calculating Eq. (16) using the diagrams $K$ 1–$K$ 8, the rules $O$ 1–$O$ 16 formulated above must be supplemented by integration over all possible values of the $x$ component of the quasimomentum $k_x^{+}$ of the particles in the subbarrier region. Indeed, the decaying wave $\exp(-\varkappa x)$ (for $x>0$) is a wave packet with probability amplitude $a(k_x')$ for having an $x$ component $k_x'$. For example, for an electron with the momentum $\mathbf{k}_1$

$$a(k_{x1}')=2\cos\delta_1e^{i\delta_1}A_{\mathbf{k}_1}\times\int_0^{\infty}\exp(-\varkappa_1x-ik_{x1}'x)dx=\frac{2\cos\delta_1e^{i\delta_1}}{\varkappa_1+ik_{x1}'}A_{\mathbf{k}_1}, \tag{18}$$

where $\delta_1$ is the phase of the reflected wave (5). It is obvious that the integration over all possible values of $k_{xj}'$ actually removes the delta functions for the $x$ components of the quasimomentum. This in turn removes the restrictions imposed by momentum conservation on the Auger decay processes in $K$ 1–$K$ 8. Thus $M_1^{(K)}$ describes no-threshold Auger recombination processes, which do not contain the exponential temperature dependence observed in volume processes of the type $M_1^{(O)}$.

Subbarrier Auger decay processes correspond to the matrix element

$$M_1^{(P)}=2\pi^3\delta(\mathbf{q}_1+\mathbf{q}_2-\mathbf{q}_h-\mathbf{q}_f)I_{cc}(\mathbf{k}_1^{+},\mathbf{k}_f^{+})I_{cv}(\mathbf{k}_2^{+},\mathbf{k}_h^{+})$$

$$\times 2\cos\delta_1e^{-i\delta_1}\cdot 2\cos\delta_2e^{-i\delta_2}$$

$$\times 2\cos\delta_he^{i\delta_h}\frac{4\pi e^2d_f}{\varkappa_0(\varkappa_2+\varkappa_h-|\mathbf{q}_f-\mathbf{q}_h|)}$$

$$\times\left[\frac{1}{(|\mathbf{q}_f-\mathbf{q}_1|+\varkappa_1-ik_{fx}^{+})|\mathbf{q}_f-\mathbf{q}_1|}\right.$$

$$\left.-\frac{2}{(\varkappa_h+\varkappa_2+|\mathbf{q}_f-\mathbf{q}_1|)(\varkappa_1+\varkappa_2+\varkappa_h-ik_{fx}^{+})}\right]\frac{1}{\pi}. \tag{19}$$

In principle, the matrix element $M_1 = M_1^{(O)} + M_1^{(K)} + M_1^{(P)}$, which we obtained above, makes it possible to calculate, after substituting the quantity $G$ into the formula (10) for the rate of the Auger process, and thus to solve the problem posed. It is easy to see that the large number [(14), (16), and (19)] of processes accompanying Auger recombination makes it impossible to obtain in the general case analytical expressions for $G$ as a function of the parameters of the heterostructure and the temperature. There is no need to do so, however, since the two types of Auger recombination processes which we established above—the threshold process (of the type $M_1^{(O)}$) and the no-threshold process ($M_1^{(K)} + M_1^{(P)}$)—play a significant role in different parametric regions. The threshold processes are the main processes occurring in large samples at sufficiently high temperatures, while the no-threshold processes predominate in small samples at low temperatures. We shall analyze below the efficiency of Auger processes and we shall determine the parametric limits of the regions where each of the Auger processes studied is efficient.

### 3. RATE OF AUGER DECAY WITH FORMATION OF A FAST PARTICLE

For many semiconductor heterostructures, parametric relations ensuring that the electron formed as a result of the Auger process has a high energy are satisfied:

$$T \ll E_g^+ - E_g^- \ll E_g^+,\ E_g^-. \tag{20}$$

The inequality (20) will be employed below to derive asymptotic expressions for the Auger recombination rate. We shall examine the volume Auger recombination rate in the narrow-gap part of the heterostructure. According to the formulas (14), we have

$$G^{(O)} = \frac{(2\pi)^9 e^4}{2\hbar\varkappa_0^2} V \int\int\int_{(k_{1z}\geqslant 0)} g(\mathbf{k}_1) g(\mathbf{k}_2) g_h(\mathbf{k}_h) \frac{d^3k_1}{(2\pi)^3} \frac{d^3k_2}{(2\pi)^3} \frac{d^3k_h}{(2\pi)^3}$$

$$\times \Bigg\{ \frac{|I_{cv}(\mathbf{k}_2, \mathbf{k}_h)|^2}{[(\mathbf{q}_2 - \mathbf{q}_h)^2 + (k_{xh} - k_{x2})^2]^2}$$

$$\times [(1 + f^2 + 2f\cos 2(\delta_h - \delta_2 - \delta_1))\, |I_{cc}(\mathbf{k}_1; \mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_h)|^2$$

$$\times \delta(E(\mathbf{k}_1) + E(\mathbf{k}_2) - E_h(\mathbf{k}_h) - E(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_h))$$

$$+ (1 + f^2 + 2f\cos 2(\delta_h - \delta_2 + \delta_1))$$

$$\times |I_{cc}(\mathbf{k}_{1r}; \mathbf{k}_{1r} + \mathbf{k}_2 - \mathbf{k}_h)|^2 \delta(E(\mathbf{k}_1)$$

$$+ E(\mathbf{k}_2) - E(\mathbf{k}_{1r} + \mathbf{k}_2 - \mathbf{k}_h))]$$

$$+ \frac{|I_{cv}(\mathbf{k}_2, \mathbf{k}_{hr})|^2}{[(\mathbf{q}_2 - \mathbf{q}_h)^2 + (k_{x2} + k_{xh})^2]^2} [(1 + f^2 + 2f\cos 2(\delta_h - \delta_1 + \delta_2))$$

$$\times |I_{cc}(\mathbf{k}_1; \mathbf{k}_1 + \mathbf{k}_{2r} - \mathbf{k}_h)|^2 \delta(E(\mathbf{k}_1) + E(\mathbf{k}_2)$$

$$- E_h(\mathbf{k}_h) - E(\mathbf{k}_1 + \mathbf{k}_{2r} - \mathbf{k}_h))$$

$$+ {}^1/_2 (1 + f^2 + 2f\cos 2(\delta_h + \delta_1 + \delta_2))\, |I_{cc}(\mathbf{k}_1; \mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_{hr})|^2$$

$$\times \delta(E(\mathbf{k}_1) + E(\mathbf{k}_2) - E_h(\mathbf{k}_h) - E(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_{hr}))] \Bigg\}. \tag{21}$$

The Auger recombination rate is proportional to the volume of the specimen $V = SL$ ($S$ is the area of the heterojunction and $L$ is the width of the narrow-gap specimen), appearing when the squared delta functions of the matrix element $|M_1^{(O)}|^2$ are integrated. In the formula (21) are dropped small terms $\sim 1/Lk_f$, which appear when the products of delta functions (14) with different arguments are integrated. The dropped terms represent processes whose phase volumes are smaller than the main processes (for example, the product of delta functions from the first two rows of (14) corresponds to the process with $k_{x1} = 0$, which signifies the indicated smallness for a system of finite size). The specific recombination rate $g^{(O)} = G^{(O)}/V$ is a volume-independent characteristic. Analysis of the expression (21) shows that there appear interference terms owing to the presence of the heteroboundary. The phases $\delta_i$ are, in principle, functions of the quasimomenta $k_{xi}$, and averaging over the quasimomenta smoothes the interference terms. In addition, and most importantly, in a real system the quasiparticle momenta are reoriented at distances from the heteroboundary of the order of the mean-free path length, i.e., the phase relations between the incident and reflected waves are lost and hence complete averaging of the interference pattern occurs.

For the cases (20) studied above the reflection amplitude is small, $f \sim (\Delta_+ - \Delta_-)/(\Delta_+ + \Delta_-) \ll 1$, and the corresponding terms in Eq. (21) can be neglected. Analysis of the integrand in $G^{(O)}$ shows that the last term in Eq. (21) makes the main contribution to the process of Auger ejection of electrons from a narrow-gap semiconductor into the wide-gap semiconductor. We shall demonstrate for the example of the calculation of this contribution the method of analysis employed in the case of structures with a high barrier (20): $T \ll (\Delta_+ - \Delta_-)$. The integrand in Eq. (21) contains two "steep" dependences: The energy delta function and a Boltzmann distribution function:

$$g(\mathbf{k}_1) g(\mathbf{k}_2) g_h(\mathbf{k}_h) \propto \exp\{-[\varepsilon(\mathbf{k}_1) + \varepsilon(\mathbf{k}_2) + \varepsilon_h(\mathbf{k}_h)]/T\},$$

$$\varepsilon(\mathbf{k}_{1,2}) = E(\mathbf{k}_{1,2}) - \Delta_-, \quad \varepsilon_h(\mathbf{k}_h) = -\Delta_- - E_h(\mathbf{k}_h), \tag{22}$$

where $\varepsilon(\mathbf{k}_{1,2})$ and $\varepsilon_h(\mathbf{k}_h)$ are the kinetic energies of the electrons and holes, the relation between which is determined by the laws of conservation of energy and momentum. The region where the kinetic energy of the initial electrons and holes (22) is a minimum under the condition that the conservation laws are satisfied makes the main contribution to the integral:

$$\varepsilon(\mathbf{k}_1) + \varepsilon(\mathbf{k}_2) + \varepsilon_h(\mathbf{k}_h) = \min,$$

$$\varepsilon(\mathbf{k}_1) + \varepsilon(\mathbf{k}_2) + \varepsilon_h(\mathbf{k}_h) + E_g^- = \varepsilon(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_{hr}). \tag{23}$$

The conditional extremum (23) with respect to $\mathbf{k}_{1,2}$ and $\mathbf{k}_h$ can be calculated by the method of indeterminate Lagrange multipliers for arbitrary dispersion relations. Next, expanding the integrand near the extremum $\mathbf{k}_{10}$, $\mathbf{k}_{20}$, $\mathbf{k}_{h0}$) in the exponential (22) up to second-order infinitesimals, we reduce the problem to calculation of Gaussian integrals. All terms in (21) can be analyzed analogously. The minimum value of the energy in Eq. (23) gives the kinetic threshold of the reaction. For information we present the computed val-

ues of the specific recombination rate $g_{hl}^{(O)} = G_{hl}^{(O)}/V$ for quasiparticles with a quadratic dispersion law and $m_e = m_{hl}$:

$$g_{hl}^{(O)} = 18\pi^8(2\pi)^{1/2}\frac{E_B}{\hbar}\left(\frac{E_g}{T}\right)^{1/2}\left(\frac{\hbar^2}{m_eE_g}\right)^3\exp\left(-\frac{E_g}{2T}\right)n_e^2n_h, \tag{24}$$

where $E_B = m_e e^4/2\hbar^2\varkappa_0^2$ is the characteristic (Bohr) interelectron interaction energy; $n_{e,h}$ is the carrier density in the narrow-gap semiconductor; and, $E_g^+ \approx E_g^- = E_g$. We note that the threshold process corresponds to all quasimomenta being parallel to one another, and in addition $|k_{10}| = |k_{20}| = |k_{h0}| = [m_cE_g/3\hbar^2]^{1/2}$.

For the processes which we are studying in semiconductor crystals the process of volume Auger recombination with participation of a heavy hole has the lowest kinematic threshold. For the case when the difference of the effective masses of the particles is large ($m_e \ll m_h$) in the threshold process all initial particles (electrons and heavy hole) have the same velocities. It is important to note that the final fast Auger electron has a significantly nonparabolic dispersion relation $\varepsilon(\mathbf{k}_f)$. On the basis of this fact, and expanding in the small parameter $\beta = m_e/m_h \ll 1$, we calculate by the previously indicated method the rate of Auger recombination with participation of a heavy hole:

$$g_h^{(O)} \approx 48\pi^8(2\pi\beta)^{1/2}\frac{E_B}{\hbar}\left(\frac{T}{E_g}\right)^{1/2}\left(\frac{\hbar^2}{m_eE_g}\right)^3\exp\left(-2\beta\frac{E_g}{T}\right)n_e^2n_h. \tag{24a}$$

To find the specific rate (24a) the overlap integrals $I_{cv}$ were calculated using the well-known rules.[5,8] It is interesting to note that the factor of 2 in the exponent of the threshold exponential is due to the "Dirac" nature of the energy spectrum of the electrons, i.e., owing to the nonparabolicity of the dispersion relation. Thus the nonparabolicity of the spectrum doubles the kinematic threshold of the Auger process with participation of a heavy hole.

## 4. LONG-RANGE COULOMB EFFECTS IN THE COMBINED AUGER PROCESS

A feature of the volume recombination processes examined above is that the momentum transferred to the Auger electron in an electron-electron collision must be large: $\Delta k \sim |\mathbf{k}_f - \mathbf{k}_1| \sim k_f$; this is possible only at small distances between the colliding electrons. This has the consequence that the probability of an Auger process in the volume has an additional power-law smallness. The volume process is suppressed primarily by the exponential threshold dependence. We shall calculate the recombination rate for the no-threshold process in the range of values of the parameters where this process predominates. For this we derive an analytic expression for the matrix element $M_1^{(K)}$ when the conditions (20) are satisfied. We write out the values of the quantities appearing in $M_1^{(K)}$ [see Eqs. (16)] at low temperatures (20), recognizing that because of the absence of a kinematic threshold (law of conservation of momentum) the small quasimomenta $\mathbf{k}_{1,2}$ and $\mathbf{k}_h$ of the particles make the main contribution to the recombination process (16):

$$\delta_i \approx \frac{\pi}{2} + \frac{k_{ix}}{\varkappa_i}, \quad k_{ix}/\varkappa_i \ll 1 \quad (i=1,2),$$
$$\cos\delta_i \approx -k_{ix}/\varkappa_i, \quad \cos\delta_h = \frac{\varkappa_h\cos\theta}{(\varkappa_h^2\cos^2\theta + k_h^2)^{1/2}}, \tag{25}$$

where $\theta$ is the angle between the quasimomentum of the hole and the normal to the boundary between the semiconductors. The transmission coefficient $d_f \approx 1$ (since $f \sim (\Delta_+ - \Delta_-)/(\Delta_+ + \Delta_-) \ll 1$). In the matrix element $M_1^{(K)}$ the terms that make the main contribution to the Auger process under the conditions being studied can be separated: These are Coulomb matrix elements, in which the momentum transferred in a Coulomb interaction is small, i.e., the first and second terms in Eq. (16) (the diagrams $K1$, $K4$ and $K2$, $K3$, respectively). They contain the long-range part of the Coulomb interaction, so that the squared matrix element

$$|M_1^{(K)}|^2 \propto [(\mathbf{q}_2 - \mathbf{q}_h)^2 + (k_{x2} - k_{xh})^2]^{-2};$$
$$[(\mathbf{q}_2 - \mathbf{q}_h)^2 + (k_{x2} + k_{xh})^2]^{-2}.$$

The last two terms in Eq. (16) correspond to processes in which the transferred momentum is large (the denominators in the expression for the Coulomb interaction contain the momentum $k_f \gg k_i$), so that their contribution is small in the parameter $T/E_g \ll 1$. Of the two main terms in the expression for $M_1^{(K)}$ the first term, containing the factor $4\pi/|\mathbf{k}_2 - \mathbf{k}_h|^2$ and diverging for small values of the relative quasimomentum of the electron $\mathbf{k}_2$ and hole $\mathbf{k}_h$, is the principal and truly long-range term. The denominator of the second term contains the difference of the quasimomenta of the incident and reflected waves and approaches zero at bounded points of the phase space, when $k_{xh}, k_{x2} \to 0$ simultaneously. On the basis of all these circumstances and substituting into Eq. (16) the corresponding values of the overlap integrals,

$$I_{cc}(\mathbf{k}_1^+, \mathbf{k}_f^+) = A_{\mathbf{k}_f^+}A_{\mathbf{k}_1^+}^*\left[1 + \frac{\gamma^2(\mathbf{q}_1\mathbf{q}_f - i\varkappa_1 k_{xf}^+)}{(E_1+\Delta_+)(E_f+\Delta_+)}\right] \approx 6^{-1/4},$$
$$I_{cv}(\mathbf{k}_2, \mathbf{k}_h) = A_{\mathbf{k}_h}A_{\mathbf{k}_2}^*\left[1 + \frac{\gamma^2\mathbf{k}_2\mathbf{k}_h}{(E_2+\Delta_-)(E_h+\Delta_-)}\right] \tag{26}$$
$$\approx \frac{\hbar k_h}{2(2m_eE_g)^{1/2}}\left(1 - \frac{\mathbf{k}_2\mathbf{k}_h}{k_h^2}\right),$$

we obtain

$$M_1^{(K)} = 2\pi^2\delta(\mathbf{q}_1 + \mathbf{q}_2 - \mathbf{q}_h - \mathbf{q}_f)\frac{4\pi e^2\exp[i(\delta_h - \delta_2 - \delta_1)]}{\varkappa_0|\mathbf{k}_h - \mathbf{k}_2|^2}$$
$$\times\frac{k_{x1}\hbar k_h}{\varkappa_1(3m_eE_g)^{1/2}}\left(1 - \frac{\mathbf{k}_2\mathbf{k}_h}{k_h^2}\right)\frac{\cos(\delta_h - \delta_2)}{(\varkappa_1 - ik_{xf})}. \tag{27}$$

In systems with sufficiently high electron density the long-range action in the Coulomb matrix element (27) is limited by screening. This can be taken into account by introducing the screening radius $r_0$ in the Coulomb matrix element: $|\mathbf{k}_h - \mathbf{k}_2|^{-2} \to \{|\mathbf{k}_h - \mathbf{k}_2|^2 + r_0^{-2}\}^{-1}$. However, when cal-

culating the recombination rate $G^{(K)}$ in an Auger process, the "divergence" typical for Coulomb interaction at small values of the relative quasimomentum $\mathbf{k}_{rel} = \mathbf{k}_2 - \mathbf{k}_h$ is eliminated because the corresponding overlap integral[2)] (26) approaches zero rapidly $|I_{cv}|^2 \propto |\mathbf{k}_h - \mathbf{k}_2|^2$, and thus the region of quasimomenta of the order of thermal momenta $k_{2,h} \sim (2m_c T)^{1/2}/\hbar$ plays a significant role. The rate of the combined Auger process is, according to Eqs. (25)–(27),

$$G^{(K)} = 8\pi^6 \frac{E_B}{\hbar}\left[\frac{\hbar}{(m_r E_g)^{1/2}}\right]^5 \frac{\hbar^2 T}{m_e (E_g^+ - E_g^-)^2} S n_e^2 n_h. \tag{28}$$

It is obvious that the rate of the combined process $G^{(K)}$ does not depend on the width $L$ of the narrow-gap sample, but contains rather only the area $S$ of the heterojunction. The absence of a threshold exponential temperature dependence in the rate of combined Auger processes (28), as we have already mentioned, is a consequence of the fact that the restrictions imposed by momentum conservation are removed, i.e., it is consequence of the existence of the heteroboundary as such. We recall that the formula for $G^{(K)}$ was derived in the case $T \ll E_g^+ - E_g^-$ and depends significantly on the height of the heterobarrier $E_g^+ - E_g^-$. It is obvious that in contrast to the volume process the combined process is very sensitive to the phases of the reflected waves. The values of the phases determine the probabilities of finding in the subbarrier region phases proportional to $|\cos\delta_1|^2$, and the value of the squared Coulomb matrix element (19) (for example, the factor $\cos^2(\delta_h - \delta_1) \approx \sin^2(\delta_h - k_{x2}/\varkappa_2)$ characterizes the phase shift of the electron and hole waves near the heteroboundary). The formula (28) can be rewritten in a form that is more convenient for comparing the relative intensity of volume and combined processes. Introducing the characteristic energy of size quantization in a narrow-gap semiconductor, $\varepsilon_L = \hbar^2/m_e L^2$, we obtain for $G^{(K)}$ (28)

$$G^{(K)} = 8\pi^6 \frac{E_B}{\hbar}\left[\frac{\hbar}{(m_e E_g)^{1/2}}\right]^6 \frac{T(\varepsilon_L E_g)^{1/2}}{(E_g^+ - E_g^-)^2} V n_e^2 n_h. \tag{29}$$

We shall make a qualitative and quantitative comparison of formulas (28) and (29) with the formulas for volume Auger recombination, for example Eq. (24), by investigating the ratio $G^{(O)}/G^{(K)}$ for light holes:

$$\eta = \frac{G^{(O)}}{G^{(K)}} = \frac{V g^{(O)}}{G^{(K)}} = \frac{9}{2}\left(\frac{\pi}{2}\right)^{1/2} \pi^2 \left(\frac{E_g}{T}\right)^{3/2} \left(\frac{E_g^+ - E_g^-}{E_g}\right)^2 \times \left(\frac{E_g}{\varepsilon_L}\right)^{1/2} \exp\left(-\frac{E_g}{2T}\right), \tag{30}$$

where $E_g \approx E_g^+$, $E_g^-$ is the average width of the band gap. The factor (30) has a clear physical meaning: The quantity $(E_g/\varepsilon_L)^{1/2} \sim k_f^+ L \gg 1$ characterizes the relative probabilities of finding the ejected Auger electron in the volume of the narrow-gap semiconductor and in the subbarrier region significant for the combined Auger process. The size of this region is of the order of $\lambda_f \sim 1/k_f^+ \sim \hbar/(m_c E_g)^{1/2}$ and is significantly smaller than the characteristic dimensions $\varkappa^{-1}$ of the wave function beneath the barrier, since $k_f^+ \gg \varkappa$ in the approximation (20) we are considering. Actually the relation (30) conveys the main characteristic features of the competition between the volume and combined processes, even in a wider range of values of the structural parameters than (20). Indeed,

$$\frac{G^{(O)}}{G^{(K)}} \sim \frac{\exp(-E_g/2T)}{(k_f^+ L)^{-1}} = \left(\frac{E_g}{\varepsilon_L}\right)^{1/2} \exp\left(-\frac{E_g}{2T}\right).$$

The realization of a combined process with a rate $G^{(K)}$ is always advantageous owing to the absence of a threshold exponential, but the required probability that one of the particles is located in the subbarrier region is small, in the ratio $\lambda_f/L$. Therefore, at sufficiently low temperatures, specimens with a narrow narrow-gap layer will be sources of fast Auger electrons, which are ejected into the wide-gap part. Electrons are ejected in a direction perpendicular to the heteroboundary in a narrow range of angles $\Delta\vartheta \sim (T/E_g)^{1/2} \ll 1$. The intensity of the Auger-electron flux arising as a result of the combined process depends quite weakly (linearly) on the temperature.

Completing this qualitative review of the results, we note that Auger recombination occurring exclusively in the subbarrier region ($M_1^{(P)}$ process) likewise does not have a threshold. But it is easy to show that this process has an additional, as compared with the combined process, smallness in the ratio $\lambda_f/L \approx (\varepsilon_L/E_g)^{1/2} \ll 1$. This smallness arises because the two electrons and the hole must be located in the subbarrier region. Thus there is no special reason for writing out the explicit form of $G^{(P)}$.

Using our expressions for the ratio $\eta$ of the rates of Auger recombination in the volume process $G^{(O)}$ and the combined process $G^{(K)}$, it is possible to determine the range of the parameters of heterostructures where one or the other mechanism predominates. Indeed, from the condition $\eta = 1$ we obtain a transcendental equation for the temperature $T^*(L)$ at which the rate of the volume process is equal to the rate of the combined process:

$$3\ln y - y = \ln\left(\frac{\varepsilon_L}{E_g}\right) + 4\ln\left(\frac{E_g}{E_g^+ - E_g^-}\right) - 8.04, \tag{31}$$

where $y = E_g/T$. The solution of Eq. (31) gives the characteristic curve $T^*(L)$, shown in Fig. 2 for three specimens with typical band parameters of the heterostructures. At temperatures below $T^*(L)$ the main Auger recombination mechanism is the combined process with the weakly temperature dependent rate $G^{(K)}$, while for $T > T^*(L)$ the threshold volume process predominates. For small specimens, when spatial quantization becomes important, nothreshold Auger processes become dominant at any reasonable temperature; this is observed for quantum microcrystals.[3]

For the record, we present the characteristic values of the rate $G^{(K)}$ of the combined processes at room temperature $T = 300$ K for carrier densities $n_e = n_h = 10^{18}$ cm$^{-3}$. For GaSb and GaAs heterostructures

$$G^{(K)}_{GaSb}(T = 300\,\text{K}) \approx 1{,}2 \cdot 10^{20}\ \text{s}^{-1},$$
$$G^{(K)}_{GaAs}(T = 300\,\text{K}) \approx 1{,}3 \cdot 10^{19}\ \text{s}^{-1}. \tag{32}$$

The numbers obtained are in good agreement with the experimentally measured values of the rate of nonradiative Auger recombination in GaSb heterostructures under the

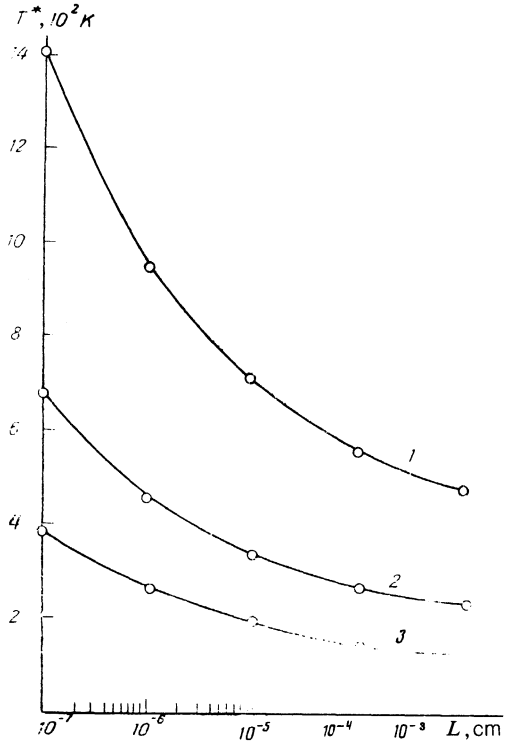



FIG. 2. The characteristic temperature $T^*$ as a function of the width $L$ of the narrow-gap semiconductor. The curves 1, 2, and 3 correspond to GaAs, GaSb, and InAs heterostructures, respectively. The region above the characteristic curves corresponds to volume Auger recombination and the region below them corresponds to the combined process.

same conditions: $G_{GaSb}^{exp} \approx 3 \cdot 10^{20}\ s^{-1}$ (Ref. 9). We note that the experimentally observed rate of nonradiative recombination is determined almost entirely by the combined Auger process investigated here. The combined Auger process significantly reduces the internal quantum yield in the system, since the rate of radiative recombination $R \approx 1 \cdot 10^{19}\ s^{-1}$ (Ref. 9) and $R \ll G^{(K)}$. The quantum yield, in this case, reaches several percent. According to the theoretical estimates (32) presented above, for GaAs structures the rate of nonradiative Auger transitions $G_{GaAs}^{(K)}$ is comparable, in order of magnitude, to the rate of radiative recombination $R_{GaAs} \approx 4.8 \cdot 10^{19}\ s^{-1}$ (Ref. 10). Thus the combined Auger process can significantly affect the quantum yield of the GaAs heterostructures employed.

In conclusion we call attention to the role of combined Auger decay processes in the formation of electrooptical properties of heterostructures such as superlattices, quantum wells, etc. The existence of an efficient no-threshold mechanism of Auger decay can significantly affect the transfer of nonequilibrium carriers from one quantum well into another. Charge is transferred by ejection of fast Auger electrons over quantum barriers even at very low temperatures by means of the combined process, especially since the predominant orientation of the momenta $\mathbf{k}_f$ of the ejected Auger particles is concentrated in a narrow solid angle along the normal to the heteroboundary. Combined Auger decay can have a pronounced effect on the quantum yield and threshold characteristics of semiconductor heterostructure lasers. In our opinion, effects analogous to those studied in this paper should also occur near a solid-vacuum boundary; this could be manifested in a dependence of the Auger decay rate on the size of the objects under study.

We thank R. A. Suris for a helpful discussion of the results.

[1)] The effect of size quantization on Auger recombination is discussed in Refs. 3 and 4.

[2)] Another factor, which compensates for the divergence of the squared matrix element in the calculation of $G^{(K)}$, is present in the phase volume of the relative momentum: $d^3k_{rel} = |\mathbf{k}_2 - \mathbf{k}_h|^2 d|\mathbf{k}_2 - \mathbf{k}_h|$.

[1] R. I. Taylor and R. A. Abram, Semicond. Sci. Technol. **3**, 859 (1988).
[2] P. Roussignol, Appl. Phys. Lett. **51**, 1882 (1987).
[3] D. I. Chepic, A. L. Efros, A. E. Ekimov *et al.*, J. Luminescence **47**, 113 (1990).
[4] A. P. Silin, Usp. Fiz. Nauk **147**, 485 (1985) [Sov. Phys. Usp. **28**(11), 972 (1985)].
[5] B. L. Gel'mont, Zh. Eksp. Teor. Fiz. **75**, 536 (1978) [Sov. Phys. JETP **48**(2), 268 (1978)].
[6] E. O. Kane, J. Phys. Chem. Solids **1**, 249 (1956).
[7] A. V. Sokol'skii and R. A. Suris, Fiz. Tekh. Poluprovodn. **21**, 866 (1987) [Sov. Phys. Semicond. **21**(5), 529 (1987)].
[8] B. L. Gel'mont, V. A. Kharchenko, and I. N. Yassievich, Fiz. Tekh. Poluprovodn. **29**, 2351 (1987) [Sov. Phys. Semicond. **29**(8), 1355 (1987)].
[9] A. A. Andaspaeva, A. N. Baranov, B. L. Gel'mont *et al.*, Fiz. Tekh. Poluprovodn. **25**, 394 (1991) [Sov. Phys. Semicond. **25**, 240 (1991)].
[10] H. Casey and M. Panish, *Heterostructure Lasers*, Academic, N. Y., 1978, Vol. 1.

Translated by M. E. Alferieff