3

Kinetic description of luminescence processes

A multitude of characteristic features of luminescence can described and understood using simple kinetic equations, which, from the point of view of mathematics, represent first-order differential equations in time. In this chapter, we shall first introduce important parameters called the probability of radiative and non-radiative transition and the luminescence quantum yield. We shall discuss their relation to the experimentally established luminescence decay time. Next, we shall formulate kinetic equations for several simple luminescence phenomena. Solving the equations, we shall proceed firstly to explicate transient luminescence phenomena (i.e. the form of luminescence decay or possibly rise in time) and, secondly, we shall obtain the functional dependence of emission intensity on the optical excitation intensity in the steady state.

3.1 Radiative and non-radiative recombination. Luminescence quantum yield

An excited luminescence centre is in a non-equilibrium thermodynamic state and it loses its electron excitation energy after a certain relaxation time and goes into the ground state. The transition can occur in two ways. The excitation energy is either radiated away in the form of a luminescence photon after the mean lifetime τ_r or the excess energy is transmitted in the form of heat (vibrations) to the crystal lattice after time τ_{nr} or it can possibly cause photochemical changes in the crystal matrix or generate a lattice defect. The first kind of transition is called a *radiative transition*, the time τ_r is the radiative recombination time (radiative lifetime), transitions of the second kind are called *non-radiative* and τ_{nr} is the non-radiative recombination time (non-radiative lifetime). The inverse values $(\tau_r)^{-1}$ and $(\tau_{nr})^{-1}$ then represent the probability of the corresponding transition per unit time (the recombination rate). The total probability of transition to the ground state is given by

$$\frac{1}{\tau} = \frac{1}{\tau_r} + \frac{1}{\tau_{nr}}; \tag{3.1}$$

we can imagine this in analogy to electromagnetism as two parallel conductors, the resulting conductivity of which is given by the sum of the inverse values of their individual resistances (Fig. 3.1).

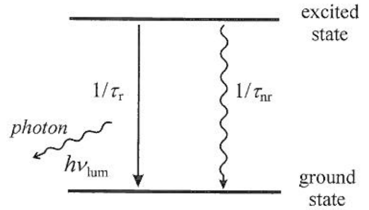

Fig. 3.1
Illustration of the definition of luminescence quantum efficiency.

We define the luminescence *quantum yield* or *quantum efficiency*, already mentioned in Section 2.1, as

$$\eta = \frac{1/\tau_r}{1/\tau_r + 1/\tau_{nr}} \leq 1 \tag{3.2}$$

i.e. as the ratio of the radiative recombination rate to the total recombination rate. For brightly luminescent materials $\tau_r \ll \tau_{nr}$ holds true, which means the given excited centre is much more likely to go into the ground state by radiating a photon $h\nu_{\text{lum}}$ than by a non-radiative transition. Neglecting $1/\tau_{nr}$ in (3.2), we get $\eta \to 1$ (we also say the quantum yield is 100%). However, such materials almost never occur in nature. In fact, we speak of strongly luminescent materials already if η is of the order of 0.1.

The word 'quantum' in the definitions given here implicitly points to the fact that both luminescence excitation and light emission take place through one-photon elementary steps; e.g. $\eta = 0.5$ means there is one emitted photon to every two absorbed photons. Thus, multiple-photon phenomena, although feasible in general, are not considered here. Likewise, definition (3.2) does not allow for the energy difference between the exciting $h\nu_{\text{ex}}$ and emitted $h\nu_{\text{lum}}$ photons—in this sense, we sometimes introduce the *power efficiency* $\eta_P = (h\nu_{\text{lum}}/h\nu_{\text{ex}})\eta$. Due to Stokes' law, we have $\eta_P \leq \eta$. For injection electroluminescence, η relates not to the number of photons absorbed but to the number of electron–hole pairs injected.

What is the physical meaning of the time τ given by expression (3.1)? It characterizes the emptying of the excited-state level in a large ensemble of luminescence centres by radiative and non-radiative transitions simultaneously. Yet, the population decay of an excited level is reflected just in a decrease of the luminescence intensity as it decays after the excitation is turned off. Therefore, the time τ is (in accordance with the notation introduced in Section 2.9) an experimentally accessible quantity—the luminescence decay time. By means of this decay time, we can rewrite the expression for quantum efficiency (3.2) in the form

$$\eta = \frac{\tau}{\tau_r} (\leq 1). \tag{3.3}$$

Let us point out straightaway the occasional erroneous use of the terms we have just introduced. The radiative lifetime (τ_r) is not the same as the luminescence decay time (τ) although the admitted similarity of the terms may tempt us to confuse them! (On the contrary, we shall show shortly that, in a special case, the luminescence decay time may be identical to the non-radiative recombination time!)

From the point of view of mathematics, relations (3.1), (3.2), and (3.3) are just trivial fractions but, surprisingly enough, they involve a wide spectrum of physical information. So firstly, it is interesting that expression (3.3) relates two measurable quantities (η, τ) to an important quantum-mechanical parameter characterizing optical transitions from a theoretical perspective (the reciprocal value of the radiative lifetime τ_r^{-1} is proportional to the square of the modulus of the corresponding transition matrix element [1–3]). The magnitude of the matrix element basically cannot be obtained by direct measurements and it is neither simple nor reliable to calculate it. Therefore, the simple formula (3.3) appears to provides a welcome opportunity to obtain τ_r^{-1} from two experimentally available quantities and to compare the result to the theory (in a semiconductor with a direct bandgap, τ_r is of the order of 10^{-9} s, and with an indirect bandgap, $\tau_r \approx 10^{-4}$ to 10^{-3} s). However, while it is very easy to measure τ—and we have discussed several of the most common experimental methods in Section 2.9—it is, in contrast, extremely difficult to measure the quantum efficiency η with sufficient accuracy, as we have, for that matter, also mentioned before. Hence, we should not overestimate the significance of expression (3.3) in this sense; nonetheless, it is very important that by measuring the temperature dependence of the luminescence intensity (which we can consider, to a good approximation, proportional to the quantum yield η) and of the luminescence decay time τ we can obtain from (3.3) information on the temperature dependence of the radiative recombination time τ_r, albeit in relative units.

In addition, relations (3.1)–(3.3) enable us to asses quickly a number of situations we may encounter in practice. Let us first consider the case $\tau_{nr} \ll$ $\tau_r < \infty$. Then the non-radiative recombination rate, τ_{nr}^{-1}, is much higher than the radiative transition rate; nonetheless, some weak luminescence radiation may still be present. Naturally, the quantum yield η will be very low ($\eta \approx$ $\tau_{nr}/\tau_r \ll 1$ as follows directly from (3.2)) and what is particularly interesting is that—as an immediate consequence of (3.1)—we get $\tau \cong \tau_{nr}$. Thus, measuring the luminescence decay time, we are now actually measuring the *non-radiative lifetime*. Or, put in still another way, if we reveal in an experiment a weak luminescence decaying very quickly then there will very likely be a strong influence of non-radiative transitions.

The opposite case reads $\tau_r \ll \tau_{nr}$. The quantum yield η is close to unity according to (3.2) and $\tau \cong \tau_r$ holds. Although we encounter this situation much less often than the previous case we must still mention the following possibility. Let us consider a material with a relatively large τ_r, which, in itself, is thus not a very efficient phosphor. Yet, if we are able to 'forbid' somehow all non-radiative transitions then $\tau_{nr} \to \infty$, $\tau_r \ll \tau_{nr}$ and the electronically excited system has no other choice but to return—albeit slowly—to the ground state exclusively via radiative transitions. Hence, in this way, we can obtain a material featuring efficient luminescence ($\eta \cong 1$). The total number of photons emitted per unit of time will be relatively low; nevertheless, this mechanism is thought to be one of the factors contributing to the intense luminescence of so-called porous silicon and of silicon nanocrystals. Silicon is a poor phosphor as a result of its indirect bandgap. The radiative lifetime is long ($\tau_r \approx 10^{-3}$ s) and almost all transitions to the ground state occur via much more efficient non-radiative processes. In nanocrystals, however, the non-radiative transitions take place mostly on the surface of the nanocrystal and they can be prevented through efficient passivation of unoccupied surface bonds by attaching, e.g., hydrogen or oxygen atoms.

3.2 Monomolecular process

Let us consider a *localized luminescence centre*[1] in a crystal or possibly in an amorphous matrix, typically represented by an impurity atom. Let us assume the centre can be brought into an electronic excited state by either absorbing light directly or via energy transfer from the optically excited surrounding matrix. Let both processes be described by a quantity G, which is the number of photons absorbed in unit volume per unit time. Subsequently, centres in an excited state return to the ground state, partly radiatively and partly non-radiatively. If the concentration of the excited centres is $n(t)$, the processes under consideration are described by the equation

$$\frac{\mathrm{d}n}{\mathrm{d}t} = G - \frac{n}{\tau_r} - \frac{n}{\tau_{nr}} = G - \frac{n}{\tau}, \tag{3.4}$$

where $\mathrm{d}/\mathrm{d}t$ denotes the derivative with respect to time and G is called the generation term. A schematic is shown in Fig. 3.2(a). We shall now analyse eqn (3.4) in two cases: (a) luminescence decay and (b) steady state.

(a) Luminescence decay is characterized by the excitation being switched off ($G = 0$) at time $t = 0$. We are interested in the fall-off of the luminescence intensity $i(t)$ at times $t \geq 0$. Equation (3.4) has the form

$$\frac{\mathrm{d}n}{\mathrm{d}t} = -\frac{n}{\tau}$$

and its solution thus reads $n(t) = n(0)\exp(-t/\tau)$. The luminescence intensity $i(t)$ (here, the number of photons emitted from unit volume per unit time) is given by the ratio $n(t)/\tau_r$, and hence the desired intensity fall-off is

$$i(t) = \frac{n(t)}{\tau_r} = \frac{n(0)}{\tau_r}\mathrm{e}^{-t/\tau} = i(0)\mathrm{e}^{-t/\tau}, \quad t \geq 0. \tag{3.5}$$

Here, $i(0)$ denotes the intensity at time $t = 0$. We can see that in this case (the so-called *monomolecular process* or a localized centre) the luminescence exhibits strictly single-exponential decay governed by the total lifetime τ, as we have already described in Section 3.1 and as plotted in Fig. 3.2(b).

(b) The steady state—when the excitation G is turned on and when, after the initial increase in $i(t)$, the intensity of luminescence settles at a constant value—is determined by the condition $\mathrm{d}n/\mathrm{d}t = 0$. From eqn (3.4) and with regard to Fig. 3.2(b), we have

$$G = \frac{n(0)}{\tau_r} + \frac{n(0)}{\tau_{nr}} = i(0) + \frac{n(0)}{\tau_{nr}}. \tag{3.6}$$

The generation term G is closely related to the excitation radiation intensity I_{ex} (here, the number of photons incident upon unit surface area per unit time). Let us consider the common case when all excitation radiation

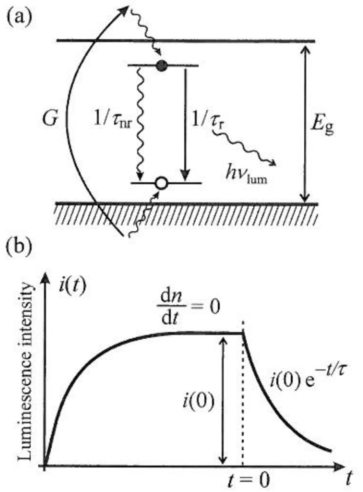

Fig. 3.2
(a) The monomolecular luminescence process. Wavy lines represent non-radiative thermalization of the carriers, their capture, and possibly non-radiative recombination. (b) The kinetics of monomolecular luminescence upon turning the excitation source on and off.

---

[1] By this we shall mean an electron excitation localized at a certain impurity or defect. Free quasi-particles do not participate in recombination phenomena.

is absorbed in the sample (we neglect reflection for the sake of simplicity). Absorption occurs in a thin subsurface layer, the depth of which is roughly equal to the reciprocal value of the absorption coefficient of the excitation light α^{-1}. If the area of the excited luminescence surface is ΔS then there are $I_{\mathrm{ex}}\Delta S$ photons incident per unit time and the photons are thus absorbed in a volume $\sim \Delta S\alpha^{-1}$. The density of the absorbed photons, or the number of centres excited per unit volume per unit time, is thus $G \approx (I_{\mathrm{ex}}\Delta S)/\Delta S\alpha^{-1} = \alpha I_{\mathrm{ex}}$.

From eqn (3.6) we then get

$$i(0) = G - \frac{n(0)}{\tau_{\mathrm{nr}}} = \alpha I_{\mathrm{ex}} - \frac{\tau_{\mathrm{r}}}{\tau_{\mathrm{nr}}} i(0)$$

and after some simple algebra and considering definition (3.2), we immediately have

$$i(0) = \eta\alpha I_{\mathrm{ex}}. \tag{3.7}$$

In the case of monomolecular recombination, the stationary value of the luminescence intensity, $i(0)$, is thus directly proportional to the excitation intensity I_{ex}.

Hence, we have obtained two simple and important guidelines facilitating the interpretation of experimental data. Observing a single-exponential decay together with a linear increase in luminescence intensity are rather strong indications that the microscopic origin of the examined luminescence might lie within a localized centre. This is not unambiguous evidence—similar features can be seen in, for example, the luminescence of a kind of free quasi-particle, the free exciton—and yet, such an observation helps us to exclude a number of other types of recombination processes.

Let us note further that the presence of two kinds of localized luminescence centres will show itself via a double-exponential decay

$$i(t) = a_1 e^{-t/\tau_1} + a_2 e^{-t/\tau_2}, \tag{3.8}$$

where $a_i (i = 1, 2)$ denotes the amplitudes of the individual components. Naturally, the prerequisite of observing a dependence of the form (3.8) is that we detect spectrally unresolved luminescence of both centres, e.g. when the emission wavelengths of both centres are close to each other and their spectra overlap. Today, computer-aided fitting of the experimental decay curve enables us to determine the parameters a_1, a_2, and τ_1, τ_2 quickly and reliably even in case we need to use deconvolution methods (see (2-64) and Fig. 2.50).

One cannot help thinking of generalizing relation (3.8) to a sum of multiple exponential functions. However, when attempting to fit the luminescence decay curve using the expression $i(t) = \sum_{i=1}^{j} a_i e^{-t/\tau_i}$ $(j \geq 3)$, we must be very careful. Of course, from the point of view of mathematics, such a fit is feasible and we usually achieve very good agreement between our calculation and experiment, yet, from the point of view of physics, such a result in itself can hardly be considered as evidence for the presence of three or more luminescence centres. This is because, in principle, any of the characteristic forms of decay curve, which we shall discuss below, can be modelled very nicely as a sum of three exponentials with six fitting parameters $(a_i, \tau_i, i = 1, 2, 3)$; the informative value of such a fit, however, is highly questionable.

3.3 Bimolecular process

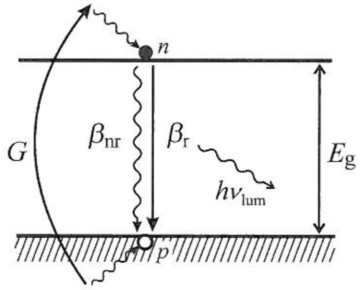

Fig. 3.3
An example of a bimolecular luminescence process.

Let us consider the process shown in Fig. 3.3. If a semiconductor absorbs a photon of energy above the bandgap, E_{g}, then free electron–hole pairs are produced. Let us denote the concentration of electrons by n and the concentration of holes by p. If there is no impurity atom or possibly a lattice defect nearby where localization of the created photocarriers could occur then electrons and holes will diffuse along the edges of the respective bands and, eventually, electrons will recombine with the nearest holes (as their wavefunctions overlap most), whether radiatively or non-radiatively. The above processes are described by the equation

$$\frac{\mathrm{d}n}{\mathrm{d}t} = G - \beta\, n\, p, \tag{3.9}$$

where G again denotes the generation term. The recombination term proportional to the product *np* represents the fact that recombination requires a free electron to encounter a free hole. The coefficient β is called the *bimolecular recombination coefficient* and can be expressed as the sum of bimolecular radiative (β_{r}) and bimolecular non-radiative (β_{nr}) recombination coefficients: $\beta = \beta_{\mathrm{r}} + \beta_{\mathrm{nr}}$.[2] If we consider, for the sake of simplicity, an intrinsic semiconductor at low temperatures, then $n = p$ and eqn (3.9) can be rewritten as

$$\frac{\mathrm{d}n(t)}{\mathrm{d}t} = G - \beta_{\mathrm{r}} n^2 - \beta_{\mathrm{nr}} n^2 = G - \beta n^2. \tag{3.10}$$

Here, the term $\beta_{\mathrm{r}} n^2$ clearly represents the intensity of luminescence generated by bimolecular recombination (order-of-magnitude values of the coefficients β_{r} are $\beta_{\mathrm{r}} \sim 10^{-10}\,\mathrm{cm^3\,s^{-1}}$ for direct bandgap semiconductors and $\beta_{\mathrm{r}} \sim 10^{-15}$–$10^{-14}\,\mathrm{cm^3\,s^{-1}}$ in indirect bandgap semiconductors). We shall again analyse both the luminescence decay and steady state.

(a) Similarly to the monomolecular process, let us turn off the excitation ($G = 0$) at time $t = 0$ and follow the decay. Then, relation (3.10) reduces to

$$\frac{\mathrm{d}n(t)}{\mathrm{d}t} = -\beta n^2,$$

---

[2] It should be clear by now that the terms 'monomolecular' and 'bimolecular' historically originated from the analogous chemical reactions.

which is an equation we can solve easily if, for example, we separate the variables, use the substitution $1/n = x$, and apply the boundary condition $n(t = 0) = n(0)$. We get

$$n(t) = \frac{n(0)}{(\beta n(0)t + 1)}$$

and the luminescence intensity $i(t) = \beta_r n^2(t)$ comes out immediately as

$$i(t) = \frac{\beta_r n^2(0)}{(\beta n(0)t + 1)^2} = \frac{i(0)}{(\beta n(0)t + 1)^2} = \frac{i(0)}{(\gamma \sqrt{i(0)}t + 1)^2}, \tag{3.11}$$

where $i(0)$ is the intensity at time $t = 0$ and $\gamma = \beta/\sqrt{\beta_r}$. We can see that the form of the bimolecular luminescence decay differs substantially from the exponential function. It is a power-law function of time and, for late times, we can approximate $i(t) \approx t^{-2}$.

(b) The steady state is characterized by the condition $\mathrm{d}n/\mathrm{d}t = 0$. From (3.10) we get

$$i(0) = \beta_r n^2(0) = G - \beta_{nr} n^2(0) = G - \beta_{nr} \frac{i(0)}{\beta_r},$$

which can be transformed into the form

$$i(0) = \left(\frac{\beta_r}{\beta_r + \beta_{nr}}\right) G = \left(\frac{\beta_r}{\beta_r + \beta_{nr}}\right) \alpha I_{ex}. \tag{3.12}$$

Therefore, in the case of bimolecular recombination, the steady-state luminescence intensity remains a linear function of the excitation intensity I_{ex}.

However, let us stress again that this result is valid only if the steady state has been reached, that is, in particular when using continuous excitation radiation of an incandescent lamp, or possibly of a gas or a semiconductor laser. Nevertheless, bimolecular recombination requires a rather high density of electron–hole pairs since its probability scales like the square of this density. This is more likely to occur when using excitation by powerful laser pulses. And under a pulsed excitation, the intensity dependence of bimolecular luminescence can be governed by a law different from the linear relation of the type (3.12) as we now proceed to show.

Let us first solve a general differential equation of the form (3.10). It is easy to check that the solution reads

$$n(t) = \sqrt{\frac{G}{\beta}} \frac{1 - \mathrm{e}^{-2\sqrt{G\beta}t}}{1 + \mathrm{e}^{-2\sqrt{G\beta}t}} + \text{const.} \tag{3.13}$$

Let us now choose a boundary condition different from the one we have used up to now. Let $n(0) = 0$ at time $t = 0$ and now we promptly turn on the excitation G in the form of a rectangular pulse of duration t_p (Fig. 3.4(a)). If $2\sqrt{G\beta}t \ll 1$ holds for all $0 \leq t \leq t_p$ (i.e. the excitation pulse is sufficiently short), then we can approximate the exponential terms in (3.13) using a series expansion $\exp(-x) \approx 1 - x$ which yields

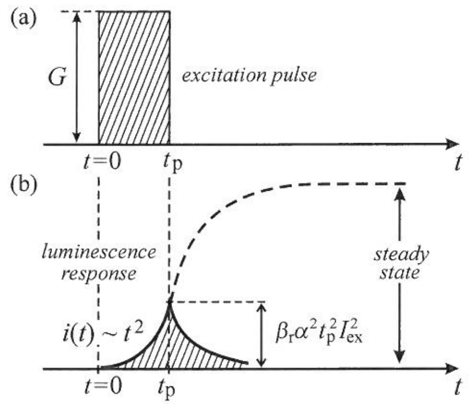

Fig. 3.4
Time-dependent bimolecular luminescence under a pulsed excitation. (a) A short excitation pulse $t_p \ll 1/2\sqrt{G\beta}$; (b) luminescence response pulse. The dashed curve represents the evolution of luminescence intensity towards the steady state provided the excitation lasts long enough ($t_p \to \infty$).

$$n(t) \approx \sqrt{\frac{G}{\beta}} \frac{2\sqrt{G\beta}t}{2(1 - \sqrt{G\beta}t)} \approx Gt. \tag{3.14}$$

Then, the luminescence intensity $i(t) = \beta_r n^2(t)$ is given by

$$i(t) = \beta_r n^2(t) \approx \beta_r G^2 t^2 \approx \beta_r \alpha^2 t^2 I_{ex}^2. \tag{3.15}$$

Hence, the amplitude of the luminescence response pulse is proportional to the square of the excitation intensity I_{ex} (Fig. 3.4(b)). We leave it to the reader to confirm that we can never get a similar quadratic dependence for a monomolecular process even if we excite using an arbitrarily short pulse $t_p \ll \tau$ (Problem 3/2).

When is the excitation pulse 'short enough' so that the quadratic dependence (3.15) can reveal itself? To get a numerical estimate, let us consider, for example, excitation by using a rather weak pulsed lamp with an excitation power density of $100\,\mathrm{W/cm^2}$, bimolecular recombination coefficient $\beta = 10^{-15}\,\mathrm{cm^3\,s^{-1}}$, and the excitation photon energy $h\nu_{ex} = 3.2\,\mathrm{eV}$. Then the excitation intensity in terms of the relevant photon flux is $I_{ex} = 100\,(\mathrm{J/cm^2\,s})/h\nu_{ex} \approx 2 \times 10^{20}\,\mathrm{photon/cm^2\,s}$. With absorption coefficient $\alpha = 10^4\,\mathrm{cm^{-1}}$ we have $G = I_{ex}\alpha \approx 2 \times 10^{24}\,\mathrm{cm^{-3}\,s^{-1}}$ and $\sqrt{G\beta} = \sqrt{2 \times 10^{24} \times 10^{-15}} \approx 4.5 \times 10^4\,\mathrm{s^{-1}}$. Thus, the relation $t_p \ll (1/2\sqrt{G\beta}) \approx 1 \times 10^{-5}\,\mathrm{s}$ will hold for pulses of duration of roughly 1 µs and less. Of course, under a higher-intensity excitation G and higher values of β, the duration of pulses leading to the quadratic dependence (3.15) will get substantially shorter; nevertheless, it is still easily achievable with excitation by common laser pulses with durations of the order of nanoseconds or tens of picoseconds.

3.4 Stretched exponential

In practice, we observe certain cases when the luminescence decay curve is neither governed by the exponential law (3.5) nor by a dependence of the type (3.11) but, instead, it can be described by the so-called stretched exponential

$$i(t) = i(0) \exp[-(t/\tau)^{\delta}], \tag{3.16}$$

where τ is called the decay time and δ is the dispersion factor with $0 < \delta < 1$. It is obvious that for the limiting value $\delta = 0$, eqn (3.16) represents a constant function, and for $\delta = 1$ relationship (3.16) reduces to the standard exponential decay. Between the limiting values of the parameter δ, the function (3.16) interpolates these two extremes as shown in Fig. 3.5. This also hints at the origin of the name—as if we 'stretched' an exponentially decreasing function towards later times.

This decay law is often observed in disordered systems and it is believed to be due to the so-called dispersion diffusion of photoexcited carriers. In particular, eqn (3.16) describes very precisely luminescence decay in amorphous semiconductors or silicon nanostructures. Nevertheless, it is important to realize that, in itself, the presence of the stretched exponential function in luminescence decay does not allow us to identify the luminescence centre in microscopic detail. There exist several plausible explanations—e.g. in a silicon nanocrystal system, this can be the case for photocarrier migration between various localized states either within a single nanocrystal or along multiple closely packed nanocrystals. Moreover, the diffusion itself can occur due to carriers being excited from localized to free states (multiple trapping and release) or hopping directly between the localized states. The degree of the dispersion of motion is measured by δ, which depends on the trap density, the distribution of their depths, etc. Moreover, both δ and τ depend also on the temperature and the wavelength at which the decay is measured.

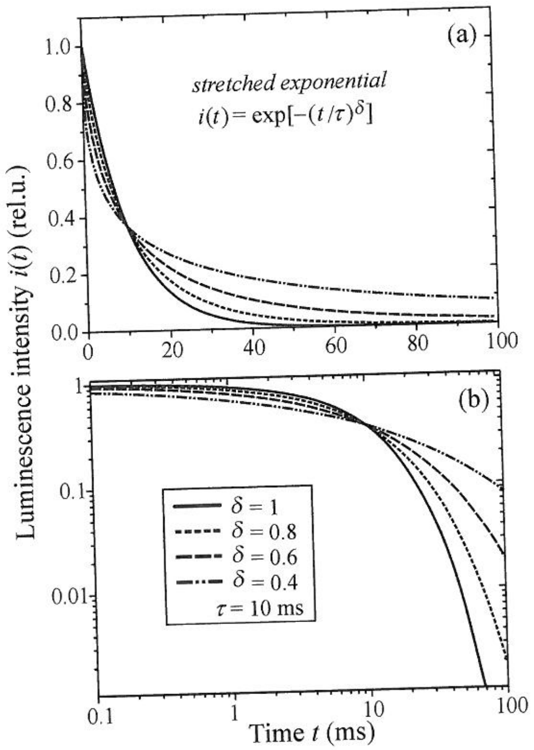

Fig. 3.5
A plot of the stretched exponential $i(t) = i(0)\exp[-(t/\tau)^{\delta}]$ for $i(0) = 1$, $\tau = 10$ ms and several values of the parameter δ. (a) A lin–lin scale; (b) a log–log scale often used to plot the stretched exponential.

3.5 Multiple processes present simultaneously

Let us now consider a somewhat more complex situation that, however, comes closer to that from a large variety of possible experimental events. We shall deal with two parallel recombination channels: bimolecular (intrinsic) recombination of free electron–hole pairs produced by optical excitation $G = \alpha I_{\mathrm{ex}}$ and monomolecular (extrinsic) recombination of impurity atoms present in a total concentration N_0. These impurity centres are excited primarily via the same optical excitation that produces free electrons and holes. All processes involved are shown in Fig. 3.6. We shall be mainly interested in the character of the dependence of both luminescence lines on the excitation intensity in the steady state. We shall show that in case there are both monomolecular and bimolecular luminescence processes present in the material simultaneously, they can be distinguished by their typical intensity dependences.

We divide the free photoelectrons and holes into the so-called 'optical', n_0, and 'thermal', n_{t}, electrons and 'optical', p_0, and 'thermal', p_{t}, holes. Optical electrons and holes represent a population of carriers occupying almost all available states in the conduction and valence bands.[3] They can mutually recombine radiatively with a bimolecular recombination coefficient β, emitting a photon $h\nu_{\mathrm{B}} \geq E_{\mathrm{g}}$. By thermal photocarriers n_{t} or possibly p_{t}, we shall mean electrons and holes in the lowest energy states just above the bottom of the conduction or below the top of the valence band, respectively. These quasiparticles originate from the population of optical carriers via thermalization, which process is described using a rate constant k. They have the highest chance to be captured by an impurity centre, promoting this into an excited state. Let us denote the concentration of the excited centres by N. Let the process of both electron and hole capture be described by the same rate constant a; its occurrence probability is proportional to the product of the

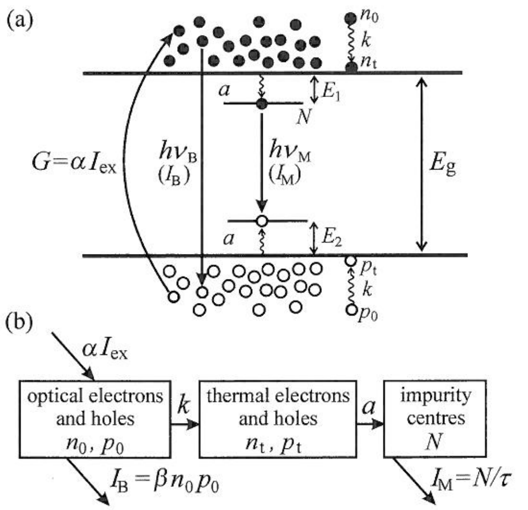

Fig. 3.6
(a) The band scheme of a semiconductor showing the processes of bimolecular (intrinsic) and monomolecular (extrinsic) photoluminescence—see the main text and eqns (3.17)–(3.19). The respective photon energies are $h\nu_{\mathrm{B}} \approx E_{\mathrm{g}}$, $h\nu_{\mathrm{M}} \approx E_{\mathrm{g}} - (E_1 + E_2)$. (b) The corresponding block diagram of the transitions.

---

[3] We can presume their energy distribution to be described by the classical Boltzmann tail of the Fermi–Dirac distribution function with the Fermi energy level being roughly amid the bandgap.

concentration of impurity centres in the ground state (N_0–N) with the concentration of thermal electrons n_t, or holes p_t, respectively. The excited impurity atoms then luminesce (emitting a photon of energy $h\nu_M \approx E_g$–$(E_1 + E_2)$), the number of recombination acts per unit volume being given by N/τ. Let us note that, for the sake of simplicity and without loss of universality of our arguments, we shall assume 100% quantum efficiency in both luminescence processes, i.e. $\beta = \beta_r$ and $\tau = \tau_r$.

All the above processes can be described by the following kinetic equations:

$$\frac{dn_0}{dt} = \alpha I_{ex} - \beta n_0 p_0 - k\,n_0, \tag{3.17}$$

$$\frac{dn_t}{dt} = k\,n_0 - a(N_0 - N)n_t, \tag{3.18}$$

$$\frac{dN}{dt} = a(N_0 - N)n_t - \frac{N}{\tau}; \quad N \neq N_0, N \neq 0, \tag{3.19}$$

$$n_0 = p_0, n_t = p_t.$$

In the steady state, $d/dt = 0$, we obtain

$$\beta n_0^2 + k\,n_0 - \alpha\,I_{ex} = 0, \tag{3.20}$$

$$a(N_0 - N)n_t - k\,n_0 = 0, \tag{3.21}$$

$$a(N_0 - N)n_t = N/\tau. \tag{3.22}$$

We rewrite equation (3.22) in the form

$$\frac{N}{\tau} = \frac{a\,n_t N_0}{(1 + a\,n_t\tau)}, \tag{3.23}$$

having thus obtained a parametric expression for the monomolecular luminescence intensity, $I_M = N/\tau$; for the bimolecular luminescence intensity, we have in accordance with our previous notation $I_B = \beta\,n_0^2$.

Solving the quadratic equation (3.20), we get the equilibrium concentration of optical electrons, n_0, and hence immediately also I_B; using eqns (3.21) and (3.23), we obtain n_t and thus also I_M.

The solution to (3.20) reads

$$n_0 = \frac{k}{2\beta}\left(\sqrt{1 + \frac{I_{ex}}{I_0}} - 1\right), \qquad I_0 = \frac{k^2}{4\alpha\beta}. \tag{3.24}$$

Substituting this expression into (3.21) and taking advantage of (3.23), we get after some simple algebra

$$a\,n_t\tau = \frac{\frac{k^2}{2\beta}\left(\sqrt{1 + \frac{I_{ex}}{I_0}} - 1\right)}{\frac{N_0}{\tau} - \frac{k^2}{2\beta}\left(\sqrt{1 + \frac{I_{ex}}{I_0}} - 1\right)}. \tag{3.25}$$

In relations (3.24) and (3.25), $I_0 = k^2/4\alpha\beta$ represents a sort of characteristic excitation intensity governed exclusively by the material parameters. We now distinguish two excitation modes, weak excitation and strong excitation, characterized by the relations $I_{ex} \ll I_0$ and $I_{ex} \gg I_0$, respectively.

(a) Weak excitation ($I_{ex} \ll I_0$). In this case we apply the familiar approximation $\sqrt{1+x} \approx 1 + x/2$ valid for $x \ll 1$. From eqn (3.24) we then get

$$n_0 \approx \frac{\alpha\,I_{ex}}{k}, \qquad I_B = \beta\,n_0^2 = \frac{\alpha^2\beta}{k^2} I_{ex}^2. \tag{3.26}$$

Therefore, bimolecular luminescence varies quadratically with excitation intensity. From eqn (3.25), we have

$$a\,n_t\tau \approx \frac{\alpha\,I_{ex}}{\frac{N_0}{\tau} - \alpha\,I_{ex}}, \qquad a\,n_t N_0 \approx \frac{N_0\alpha I_{ex}}{N_0 - \alpha\tau I_{ex}} \tag{3.27}$$

and, finally, substituting the above relations into (3.23) yields, for the monomolecular luminescence,

$$I_M = \alpha\,I_{ex}; \tag{3.28}$$

thus, I_M increases linearly with the excitation.

Unlike the situation when there is just a single active luminescence process in a semiconductor (Sections 3.2 and 3.3), the steady-state bimolecular and monomolecular luminescences now depend on the excitation intensity in a markedly different way: I_B scales quadratically while I_M scales linearly with excitation intensity. This corresponds to the intuitive and rather widespread scheme; we stress, however, that the necessary prerequisite is the parallel occurrence of both phenomena, which is sometimes overlooked just as the weak excitation condition. (Short-pulse excitation, discussed in Section 3.3, is another example of a quadratic demonstration of bimolecular recombination.)

It is not difficult to understand what goes on in our model material during weak excitation. Comparing (3.28) with the starting eqn (3.17), we see that basically all the absorbed energy αI_{ex} is radiated in the form of monomolecular impurity luminescence. Thus, under weak excitation, bimolecular emission is substantially weaker than monomolecular emission ($I_B \ll I_M$), which makes sense since the density of the 'optical' electron–hole pairs is low and it is difficult for the electrons and holes to find one another to recombine. Almost all the photoelectrons are trapped at the impurity centres, which subsequently return to the ground state, emitting an extrinsic luminescence photon $h\nu_M \approx E_g$–$(E_1 + E_2)$.

This enables us to formulate now the weak excitation condition alternative to the relationship $I_{ex} \ll I_0$. Under the circumstances, the low excitation intensity is physically equivalent to the fact that by far not all impurity atoms are brought into the excited state. In other words, the number of all available recombination impurity events, N_0/τ, is much larger than the density of photoexcitation events per second: $N_0/\tau \gg \alpha I_{ex}$. This, of course, then reduces relation (3.27) to the form

$$a\,n_t\tau \approx \frac{\alpha\tau\,I_{ex}}{N_0} \ll 1. \tag{3.29}$$

Indeed, neglecting $an_t\tau$ as compared to unity in the denominator of eqn (3.23), we immediately obtain $I_M = N/\tau \approx an_tN_0$. This, considering (3.29) again yields $I_M \approx (\alpha I_{ex}/N_0)N_0 = \alpha I_{ex}$, which is in accordance with (3.28).

Therefore, inequality (3.29) is an alternative expression for the weak excitation condition.

(b) Strong excitation. It appears beneficial to formulate the strong excitation condition as a relation opposite to (3.29):

$$a n_t \tau \gg 1.$$

Equation (3.23) then immediately informs us of the monomolecular luminescence intensity

$$I_M = \frac{N}{\tau} = \frac{a\, n_t\, N_0}{(1 + a\, n_t \tau)} \approx \frac{a\, n_t\, N_0}{a\, n_t \tau} = \frac{N_0}{\tau}. \qquad (3.30)$$

This is clearly the maximum achievable intensity in the case when all impurity centres are excited. Its value no longer depends on the excitation intensity. Therefore, the extrinsic radiative recombination channel becomes *saturated*.

What dependence of the bimolecular intrinsic luminescence on the excitation intensity do we get now? Using the strong excitation condition in the form $I_{ex}/I_0 \gg 1$, we obtain from eqn (3.24)

$$n_0 = \frac{k}{2\beta}\sqrt{\frac{I_{ex}}{I_0}}$$

or

$$I_B = \beta\, n_0^2 \approx \frac{\beta\, k^2}{4\,\beta^2}\,\frac{I_{ex}}{I_0} = \alpha\, I_{ex}. \qquad (3.31)$$

Hence, bimolecular luminescence increases linearly in the strong excitation limit and, in a sense, it basically swaps roles with monomolecular luminescence as can be seen by comparing with eqn (3.28). The overall behaviour of the system on a wide scale of excitation intensities is summarized by the plot in Fig. 3.7(a).

We can formulate a system of kinetic equations analogous to eqns (3.17)–(3.19) for a large number of various luminescence centres and photocarrier trap levels. An example of a very successful model is the kinetic model of exciton and biexciton optical creation, which introduces, among other things, exactly the concept of 'optical' and 'thermal' excitons to explain the experimental data [4]. An older monograph by Antonov-Romanovskij [5] even attempts to cover the basis of all luminescence phenomena in semiconductors by making use of various combinations of unspecified radiative and non-radiative transitions and a large number of trapping levels. From today's perspective, however, this approach appears obsolete.

To conclude this section, a few more general notes may be in order. The kinetic approach to luminescence phenomena may be useful when analysing experimental data, in the sense that it is able to suggest a basic idea or a hint about what type the examined luminescence radiation might be. Here, both the character of the decay curve and the measured intensity dependence can help us. A specific model can give us a very good picture of the experimental results in a number of cases. For example, the basic features of the curves in Fig. 3.7(b), representing the intensity dependence of intrinsic and impurity low-temperature luminescence in InSb [6], are qualitatively absolutely identical to the attributes of our model curves from Fig. 3.7(a) (the quadratic dependence of intrinsic luminescence and the linear dependence of impurity luminescence in the case of weak excitation; a marked decrease in the slopes and possibly a tendency to saturation at high excitation intensities).

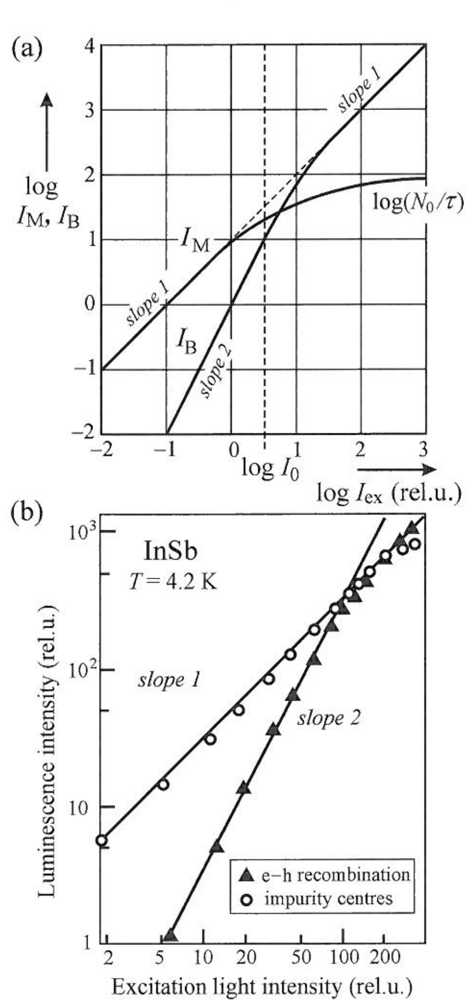

Fig. 3.7
(a) A schematic intensity dependence of monomolecular (I_M) and bimolecular (I_B) luminescence for the model shown in Fig. 3.6. (Using the log-log scale, the plot of a linear function is a straight line with slope equal to 1, the plot of a quadratic function is a straight line with slope equal to 2.) (b) The intensity dependence of bimolecular luminescence (electron–hole intrinsic recombination, triangles) and monomolecular luminescence (localized impurity centre, circles) in InSb at temperature $T = 4.2$ K. Adapted from Mooradian and Fan [6]. Compare to panel (a).

On the other hand, we must not overestimate the significance of kinetic equations. First of all, they contain no information about the spectral form of the individual luminescence lines. In addition, based on these equations alone we cannot usually draw unambiguous conclusions in the sense of a bilateral mathematical implication. For example, an experimentally established quadratic intensity dependence is evidence (if we disregard possible nonlinear optical phenomena and very high excitation effects) for the presence of a bimolecular recombination mechanism, yet the absence of such a dependence does not necessarily imply the absence of a bimolecular process—see Fig. 3.7(a) on the side of high excitation intensities. Finally, we must keep in mind that the effects of very high excitation accompanied by the bands filling (being reflected e.g. in spontaneous emission in an electron–hole system of high density, n_0, $p_0 \geq 10^{17}\,\mathrm{cm}^{-3}$, occurring for instance in LED operation or near the threshold condition of a semiconductor laser) may modify the formulation of the kinetic equations very substantially; e.g. the simple relation (3.10) is no longer valid, etc. [7]. After all, in the case of very high excitation, brand-new luminescence systems appear (they are discussed in Chapter 8), for which it often does not make much sense to distinguish mono- and bimolecular luminescence phenomena.

3.6 Problems

3/1: Show that even if we introduce quantum yields of monomolecular and bimolecular luminescence ($\eta = \tau/\tau_r < 1$, $\beta_r/(\beta_r + \beta_{nr}) = (\beta_r/\beta) < 1$) into eqns (3.17)–(3.19), all the salient features of the model shown in Fig. 3.6 remain qualitatively preserved.

3/2: Check that for a monomolecular process we cannot (as against a bimolecular process) obtain a quadratic dependence of the luminescence response to the excitation intensity even if we excite the system by an arbitrarily short pulse $t_p \ll \tau$, where τ is the luminescence decay time.

3/3: Prove that the solution to the equation $\mathrm{d}n/\mathrm{d}t = -\beta n^2 - n/\tau$ is a decay curve $n(t)$ of the form

$$n(t) = \frac{n(0)\mathrm{e}^{-t/\tau}}{1 + n(0)\,\beta\,\tau\,(1 - \mathrm{e}^{-t/\tau})}.$$

Show that at later times after the excitation was turned off (when the density of the recombining photocarriers, $n(t)$, has already dropped substantially) the total luminescence decay is governed by a sum of exponential curves. What condition must be satisfied for the decay to be described by a single exponential $\sim \exp(-t/\tau)$?

3/4: Show that (a) the solution to eqn (3.10) is (3.13) and that (b) for long excitation pulses and after a transient initial increase in luminescence intensity, $i(t)$, a steady-state value of intensity is established, which is in accord with (3.12). Plot the curve $i(t)$, from the time we turn on the excitation until the emission intensity settles.

3/5: Discuss the kinetic model of exciton and biexciton creation according to [4].

References

1. Bebb, H. B. and Williams, E. W. (1972). *Photoluminescence I. Theory*. In *Semiconductors and Semimetals* (ed. R. K. Willardson and A. C. Beer), Vol. 8, p. 181. Academic Press, New York.

2. Rice, T. M. (1977). *The electron-hole liquid in semiconductors: Theoretical aspects*. In *Solid State Physics* (ed. H. Ehrenreich, F. Seitz, and D. Turnbull), Vol. 32, Chap. 1. Academic Press, New York.
3. Yu, P. Y. and Cardona, M. (1996). *Fundamentals of Semiconductors*. Springer, Berlin
4. Knox, R. S., Nikitine, S., and Mysyrowicz, A. (1969). *Optics Comm.*, **1**, 19.
5. Antonov-Romanovskij, V. V. (1966). *Kinetics of photoluminescence of crystalline phosphors* (in Russian: *Kinetika fotoljuminescencii kristalofosforov*). Nauka, Moskva.
6. Mooradian, A. and Fan, H. Y. (1966). *Phys. Rev.*, **148**, 873.
7. Bourdon, G., Robert, I., Sagnes, I., and Abram, I. (2002). *J. Appl. Phys.*, **92**, 6595.