TUNNEL-EFFECT RADIATIVE RECOMBINATION IN p-n JUNCTIONS

A. É. YUNOVICH and A. B. ORMONT

Moscow State University

Submitted to JETP editor Mary 27, 1966

J. Exptl. Theoret. Phys. (U.S.S.R.) 51, 1292–1305 (November, 1966)

The mechanism of interband radiative recombination in the strong electric field of a p-n junction is considered for the case in which electrons and holes can tunnel across the potential barrier. The radiation intensity I as a function of quantum energy $\hbar\omega$ and applied voltage U is calculated for a uniform field and simple parabolic bands. The occupation probabilities of states are taken into account for a semiconductor that is degenerate on both sides of the p-n junction. The emission band $I(\hbar\omega)$ associated with tunneling is investigated experimentally in GaAs, GaSb, and InP p-n junctions. It is shown that the position of the emission peak $\hbar\omega_m$ varies with voltage in accordance with the theory. In the GaSb and InP samples the emission intensity $I(U)$ for $\hbar\omega = \text{const}$ decreases as U increases, in agreement with the theory. The discrepancies between the experimental data and the calculations can be attributed to participation by the density-of-state tails and by local centers in tunneling recombination; this resembles the explanation of excess currents in tunnel diodes.

INTRODUCTION

AN investigation of the emission spectra from p-n junctions in GaAs has shown that the observed radiation bands can be divided into three groups: a) The main band, which is dominant for high currents, has energy close to the forbidden gap width, and is associated either with interband recombination or, in most cases, with the recombination of injected electrons via shallow acceptors; b) a radiation band whose maximum changes its position over a wide range as the bias of the sample is varied, is dominant for low currents in highly doped samples, and is associated with tunneling; c) the bands whose maximum energies are considerably smaller than the forbidden gap width and are associated with radiative recombination via impurity centers.[1-6] Similar classifications are applicable to p-n junctions in GaSb[7,8,31] and InP.[9,10]

In the present work we have investigated the radiation band that is associated with tunneling. This phenomenon was discussed in[2], where it was called "diagonal" tunneling or tunneling-assisted interband radiative recombination. It was shown in[2] that the probability of this effect is given by the formula of Keldysh for optical transitions in high electric fields.[11] Several investigators[2-6,12,13] have also discussed how the tails of state density in the forbidden gap of a highly doped semiconductor participate in the radiative recombination process. The results obtained in[5,7,10] have made it clear that a detailed analysis of the interband tunneling model is required for the purpose of discriminating this mechanism and to determine the extent to which participation by density-of-state tails in recombination affects the emission spectra.

The first part of this paper presents a theoretical analysis of the diagonal tunneling model under relatively simple assumptions; the second part gives experimental results obtained mainly by investigating highly doped GaSb and InP diodes; the third part compares the experimental findings with the theory and discusses the cases that indicate a complex pattern of optical transitions in a highly doped semiconductor.[14-16]

1. THEORY

1. Statement of the Problem

Let us consider the band diagram of a p-n junction that is degenerate on both sides, when a forward biasing voltage is applied (Fig. 1). In a sufficiently narrow junction, electrons and holes can tunnel across the potential barrier and emit light quanta $\hbar\omega$ as they recombine. We shall be interested in the emission intensity $I(\hbar\omega, U)$ as a function of the energy $\hbar\omega$ and the bias U on the p-n junction. This intensity will depend on the probability of an elementary recombination event,

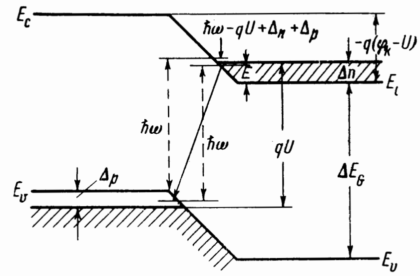

FIG. 1. Band diagram of a p-n junction that is degenerate on both sides, showing diagonal tunneling.

the numbers of states in the bands that can participate in a transition, and the probability of state occupation (by electrons in the conduction band and by holes in the valence band).

In [11,17] the probability of an interband optical transition was calculated under the assumptions of a) a uniform electric field and b) a determinate dependence of electron energy on quasimomentum. Keldysh restricted himself to the case of simple parabolic bands, while Franz also considered a state density that falls off exponentially in the gap. The fulfillment of both conditions does not occur, as a general rule, in the case that interests us here.

For the sake of simplicity we shall assume a uniform electric field within the p-n junction, depending on the applied voltage. We shall assume quadratic dependence of the energy on quasimomentum in both bands. The justification for these simplifications is found in the fact that the solution of the analogous simplified problem for ordinary nonradiative tunneling transitions accounts well for many qualitative properties of tunnel-diode current-voltage characteristics [18-31]. The foregoing assumptions may possibly be incorrect for transitions at the edge of a band, where the density-of-state tail can be important. The assumptions appear to be justified, however, in the interior of a band, near the quasi-Fermi levels. We shall assume that optical interband optical transitions are of the direct type, for which quasimomentum is conserved.

We shall make use of Keldysh's results under the given conditions. In accordance with (5)—(7) of [11] the probability of a transition accompanied by the emission (absorption) of one quantum in unit time is

$$W(\mathbf{p}, \hbar\omega, \mathscr{E}) = \left(\frac{q}{m}\right)^2 \frac{2\pi\hbar}{\omega} \frac{2}{\varepsilon} a \times \frac{|\mathbf{eM}(p_\perp)|^2 m_t^*}{[2m_t^*(\Delta E_G - \hbar\omega + p_\perp^2/2m_t^*)]^{1/2}} \times \exp\left[-\frac{4}{3}\frac{(2m_t^*)^{1/2}}{q\hbar\mathscr{E}}\left(\Delta E_G - \hbar\omega + \frac{p_\perp^2}{2m_t^*}\right)^{3/2}\right], \tag{1}$$

where $p_\perp$ is the component of the quasimomentum $\mathbf{p}$ that is perpendicular to the electric field $\mathscr{E}$; q and m are the electron charge and mass; $2/\epsilon$ is the factor that takes into account the dielectric constant ϵ; $|\mathbf{e} \cdot \mathbf{M}(p_\perp)|^2$ is the squared matrix element for an interband dipole transition and a quantum having the polarization vector $\mathbf{e}$; ΔE_G is the width of the energy gap in the semiconductor; a is the lattice constant of a cubic crystal; m_t^* is the reduced effective mass: $1/m_t^* = 1/m_c^* + 1/m_v^*$. We note that m_v^* should be averaged over the bands of heavy and light holes. However, such averaging requires the solution of a very complicated problem that is outside the scope of the present work.

The probability of spontaneous emission in a unit of volume is proportional to the number of quantum states in the interval $d(\hbar\omega)$:

$$\frac{4\pi}{(2\pi)^3}\frac{N^3\omega^2}{c^3\hbar} d(\hbar\omega), \tag{2}$$

where c is the velocity of light and N is the refractive index.

The probability that an electron with energy E will go from the conduction band to a state of energy $E - \hbar\omega$ in the valence band is proportional to the probability that these states will be occupied:

$$f_c(E)\,[1 - f_v(E - \hbar\omega)], \tag{3}$$

where f_c and f_v, the corresponding occupation probability functions for electrons, depend on the positions of the quasi-Fermi levels Δ_n and Δ_p relative to the boundaries of the respective bands. This resembles Esaki's treatment of ordinary tunneling transitions.[18]

The total number of transitions in the energy interval dE and quasimomentum interval $dp_\perp$ is also proportional to the number of levels in these intervals:

$$dE / q\mathscr{E}a, \tag{4}$$

$$2 \cdot 2\pi p_\perp dp_\perp / (2\pi\hbar)^2. \tag{5}$$

2. Spectral Intensity of the Radiation

The total number $I(\hbar\omega, U)$ of quanta in the interval $d(\hbar\omega)$ that are emitted in unit time per unit area of a p-n junction is derived by multiplying together (1)—(5) and integrating first over all possible states with the quasimomentum $p_\perp$ at a given energy E, and then over all possible energies. This is similar to the procedure for calculating the current-voltage characteristic of a tunnel diode.[18-21] We shall assume that the squared dipole matrix element is only slightly dependent on $p_\perp$ and E and we therefore remove its averaged value $\langle |\mathfrak{M}|^2 \rangle$ from the integrand. The upper limit of $p_\perp$ integration is derived from the condition that for a band-to-band transition $p_\perp$ cannot exceed the smaller of its two maximum values—in the conduction and valence bands—for a given value of E. When $p_\perp$ is small, i.e., when

$$p_\perp^2 / 2m_t^* (\Delta E_G - \hbar\omega) \ll 1, \qquad (6)$$

the terms in (1) that depend on $p_\perp$ can be expanded in a series. Restricting ourselves to the first term of this expansion, we can integrate explicitly over $p_\perp$ with the accuracy represented by (6). This condition signifies that we are limiting ourselves to transitions that are not too close to the gap width. When this condition is violated electron tunneling becomes less likely, on the whole, than carrier injection through the p-n junction.

When (3) is expressed explicitly as a function of E we obtain

$$I(\hbar\omega, U)\, d(\hbar\omega) = \frac{q^2 N m_t^* \langle |\mathfrak{M}|^2 \rangle}{m^2 \hbar^2 c^3 \cdot 2\pi^2} \frac{\hbar\omega\, d(\hbar\omega)}{\Delta E_G - \hbar\omega}$$
$$\times \exp\left[-\frac{4}{3}\frac{(2m_t^*)^{1/2}}{q\hbar\mathscr{E}(U)}(\Delta E_G - \hbar\omega)^{3/2}\right]$$
$$\times \int\limits_0^{\hbar\omega - qU + \Delta_n + \Delta_p} dE \left\{\left[1 - \exp\left(-\frac{2(2m_t^*)^{1/2}(\Delta E_G - \hbar\omega)^{1/2}}{q\hbar\mathscr{E}(U)} E^*\right)\right]\right.$$
$$\times \left[1 + \exp\left(\frac{E - \Delta_n}{kT}\right)\right]^{-1}$$
$$\left.\times \left[1 + \exp\left(-\frac{E - \hbar\omega + qU - \Delta_n}{kT}\right)\right]^{-1}\right\}. \qquad (7)$$

Here the upper limit of integration is defined as the maximum energy level of the conduction band from which an electron can drop to the top of the valence band, emitting a quantum $\hbar\omega$ (Fig. 1). E^* is a minimum:

$$E^* = \min\begin{cases} (m_c^*/2m_t^*)E \\ -(m_v^*/2m_t^*)(E - \hbar\omega + qU - \Delta_n) \end{cases}, \qquad (8)$$

and is derived from the selection rules for $p_\perp$. A similar procedure is discussed thoroughly in [21].

In Eq. (7) the electric field $\mathscr{E}$ is a function of the applied voltage U; in the case of a narrow p-n junction it has the explicit form

$$\mathscr{E} = \mathscr{E}_0[(\varphi_k - U)/\varphi_k]^{1/2}, \qquad (9)$$

where φ_k is the contact potential difference, which in our simple model (Fig. 1) is computed from

$$q\varphi_k = \Delta E_G + \Delta_n + \Delta_p; \qquad (10)$$

$\mathscr{E}_0$ is the field in the absence of an external bias. For a smooth linear distribution of impurities the exponent in (9) should be changed to $1/3$. The spectral intensity $I(\hbar\omega, U)$ then becomes a function of two variables, while depending on the parameters ΔE_G, Δ_n, Δ_p, $\mathscr{E}_0$, and φ_k.

The foregoing parameters are not independent, as a general rule, but are related through (9) and (10), and can also be determined from independent experiments. However, the value of ΔE_G determined from optical experiments with a relatively pure semiconductor may not be correct for heavily doped material. The values of Δ_n and Δ_p determined from the electron and hole concentrations on the basis of Fermi integrals in a simple parabolic model[22] can also differ from the true values for a heavily doped semiconductor. It is known from tunnel diode experiments that (10) is not fulfilled[23] in a p-n junction that is degenerate on both sides. It may therefore be difficult to compare theory and experiment quantitatively. In that case $\mathscr{E}_0$, Δ_n, and Δ_p may be taken as free parameters of the theory.

We shall now note the qualitative characteristics of the spectral intensity $I(\hbar\omega, U)$ that can be determined by analyzing (7). The energies $\hbar\omega$ and qU appear symmetrically in the upper limit of the integral in (7). To the extent that the behavior of the function is determined by this integral $I(\hbar\omega, qU)$ must therefore be a symmetric function of $\hbar\omega$ and qU (except for their signs).

If we reach fulfillment of the equality

$$\hbar\omega - qU + \Delta_n + \Delta_p = 0 \qquad (11)$$

as $\hbar\omega$ decreases or U increases, then the integral vanishes and the intensity $I(\hbar\omega, U)$ that is calculated from our model should drop to zero. This is understandable because we have assumed sharply bounded bands. Therefore, for a given value of U the lowest photon energy that can be emitted in interband tunneling transitions equals the difference between E_C in the n-region and E_V in the p-region. For a given value of $\hbar\omega$ the maximum bias permitting emission is also determined from (11). We note that under this condition the initial theoretical assumptions may be unfulfilled, because the band boundaries should be affected by the density-of-state tails in a semiconductor that is highly doped with impurities having low ionization energies.

At low temperatures the strongest energy dependence is exhibited by the distribution functions (3) that determine the denominator in the integrand. If the electric field $\mathscr{E}_0$ is sufficiently high the numerator of the integrand varies slowly and can be taken outside of the integral at the midpoint of the integration region, where the denominator reaches its minimum. The integral then assumes an explicit form. In the given approximation, using (9) and (10), we obtain

$$I(\hbar\omega, qU) = \text{const}\cdot\frac{\hbar\omega}{\Delta E_G - \hbar\omega}\exp\left[-\frac{1}{E_0}\frac{(\Delta E_G - \hbar\omega)^{3/2}}{[q(\varphi_k - U)]^{1/2}}\right]$$
$$\times\left\{1 - \exp\left[-\frac{1}{E_0}\frac{(\Delta E_G - \hbar\omega)^{1/2}(\hbar\omega - qU + \Delta_n + \Delta_p)}{{}^4/_3[q(\varphi_k - U)]^{1/2}}\right]\right\}$$
$$\times\left[\exp\left(\frac{\hbar\omega - qU}{kT}\right) - 1\right]^{-1}\ln\left[\text{ch}\frac{\hbar\omega - qU + \Delta_n}{2kT}\right.$$
$$\left.\times\text{ch}\frac{\hbar\omega - qU + \Delta_p}{2kT}\text{ch}^{-1}\frac{\Delta_n}{2kT}\text{ch}^{-1}\frac{\Delta_p}{2kT}\right], \qquad (12)^*$$

where the characteristic energy $\mathscr{E}_0$ is determined by the electric field $\mathscr{E}_0$ and is the most important parameter involved in the probability of a tunneling transition:

$$E_0 = \frac{3}{4}\frac{q\hbar\mathscr{E}_0}{(2m_t^*)^{1/2}(q\varphi_k)^{1/2}}. \qquad (13)$$

We have used the approximate formula (12) rather than the exact formula (7) in our numerical analysis of $I(\hbar\omega, U)$.

For increasing values of $\hbar\omega$ (with U given) or for decreasing values of U (with $\hbar\omega$ given) the radiation intensity should fall off rapidly, but not as rapidly as $\sim\exp[(\hbar\omega - qU)/kT]$ because of the term involving the tunneling transition probability and the determined value of E_0. In this case the transitions go from levels lying above the quasi-Fermi level in the conduction band to levels lying below the quasi-Fermi level in the valence band. For these transitions the potential barrier is lower and the tunneling probability is higher than for transitions between the quasi-Fermi levels. In this region the parabolic approximation should provide a good description of the energy bands.

Between regions of decrease the spectral intensity reaches its maximum at

$$\hbar\omega = qU + \Delta \qquad (14)$$

with Δ depending on the parameters $\mathscr{E}_0$, Δ_n, Δ_p, and kT, and can be calculated from (7) or approximately from (12). The half-maximum line width also depends on these parameters and can be calculated numerically. The dependence of the line width on $\hbar\omega$ is negligibly small when the denominator of the first term in (12) is not small.

3. Total Radiation Intensity

For the purpose of calculating the absolute total (integrated) radiation intensity we must first know the value of the matrix element $\langle|\mathfrak{M}|^2\rangle$. We use Kane's value[24], which was derived for an InSb semiconductor with a narrow forbidden gap:

$$\langle|\mathfrak{M}|^2\rangle = \frac{1}{\hbar^2}\frac{m^2\Delta E_G}{12m_c^*}\frac{\Delta E_G + \Delta_{so}}{\Delta E_G + \frac{2}{3}\Delta_{so}}; \qquad (15)$$

here Δ_{so} is the spin-orbit splitting constant. This formula was used in[25] to calculate the probability of radiative recombination in a relatively wide-gapped (GaAs) semiconductor, and yielded results that agreed with experiment. This provides a justification for the use of (15) in our case.

If it is assumed as an approximation that the spectral maximum is located at $\hbar\omega = qU$ and that $\Delta_n, \Delta_p \ll \Delta E_G$, i.e., $q\varphi_k \approx \Delta E_G$, then the value J_m of the integral in (7) at that maximum point does not depend on the applied voltage. Since the spectral line width is only slightly dependent on U, we can use (15) to evaluate the total intensity:

$$I(U) \approx \frac{q^2}{24\pi^2\hbar^4c^3}\frac{m_t^*}{m_c^*}\frac{\Delta E_G + \Delta_{so}}{\Delta E_G + {}^2/_3\Delta_{so}}\frac{qU\Delta E_G}{\Delta E_G - qU}$$
$$\times\Delta_{1/2}J_m\exp\left[-\frac{q(\varphi_k - U)}{E_0}\right] \qquad (16)$$

where $\Delta_{1/2}$ is the line width at half-maximum. The product $\Delta_{1/2}J_m$ must be computed before numerical estimates are possible. As a function of the parameters E_0, kT, Δ_n, and Δ_p this product has the following limits for $E_0 > kT$:

$$(kT)^2 < \Delta_{1/2}J_m < (\Delta_n + \Delta_p)^2. \qquad (17)$$

By multiplying I(U) of (16) and the charge q we obtain the current density j(U) that is associated with tunneling-assisted radiative recombination.

II. EXPERIMENT

1. Investigated Diodes

We investigated luminescent diodes prepared out of GaAs, InP, and GaSB semiconductor materials, where the valence band maximum and the conduction band minimum are located at the center of the Brillouin zone and quasimomentum is conserved in optical interband transitions. The band structures are described in[26].

The carrier concentrations in the original semiconductors were $(1-2)\times 10^{18}$ cm^{-3}. The p-n junctions were prepared by Zn diffusion in n-GaAs and N-InP, by Be diffusion in n-GaAs, or by pulling from the melt in the case of GaSb. The thicknesses 300–600 Å of the space-charge layer in the selected diodes were determined from the capacitances of the diodes at zero bias and room

*ch ≡ cosh.

temperature. These were conventional Fabry-Perot laser diodes with p-n junctions having areas of about 3×10^{-3} cm^2. The radiation was observed through a reflecting face in the plane of the p-n junction.

2. Procedure

The diode emission was investigated in the spectral region 0.6—1.5 eV using an IKS-12 spectrometer and a glass prism. The detector was a PbS photoconductive cell whose signals were amplified by a tuned amplifier with a synchronous detector having a ~ 15-sec time constant. For some spectral measurements a logarithmic amplifier with a 60-db (three orders of magnitude) range was inserted between the tuned amplifier and the synchronous detector.[27]

Most measurements were made at constant current J (10^{-2}—10^2 mA) flowing through the diode at 9, 77, and 293°K. The radiation was interrupted at a frequency of 800 Hz. In addition to the emission spectra, we investigated the voltage-current characteristics. The p-n junction voltage U was determined from the diode voltage V with allowance for the voltage drop (at large J) on the series resistance of the diode JR_s (R_s = 0.05—1 ohm). These corrections determined the accuracy of the measurement of U (in the best case, at J = 100 mA, the accuracy was ±0.01 V).

3. Experimental Results

A. Spectra of radiation $I(\hbar\omega)$ at U = const. Each diode emitted a characteristic band with a maximum $(\hbar\omega)_m$ whose position depended on the applied voltage within a broad interval. This band appeared at diode biasing corresponding to 0.13—0.5 eV less than the forbidden gap width, when other peaks and bands were still absent from the spectra.

Figure 2 shows the logarithmic emission spectra of an InP diode at 77°K for several applied voltages and the corresponding maxima. The intensity decreases exponentially as $\hbar\omega$ decreases: $I(\hbar\omega) \sim \exp(\hbar\omega/E_0)$. The characteristic values of E_0 for our samples were 45—50 mV for InP and ~ 20 meV for GaSb. The intensity falls off more steeply on the short-wave side.

At higher diode biases (e.g., curve 1 in Fig. 2 for an InP diode) the long-wave region exhibits a band that is associated with radiative recombination via deep traps. For GaSb diodes at 77°K the impurity emission peak was relatively high; its maximum near 0.73 eV was independent of the bias. With increasing voltage the "moving" peak passed across this fixed peak, and at high voltages the maximum of the moving peak was located above 0.73 eV.

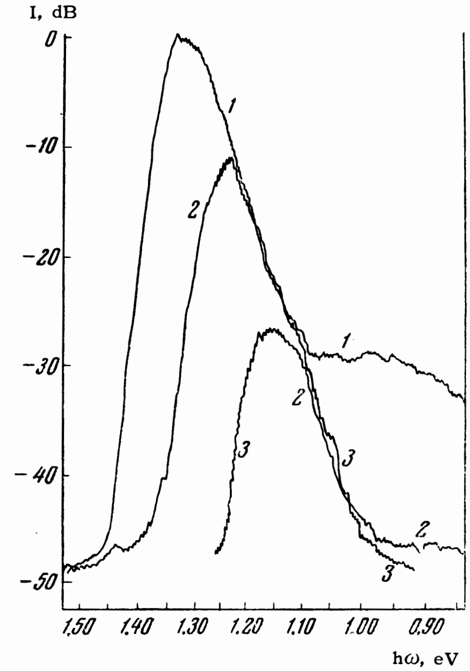

FIG. 2. Emission spectra of an InP diode recorded with a logarithmic amplifier for different p-n junction biases U: 1 – 1.413 V, 2 – 1.351 V, 3 – 1.235 V (T = 77°K).

Figure 3 shows how the maxima of different bands vary with the voltage for GaAs, InP, and GaSb diodes. The range of location of the moving peak maximum $\hbar\omega_m$ can be seen, as well as the fact that with increasing voltage this maximum approaches the main band that predominates for high currents.

The observed dependence of $\hbar\omega_m$ on the voltage is represented by

$$\hbar\omega_m = qU + \Delta, \tag{14a}$$

where Δ is negative and of the order 30 meV at 77° and 9°K. In the case of InP diodes, for which the moving peak was also observed at 293°K, Δ was positive at that temperature. The half-maximum width of the moving peak was considerably greater than kT, independently of the voltage.

B. I(U), spectral intensity versus voltage, for $\hbar\omega$ = const. This functional dependence was recorded under the conditions for which the moving peak was dominant, and with the minimum possible spectrometer gap width. It was the principal purpose of these experiments to determine whether the theoretically predicted decrease of radiation intensity with increasing voltage can be observed experimentally.

Figure 4a shows I(U) for GaAs at 77°K in the cases of two different photon energies. Decreasing

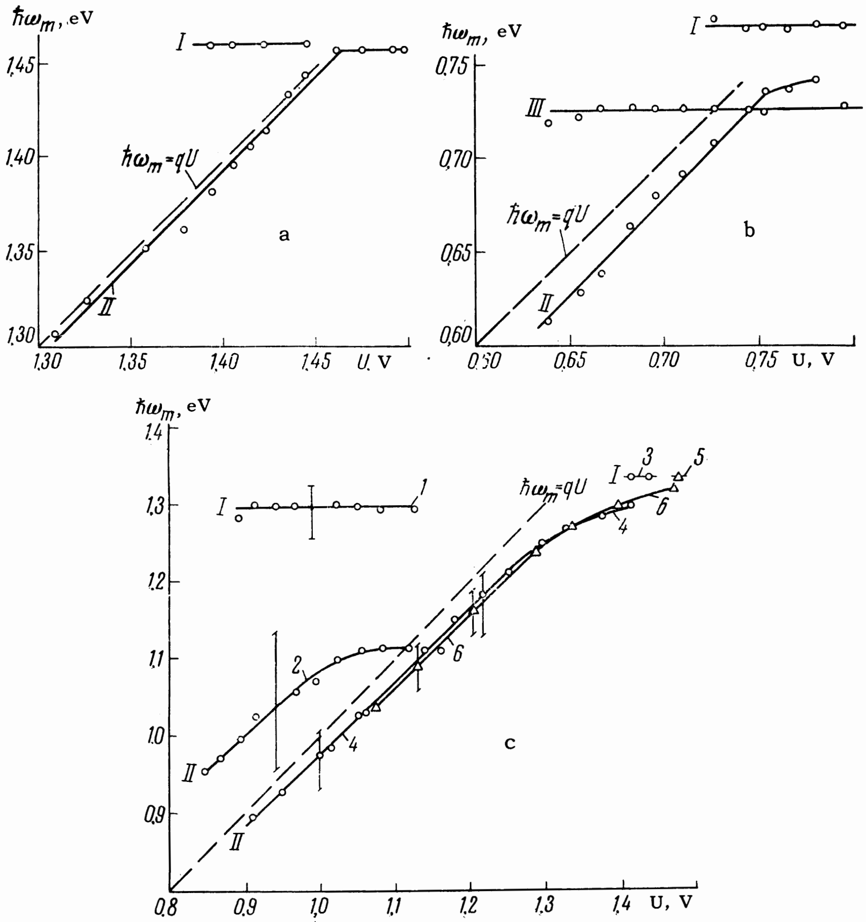

FIG. 3. Position of emission band maximum versus p-n junction bias for different semiconductor diodes. a – GaAs, T = 77°K; b – GaSb, T = 77°K; c – InP: curves 1 and 2 – T = 293°K; curves 3 and 4 – T = 77°K; curves 5 and 6 – T = 9°K. The vertical bars denote the half-maximum band widths. I – main band, II – moving band, III – impurity band.

intensity with growing current had been observed previously[3] in GaAs tunnel diodes with $\sim 2 \times 10^{19}/\text{cm}^3$ carrier concentration, because a moving peak was observed at 0.9–1.1 eV. It was reported in[4,5] that decreasing emission as U increases is not observed in GaAs for $\hbar\omega > 1.2$ eV. However I(U) approaches saturation for voltages $\gtrsim (\hbar\omega + \Delta)/q$. When $\hbar\omega$ is reduced the saturation region of emission intensity is broadened. When the voltage approaches the contact potential difference the intensity again rises sharply.

Figure 4b shows similar behavior for GaSb at 9° K. The emission intensity clearly passes through a maximum at a voltage somewhat higher than $\hbar\omega$; this maximum becomes more pronounced as $\hbar\omega$ decreases. A 30–40% decrease of intensity appears in this figure. With further increase of the voltage I(U) passed through a minimum and rose again. (The last-mentioned branch of the curve is not shown for curves pertaining to 9°K, because high currents were accompanied by considerable heating at this temperature.)

The described effects were recorded most clearly for InP samples, where the intensity was observed to decrease by a factor of three to five after the maximum had been passed (Fig. 4c). An

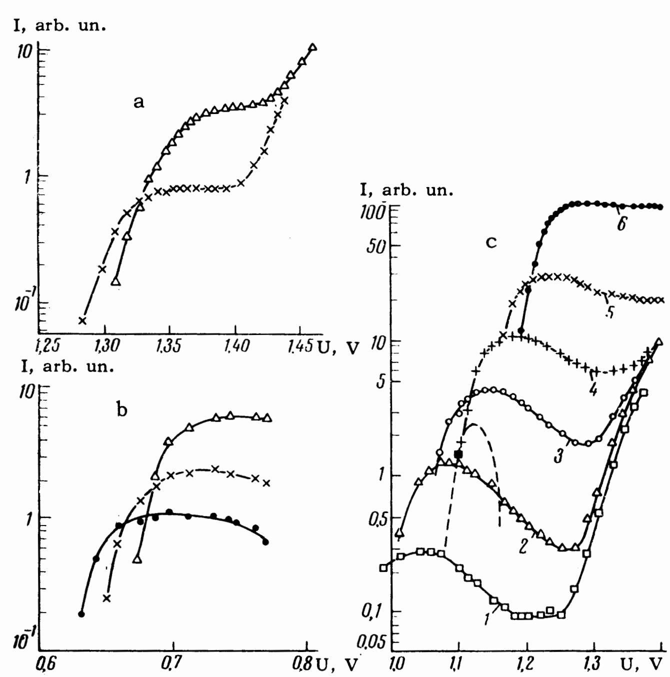

FIG. 4. Spectral intensity I(U) versus voltage applied to p-n junction, for $\hbar\omega$ = const. a) GaAs: curve 1 – $\hbar\omega$ = 1.300 eV, 2 – $\hbar\omega$ = 1.335 eV (T = 77°K); b) GaSb: curve 1 – $\hbar\omega$ = 0.62 eV, 2 – $\hbar\omega$ = 0.64 eV, 3 – $\hbar\omega$ = 0.66 eV (T = 9°K); c) InP: curve 1 – $\hbar\omega$ = 0.95 eV, 2 – $\hbar\omega$ = 1.00 eV, 3 – $\hbar\omega$ = 1.05 eV, 4 – $\hbar\omega$ = 1.10 eV, 5 – $\hbar\omega$ = 1.15 eV, 6 – $\hbar\omega$ = 1.20 eV (T = 77°K).

obvious resemblance appears in this case between the voltage dependence of light intensity and the current-voltage characteristics of tunnel diodes.

C. Current-voltage characteristics and total radiation intensity. In all the investigated samples the current-voltage characteristic for low forward currents was exponential,[5,7,10] the current being proportional to exp (qU, E_1), while the characteristic energy E_1 was independent of the temperature. The correlation between E_1 and the p-n junction width w ($E_1 \sim 1/w$) was verified for GaAs by Dumin and Pearson[28] and by the present authors.[5] We therefore affirm that the right-hand branch of the current-voltage characteristic was determined for our diodes by the tunneling component of the current.

The total emission intensity in a moving peak also varied exponentially with the voltage, being proportional to $\exp(qU/E_0)$. Here E_0 was approximately equal to the characteristic energy in the exponential spectral function $I(\hbar\omega) \sim \exp(\hbar\omega/E_0)$ for the long-wave wing. The value of E_0 was smaller than E_1 by a factor 1.2—2; thus the dependence of the total intensity on the current obeyed a power law with a power exponent exceeding unity.

The absolute emission intensity of the investigated diodes was estimated roughly by comparison with the emission intensity of a GaAs diode with a known external quantum yield in the case of high currents. These calculations indicated a small internal quantum yield in a moving peak, of the order 10^{-3}, i.e., the total current is controlled by some nonradiative mechanism that is dependent on tunneling.

III. DISCUSSION OF RESULTS

The experimental results leave no doubt about the existence of the same radiative recombination mechanism in the moving peaks for GaAs, GaSb, and InP p-n junctions. This is shown by the spectral dependence of the radiation and by the position of its maximum with respect to the width of the forbidden gap. Equation (14a) represents the characteristic linear voltage dependence exhibited by the position of the maximum in p-n junctions. This mechanism appears under the same conditions in each instance—a narrow p-n junction in a semiconductor that is degenerate on both sides.

We consequently apply the model of tunneling-assisted interband radiative recombination that was proposed in[2], and that we have analyzed in Sec. I, to the foregoing experimental results. The theory provides a good description of the qualitative spectral characteristics—the rapidly falling-off in the short-wave region and the slower decline in the long-wave region, the relation between the line width and the degree of semiconductor degeneracy, and the displacement of the maximum relative to the point $\hbar\omega = qU$.

It is especially important that the theory agrees qualitatively with the experimental finding that the low-energy spectral intensity diminishes with increasing voltage. We can thus consider that an analogy has been definitely established between the voltage dependence of spectral intensity for for tunneling-assisted radiative recombination and the voltage dependence of ordinary tunnel diode currents.

The experimental estimates of absolute radiation intensity do not conflict with calculations based on Eq. (16).

In view of our comments in Sec. I care must be exercised in making a quantitative comparison of theory and experiment. For the purpose of bringing the theory into agreement with experiment we selected the field $\mathscr{E}_0$ that would result in coincidence between the observed value of E_0 and its value derived from the voltage dependence of the long-wavelength band edge [$\sim \exp(\hbar\omega/E_0)$]. This value of $\mathscr{E}_0$ did not exceed twice its value as calculated from (13). The discrepancy is understandable because $\mathscr{E}_0$ was calculated from capacitance measurements using the formula for a narrow p-n junction, whereas the true field can locally exceed the calculated value. We determined Δ_n and Δ_p from the known concentrations on both sides of the junction, using the nomograms of[22].

The spectral intensity for one InP sample was calculated from (12) with the following parameters: $T = 77°K$, $E_0 = 4.5 \times 10^{-2}$ eV, $\Delta E_G = 1.41$ eV, $\Delta_n = 7.0 \times 10^{-2}$ eV, $\Delta_p = 3.0 \times 10^{-2}$ eV; the results are consistent with the experimental data represented in Fig. 5. The curves intersect at the point $\hbar\omega = qU$; the slopes agree on the short-wave side. The maximum is shifted from the point $\hbar\omega = qU$ in the longer wavelength direction; the observed shift Δ is larger, but of the same order of magnitude as the calculated value. The theoretical right-hand slope (the longer-wavelength side) is exponential, with the characteristic energy E_0, in agreement with experiment. However, the theory predicts zero intensity at $\hbar\omega = qU - (\Delta_n + \Delta_p)$, but this is not supported by experiment. The calculated half-maximum line width considerably exceeds kT, but is still below the experimental result. The calculated voltage dependence of the spectral shift describes satisfactorily the experimental observations.

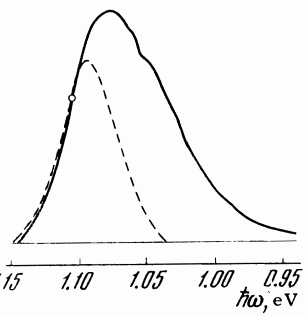

FIG. 5. InP diode emission consisting of a single moving peak. The theoretical peak is represented by the dashed curve. The linear ordinate scale represents relative intensity. T = 77°K, U = 1.10 V.

In Fig. 4c the theoretical spectral dependence on U for the same InP sample is represented by the dashed curve, which crosses the experimental curve at $\hbar\omega = qU$. At low voltages the theoretical curve is a good fit of the experimental points. With increasing voltage the experimental curve maximum moves away from the theoretical maximum, and the theoretically predicted zero intensity at $qU = \hbar\omega + (\Delta_n + \Delta_p)$ is not observed.

The foregoing quantitative comparison can be performed for many samples at different temperatures through a variation of the parameters; this requires a large amount of calculation. Therefore Eq. (7) was put into dimensionless form and used in computer calculations which, when compared with experiment, appear to indicate the limits within which quantitative agreement is observed.

The divergence of experiment and theory for small $\hbar\omega$ or large U values should not be surprising. In these cases the transitions go from the edge of the conduction band to the edge of the valence band in a highly doped semiconductor. It has been shown many times[3-6,12,13,29] that one must take account of the density-of-state tail in GaAs in order to account for the emission and absorption spectra and for the properties of semiconductor diode lasers.[30] At the band edges of a highly doped semiconductor both the parabolic approximation and the concept of a definite dispersion law E(p) may prove to be inapplicable.[14]

To describe the long-wave slope of a tunnel radiation peak we must therefore consider transitions involving impurity states close to band edges, - states that spread into an impurity band or overlap the main semiconductor band and form a tail.

The main emission band, which is dominant at high currents in GaAs and InP, is associated with the radiative transitions of electrons, injected into the p region, to a shallow acceptor level.[1-6] This mechansim has also been proposed for GaSb,[7-31] although some investigators have associated the main peak with interband transitions.[8] Since the tunneling band merges with the main peak at large currents we can assume that shallow acceptor levels participate in tunneling-assisted radiation.

The suggestion of a final relatively deep acceptor level for the tunneling transition that leads to the moving peak in GaSb[7] is apparently unsound. The experiments reported in[7] were performed with relatively low resolution; the tunneling-assisted radiation band merged with the impurity band and it was impossible to follow the motion of the former at energies above 0.73 eV (at 77°K).

However, impurity centers with deep-lying levels can also participate in tunneling-assisted radiative recombination. An analogy has been established between the voltage dependence of the spectral intensity at $\hbar\omega$ = const and the current-voltage characteristics of tunnel diodes. Therefore emission in the region $\hbar\omega < qU - (\Delta_n + \Delta_p)$ can be associated with all mechanisms that serve to account for excess currents in tunnel diodes while taking account of the quantum ($\hbar\omega$) emission in recombination. For example, this process can be electron tunneling from the conduction band to an impurity level followed by a radiative drop into the valence band, or virtual tunneling from the conduction band with a radiative transition to an acceptor level, or analogous tunneling transitions of holes from the valence band etc.

The particpation of a deep impurity level in radiative recombination in InP can be traced in Fig. 4c. The increase of intensity I(U) for $\hbar\omega$ = const after passing through the minimum is especially prominent at 1.02 eV. This energy corresponds to the maximum of the radiation band that is associated with unknown deep-lying levels occurring in relatively lightly doped crystals.[9,10] Investigation is still required to determine whether in this case a tunneling transition occurs or injection becomes important at high voltages.

The fact that the ratio between the maximum and the minimum of the $I|_{\hbar\omega = \text{const}}(U)$ curve is not identical for all semiconductors is associated not only with the sharpness of smearing of band edges, but also with the participation of different impurity levels in tunneling-assisted radiative recombination, and should depend strongly on the technology of diode fabrication. We know that the ratio between the maximum and minimum currents for the current-voltage characteristics of tunnel diodes depends strongly on technological factors. It can therefore be expected that in future investigations the decrease of emission intensity with increasing voltage, an effect that is associated with tunneling, will be manifested clearly in the p-n junctions of other semiconductors (in addition to GaAs, GaSb, and InP).

We must also refer to the possibility of refining the theory by taking into account the nonuniformity of the field, in analogy with [20]. A quantitative comparison of theory and experiment will be possible when one has overcome the methodological difficulties of exact capacitance measurements in the case of a relatively small differential resistance of diodes, and of the exact determination of impurity distributions near a p-n junction.

An analogy has now been established, both theoretically and experimentally, between optical effects in highly doped p-n junctions and the phenomena in tunnel diodes that determine the negative portion of the current-voltage characteristic and excess currents.

The authors are deeply indebted to V. L. Bonch-Bruevich, V. S. Vavilov, and L. V. Keldysh for discussions of the results and for valuable suggestions. They are also grateful to É. A. Poltoratskiĭ and V. M. Stuchebnikov for providing GaAs diodes, to A. R. Silin' for providing GaSb diodes, and to P. G. Eliseev and I. Ismailov for providing InP diodes.

[1] M. I. Nathan and G. Burns, Proc. IEEE **52**, 770 (1964); M. I. Nathan, Solid State Electronics **6**, 425 (1963).

[2] R. C. C. Leite, F. C. Sarace, D. H. Olson, B. G. Cohen, J. M. Whelan, and A. Yariv, Phys. Rev. **137**, A1583 (1965).

[3] A. A. Rogachev and S. M. Ryvkin, FTT **6**, 3188 (1964), Soviet Phys. Solid State **6**, 2548 (1965).

[4] A. N. Imenkov, M. N. Kozlov, S. S. Meskin, D. N. Nasledov, V. N. Ravich, and B. V. Tsarenkov, FTT **7**, 634 and 1480 (1965), Soviet Phys. Solid State **7**, 504 and 1190 (1965).

[5] A. B. Ormont, E. A. Poltoratskiĭ, and A. É. Yunovich, FTT **8**, 546 (1966), Soviet Phys. Solid State **8**, 431 (1966).

[6] A. É. Yunovich and P. G. Eliseev, Report at the Inter-University Conference on the Investigation of GaAs, Tomsk, 1965.

[7] V. S. Vavilov, I. A. Nakhodnova, A. R. Silin', and A. É. Yunovich, FTT **8**, 2330 (1966), Soviet Phys.-Solid State **8**, 1859 (1966).

[8] I. V. Kryukova, M. S. Mirgalovskaya, V. G. Karnaukhov, A. M. Baranova, and I. A. Strel'nikova, FTT **7**, 3421 (1965) and **8**, 1028 (1966), Soviet Phys. -Solid State **7**, 2757 (1966) and **8**, 822 (1966).

[9] K. Weiser, R. S. Levitt, M. I. Nathan, G. Burns, and J. Woodall, Trans. Met. Soc. AIME **230**, 271 (1964); K. Weiser and R. S. Levitt, Appl. Phys. Letters **2**, 178 (1963).

[10] P. G. Eliseev, I. I. Ismailov, A. B. Ormont, and A. É. Yunovich, FTT **8**, 3383 (1966), Soviet Phys. Solid State **8**, (1967), in press.

[11] L. V. Keldysh, JETP **34**, 1183 (1958), Soviet Phys. JETP **7**, 788 (1958).

[12] J. I. Pankove, Phys. Rev. Letters **9**, 283 (1962).

[13] D. F. Nelson, M. Gershenzon, A. Ashkin, L. A. D'Asaro, and J. C. Sarace, Appl. Phys. Letters **2**, 182 (1963).

[14] V. L. Bonch-Bruevich, FTT **4**, 2660 (1962) and **5**, 1852 (1963), Soviet Phys. Solid State **4**, 1953 (1963) and **5**, 1353 (1964); V. L. Bonch-Bruevich and R. Rozman, FTT **5**, 2890 (1963), Soviet Phys. Solid State **5**, 2117 (1964).

[15] V. S. Bagaev, Yu. N. Berozashvili, B. M. Vul, E. I. Zavaritskaya, L. V. Keldysh, and A. P. Shotov, FTT **6**, 1399 and 3188 (1964), Soviet Phys. Solid State **6**, 1093 and 2548 (1964).

[16] L. V. Keldysh, in Proceedings of the International Conference on the Physics of Semiconductors, Exeter, 1962 (The Institute of Physics and the Physical Society, London, 1963), p. 484.

[17] W. Franz, Z. Naturforsch. **13a**, 484 (1958).

[18] L. Esaki, Phys. Rev. **109**, 630 (1958) [Russ. transl. in Tunnel'nye diody (Tunnel Diodes), IIL, 1960].

[19] E. O. Kane, J. Appl. Phys. **32**, 83 (1961).

[20] V. L. Bonch-Bruevich and P. S. Serebrennikov, Radiotekh. i élektron, (Radio and Electronics) **6**, 2041 (1961).

[21] G. E. Pikus, Osnovy teorii poluprovodnikovykh priborov (Elements of the Theory of Semiconductor Devices), Nauka Press, 1965, p. 294.

[22] J. I. Pankove and E. K. Annavedder, J. Appl. Phys. **36**, 3948 (1965).

[23] A. N. Kovalev, Candidate dissertation, Inst. of Radio and Electronics, Acad. Sci. USSR, 1964.

[24] E. O. Kane, J. Phys. Chem. Solids **1**, 249 (1957).

[25] G. Lasher and F. Stern, Phys. Rev. **133**, A553 (1964).

[26] C. Hilsum, in Proc. Intern. Conf. Semicond., Paris, 1964, p. 1152.

[27] A. B. Ormont and E. K. Pervoĭ, PTÉ, 1967 (in press).

[28] D. J. Dumin and G. L. Pearson, J. Appl. Phys. **36**, 3418 (1965).

[29] J. I. Pankove, Phys. Rev. **140**, 2059 (1965).

[30] J. I, Pankove, in Proc. Intern. Conf. Semicond., Rad. Rec. Symp., Dunod, Paris, 1964, p. 201.

[31] T. Deutsch, R. C. Ellis, and D. M. Warschauer, Phys. Stat. Solidi **3**, 1001 (1963); A. R. Calawa, J. Appl. Phys. **34**, 1660 (1963).

Translated by I. Emin
153