Effect of Auger recombination on thermal processes in InGaAs and InAsSbP IR-emitting diodes


Institute of Semiconductor Physics, National Academy of Sciences of Ukraine, 45 Prospect Nauki, 03028 Kiev 28, Ukraine

Received 10 August 1999

Abstract

We investigated IR-emitting diodes (IREDs) based on the lattice-mismatched structures with stressed InGaAs/InAs layers and on near-matched InAsSbP/InAs structures. The emission wavelength range was from 2.5 to 5.0 µm. Both the processes of excess energy relaxation and mechanisms responsible for the active area overheating were studied. For In$_{1-x}$Ga$_x$As-based IREDs, it was shown that the active area overheat temperature, $\Delta T$, is related to the Auger recombination processes. When $x$ is increased from 0 to 0.09, the Auger recombination efficiency decreases, thus, favoring an abrupt drop of $\Delta T$. At $x = 0.09–0.22$, the efficiency of CHHS Auger processes exponentially decreases. However, owing to an increase in the dislocation density (due to a considerable, $\sim$6.9%, lattice mismatch parameter $\Delta a/a$), the $\Delta T$ value increases slowly and less markedly. At $x > 0.2$, the $\Delta T$ value becomes constant. For the InAsSbP-based IREDs, the active area overheat due to the current flow was less pronounced. © 2000 Elsevier Science B.V. All rights reserved.

PACS: 73.40.Kp; 73.50.Gr; 73.50.Yg
Keywords: IR-emitting diodes; InGaAs; InAsSbP; Heterostructures; Overheat temperature; Dislocations; Recombination

It is well known that the number of semiconductor materials emitting in the wavelength range of $\lambda = 2.5–5.0 \mu$m is limited. Investigations of such materials are very important because the absorption bands of the majority of industrial (harmful, toxic and explosive) gases belong to this region.

Among the promising materials that emit in this spectral region are InSb [1], quantum cascade systems [2], II-type interband cascade systems [3], single and double heterostructures based on multicomponent III–V heterostructures on InSb [4] and GaSb [5] substrates. It seems to us that the semiconductor solid solutions based on InGaAs/InAs and InAsSbP/InAs heterostructures are also promising. Such materials enable us to smoothly change the emission wavelength in this spectral region. Besides, the low-cost liquid-phase technology of obtaining epitaxial p–n junctions has been developed for these materials. Investigations of their structural, electrophysical, optical and recombination properties were carried out [6,7].

Of particular interest are the IR-emitting diodes (IRED) with weak temperature dependence of the emission wavelength. They seem quite promising for application in metrology, as reference and
standard emission sources, and elements of analog converters. The effect of ambient temperature on the luminescence of InAsSbP and InGaAs solid solutions was studied in Refs. [8,9]. In Refs. [5,10], the presence of a strong overheat due to the current flow was found for single and double heterostructures based on lattice-matched InGaAsSb/InSb compounds that emit in the $\lambda = 1.7$–2.4 $\mu$m range. Such an overheat leads to undesired changes in the emission, threshold and other characteristics of the devices. The interrelation between the thermal and recombination (including the Auger-type ones) processes, as well as the role of energy barriers in thermal processes, deserves to be analyzed.

Here we present some results of our investigations of the effect of current flow on the excess energy relaxation, as well as the role of the Auger-recombination processes in the active area overheat, for the IREDs based on InGaAs/InAs and InAsSbP/InAs structures. We studied heterostructures of two types, namely, (i) lattice-mismatched ones (with stressed layers) based on In$_{1-x}$Ga$_x$As (0.02 $< x < 0.25$) that emit in the $\lambda = 2.5$–3.6 $\mu$m range and (ii) lattice-matched heterostructures based on InAs$_{1-x-y}$Sb$_x$P$_y$ ($x = 0.07$–0.12, $y = 0$–0.01) that emit in the $\lambda = 3.8$–5.0 $\mu$m range. IREDs based on p-In$_{1-x}$Ga$_x$As/n-In$_{1-x}$Ga$_x$As/n-InAs and p-InAs$_{1-x-y}$Sb$_x$P$_y$/n-InAs$_{1-x-y}$Sb$_x$P$_y$ heterostructures were fabricated and studied.

The n-InGaAs and n-InAsSbP layers on the n-InAs wafers of (1 1 1) orientation were grown in a standard growth method. The n-InGaAs/Sb layers on the n-InAs/Sb wafers of (1 1 1) orientation were grown in a standard growth method. The n-InGaAs/Sb layers on the n-InAs/Sb wafers of (1 1 1) orientation were grown in a standard growth method. The n-InGaAs/Sb layers on the n-InAs/Sb wafers of (1 1 1) orientation were grown in a standard growth method.

The overheat temperature in the diode active region, $\Delta T = T_{p-n} - T_{env}$ (here $T_{p-n}$ and $T_{env}$ are the temperature of the p–n junction and the ambient temperature, respectively) was determined from the spectral peak shift with current for the near-edge emission band [11]. Firstly the gap $E_g$ dependence on temperature $T$ was calibrated. To do this, we have found experimentally how the photon energy at the emission peak ($h\nu = E_g$) for the IRED active area is related to the static external temperature $T$ when the latter was changed in the 300–400 K range. Then we have found the gap temperature factor $\alpha$ from the $E_g(T)$ dependence for different molar compositions of In$_{1-x}$Ga$_x$As and In$_{1-x-y}$Sb$_x$P$_y$ compounds. This factor was used when calculating $\Delta T$.

In another experimental method of $\Delta T$ determination both stationary and pulsed (at a relative pulse duration $Q > 100$) forward $I$–$V$ curves of p–n junctions were first measured at different external temperatures in the 300–400 K range. Then the temperature coefficient of voltage (TCV) was determined from them at various values of the stationary forward current flowing through the p–n junction. Having TCV value, one can calculate $\Delta T$ from an analysis of stationary and pulse $I$–$V$ curves [10].

Fig. 1 shows the $\Delta T$ vs sinusoidal current amplitude curves obtained from the two experiments. For In$_{1-x}$Ga$_x$As heterostructures (2–4) their feature is the presence of two portions with different slopes. It was found that the heterostructures with different $x$ have the same slope when $I < 50$ mA. But for $I > 50$ mA, the slopes of curves corresponding to heterostructures of different compositions were different. The heterostructures with bigger $x$ ($x = 0.026$) had the largest slope, whereas those with smaller $x$ ($x = 0.09$) had the smallest one. The slope for the InAsSbP heterostructures did not change. The error magnitude when determining $\Delta T$ did not exceed 0.2 K over the whole temperature range.
Fig. 2 presents experimental relationships between the temperature coefficient of the gap, \(a\), and the gap, \(E_g\), for InGaAs and InAsSbP of different compositions. One can see that for the quaternary compounds \(a \approx 0.75 - 0.95\) \(\times 10^{-4}\) eV K\(^{-1}\). It practically does not depend on the molar fractions of \(x\) and \(y\) components over the whole range of the compositions studied. For In\(_{1-x}\)Ga\(_x\)As compounds, the value of \(a\) decreases with \(x\) from 2.65 \(\times 10^{-4}\) eV K\(^{-1}\) (\(x = 0.026\)) to 1.65 \(\times 10^{-4}\) eV K\(^{-1}\) (\(x = 0.22\)). This behaviour of \(a(T)\) qualitatively coincides with the calculations of the Varshni function for InGaAs compounds of the same compositions [8]. It should be noted that the experimental \(a\) values obtained by us are 20%–30% lower than those calculated for the structures based on In\(_{1-x}\)Ga\(_x\)As (0 < \(x < 0.25\)) [8] (the discrepancy lies within the approximation error) and 25% lower than those estimated based on the results of [9] for In\(_{1-x}\)Ga\(_x\)As compositions with \(x = 0.045\) (\(E_g = 0.3763\) eV). The estimation of \(a\) for InAsSbP compounds [9] gave a value twice as large as in our experiments.

Concurrent analysis of the \(\Delta T(I)\) and emission power \(W\) dependencies on \(I\) and \(T\) made for InAsSbP compounds showed that in the vicinity of \(T = 300\) K, the emitters studied are highly stable to temperature fluctuations. Such IREDs have weak temperature dependence of both \(k\) and \(W\), i.e. they may be applied in reference and standard emission sources working at varying temperatures. It seems to us that the stability of InAsSbP IREDs may be due to the following reasons. First, a rather low \((p_0 < 3 \times 10^{16}\) cm\(^{-3}\)) concentration of the majority charge carriers does not favor the Auger processes. Second, the suppression of other alternative mechanisms of non-radiative recombination (related to defects, dislocations and dopants) is surely determined by the stress of lattice-matched layers.

To determine the inclined dislocation density in the p-InAsSbP samples, we have measured the etch pits in the (111)\(A\) and (111)\(B\) planes after selective chemical etching in HF + CrO\(_3\) + AgNO\(_3\) + H\(_2\)O = 2 ml: 1 g: 8 mg: 1 ml. The densities of etch pits in \(A\) and \(B\) planes correlated with the index
equal to 1, thus confirming the dislocation nature of the etch pits. The accuracy of the dislocation density determination was ±40%. The misfit $\Delta a$ between the lattice constants of epilayers and the wafer was measured using the diffraction of a wide-spread X-ray beam from the cleavage plane perpendicular to the (1 1 1) surface. The accuracy of $\Delta a$ determination was not worse than 0.001%.

Fig. 3 shows the dependence of the inclined dislocation density on the mismatch parameter $\Delta a/a_0$ for InAsSbP epitaxial layers and InAs substrate. It is evident that in the 0–0.15% range of $\Delta a/a_0$ (it is just the IREDs based on such optimized structures that we investigated) the dislocation density does not exceed $5 \times 10^3$ cm$^{-2}$. As a result, the emissive power does not depend on the dislocation density because an average spacing between dislocations exceeds the diffusion length of minority charge carriers in the active area of p-InAsSbP structure. At the same time at $\Delta a/a_0 = 0.5\%$, the dislocation density is $\sim 10^7$ cm$^{-2}$ and the emission power of such IR diodes is lower.

Let us analyse the difference between the $\Delta T = f(I)$ curves for the heterostructures based on p-In$_{1-x}$Ga$_x$As with different $x$ taking into account the Auger recombination processes. The latter limits the charge carrier lifetime $\tau$ in the narrow-band materials with high concentration of the majority charge carriers in the IRED active area.

The fact that for all the three structures based on In$_{1-x}$Ga$_x$As, the slope coefficients of the $\Delta T = f(I)$ curves at $I < 50$ mA are the same (taking into account thermal action of the current, the effective current value was $I_{eff} < 35$ mA) is an indication that the same excess energy relaxation mechanism is followed in all these cases. The heating mechanism is as follows [5,7]: at low and moderate currents, the electron gas temperature (that is stabilized by the electron–electron and electron–hole interactions) is practically the same as the lattice temperature $T_I = T_{p-n}$. In this case ($I < 50$ mA) the momentum of heated electrons is less than momenta of both equilibrium and non-equilibrium phonons. So the main ($\approx 15\%$) contribution to IRED active area warming up comes from the lattice component only. Since the In$_{1-x}$Ga$_x$As ($x \leq 0.22$) lattice type is defined by that of the prevailing component (InAs), the slopes $d(\Delta T)/dI$ coincide in this current region. At $I > 50$ mA, the steady-state temperature of the electron gas exceeds the lattice temperature (the time of temperature stabilization is $\sim 10^{-14}–10^{-13}$ s). The excess energy of electron gas is transferred to the lattice, firstly, through the interaction with the long-wave (and then with all) optical phonons and, secondly, through the interaction between the non-equilibrium and equilibrium phonons. In the second case (at $I > 50$ mA), the whole phonon system in the device active area is warming up, and different slopes of $\Delta T = f(I)$ curves were observed for different In$_{1-x}$Ga$_x$As compositions.

The dissipation of the excess electron gas energy occurs both immediately to the ambience and through the contact areas. In the latter case, the excess energy relaxation time is greater. An estimation of the characteristic time $\tau_\Sigma$ for the excess energy relaxation through contact areas gave $\tau_\Sigma \approx 3 \times 10^{-9}$ s. $\tau_\Sigma$ is inversely proportional to the product of sound speed on the probability of phonon traveling through the metal–semiconductor interface. This is less than the time of electron–phonon interaction ($\tau \approx 10^{-8}$ s) by a factor of 3–4. Thus, the main channels of energy dissipation are determined by the phonon–phonon and electron–phonon interactions.

The difference between the $\Delta T = f(I)$ slopes for In$_{1-x}$Ga$_x$As structures with different $x$ at
$I > 50$ mA indicates a change of both the energy dissipation mechanism (the interaction of charge carriers with optical and acoustic phonons, as well as the phonon-phonon interaction, that determine energy transfer from the electron system to the lattice) and the recombination mechanism. The latter is related to a rate increase for both radiative, as well as Auger, recombination (both rates are proportional to $\Delta n$), while the Shockley–Reed–Hall recombination rate remains the same.

It is well known that for In$_{1-x}$Ga$_x$As compounds with $x \lesssim 0.22$ the radiative and Auger processes compete in 200–400 K temperature range [12,13]. This competition determines the recombination process type [12,13]. At $T \sim 130–160$ K the contributions of both processes are comparable. And at $T = 300$ K for In$_{1-x}$Ga$_x$As structures with $x = 0–0.24$, the probability of Auger recombination exceeds that of radiative one by a factor of 2–3. At $x \approx 0.35$, the contributions from these processes become comparable. Thus, in In$_{1-x}$Ga$_x$As structures studied, the efficiency of non-radiative Auger recombination channel is somewhat over that of the radiative recombination channel.

Excess of the overheat temperature for the active areas of IREDs with different $x$ above that for the sample with optimal composition is shown in Fig. 4. One can conclude that the optimal composition is that of the IRED with minimum overheat temperature ($x = 0.09$). It is obvious that at maximal amplitude of the sinusoidal current ($I \approx 300$ mA), this difference $\delta T$ is not higher than 10 K. As was shown in Ref. [12], within the four-band Kane model, the traditional classical channel of Auger recombination in p-InAs and similar compounds of p-In$_{1-x}$Ga$_x$As-type, CHHV (i.e., an electron and two holes take part in recombination, with an energy transfer by one of them in the same band) is not essential. The main recombination mechanism for the non-equilibrium carriers at $T > 100$ K is the CHHS-Auger process involving an electron and two holes, with the ejection of one of them (the heavy hole whose effective mass is $m_h$) to the spin-split subband. The rate of such process is $\propto R_p \hat{A}_n^2$, and the time is $\tau_A = R_A^{-1} \hat{p}_0^{-2}$, where $R_A$ is the Auger recombination coefficient according to CHHS mechanism. For In$_{1-x}$Ga$_x$As material with $x = 0–0.2$ $R_A$ is $(2.0 \pm 2.2) \times 10^{-27}$ cm$^6$s$^{-1}$ [13]. The laws of energy and momentum conservation do not put restrictions on the energies of the recombining particles.

We found that the interband Auger recombination time was $\tau = 5–55$ ns. This value is in good agreement with both the experimental ones obtained by measurements of photoconductivity and photoelectromagnetic effect ($\tau = 5–10$ ns [14]) and that calculated theoretically [13] ($\tau = 11–45$ ns) for p-In$_{1-x}$Ga$_x$As layers ($x \lesssim 0.1$) with the charge carrier concentration $\sim (1–2) \times 10^{17}$ cm$^{-3}$. The Auger process efficiency is maximal and the internal quantum yield $\eta_i$ is minimal for p-InAs and p-In$_{1-x}$Ga$_x$As with $x \to 0$. This situation is due to the fact that the gaps $E_g$ in these materials are approximately equal to the spin–orbital splitting $\Delta$ of the valence band (“band resonance” effect). This fact favours luminescence quenching in the above materials and maximal overheating of the structures based on them [5]. According to [13], when the condition

$$\frac{E_g - \Delta}{T} = \frac{2 m_h}{m_{\text{lo}}}$$

holds at room temperature, the Auger coefficient $R_A$ is maximum and equals $2.2 \times 10^{-27}$ cm$^6$s$^{-1}$. 

Fig. 4. The excess, $\delta T$, of the active area overheat temperature over that for the In$_{1-x}$Ga$_x$As-based IREDs of optimum ($x = 0.09$) composition as a function of composition; ($T = 298$ K, $I = 300$ mA).
Moreover, it does not depend on the majority carrier concentration in a non-degenerated case. In expression (1), \( m_g = 0.41 - 0.42 \) \( m_0 \) and \( m_{s0} = 0.14 \) \( m_0 \) [12,13] are the effective masses of a heavy hole and a hole in a spin-split band, respectively; \( m_0 \) is the electron mass.

The 0–0.09 range of \( x \) is characterized by the efficiency decrease for the non-radiative channel of CHHS-type Auger recombination due to an increase in \( E_g - \Delta \). For example, at \( x = 0.1 \) \( E_g - \Delta \approx 40 \) meV. In this case, the gap \( E_g \) increases by 40 meV as compared to InAs, and \( \Delta \) increases only by 20 meV [13]. This leads to growth of recombination through the alternative radiative channel and decrease of \( \delta T \) (Fig. 4) and to a minimum \( \Delta T \) value at \( x \approx 0.09 \) (Fig. 1). As a result, an abrupt increase of the radiative power in the p-InGaAs active area at \( E - \Delta \) variation over the 30–45 meV range may be observed. As compared to p-InAs, the radiative power of In\(_{1-x}\)Ga\(_x\)As compound with \( x = 0.082 \) at \( T = 77 \) K is twice as much [12].

For further increases in \( x \) the probability of CHHS-Auger process goes down exponentially despite the fact that the density of the final states, \( \propto (m_{s0}T)^{3/2} \), grows. This is because only those charge carriers that belong to the “tail” of distribution function take part in recombination. It is determined by the increase of \( E_g \) and, particularly, the \( E_g - \Delta \) difference [9,13]. To illustrate, for \( x = 0.2 \), \( E_g \) grows by 60 meV and \( E_g - \Delta \) value increases by 80 meV as compared to the corresponding values for the In\(_{0.9}\)Ga\(_{0.1}\)As composition [13]. In this case, \( \delta T \) should go down even more. However, in this composition range, the alternative channels of non-radiative recombination come into action. (They are stimulated by an abrupt increase in the density of the inclined dislocations due to substantial (∼6.9%) misfit between the GaAs and InAs lattices.) These may be the Shockley–Reed–Hall processes through the deep levels in the heterostructure bulk and interface regions related to the presence of dislocations and their atmosphere [15]. Some other Auger recombination processes that have low probability under usual conditions may also occur. Among them are classical CHCC process (i.e., recombination involving a hole and two electrons, with energy transfer to one of them in the conduction band) and CHHV or CHHL processes (i.e., recombination involving an electron and two heavy holes, with transfer of a hole to the light hole band). Their probabilities increase, contrary to that of the CHHS processes, at the presence of a great number of photons and impurities. This removes the restrictions imposed on these processes by the momentum conservation law [16]. Such processes result in a decrease of both \( \tau \) and \( \eta \), and, consequently, a slow (with flattening out) \( \delta T \) growth.

It should be noted that in p-In\(_{1-x}\)Ga\(_x\)As the inclined dislocation density changed within the (1–5) × 10\(^6\) cm\(^{-2}\) range. The dislocations serve as the source for an efficient stable channel of non-radiative recombination. It ensures the \( \delta T(x) \) growth in the region where the probability of CHHS Auger processes drops sharply.

A contribution to the non-radiative recombination comes from dislocations and surrounding point defects (whose average concentration is ∼5 × 10\(^{17}\) cm\(^{-3}\)). Moreover, the misfit dislocations weaken the stress at the interface and form a great number (∼3 × 10\(^4\) cm\(^{-2}\)) of surface dangling bonds. They reveal themselves in generation–recombination and, consequently, result in thermal processes [5,15]. Fig. 5 shows various types of non-radiative recombination during thermal processes.

![Fig. 5](image-url)
occurring in IREDs based on InGaAs/InAs and InAsSbP/InAs heterostructures.

Let us now consider the problem of interrelation between the charge carrier effective lifetime, \( \tau \), and concentration of recombination centers. Shown in Fig. 5 are the \( \tau(p_0) \) curves obtained with the modified procedure discussed in [17]. Essentially, the method for \( \tau \) measurement is as follows:

1. at first, a direct current pulse (of amplitude \( I_f \), duration \( t_i \gg \tau \) and very high relative pulse duration \( Q \)) is allowed to pass through the p–n junction;
2. after this the sign of the voltage applied across the p–n junction is abruptly changed (i.e., the forward voltage gives way to reverse);
3. as a result, for some time \( T \) after switching the reverse current through the p–n junction retains its value \( I_f \) that is substantially higher than the saturation current \( I_s \) of the device. This is related to the fact that the recombination rate of excess minority charge carriers injected into the diode base is of a finite value;
4. the \( \tau \) value is then calculated from the duration \( T \) of the current saturation region. For the modified method [17] the absolute value of measurement error is \(<1.5 \text{ ns}\).

The \( \tau \) has been plotted as a function of majority charge carrier concentration \( p_0 \). It is obvious that at \( p_0 < 3 \times 10^{16} \text{ cm}^{-3} \) (InAsSbP heterostructures), the main contribution to \( \tau(p_0) \) dependence comes from the competition between the radiative recombination (for which \( \tau \propto 1/p_0 \)) and Shockley–Reed–Hall non-radiative recombination (\( \tau \) does not depend on \( p_0 \)). The slope value for the \( \tau(p_0^{-2}) \) curve (that may lie between 0 and 1) is 0.5 in our case (Fig. 5). It is evident that the charge carrier lifetime is determined by the competition between the two recombination channels, namely, that of radiative recombination (\( \tau \propto \Delta n \)) and Shockley–Reed–Hall (\( \tau \) does not depend on \( \Delta n \)).

The latter structures are not stressed. This seems to be the reason for the dislocation density minimization. For InAsSbP compounds, the concentrations of intrinsic charge carriers are higher than those for InGaAs, and \( p_0 \) values are low (\(<4 \times 10^{16} \text{ cm}^{-3}\)). This prevents observation of Auger recombination against a background of pronounced Shockley–Reed–Hall recombination.

It seems to us that Auger recombination would be observable in the InAsSbP structures with the majority charge carrier concentrations \( p_0 \), 1.5 to 2 orders higher. The technological development and experimental studies in this line could make the situation clearer.

At \( p_0 > 9 \times 10^{16} \text{ cm}^{-3} \) (InGaAs heterostructures), the Auger processes make the main contribution to \( \tau(p_0) \). This is confirmed by the similarity of theoretical and experimental curves for \( \tau \) that is quadratic in \( p_0 \). The experimental curve depression may be due to the presence of some other recombination mechanisms.

In this paper, we presented the results of our investigations dealing with the overheat temperature of the active areas in IR diodes based on In\(_{1-x}\)Ga\(_x\)As materials of various compositions (0.02 < \( x < 0.25 \)) (emitting in the \( \lambda = 2.5–3.6 \mu \text{m} \) range) and InAs\(_{1-y}\)Sb\(_{y}\)P\(_y\) materials (\( x = 0.07–0.12, \ y = 0–0.1 \)) (emitting in the \( \lambda = 3.8–5.0 \mu \text{m} \) range). It was found that in IREDs based on In\(_{1-x}\)Ga\(_x\)As the overheat of the structures with \( x = 0–0.09 \) is due to CHHS Auger-processes (at \( x \rightarrow 0 \) and \( I = 300 \text{ mA} \); \( \delta T \) value is \(~\sim 10 \text{ K} \)). For the structures with \( x > 0.09 \), the overheat is due to the Shockley–Reed–Hall recombination through the deep centers. In the IREDs based on the non-stressed InAs\(_{1-y}\)Sb\(_y\)P\(_y\) layers where the dislocation densities and probability of Auger processes are low the excess overheat of the active area is minimal.

High-stable IREDs based on lattice-matched non-stressed InAsSbP heterostructures have been developed and produced.

References