These effects lead to very dissimilar force constants for different vibration types, to a low energy of TO-phonons at the L point (Fig. 4.6), and to the inverted character of TO- and TA-phonons near the L point. Discuss these phenomena according to Dorner *et al.* [15].

4/5: Confirm that at $T = 0\,\mathrm{K}$ the Debye–Waller factor u_{DW} and the Huang–Rhys factor S are related by $u_{\mathrm{DW}} = \exp(-S)$. (Hint: Use eqns (4.15) and (4.16).)

References

1. Dekker, A. J. (1963). *Solid State Physics*. Chap. 2. Prentice-Hall, Englewood Cliffs, N.J.
2. Peyghambarian, N., Koch, S. W., and Mysyrowicz, A. (1993) *Introduction to Semiconductor Optics*. Chap. 4. Prentice Hall, Englewood Cliffs, N.J.
3. Dargys, A. and Kundrotas, J. (1994). *Handbook on Physical Properties of Ge, Si, GaAs and InP*. Science and Encyclopedia Publishers, Vilnius.
4. Hayes, W. and Stoneham, A. M. (1985). *Defects and Defect Processes in Non-metallic Solids*. Chap. 1. John Wiley, New York.
5. Yu, P. Y. and Cardona, M. (1996). *Fundamentals of Semiconductors*. Springer, Berlin.
6. Song, K. S. and Williams, R. T. (1996). *Self-Trapped Excitons* (Springer series in solid state sciences Vol. 105). Springer, Berlin.
7. Böer, K. W. (1990). *Survey of Semiconductor Physics*. Van Nostrand Reinhold, New York.
8. Vij, D. R. (ed.) (1998). *Luminescence of solids*. Plenum Press, New York; Yamamoto, H. (1999). *Fundamentals of Luminescence*. In *Phosphor Handbook* (ed. S. Shionoya and W. M. Yen), p. 35. CRC Press, Boca Raton.
9. Pelant, I., Hála, J., Ambrož, M., Vácha, M., Valenta, J., Adamec, F., Kohlová, V., and Matoušková, J. (1990). *Impurity assessment in Si wafers by photoluminescence method V*. Research report for Tesla Rožnov. Charles University in Prague, Faculty of Mathematics & Physics, Prague.
10. Wassmuth, W., Stolz, H., and von der Osten, W. (1990). *J. Phys. C: Cond. Matter*, **2**, 919.
11. Kanzaki, H. and Sakuragi, S. (1969). *J. Phys. Soc. Japan*, **27**, 109.
12. Shah, J. (1972). *Phys. Rev. B*, **6**, 4592.
13. Pelant, I. and Hála, J. (1991). *Solid State Comm.*, **78**, 141.
14. Williams, F. E. (1948). *The mechanism of rate processes in the luminescence of solids*. In *Preparation and Characteristics of Solid Luminescent Materials*. Cornell Symposium 1946, p. 337. John Wiley, New York; Chapman & Hall, London.
15. Dorner, B., von der Osten, W., and Bührer, W. (1976). *J. Phys. C: Solid State Phys.*, **9**, 723.

Channels of radiative recombination in semiconductors

5

Various radiative recombination processes have already been mentioned several times in the previous chapters. These processes were, however, not discussed from the point of view of the underlying physics; more emphasis was put on the introduction (or reminder) of the basic concepts and terminology, which will be indispensable later in the book. The configurational coordinate model, which was described in the previous chapter, for example, is broadly applicable. In this chapter, we are going to show in which particular aspects of luminescence in semiconductors the application of this model is completely straightforward, when, on the other hand, it does not make any sense at all, or when it can be used only formally and with prudence.

Now that everything is ready we can move on to giving systematic explanations of known channels of radiative recombination in semiconductors. First, we will list an overview of these channels. Secondly, we will focus in more detail on the spectral shape of the emission line of recombining free electron–hole pairs, and of related recombinations of a free electron with a neutral acceptor or a free hole with a neutral donor. Then we will talk about the shape of the emission spectra of recombining donor–acceptor pairs and finally we will mention the particularity of exciting luminescence via a two-photon absorption process.

5.1 Overview of luminescence processes in crystalline semiconductors

Let us start with the widely accepted classification of luminescence processes according to the intensity of excitation. The *low-fluence* or *weak excitation* processes (i.e. taking place when a gas-discharge lamp, an incandescent lamp or a continuous-wave gas laser with the output power of the order of 0.01–$10\,\mathrm{W/cm^2}$ are used for excitation) comprise:

- radiative recombination of free excitons (often denoted as FE for a free exciton, or X);
- radiative recombination of free excitons with simultaneous emission of an LO phonon (FE–LO, X–LO);

- radiative recombination of a bound exciton (BE), which can be further subdivided into
  - radiative decay of an exciton bound to a neutral donor (D^0, X) or (D^0–X);
  - radiative decay of an exciton bound to a neutral acceptor (A^0, X) or (A^0–X);
  - radiative decay of an exciton bound to an ionized donor (D^+, X) or (D^+–X);
  - radiative decay of an exciton bound to an ionized acceptor (A^-, X) or (A^-–X);
  - radiative decay of an exciton bound to an isoelectronic impurity;
- radiative recombination of donor–acceptor pairs (D^0–A^0);
- radiative recombination of a free hole with a neutral donor (h–D^0) or a free electron with a neutral acceptor (e–A^0);
- radiative recombination of free electron–hole pairs (e–h), which can also be referred to as band-to-band recombination.

At *high-fluence (strong) excitation*, usually by means of a pulsed laser with excitation intensity of the order of 1 kW–10 MW/cm^2, additional processes occur:

- radiative decay of an excitonic molecule (EM or XX);
- luminescence coming from inelastic collisions of excitons (X–X collisions);
- luminescence of electron–hole liquid (EHL);[1]
- luminescence of electron–hole plasma (EHP);
- Bose–Einstein condensation of excitons or excitonic molecules and luminescence coming from the condensate.

Obviously, two or more types of these processes can occur simultaneously in a material. In addition, most of these processes can also be observed in low-dimensional semiconductor structures, although the emission of light from low-dimensional structures is in some aspects specific. The emission of light from amorphous semiconductors, on the other hand, is governed by significantly different processes.

Later in this chapter, the most important spectral features connected with luminescence processes occurring at low-fluence excitation levels will be treated, deliberatly neglecting any luminescence manifestation of excitons, to which Chapter 7 is devoted. The high-excitation luminescence spectral features will be discussed in Chapter 8 and the luminescence of amorphous and low-dimensional semiconductors will be described in Chapters 9 and 12.

5.2 Recombination of free electron–hole pairs

Even though, in reality, this process is not encountered very often, the explanation of the shape of a luminescence emission spectrum will begin here; the reasons for choosing this particular process will become evident later. (We would like to note here that this process was called 'bimolecular' in Chapter 3, where it was treated from the point of view of its kinetics.)

[1] This process is sometimes not considered typical for high-level excitations because it can often be observed even when a continuous source of radiation is used.

5.2.1 Direct bandgap

The process is schematically shown in Figure 5.1. An electron with energy E_2 situated close to the minimum of the conduction band spontaneously recombines with a hole situated at the top of the valence band with energy E_1 and wavector $\mathbf{k}_0$ via a direct vertical transition. The physical quantity of interest here is the probability density $I_{sp}(h\nu)$ of the spontaneous emission of a photon with energy $h\nu = E_2 - E_1$. This is obviously influenced by three factors:

1. the probability of a suitable configuration of occupancy of conduction and valence band states $f_e(\nu)$;
2. the quantum-mechanical probability of this transition in the form of the square of the absolute value of the matrix element of the transition (in the dipole approximation) $|M|^2$;
3. the joint density of electron and hole states $\rho(\nu)$ in the corresponding bands.

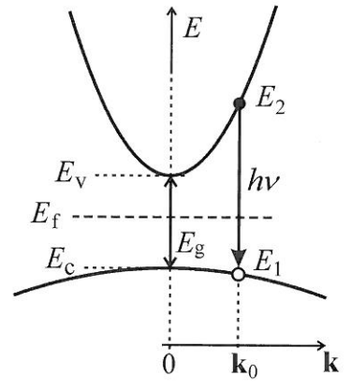

Fig. 5.1
Free electron–hole recombination in a direct bandgap semiconductor. E_f stands for the Fermi level.

Thus

$$I_{sp}(h\nu) \approx |M|^2 f_e(\nu)\rho(\nu) \approx \frac{1}{\tau_r} f_e(\nu)\,\rho(\nu). \tag{5.1}$$

Ad 1. For the radiative recombination to take place, the corresponding energy state with energy E_2 in the conduction band has to be occupied by an electron while the state with energy E_1 in the valence band has to be empty (i.e. occupied by a hole), as is indicated in Fig. 5.1. Consequently, $f_e(\nu)$ is given as the product of the occupancy factors f_c and f_v:

$$f_e(\nu) = f_c\, f_v = f(E_2)[1 - f(E_1)], \tag{5.2}$$

where $f(E) = \{\exp[(E - E_f)/k_BT] + 1\}^{-1}$ is the Fermi–Dirac distribution function. Let us assume that the semiconductor is intrinsic, has a sufficiently wide bandgap E_g and the process takes place at low temperature. At low-fluence excitation, the semiconductor can be treated as if it is still in thermal equilibrium, i.e. the Fermi level stays localized in the middle of the bandgap at low temperature (Fig. 5.1). As long as the condition of low temperature and wide bandgap $k_BT \ll (E_g/2) \approx E_2 - E_f \approx E_f - E_1$ is fulfilled, one can write

$$f(E_2) = \frac{1}{e^{(E_2 - E_f)/k_BT} + 1} \approx \exp\left[-(E_2 - E_f)/k_BT\right],$$

$$[1 - f(E_1)] = 1 - \frac{1}{e^{(E_1 - E_f)/k_BT} + 1} \approx \frac{e^{(E_1 - E_f)/k_BT}}{e^{(E_1 - E_f)/k_BT} + 1}$$

$$\approx \exp[-(E_f - E_1)/k_BT],$$

which transforms (5.2) into the simple form

$$f_e(\nu) \approx \exp(-h\nu/k_BT). \tag{5.3}$$

Ad 2. The matrix element of the optical transition M can be considered to be constant to a very good approximation over a wide range of wavevectors close to the extremes of the bands. In relation to luminescence, the radiative lifetime τ_r will be treated as inversely proportional to the square of the matrix element $\tau_r \sim |M|^{-2}$. (This simple relation is strictly valid only for atomic transitions. When it comes to semiconductors, the application of this relation is a bit less straightforward; the parameters of the particular band structure as well as the bandgap width come into play. A detailed theoretical discussion can be found in [1]. This approximation, however, is fully sufficient within the framework of this textbook.)

Ad 3. The density of states $\rho(\nu)$ is pertinent to the difference energy band $E_c(\mathbf{k})$–$E_v(\mathbf{k})$, which is obtained—mathematically speaking—by subtracting $E_v(\mathbf{k})$ from $E_c(\mathbf{k})$. Since these transitions are vertical the subtraction makes sense in this context. The densities of states in the individual bands, i.e. in the conduction band $\rho_c(E)$ and the valence band $\rho_v(E)$, are well known (see, e.g., [2, 3]):

$$\rho_c(E) = \frac{(2m_e)^{3/2}}{2\pi^2\hbar^3}(E - E_c)^{1/2},\ E \geq E_c, \tag{5.4a}$$

$$\rho_v(E) = \frac{(2m_h)^{3/2}}{2\pi^2\hbar^3}(E_v - E)^{1/2},\ E \leq E_v, \tag{5.4b}$$

where m_e and m_h stand for the effective masses of an electron and a hole, respectively, connected with the curvatures of the corresponding bands close to the extremes. Thus, one can easily get (using the incremental relation $\rho_c(E_2)\mathrm{d}E_2 = \rho(\nu)\mathrm{d}\nu$)

$$\rho(\nu) = \frac{(2m_r)^{3/2}}{\pi\hbar^2}(h\nu - E_g)^{1/2};\quad h\nu \geq E_g. \tag{5.5}$$

Relation (5.5) includes the reduced mass of the electron and hole given by the expression

$$\frac{1}{m_r} = \frac{1}{m_e} + \frac{1}{m_h}. \tag{5.6}$$

To derive $\rho(\nu)$ by making use of the aforementioned incremental relation, the following expressions, which stem directly from Fig. 5.1 (see Problem 5/1), are employed:

$$E_2 = E_c + \frac{m_r}{m_e}(h\nu - E_g), \tag{5.7}$$

$$E_1 = E_v - \frac{m_r}{m_h}(h\nu - E_g). \tag{5.8}$$

One more thing worth mentioning is that the density $\rho(\nu)$ is given per unit frequency, not per unit energy, which causes different fractional prefactors to appear in (5.4) and (5.5).

Now, it is possible to get the desired formula for the spectral shape of luminescence coming from the radiative recombination of free electron–hole pairs in a semiconductor with direct bandgap by combining (5.3), (5.5) and (5.1):

$$I_{sp}(h\nu) \approx D_0(h\nu - E_g)^{1/2}\exp\left[-(h\nu - E_g)/k_BT\right], \tag{5.9a}$$

where $D_0 = [(2m_r)^{2/3}/\pi\hbar^2\tau_r]\exp(-E_g/k_BT)$ is a parameter independent of the photon energy $h\nu$. Disregarding this parameter, $I_{sp}(h\nu)$ is given by the product of a factor characterizing the joint density of states $(h\nu–E_g)^{1/2}$ with a term describing basically the occupancy of these states, i.e. $\exp[-(h\nu - E_g)/k_BT]$. The shape of the luminescence spectrum (5.9a) is sometimes also called the *Maxwell–Boltzmann distribution*. It is interesting to note that the formula describing the radiative band-to-band recombination contains solely the reduced electron–hole mass m_r, not the individual effective masses of an electron and a hole m_e, m_h. In other words, it is impossible to determine the values of the effective masses of the carriers using only luminescence (or absorption) measurements.

The spectral shape $I_{sp}(h\nu)$ is shown in Fig. 5.2(a). The curve is asymmetric with a tail towards the high-energy photon side. The asymmetry stems from the kinetic energy of recombining free electrons and holes, which can be described by an exponential distribution function.[2] The full-width

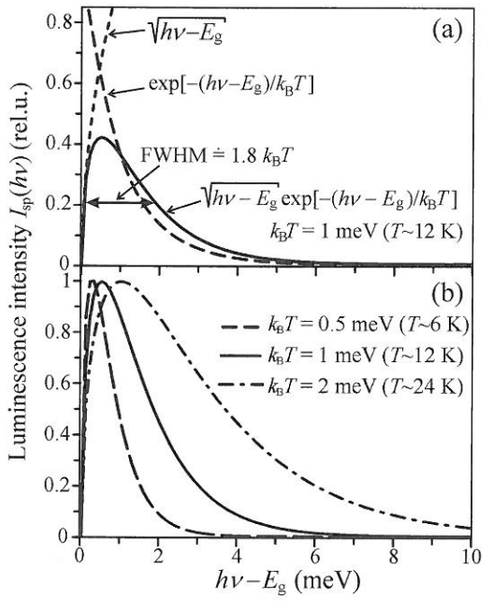

Fig. 5.2
(a) Maxwell–Boltzmann lineshape $I_{sp}(h\nu) \approx \sqrt{h\nu - E_g}\exp[-(h\nu - E_g)/k_BT]$ at $T \approx 12$ K. (b) Asymmetric linewidth broadening with increasing temperature. Curves are normalized.

[2] It may be of interest to recall that the formula for the number of molecules of an ideal gas Δn (out of the total number n), whose kinetic energy falls in the interval $(E_k, E_k + \Delta E_k)$, reads

$$\frac{\Delta n}{n\,\Delta E_k} = \frac{2}{\sqrt{\pi}\,(k_B\,T)^{3/2}}\sqrt{E_k}\exp(-E_k/k_BT).$$

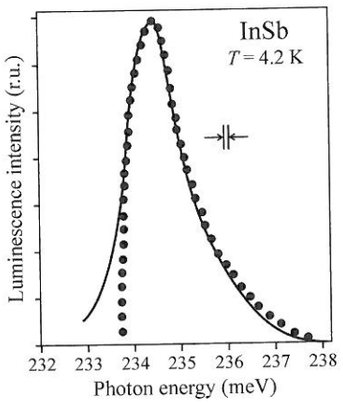

Fig. 5.3
Emission spectrum of the recombination of free electron–hole pairs in InSb at $T = 4.2$ K. Solid curve—experiment; circles—theoretical lineshape given by (5.9). After Mooradian and Fan [4].

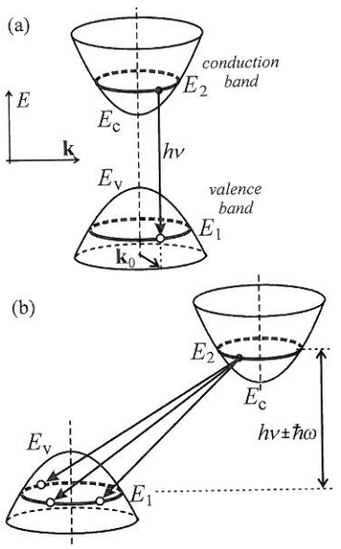

Fig. 5.4
Schematic drawing of radiative recombination of electron–hole pairs in the energy band struture diagram $E(\mathbf{k})$ for a semiconductor with (a) a direct and (b) an indirect bandgap. $h\nu$ stands for the energy of a photon, $\hbar\omega$ for the energy of a phonon.

at half maximum (FWHM) of this curve can then be shown to be directly proportional to temperature

$$\text{FWHM} \sim 1.8\, k_B T.$$

Thus, the linewidth broadens with rising temperature, as can be seen in Fig. 5.2(b). On the whole, however, the line is very narrow: temperatures T between 4 and 100 K correspond to FWHM $\approx$ 1–20 meV, which amount to only tenths of nanometres, or nanometres at most, in the visible spectral region. The physical reason underlying the linewidth broadening is obvious: the Boltzmann tail of the Fermi–Dirac distribution broadens with increasing temperature, or in other words, at higher temperatures photoexcited electrons and holes occupy states further from the extremes of the bands. Thus, photons with higher energy are generated by these higher-energy electron–hole pairs, while the low-energy onset of the spectral line is determined by E_g and its position remains fixed. (The possible shift of E_g with changing temperature leads to a shift of the emission line as a whole.)

Figure 5.3 demonstrates the experimentally observed luminescence of electron–hole pairs in InSb at 4.2 K [4]. The dotted curve indicates the theoretical lineshape given by (5.9a) and agrees very well with experiment except for the low-energy side. This slight difference can be due to several factors, including the influence of the slit of the monochromator as described in Section 2.8, which smears the sharp onset of luminescence. As has already been mentioned, the radiative recombination of free electrons and holes is a scarcely encountered process. Indium antimonide InSb is the exception that proves the rule. Its very narrow bandgap ($E_g \sim 0.23$ eV at $T = 4.2$ K) and consequently also the low exciton binding energy ensure that even at low temperatures excitons are thermally dissociated. In nearly all other semiconductive compounds, low-temperature luminescence is dominated by excitonic effects, which means that free electrons and holes and their luminescence are practically absent.

In the literature, formula (5.9a) can sometimes be written in the slightly modified form

$$I_{sp}(h\nu) \approx \nu^2\, (h\nu - E_g)^{1/2} \exp\left[-(h\nu - E_g)/k_B T\right]. \qquad (5.9b)$$

The factor ν^2 comes from the *photon density of states* in three dimensions, which equals $4\nu^2/c^3$ in free space. Taking into account the already mentioned narrow spectral width and high optical frequencies (FWHM/$h\nu$) ≤ 0.01, however, makes it possible to consider ν^2 a constant, and thus to neglect the influence of this term on the spectral lineshape.

5.2.2 Indirect bandgap

A fundamental difference between the recombination of electron-hole pairs in a semiconductor with a direct and indirect bandgap is shown in Fig. 5.4.

This formula implies that the Maxwell–Boltzmann distribution is a fully valid name for (5.9a) and, moreover, the comparison of these two formulas suggests that at low concentrations of free quasi-particles (= 10^{17} cm^{-3}, i.e. for weak optical excitation) one can speak of a 'gas' of free electrons, holes or excitons in semiconductors.

In a direct-bandgap semiconductor, the recombination takes place between electrons and holes with the same wavevector $\mathbf{k}_0$, whereas in an indirect-bandgap semiconductor, an excited electron with energy E_2 can recombine with any hole with energy E_1 and varying values of $\mathbf{k}$, as long as the energy-conservation rule E_2–$E_1 = h\nu \pm \hbar\omega$ is fulfilled. The energy $\hbar\omega$ then denotes the energy of a phonon, which takes part in the process in order to ensure the conservation of quasi-momentum: $k_{\hbar\omega} = |\mathbf{k}(E_2)$–$\mathbf{k}(E_1)|$. Considering the fact that phonon dispersion relations are defined for any $\mathbf{k}$ from the first Brillouin zone (see, e.g., Fig. 4.3), it is obvious that an appropriate phonon can always be found, unless the transitions are forbidden due to the symmetry of electron and/or phonon states. For the sake of simplicity, only low-temperature cases will be treated hereafter. At low temperatures, only a small number of phonons are present in the semiconductor and the recombination process is accompanied by the emission of a photon and the *emission of a phonon*, i.e. the electron excitation energy is released both in the form of luminescence and as an additional vibration energy of the lattice: E_2–$E_1 = h\nu + \hbar\omega$.

The calculation of the probability density of spontaneous emission I_{sp}^{in} of a luminescence photon with energy $h\nu = (E_2$–$E_1)$–$\hbar\omega$ is then carried out by summing all of the contributions over the curves marked in bold in Fig. 5.4(b), subsequently 'infinitesimally shifting' the difference $(E_2$–$E_1)$ on the energy axis, repeating the summation, etc. Factors which play important roles in the calculation are thus, naturally, not only the densities of states in the bands (5.4), but also the occupancy factors f_c, f_v.[3] In analogy with (5.1), the described procedure leads to the integration over all pairs of states in the conduction and valence bands, whose energies relative to the corresponding band edges will be denoted as $E_e = E_2$–E_c for the electron and $E_h = E_v$–E_1 for the hole, respectively (see Fig. 5.5(a)). The mathematical formulation of the described procedure then reads

$$\begin{aligned} I_{sp}^{in}(h\nu) &\approx |M|^2 \int_0^{h\nu+\hbar\omega-E_g} \rho_c(E_e)\, f_c(E_e)\, \rho_v(E_h)\, f_v(E_h)\mathrm{d}E_e \\ &= |M|^2 \int_0^{\overline{h\nu}} \rho_c(E_e)\, f_c(E_e)\, \rho_v(h\nu + \hbar\omega - E_g - E_e) \\ &\qquad f_v(h\nu + \hbar\omega - E_g - E_e)\mathrm{d}E_e \\ &= |M|^2 \int_0^{\overline{h\nu}} \sqrt{E_e} f_c(E_e) \sqrt{\overline{h\nu} - E_e}\, f_v(\overline{h\nu} - E_e)\mathrm{d}E_e, \end{aligned} \qquad (5.10)$$

where, as can be seen from Fig. 5.5(a), $h\nu = E_g + E_e + E_h$–$\hbar\omega$ or $E_h = h\nu + \hbar\omega$–E_g–E_e. The upper integration limit $\overline{h\nu} = h\nu + \hbar\omega$–$E_g$ reflects the above-mentioned summation of all contributions to the flux of emitted photons $h\nu$.

[3] The joint density of states (5.5) cannot be applied, since it has a reasonable meaning only for vertical transitions.

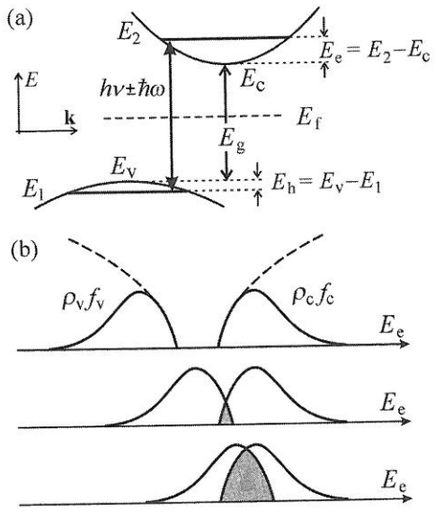

Fig. 5.5
Illustrating the calculation of the spectral lineshape of luminescence coming from the recombination of free electron–hole pairs in an indirect-bandgap semiconductor; (a) detail of band structure; (b) convolution of occupied electron and hole states according to (5.10).

On closer inspection, one can easily recognize that what (5.10) represents is the convolution of states occupied by electrons in the conduction band with the states occupied by holes in the valence band. (Let the reader compare formula (5.10) with eqn (A-5) in Appendix A.) The convolution-based calculation (5.10) is schematically shown in Fig. 5.5(b). The matrix element $|M|^2$ this time contains both vertical transitions (via virtual states with higher energies of electrons and holes at the same quasi-momentum) and also electron–phonon scattering ensuring the indispensable alteration of quasi-momentum in the recombination event. If both the photon and phonon parts of the matrix element are symmetry-allowed, this matrix element can be in a very good approximation treated as independent of both E_e and E_h and its influence on the emission lineshape can be neglected.

The calculation (5.10) can be carried out analytically without being too complicated, although one might not think so at first sight. First, we can realize that the product $f_c(E_e)\, f_v(\overline{h\nu} - E_e)$ under the integration sign is completely analogous to (5.2), but with $h\nu$ being replaced with $(h\nu + \hbar\omega)$. Similarly to (5.3) the term thus can be approximated

$$f_c(E_e)\, f_v(\overline{h\nu} - E_e) \approx \exp\left[-(h\nu + \hbar\omega)/k_B T\right]. \qquad (5.11)$$

Since the product $f_c f_v$ is thus independent of E_e, it can be placed outside the integration sign in (5.10). The formula then reads

$$I_{sp}^{in}(h\nu) \approx e^{-(h\nu+\hbar\omega)/k_B T} \int_0^{\overline{h\nu}} \sqrt{E_e}\sqrt{\overline{h\nu} - E_e}\, dE_e, \qquad (5.12)$$

where the definite integral is of the type $\int_0^a \sqrt{x}\sqrt{(a-x)}\, dx \approx a^2$ (Problem 5/4). The resulting formula for the spectral lineshape of luminescence

due to recombination of free electron–hole pairs in a semiconductor with an indirect bandgap thus acquires the form

$$I_{sp}^{in}(h\nu) \approx \left[h\nu - (E_g - \hbar\omega)\right]^2 \exp\left[-(h\nu - (E_g - \hbar\omega))/k_B T\right]. \qquad (5.13)$$

Finally, this formula is not very different from the analogous relation for a direct-bandgap semiconductor (5.9a). The first difference is represented by the red-shift of the low-energy side of the spectrum by $\hbar\omega$, the second then consists of the replacement of the square-root dependence of the pre-exponential factor in (5.9a) by a quadratic dependence in (5.13). This change in the pre-exponential factor, however, does not influence the lineshape radically; it just slightly modifies the low-energy onset of the spectrum (a precise experimental determination of which is, moreover, problematic due to spectral broadening by the slit of the dispersion device as well as by other mechanisms), while the shape of the high-energy tail is most significantly influenced by the exponential factor, which is of identical form in both (5.9) and (5.13). This similarity is documented in Fig. 5.6.

Now, it is very important to note that this luminescence process scarcely occurs in reality. The fact that the excited electrons and holes in an indirect-bandgap semiconductor live much longer than in direct-bandgap materials causes nearly all electrons and holes at low temperatures to be bound in excitons. The emission spectrum then features only the luminescence line due to the radiative recombination of excitons. Emission coming from the recombination of free electron–hole pairs can only be observed in pure crystals at slightly elevated temperatures (at about 100 K), when some of the excitons are already ionized and when, simultaneously, thermal luminescence quenching is not yet sufficiently efficient. The intensity of such luminescence is, however, very low, since the matrix element in (5.10) is of second-order correction in perturbation theory.

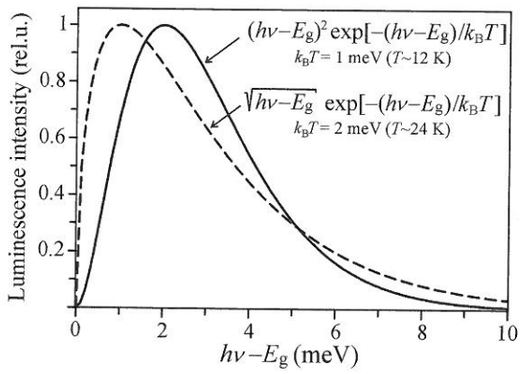

Fig. 5.6
Comparison of emission lineshapes of the recombination of free electron–hole pairs (5.9a)—dashed curve—with the lineshape described by (5.13)—solid curve. The curves are normalized both vertically and relative to the origin of the energy axis (i.e. $\hbar\omega = 0$). Typical asymmetry in the high-energy tail is present in both cases.

5.3 Recombination of a free electron with a neutral acceptor (e–A^0) and of a free hole with a neutral donor (h–D^0)

An impurity atom in a semiconductor, whose valence is smaller (larger) by one electron than that of the main constituent of the crystal lattice, is referred to as a shallow acceptor (donor). Classic examples of a shallow acceptor (donor) are boron (phosphorus) atoms in silicon. The additional hole or electron arising from the replacement of the main constituent's atom in its position in the crystal lattice by the impurity atom is weakly bound and can be thermally excited into the valence (conduction) band, thus boosting the electrical conductivity of the material. The corresponding amount of energy needed for the release of the hole or electron is referred to as the *ionization (binding) energy* of the acceptor E_A or donor E_D. The acceptor and donor states are usually represented in the band structure diagram by short segments drawn inside the bandgap, just above the maximum of the valence band or below the minimum of the conduction band. As the definition implies, the difference in energy between the acceptor (donor) level and the extreme of the valence band equals E_A (or E_D in the case of the conduction band). A neutral acceptor (donor) is a term used to describe an acceptor (donor) with a non-ionized hole (electron); in the opposite case, the impurity will be referred to as an ionized acceptor or donor. The short segment in the band structure demonstrates the fact that the hole or the electron is localized at the neutral acceptor or donor and their quasi-momentum $\hbar\mathbf{k}$ is practically zero. This situation is depicted in Fig. 5.7 for an acceptor and a donor in a direct-bandgap semiconductor.

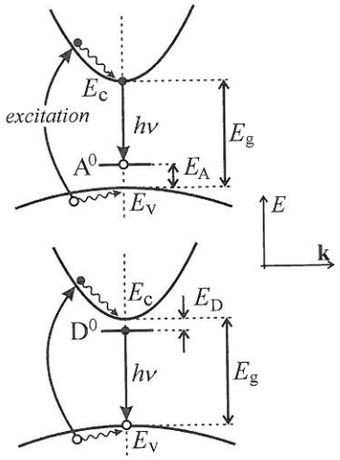

Fig. 5.7
(e – A^0) recombination (top panel) and (h – D^0) recombination (bottom panel) in a direct-bandgap semiconductor.

Figure 5.7 implies that the presence of a neutral acceptor or donor in the semiconductor can give rise to specific luminescence. Hereafter, we will focus only on acceptors (Fig. 5.7(a)). If a free electron is present at the bottom of the conduction band, it can recombine with a hole localized on the acceptor level, which leads to the emission of a luminescence photon with approximate energy $h\nu \approx E_g{-}E_A$. The luminescence process would not proceed further, however, if it were not for excitation, either optical or by employing other means, resulting in the liberation of other free electron–hole pairs, as is also demonstrated in Fig. 5.7. Excited electrons and holes quickly thermalize and 'diffuse' close to the extremes of the bands. The electron remaining after the radiative recombination at the acceptor level recombines (usually non-radiatively) with a hole at the top of the valence band, which restores the neutral state of the acceptor and the whole process of the emission of a luminescence photon can happen again.

Next, the spectral shape of luminescence (e–A^0) in a direct-bandgap semiconductor will be described. Let us start with (5.9a); this relation can be adapted so that it explicitly contains the effective masses m_e and m_h:

$$I_{sp}(h\nu) \approx (h\nu - E_g)^{1/2} \exp\left(-\frac{m_e}{m_e + m_h}\frac{h\nu - E_g}{k_B T}\right) \exp\left(-\frac{m_h}{m_e + m_h}\frac{h\nu - E_g}{k_B T}\right). \tag{5.14}$$

Assuming a high hole effective mass $m_h \to \infty$ and thus $m_h \gg m_e$ immediately yields, by making use of (5.14),

$$I_{sp}(h\nu) \approx (h\nu - E_g)^{1/2} \exp\left(-\frac{h\nu - E_g}{k_B T}\right). \tag{5.15}$$

We get back to a formula identical to (5.9a) and the whole procedure seems pointless. In reality, however, the transition from (5.14) to (5.15) only once again stresses the fact that optical transitions across the bandgap do not carry information regarding the values of effective masses m_e, m_h themselves. Importantly now, in the case of recombination of the type (e–A^0) the hole bound to an acceptor can be viewed as an immobile quasi-particle with an effective mass $m_h \to \infty$, which immediately suggests that the lineshape will again correspond to (5.9a), or (5.15), except for the fact that the low-energy onset is determined by (E_g–E_A) instead of E_g, which can also be easily deduced from Fig. 5.7(a). Consequently, the formula will read

$$I_{sp}^{(e-A^0)} \approx \left[h\nu - (E_g - E_A)\right]^{1/2} \exp\left(-\frac{h\nu - (E_g - E_A)}{k_B T}\right). \tag{5.16a}$$

Since similar reasoning (but for the substitution of electrons with holes) applies to the recombination of a free hole with a neutral donor (h–D^0), one can instantaneously write the formula for the corresponding lineshape

$$I_{sp}^{(h-D^0)} \approx \left[h\nu - (E_g - E_D)\right]^{1/2} \exp\left(-\frac{h\nu - (E_g - E_D)}{k_B T}\right). \tag{5.16b}$$

All in all, the Maxwell–Boltzmann lineshape (5.9a), (5.16) (or the analogous formula (5.13)) is shared by all luminescence processes for which at least one of the recombining quasi-particles is free. This similarity in behaviour reflects the *kinetic energy* of a free quasi-particle, which manifests itself in the asymmetry of the spectral line and the broadening of the high-energy tail with increasing temperature.[4]

An example of an experimentally observed luminescence spectrum (e–A^0) coming from a high-purity epitaxial GaAs layer at $T = 1.9\,\mathrm{K}$ is shown in Fig. 5.8 [5]. Both aspects of the experiment, i.e. the high purity and low temperature, are essential for this observation. If the semiconductor is not highly pure (high purity for GaAs stands for the concentration of unintentional (residual) acceptors and donors of the order less than $10^{15}\,\mathrm{cm}^{-3}$), several different acceptors and donors with slightly varying ionization energies are present in the sample, their emission lines thus being, according to (5.16), slightly shifted with regard to one another. This may result in the occurrence of two inhomogeneously broadened lines, one representing the unresolved acceptor transitions and the second one unresolved donor transitions, with the typical asymmetry given by (5.16) smeared and consequently lost. The sample whose spectrum is

[4] From the point of view of the configurational coordinate model, the processes (e–A^0) and (h–D^0) can be viewed as weakly localized excitation (i.e. only the zero-phonon line is observed), and broadening of the line due to kinetic energy of the quasi-particles taking part in the process goes beyond the scope of this model.

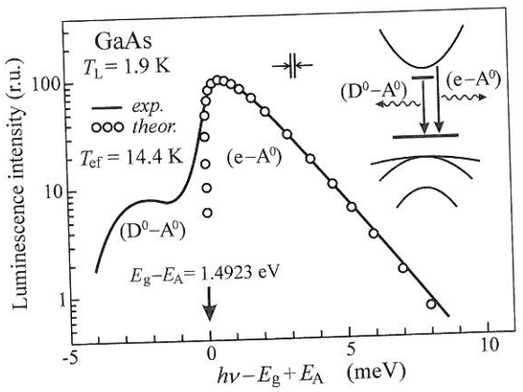

Fig. 5.8
Photoluminescence spectrum (e–A^0) of high-purity GaAs containing a single residual shallow acceptor with ionization energy $E_A \approx 27$ meV. The solid curve denotes the experimental data, and circles stand for the theoretical lineshape (5.16a). The low-energy edge reveals the occurrence of donor–acceptor pair emission (D^0–A^0), which is, however, more than an order of magnitude lower in intensity (note the logarithmic intensity scale). The bath temperature of liquid helium was $T_L = 1.9$ K, and the effective temperature of the electron gas for the fit of eqn (5.16a) was $T_{ef} = 14.4$ K. After Ulbrich [5].

shown in Fig. 5.8 contains a single type of acceptor only, which is why the typical asymmetry of the experimentally obtained curve towards the high-energy wing is well evidenced. Yet, at the low-energy edge near $h\nu \cong E_g$–E_A another emission band (D^0–A^0), which will be treated in Section 5.4, clearly stands out. The second experimental aspect, i.e. low temperatures, plays an important role not only due to the effect of thermal quenching, which has already been mentioned several times, but even more importantly to prevent the ionization of donors by the thermal energy $k_B T$.

Figure 5.8 contains our very first notion of two different temperatures in semiconductors: the lattice temperature (T_L) and the so-called *effective temperature* of a gas of free quasi-particles T_{ef}, introduced in photoluminescence measurements. The concept of the effective temperature indicates that, although the energy distribution of photoexcited free electrons and holes (or excitons) can be described by the Fermi–Dirac distribution function or by its Boltzmann tail, the temperature T_{ef} in the relevant formula is higher than the bath temperature T_L at which the sample is kept (liquid helium, nitrogen, etc.). In other words, even if the free quasi-particles are in thermal equilibrium with one another, they are not in thermal equilibrium with the subsystem of atoms (ions) vibrating in the lattice. This effect was found to be quite general and occurs not only at high-fluence excitation with high-power laser pulses, but also at conventional low-fluence excitation levels. The different temperatures of the carriers and the bath are a natural outcome of the relatively short photocarrier or exciton lifetime, the term 'relatively short' meaning the fact that the lifetime of carriers and excitons is shorter than the relaxation time necessary for the establishment of full thermodynamic equilibrium between the two described subsystems, i.e. the gas of free quasi-particles and the crystal lattice.[5] Naturally, the difference between the temperatures T_{ef} and T_L increases if a high-power excitation source is used or when the energy difference between the exciting photons and the bandgap E_g rises. Ultrafast semiconductor spectroscopy [6] is the branch of optics dealing with, among other things, the study of the rate of dissipation of this 'hot carrier' excess

[5] The exception to the rule $T_{ef} > T_L$ can sometimes be encountered in indirect-bandgap semiconductors, in which the lifetime of excitons may be sufficiently long.

energy (i.e. the decrease of T_{ef} as a function of time after the excitation pulse has ended) after picosecond or femtosecond excitation.

The last remark concerns the practical importance of (e–A^0) luminescence: it enables the ionization energy E_A to be spectroscopically determined as the energy difference between the low-energy onset of the line (E_g–E_A) and the energy corresponding to the width of the forbidden gap E_g. The bandgap energy E_g, however, needs to be obtained with sufficient accuracy from another type of measurement, since luminescence measurements rarely allow for its determination (but for the case of band-to-band recombination). It is of the essence, moreover, that a single type of acceptor is present; the alternative would inevitably lead to the smearing of the lineshape. A characteristic feature confirming the proper identification of a line being due to (e–A^0) or (h–D^0) can, besides the Maxwell–Boltzmann lineshape, be the saturation tendency of the intensity dependence, as discussed in Section 3.5 for extrinsic luminescence.

Radiative recombination (e–A^0), (h–D^0) is difficult to observe in an indirect-bandgap semiconductor due to the negligible concentration of free electrons and holes with respect to the concentration of excitons.

5.4 Recombination of donor–acceptor pairs (D^0–A^0)

If both neutral donor and acceptors are present in a semiconductor simultaneously, another type of radiative recombination might occur, as depicted in Fig. 5.9. In the approximation of effective mass, the neutral donor D^0 can be considered as an 'extra' electron orbiting all the remaining parts of the atom, i.e. the atomic nucleus, the inner atomic shells and electrons taking part in the covalent bonding with the nearest constituents of the crystal lattice. A neutral acceptor A^0 can then be treated analogously, if the extra electron is substituted with a hole. This electron–hole pair can recombine radiatively while emitting a photon $h\nu$ (Fig. 5.9(a)). The very same process is sketched in Fig. 5.9(b) in the band structure diagram, although the (E, x) system of coordinates, x being the real space coordinate, is used instead of the conventional $(E, \mathbf{k})$ layout.

Immediately after the recombination, both the donor and the acceptor become ionized (D^+, A^-), but free electron–hole pairs generated by the band-

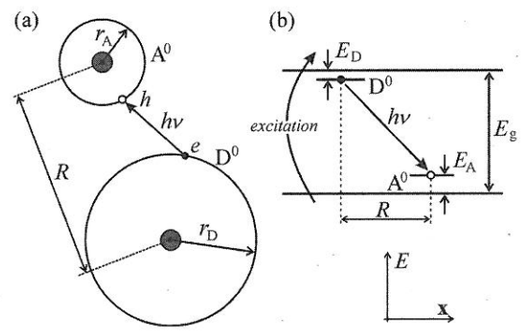

Fig. 5.9
The concept of luminescence arising from the recombination of donor–acceptor pairs (D^0–A^0): (a) in real space, (b) in the electron energy band structure as a function of the real-space coordinate x. Donor and acceptor atoms are located at different positions of the crystal lattice, separated by the distance R. The quantities r_D and r_A stand for the radius of the donor and acceptor, and E_D and E_A are their ionization energies, respectively.

to-band excitation quickly neutralize these ionized donors and acceptors and the whole process resumes.

This process is sometimes schematically described as

$$D^0 + A^0 \rightarrow D^+ + A^- + h\nu.$$

Glancing at Fig. 5.9(b) one could deduce that the energy of the emitted photon is determined as

$$h\nu(D^0 - A^0) = E_g - (E_A + E_D). \tag{5.17}$$

Nevertheless, this formula is not fully valid. Since a charged pair (D^+, A^-) is generated as a result of the recombination process, the additional energy of its mutual electrostatic attraction $e^2/4\pi\varepsilon_0\varepsilon_r R$ needs to be added to (5.17):

$$h\nu(D^0 - A^0) = E_g - (E_A + E_D) + \frac{e^2}{4\pi\varepsilon_r\varepsilon_0 R}(-J(R)). \tag{5.18}$$

In this formula, R denotes the distance between the donor and the acceptor and r_D stands for the radius of the electron shell in D^0 (and, analogously, r_A for the hole in A^0), see also Fig. 5.9(a). Formula (5.18) holds only provided the donor and the acceptor represent point charges, i.e. $R \gg r_D, r_A$. The last term $J(R)$ in (5.18) constitutes a non-Coulomb term, correcting for the overlap of the wavefunctions in the case of small distance R and decreasing the additional Coulomb energy. This term disappears when R is sufficiently large.

The (D^0–A^0) recombination manifests itself in the emission spectrum in various ways, depending on the strength of the electron–phonon coupling. First we shall discuss semiconductors with weak electron–phonon coupling (which means, according to Table 4.2, the fourth group and III-V group semiconductors, taking into account the polaron coupling constant α_e as a measure of this interaction). Since the donors and acceptors replace the main constituents of the crystal lattice in their sites, the values of R can only be discrete and, thus, so will the energies of the emitted photons, according to (5.18). When the electron–phonon interaction is weak, one can consequently expect the emission spectrum to comprise a number of narrow (zero-phonon) lines corresponding to the allowed values of R, or, in other words, to virtual spherical 'shells' of partners for radiative recombination around the particular donor or acceptor.

A prime example of such an emission spectrum is shown in Fig. 5.10 [7], in which one immediately recognizes an overwhelming number of narrow lines. The depicted luminescence spectrum is that of silicon- and sulphur-doped gallium phosphide, GaP, at liquid helium temperature (1.6 K). In this case, silicon and sulphur are the so-called substitution donors and acceptors of type I, which substitute atoms only in one of the sublattices of the crystal, in this case phosphorus: a phosphorus atom is replaced with a sulphur (donor S_P) or silicon (acceptor Si_P) atom. Zinc and oxygen, on the other hand, is a pair that can serve as an example of substitution donors and acceptors of type II, which are built into both sublattices, i.e. Zn_{Ga}–O_P.

Both the density of lines and their intensity obey general rules. The lines located at the high-energy side of the spectrum are relatively sparse and of low intensity, and in the direction towards lower energies they become more

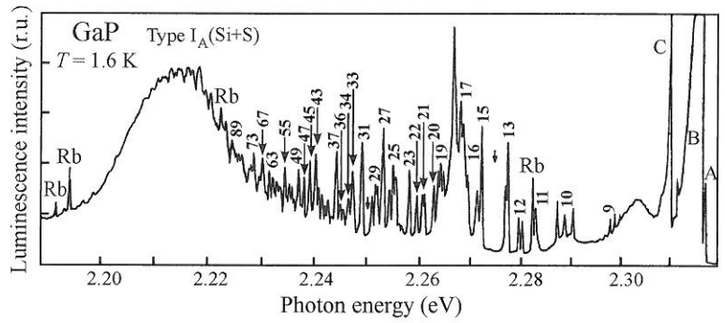

Fig. 5.10
Photoluminescence spectrum of silicon- and sulphur-doped GaP (the doping atoms are the so-called substitution impurities of type I, S_P donor and Si_P acceptor). The doping level is $|N_D - N_A| \sim 10^{18}$ cm^{-3}, and measurement temperature 1.6 K. The numbers next to the lines denote the number of the shell, lines marked with Rb come from a rubidium calibration lamp. Lines A, B, C are due to bound excitons. The whole spectrum is located between $E_g = 2.35$ eV and E_g–$(E_D + E_A) \approx 2.044$ eV. After Thomas *et al.* [7].

densely packed and their intensity increases, whereas at the very low-energy edge the intensity decreases again and the lines merge into a 'fine-structure-modulated' wider band. Such behaviour is easily understandable considering relation (5.18): high photon energies correspond to small R, the number of suitable neighbours in the nearest shell being low, though; the corresponding discrete values of R are relatively scarce, resulting in sparse and weak lines on the high-energy side of the spectrum. When the distance R increases, however, more partners can be found in the given shell, and the shells themselves are located closer to one another, which means that the lines become more intense and more densely packed. The situation changes once again when R becomes even larger, as the suitable shells nearly merge and a decrease in the probability of radiative recombination $W(R)$ comes into play owing to the fact that the electron and a hole are too far apart, for one can prove that

$$W(R) = W(0)\exp(-2R/a), \tag{5.19}$$

where $a = \max(r_D, r_A)$. This is why the low-energy end of the spectrum gets smeared into a broad band and drops to zero somewhere above the energy limit of E_g–$(E_A + E_D)$, see Fig. 5.10.

This type of emission spectrum contains so many characteristic features that the experimental identification of the (D^0–A^0) radiative channel should not pose any problems (even though some spectral similarity with the luminescence of an electron–hole liquid in a doped semiconductor can appear, see Chapter 8). Nevertheless, two additional characteristic features of this type of luminescence spectrum can be listed, the first one being an overall *blue-shift* of the broad band with increasing excitation intensity, as displayed in Fig. 5.11. This blue-shift is easily noticeable and can be qualitatively understood again on the basis of eqn (5.18): when increasing the excitation intensity, more D^0, A^0 atoms closer to one another, i.e. with lower R, get excited, which increases the influence of the Coulomb term in (5.18). The second attribute of the (D^0–A^0) luminescence spectrum is then typical behaviour in time-resolved luminescence experiments: the increasing delay after the excitation impulse gives rise to a pronounced *red-shift* of the emission spectrum. The reason for the behaviour lies in the fact that, according to (5.19), the closest pairs with the lower R are the first ones to recombine, while pairs further and further apart start to send out luminescence in longer time delays. This, in agreement with

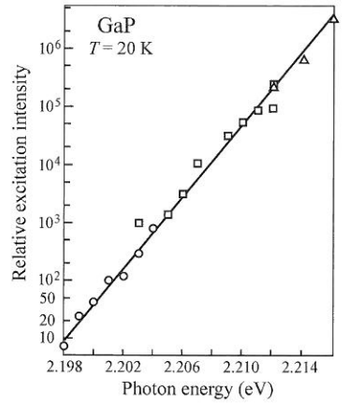

Fig. 5.11
Position of the maximum of the broad ($D^0 - A^0$) band from Fig. 5.10 as a function of excitation intensity. Bath temperature of 20 K. Different types of symbols denote different samples and different excitation wavelengths. After Thomas *et al.* [7].

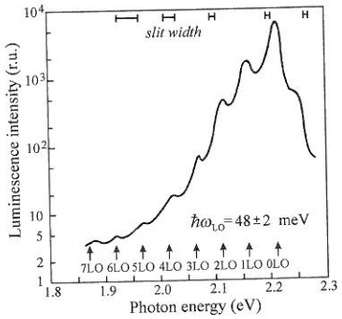

Fig. 5.12
$(D^0–A^0)$ luminescence spectrum of GaP at $T = 20$ K (the very same sample as in Figs 5.10 and 5.11). The high-energy tail at ~ 2.25 eV includes the unresolved lines of $D^0–A^0$ pairs (wide slit, higher temperature) and the peak at 2.20 eV corresponds to the broad band from Fig. 5.10. The structure is periodically repeated with period $\hbar\omega_{LO}$. Note the gradual widening of the monochromator slit for higher-energy phonon replicas and the logarithmic intensity scale. After Thomas *et al.* [7].

(5.18), signifies the shift of $h\nu(D^0–A^0)$ towards lower energies and evidently also the decrease of the overall emission intensity.

In the foregoing discussion we have put the $(D^0–A^0)$ radiative recombination in the framework of the configurational coordinate model as a special example of 'localized excitation' without the contribution of the kinetic energy of free carriers. The narrow spectral lines in Fig. 5.10 were referred to as 'zero-phonon' lines. The arguments that justify the application of such terminology are summarized in Fig. 5.12, which represents the emission spectrum of the very sample of Figs 5.10 and 5.11 measured, however, in a much wider spectral range and at slightly higher temperature (20 K). The spectrum contains an intense zero-phonon component near 2.20 eV and several satellites with decreasing intensities, shifted by multiples of the LO-phonon energy. Such behaviour is a typical manifestation of weak electron–phonon coupling. This time, the densely packed lines from Fig. 5.10 are hidden in the unresolved shoulder at 2.25 eV and the wide low-energy band originally located at 2.215 eV (Fig. 5.10) is shifted to 2.20 eV in Fig. 5.12, as a result of the thermally induced decrease in E_g.

Under what experimental conditions can the characteristic $(D^0–A^0)$ recombination luminescence with enormous numbers of narrow lines be observed in a semiconductor with weak electron–phonon coupling? The most important parameter will be a suitable level of concentration of both dopants. If the number of donors and acceptors is too low, D^0 and A^0 in the pairs will be too far apart from one another and no recombination will take place; on the other hand, when the concentration is too high, the simple Coulomb interaction of D^0 and A^0 ceases to be valid and, moreover, concentration quenching can set in. Therefore, the typical emission with densely packed lines can be observed in moderately or highly doped materials (e.g. in GaP at $|N_D–N_A| \sim 10^{17}–10^{18}$ cm^{-3}) and under slightly elevated excitation (otherwise, only donors and acceptors far apart from one another will be excited and the spectrum will feature only a band). Furthermore, low temperatures are absolutely essential and, if possible, the number of competing radiative recombination channels should be minimized. The importance of fulfilling all of these conditions becomes more obvious if we realize that GaP is, in fact, the only material in which a large number of narrow $(D^0–A^0)$ lines can be observed. For instance in silicon, whose electron–phonon coupling is even weaker than in GaP (and which, similarly to GaP, is an indirect-bandgap semiconductor), luminescence lines of this origin have never been experimentally observed. In GaAs, just to give another example, whose electron–phonon interaction is comparable to that in GaP, the $(D^0–A^0)$ emission manifests itself as a single unresolved line. There are, obviously, other factors coming into play; the problem will be further discussed in Section 7.2.

Now, let us briefly focus on semiconductors with strong electron–phonon coupling. A typical example of such a semiconductor is, taking into account the Fröhlich constant α_e in Table 4.2, direct-bandgap zinc sulphide ZnS with $E_g \sim 3.8$ eV. A particular case of doping by Cu_{Zn} acceptors and Cl_S donors leads to intense green emission, which has been well known for many years [8]. The most recent interpretation of this luminescence consists in the $D^0–A^0$ recombination. As the electron–phonon coupling in this material is particularly

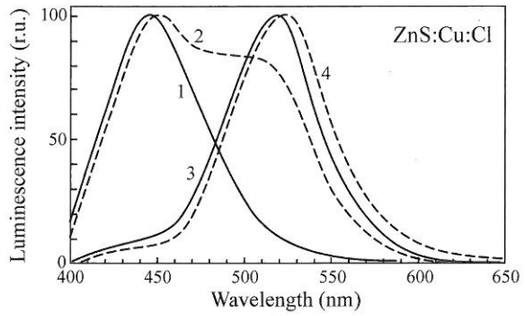

Fig. 5.13
Spectral distribution of cathodoluminescence of ZnS doped (activated) by copper with the addition of a 'fluxing agent' in the form of NaCl. This results in the luminescence of donor (Cl_S)–acceptor (Cu_{Zn}) pairs, giving rise to the emission band at ~ 520 nm. Curves enumerated 1–4 correspond to increasing concentration of copper: 1 – nominally pure ZnS, $2–10^{-3}$ % Cu, $3–5 \times 10^{-3}$ % Cu, $4–10^{-2}$ % Cu. Measured at room temperature, curves are normalized. After Espe [8].

strong, the luminescence band is broad, lacks any structure and is located far from edge emission (i.e. with large Stokes shift), near 520 nm (Fig. 5.13). The existence of other dopants, the introduction of which makes it possible to cover an important part of the visible spectral region, has been of great importance for applications in the development of phosphors for cathodoluminescence.

Here, it might be of interest to make a small historical detour: the fact that efficient luminescence in zinc sulphide can be achieved only if the material is co-doped with two different elements was empirically discovered as long ago as in the 1940s. Since the $D^0–A^0$ recombination mechanism was not yet known to be responsible for the luminescence, the metal dopant was not called an 'acceptor', but an 'activator', and the Cl^- or Br^- ions were not known as 'donors', but as 'co-activators'.

A similar type of broadband $D^0–A^0$ emission can also be encountered in ZnSe and CdS and some attempts were also made to interpret the intense low-temperature blue-green luminescence of AgCl by means of the same mechanism. From the application perspective, $D^0–A^0$ recombination is exploited as a mechanism of efficient emission from GaP light-emitting diodes. One of the proposed models describes the red electroluminescence at 1.8 eV as being due to the recombination of a compact $O_P–Zn_{Ga}$ pair in highly doped material.

5.5 Luminescence excited by two-photon absorption

In the text so far, we have taken for granted that the excitation of the photoluminescence processes takes place via the conventional absorption of a single (suitable) photon with energy $h\nu_{ex} > E_g$ which leads to the generation of electron–hole pairs or excitons and is followed by radiative recombination via one of the above-mentioned mechanisms. This kind of absorption is governed by the well-known Lambert–Beer law (2.1)

$$I_{h\nu_{ex}}(d) = I_0 \exp\left(-\alpha\left(h\nu_{ex}\right) d\right), \tag{5.20}$$

where $\alpha(h\nu_{ex})$ is the absorption coefficient and d stands for the depth under the surface of the sample to which the light penetrates. Unless the experiment has some specific requirements, the usual energy of the excitation photons

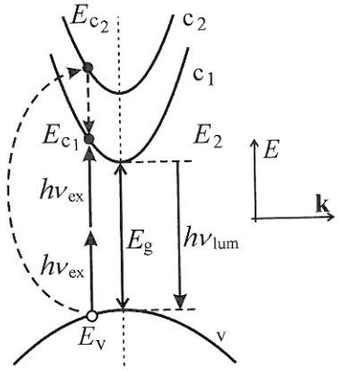

Fig. 5.14
Schematic illustration of two-photon absorption in semiconductors. The process is mediated by the presence of a higher lying conduction band c_2, but the energy-conservation law reads $E_{c1} - E_v = 2h\nu_{ex}$.

substantially exceeds E_g, which guarantees high values of the absorption coefficient and in turn boosts the luminescence intensity.

The typical values of absorption coefficient for these cases are in the order of $\alpha(h\nu_{ex}) \approx 10^4$–$10^5\,\mathrm{cm}^{-1}$, which allows the penetration of light only into depths of the order of $d \approx \alpha^{-1} = 10^{-5}$–$10^{-4}\,\mathrm{cm} = 100\,\mathrm{nm}$–$1\,\mu\mathrm{m}$. Such low values can pose a problem, because radiative recombination can be strongly influenced by surface and subsurface states, the existence of which is well known in solid-state physics. These states can either quench the luminescence, or, if the semiconductor is prone to oxidation, the surface oxide layer (or a surface layer of adsorbed molecules or radicals) can be the source of luminescence. Under such conditions, the obtained spectra *might not* in fact reflect the properties of the studied material—or, at least, an argument of this kind can be raised by any reviewer of a thesis or manuscript dealing with luminescence of semiconductors.

A simple but reliable way to confirm or refute such an argument may consist in comparing emission spectra excited via a one-photon and two-photon process. In order to understand why, it is important to give an explanation of two-photon absorption first.

This is a nonlinear optical phenomenon, in which two photons are absorbed at the same time *in a single elemental act*. The special case of absorption of two identical photons $h\nu_{ex}$ in a semiconductor is schematically represented in Fig. 5.14. A condition to be fulfilled if this process is to occur reads

$$E_g > h\nu_{ex} \geq E_g/2. \tag{5.21}$$

The right part of the inequality can be easily understood, as it reflects the energy-conservation law. The left part of eqn (5.21) results from the nonlinear nature of this process, since two-photon absorption is described in second-order perturbation theory [9] of quantum mechanics, implying that this process is several orders of magnitude weaker than the conventional one-photon absorption (first-order perturbation theory). Thus, were the condition $h\nu_{ex} > E_g$ fulfilled, intense single-photon absorption over the bandgap would set in and two-photon absorption would cease to be detectable.

The formalism of second-order perturbation theory implicitly comprises the so-called 'virtual intermediate level', which mediates the transition between the ground (E_v) and final (E_{c1}) states. The role of the virtual intermediate state in a semiconductor is taken on for example by one of the higher lying conduction bands (c_2 in Fig. 5.14). The energy-conservation law, though, is a relation between the ground and final states: $2h\nu_{ex} = E_{c1}$–E_v.

Now, we will derive a relation describing the decrease in light intensity as it passes through a semiconductor in the case when eqn (5.21) holds; in other words, we will derive a two-photon analogue of the Lambert–Beer law (5.20). Let us denote by $\alpha^{(2)}(I)$ the two-photon absorption coefficient, which evidently depends, unlike its one-photon counterpart, on the intensity of light I. It can be shown that this dependence is linear: $\alpha^{(2)}(I) = aI$ [9], where the coefficient a is a material constant. The decrease in the intensity of the light beam $\mathrm{d}I$ as a function of travel distance $\mathrm{d}x$ in a sample can then be described by the differential equation

$$\mathrm{d}I = -\alpha^{(2)}\, I\, \mathrm{d}x = -a\, I^2\, \mathrm{d}x. \tag{5.22}$$

This equation can be easily solved using separation of variables with the initial condition $I(x = 0) = I_0$. The result can be written as

$$I(d) = \frac{I_0}{(a\, I_0\, d + 1)}, \tag{5.23}$$

which is the formulation of the law we were looking for; d again stands for the depth of penetration under the surface of the sample, or its thickness. This relation was derived for the first time by Basov [10].

In contrast to the fast exponential decrease as in the case of linear one-photon absorption (5.20), the decrease of light intensity during two-photon absorption is only hyperbolic. This fundamental difference in the rate of decrease is documented in Fig. 5.15(a); the hyperbolic decrease is incomparably slower. In order to induce nonlinear optical phenomena, obviously, the application of a high-power source of light is necessary. Two-photon absorption can only be observed when relatively high-power laser beams are used as excitation sources. Taking into account the typical value of coefficient $a(\sim 10^{-3}\,\mathrm{cm/MW})$ and the power density of the incident beam $I_0 = 10\,\mathrm{MW/cm}^2$, one can immediately see that—according to (5.23)—a sample length of 1 cm(!) is needed so that the intensity could decrease by a single per cent (i.e. $I(d)/I_0) \approx 0.99$)! Two-photon excited luminescence coming from thick samples is then, without doubt, a phenomenon connected with the bulk of the material. (If semiconductor nanostructures with typical dimension of units and tens of nanometres are used as samples, however, the difference in the depth of penetration between one- and two-photon excitation is no longer so significant.)

An experimental verification of the described behaviour can be found in Fig. 5.15(b). The figure demonstrates a view through a cryostat window (dark circle) of a bulk AgCl sample at $T \cong 80\,\mathrm{K}$ (grey rectangle in the upper part of the window). The ruby laser beam ($h\nu_{ex} = 1.78\,\mathrm{eV}$) is coming from the left side and excites the luminescence in AgCl; the forbidden gap width $E_g \cong 3.2\,\mathrm{eV}$ of the silver chloride signifies that condition (5.21) is fulfilled. This luminescence can be observed as a bright trace coming through the sample in its middle part along its whole length, i.e. about 1.5 cm. Its brightness only minutely decreases towards the right side of the sample, reflecting a very modest decrease of laser intensity during its passage through the sample.

Which brings us back to the initial question: 'What does a two-photon excited, i.e. inherently coming from the bulk, luminescence spectrum look like in comparison with one-photon excitation?' Although a general and unambiguous answer to this question does not exist, a large number of experiments carried out at different laboratories (including that of the authors of this book) suggest that in the vast majority of cases the emission spectra are almost identical, at least when it comes to the salient features. This similarity is documented in Fig. 5.16, in which emission spectra of (a) a bulk AgCl sample at low temperatures [11] and (b) a thin CdS film (thickness of 1.6 μm) at room temperature [12] are compared. Similar comparisons were obtained using a large number of other semiconductors under diverse experimental conditions,

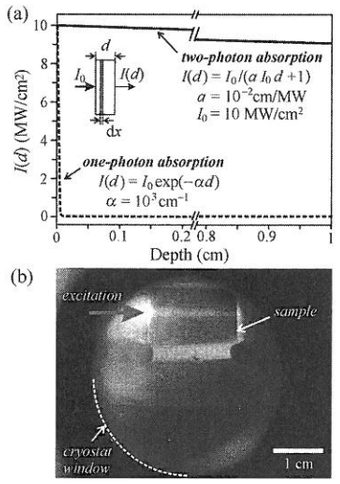

Fig. 5.15
(a) Numerical comparison of the decrease in intensity of a light beam as a function of the penetration depth during one- and two-photon absorption. An unrealistically high value of $a = 10^{-2}\,\mathrm{cm/MW}$ was chosen in order to make the decrease of intensity during the two-photon absorption process observable to the naked eye. (b) Experimental demonstration: a view through the window of a cryostat (dark circle) of a bulk AgCl sample with a horizontal trace of two-photon excited blue-green luminescence across the whole length of the sample (grey rectangle with a light strip in the middle). The sample temperature was $T \cong 80\,\mathrm{K}$. The focused beam of the excitation ruby laser 694 nm is coming from the left side, and three green optical filters with high optical densities were applied when the photograph was taken in order to block the scattered laser light.

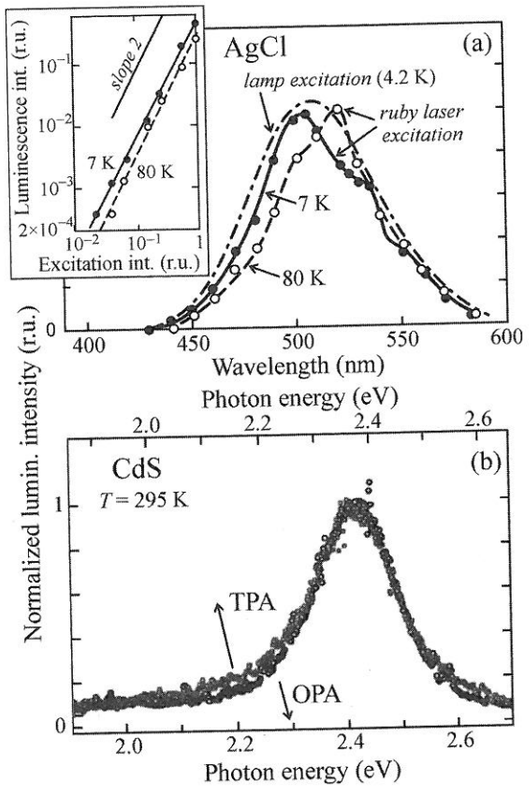

Fig. 5.16
Comparison of emission spectra acquired with one- and two-photon excitations. (a) AgCl single crystal, excited by one-photon absorption with a cw high-pressure lamp ($h\nu_{ex} = 3.4\,eV$) and by two-photon absorption with a pulsed ruby laser ($2h\nu_{ex} \cong 3.56\,eV$). After Laibowitz and Sack [11]. (b) Thin-film CdS excited at room temperature by one-photon absorption (OPA) with nanosecond pulses of the third harmonic of a Nd:YAG laser ($h\nu_{ex} \cong 3.49\,eV$) and by two-photon absorption (TPA) with 200-fs pulses of a Ti:sapphire laser ($2h\nu_{ex} \cong 3.23\,eV$). The spectra are shifted with respect to one another because the penetration depth of the femtosecond laser was larger than the thickness of the sample (1.58 μm); that is why the luminescence in this case is strongly influenced by reabsorption and consequently shifted to lower energies. After Yano and Ulbrich [12].

both in air and under vacuum. This indicates that the influence of surface states on luminescence is not so pronounced as was commonly supposed in the past. In some cases, some new features, e.g. phonon replicas, appear in the two-photon excited spectrum; these new features, however, are again more closely related to the properties of the material itself than to its surface. When in doubt, though, comparative experiments should be carried out on the material.

The last question needing clarification here is how to recognize undoubtedly that the observed luminescence is, indeed, two- and not one-photon excited. Two characteristic features can be applied, namely the quadratic intensity dependence and an 'anti-stokes' shift of the two-photon excited luminescence. The term 'anti-stokes shift' actually indicates that the spectral position of the emission spectrum violates Stokes' law, which states that the luminescence wavelength λ_{lum} is always longer than the excitation wavelength λ_{ex}. This law was formulated in the nineteenth century when one-photon excitation was the only conceivable way to excite luminescence (the very first laser was designed as late as in 1960); this mode of excitation, obviously, always leads to a Stokes shift. This shift clearly renders nothing less than the energy-conservation law $h\nu_{ex} \geq h\nu_{lum}$. During two-photon excitation, however, the

law reads $2h\nu_{ex} \geq h\nu_{lum}$ or $\lambda_{lum} \geq \lambda_{ex}/2$ and, in particular, the condition $\lambda_{ex} \geq \lambda_{lum} \geq \lambda_{ex}/2$ is very likely to occur; this happens very often indeed in semiconductors. Such behaviour is demonstrated in Fig. 5.14 depicting the excitation that takes place not too high above the bandgap, and luminescence photons have an energy very close to E_g, meaning that $\lambda_{lum} < \lambda_{ex}$. Thus, when looking at the photograph in Fig. 5.15(b), one should realize that the trace of luminescence across the sample is blue-green ($\lambda_{lum} \approx 500\,nm$), while the exciting laser beam is red ($\lambda_{ex} = 694\,nm$).[6]

The fact that two-photon excited luminescence depends quadratically on the excitation intensity can be derived from simple energy-based reasoning. The density of the excitation power absorbed in the sample of thickness d equals (if reflection losses are neglected), according to eqn (5.23),

$$\left(I_0 - \frac{I_0}{(ad\,I_0 + 1)}\right) = \frac{ad\,I_0^2}{(ad\,I_0 + 1)} \approx ad\,I_0^2; \qquad (5.24)$$

the last modification in (5.24) is justified by the fact that even when applying the highest power densities of the order of $I_0 = 10\,MW/cm^2$, still low enough not to destroy the material with the application of nanosecond laser pulses, and considering typical values of $a \leq 10^{-3}\,cm/MW$, the condition $adI_0 \ll 1$ easily holds. Making just the natural assumption, namely, that luminescence is proportional to the absorbed power, we see that luminescence is proportional to the squared incident laser intensity I_0^2 according to (5.24). This dependence is illustrated in the inset in Fig. 5.16(a). The reader surely recalls the case of bimolecular (one-photon excited) recombination which can, under certain circumstances, exhibit quadratic dependence on the excitation intensity, as was discussed in Section 3.5. Consequently, even though the experimental determination of such a dependence on its own does not prove that two-photon excitation takes place, the combination of this quadratic excitation-intensity dependence with an anti-stokes luminescence shift cannot be interpreted in any other way.

It is important to keep in mind that the realization of a two-photon experiment for a given value of bandgap is to some extent limited by the availability of high-power pulsed lasers complying with (5.21). Due to the growing numbers of tunable parametric generators this requirement is no longer as limiting as it used to be. The very last remark in this section relates to the fact that two-photon excitation can sometimes be applied to ensure excellent homogeneity of photoexcited carriers or excitons in order to compare experimental results with theoretical models, as will be mentioned later on in Chapter 8.

[6] To detect two-photon excited luminescence, the 'transmission geometry' cannot be applied. The direct laser light is many orders of magnitude more powerful than the luminescence and a direct detection of the beam could easily destroy the detector. The two-photon excited luminescence signal is rarely detected in a different way than at right angles to the excitation beam.

5.6 Luminescence from transition metal and rare earth ion impurities

Donors and acceptors are introduced in semiconductors as impurities in order to control their electrical transport properties. A number of impurities, however, are introduced into solids primarily in order to boost the luminescence efficiency and achieve a particular emission wavelength. Some of these impurities will be discussed in Chapter 7 (the so-called isoelectronic impurities); this section will deal with two types of specific impurities active in luminescence, which are exploited in countless commercially produced phosphors (e.g. for various types of reflecting surfaces, displays, luminescent signs, etc.): transition metal ions from the fourth period of the periodic table and rare earth ions.

These impurities, which cannot be easily classified from the point of view of their electrical activity as donors or acceptors, are active also in the luminescence of semiconductors. The corresponding channels of radiative recombination, however, will be mentioned only briefly, despite being both efficient and very commonly applied; they are not typical for semiconductors. They can be much more frequently encountered in various glasses, wide-bandgap ionic compounds, polymer-based materials, etc. Anyone more deeply interested in them can consult the chapters in the textbooks [13, 14].

Transition metal ions

Transition metal ions exhibit luminescence as a result of transitions of electrons between the levels of the partially filled 3d shell. Ions having the greatest importance in applications are Cr^{3+}, Mn^{4+} and Mn^{2+}. For example, $ZnS : Mn^{2+}$ has been a material commonly exploited for the preparation of electroluminescent thin film for quite some time; Cr^{3+} ions present in sapphire, Al_2O_3, were, on the other hand, the active medium for the very first ruby laser operating in 1960.

It is important to realize that $3d^{n*} \rightarrow 3d^n$ (* stands for an excited state) transitions are, from the atomic point of view, dipole-forbidden, since they do not comply with the condition $\Delta\ell = \pm 1$. When the ions are embedded into the crystal lattice, the interaction with the electric field of the crystal makes the transitions partially allowed. This interaction can be described in the framework of so-called ligand-field theory, which, among other things, shows that in some crystal lattices the emission spectra of the transition metal impurity ions can consist of both discrete lines (atomic-like emission) and broad bands. The crucial factor determining the spectral shape is the arrangement of the neighbouring ligands, the radius of the d orbitals and the distances between the built-in metal ion and the ligand. The shape of the spectrum can be predicted on the basis of the so-called Tanabe–Sugano diagram [15].

Recently, much effort has been devoted to the study of magnetic semiconductors, i.e. semiconductors with impurities of paramagnetic ions (Mn^{2+}), with the view towards applications in spintronics. Spintronics deals with the potential exploitation of spin polarization of charge carriers in informatics. Luminescence spectroscopy is presumed to be an excellent instrument for the study of these materials.

Rare earth ions

The most commonly applied rare earth ions are trivalent ions of lanthanides Ce^{3+}–Yb^{3+}, in which luminescence comes from $4f^{n*} \rightarrow 4f^n$ transitions. Similarly to the previous case, these transitions are also dipole-forbidden, while the crystal field makes them partially allowed. An important difference with respect to the previous case of $3d^{n*} \rightarrow 3d^n$ transitions, however, lies in the fact that the 4f shell is strongly screened from the influence of the crystal field by the external filled orbitals $5s^25p^6$. The influence of the crystal field is, on the one hand, strong enough to partially alter the character of the transitions, and on the other hand, the screening manifests itself by keeping the atomic-like character of the emission spectrum. In the framework of the configurational coordinate model in Fig. 4.8 this weak interaction with the lattice can be described as $\Delta Q = (Q_{e0} - Q_{g0}) \rightarrow 0$. This description implies two important consequences:

Firstly, the weak effect of the matrix suggests that the energy levels in ions, and thus also the emission (absorption) spectra, will be independent of the chemical nature and periodic arrangement of the matrix. Secondly, the activation energy E_A of the thermal quenching of luminescence (Fig. 4.8(b)) will be high, giving rise to more or less temperature-independent emission. Consequently, the emission can be expected to occur even at room temperature unless a different quenching mechanism, such as an Auger process (see Chapter 6), is present.

These aspects of rare-earth-doped semiconductor luminescence are nicely demonstrated in Fig. 5.17, depicting the luminescence of Er^{3+} ions built into various matrices [16, 17]. The emission is due to $4f^{11*} \rightarrow 4f^{11}$ transitions from the excited state $4I_{13/2}$ to the ground state $4I_{15/2}$. Comparison of the emission spectra in Fig. 5.17(a), (b) and (c) immediately confirms that the emission of these ions is represented by a relatively narrow line peaking approximately at 1530 nm, regardless of whether they are incorporated in a glass GeO_2–SiO_2 matrix or in crystalline or amorphous silicon. Comparison of the emission spectra in Fig. 5.17(c) and (d) further suggests that the intensity of luminescence of Er^{3+} in hydrogenated amorphous silicon decreases by only about 20% when the temperature increases from 2 K to 300 K. Recently, there has been considerable interest in the luminescence of Er^{3+} in silicon as the exploitation of such luminescence would allow for the operation of an optoelectronic light source on a silicon basis, a necessity for silicon optoelectronics or photonics. The emission wavelength near 1.53 μm is highly attractive because it lies in the so-called second transmission window of the glass fibres used in optical communications.

Apart from the trivalent form, rare earths can be built into the lattice also in the form of divalent ions. In the Eu^{2+} ion, for example, which is an important activator in a number of phosphors used in applications, luminescence occurs via dipole-allowed $5d \rightarrow 4f$ transitions. The reason underlying this difference against the case of Er^{3+} ions is as follows: the Eu^{2+} ground state is $4f^7$, and it can have for partners in optical transitions both the $4f^{7*}$ excited state and the $4f^65d$ excited state. The intensity of the crystal field then determines which one of these excited states lies lower in energy; the $4f^65d$ level turns out to be the lower one in most Eu^{2+}-ion-activated phosphors, resulting in

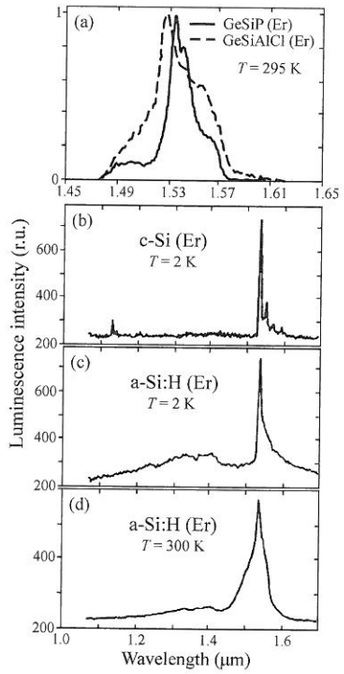

Fig. 5.17
Emission spectra of Er^{3+} ions built into various matrices and measured at different temperatures. (a) A glass GeO_2–SiO_2 matrix doped with Al_2O_3 (dashed curve) and P_2O_5 (solid curve) at $T = 295$ K. After Martucci *et al.* [16]. (b) Crystalline silicon matrix at $T = 2$ K. (c) Hydrogenated amorphous silicon (a-Si:H) matrix at $T = 2$ K. (d) a-Si:H matrix at $T = 300$ K. Except for the temperature, the experimental conditions and intensity scales are the same for (c) and (d). Graphs (b)–(d) after Bresler *et al.* [17].

$4f^6 5d \rightarrow 4f^7$ being the allowed transition underlying the luminescence in this case. One more difference can be found: the 4f shell is, similarly to trivalent ions, screened from the influence of the crystal field by the filled $5s^2 5p^6$ shell, but the excited 5d orbital is spatially located outside this shell and as such feels the interaction of the matrix much more strongly. This implies that, on the one hand, the emission spectrum can contain broad bands and, on the other hand, that its spectral position may vary with the chemical composition and spatial symmetry of the matrix.

For these reasons, Eu^{2+} ions provide a wide range of luminescence and absorption bands. The $CaF_2 : Eu^{2+}$ crystal can serve as an interesting example, since it was chosen as the very first material for the experimental demonstration of the occurrence of two-photon absorption in solids in 1961. The only laser then available was the ruby laser $\lambda = 694\,\text{nm}$ ($h\nu = 1.78\,\text{eV}$); the cubic $CaF_2 : Eu^{2+}$ crystal is transparent at this wavelength and at the same time has a strong absorption band $4f \rightarrow 5d$ at $2h\nu = 3.56\,\text{eV}$. In addition, $CaF_2 : Eu^{2+}$ was known to exhibit intense blue luminescence with a band around 420 nm ($\sim 2.95\,\text{eV}$), being due to radiative transitions after the relaxation of the excited electrons. The clever employment of these properties by Kaiser and Garrett [18] led to the actual detection of blue luminescence excited with red ruby laser pulses, with the intensity of this luminescence quadratically rising with the excitation intensity according to (5.24), see Fig. 5.18 (compare with the inset in Fig. 5.16(a)).

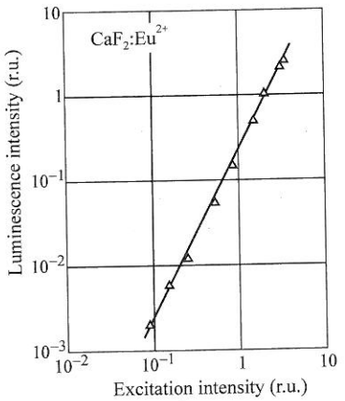

Fig. 5.18
Quadratic dependence of the intensity of blue luminescence on red laser excitation in $CaF_2 : Eu^{2+}$ in an original experiment demonstrating two-photon absorption in solids; $T = 295\,\text{K}$. After Kaiser and Garret [18].

5.7 Problems

5/1: Derive relations (5.7) and (5.8) by making use of Fig. 5.1. Employ the parabolic approximation $E_2 = E_c + \hbar^2 k_0^2 / 2 m_e$, $E_1 = E_v - \hbar^2 k_0^2 / 2 m_h$, along with $E_2 - E_1 = h\nu$.

5/2: Relation (5.9) for the luminescence lineshape of electron–hole pair recombination includes a parameter $D_0 \sim \exp(-E_g / k_B T)$, which increases with rising temperature. Does this mean that the intensity of this luminescence will also rise with temperature? Which important effect was neglected when deriving eqn (5.9)?

5/3: Prove (e.g. graphically) that the halfwidth of the Maxwell–Boltzmann lineshape (5.9) approximately equals $1.8\,k_B T$.

5/4: Calculate the integral $I = \int_0^a \sqrt{x}\,\sqrt{(a-x)}\,\text{d}x$, $a \geq 0$, in eqn (5.12). Hint: Apply integration by parts and then the substitution $x = a\,\sin^2 t$. (Result: $I = (\frac{2}{3} \int_0^{\pi/2} \sin^4 t\,\text{d}\,t) a^2$.)

5/5: Estimate what concentration of donor–acceptor pairs should occur in crystalline silicon so that the discrete D^0–A^0 lines appear in its low-temperature luminescence spectrum. The spectral position of such lines would be determined by the Coulomb term in (5.18). Using Fig. 1.1, find the corresponding spectral positions of these lines. Typical binding energies of donors and acceptors in silicon are $E_D(\text{phosphorus}) \cong E_A(\text{boron}) = 45\,\text{meV}$, and consider typical radii $r_A \cong r_D \cong 3\,\text{nm}$.

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. Kittel, C. (1976). *Introduction to Solid State Physics*, 5th edn. John Wiley, New York.
3. Saleh, B. E. A. and Teich, M. C. (1991). *Fundamentals of Photonics*. John Wiley, New York.
4. Mooradian, A. and Fan, H. Y. (1966). *Phys. Rev.*, **148**, 873.
5. Ulbrich, R. (1973). *Phys. Rev. B*, **8**, 5719.
6. Shah, J. (1999). *Ultrafast Spectroscopy of Semiconductors and Semiconductor Nanostructures*. Springer, Berlin.
7. Thomas, D. G., Gershenzon, M., and Trumbore, F. A. (1964). *Phys. Rev. A*, **133**, 269.
8. Espe, W. (1954). *Luminescent Substances in Electrical Engineering* (in Czech: *Luminiscenční látky v elektrotechnice. Výroba, vlastnosti a použití*). SNTL, Prague.
9. Yariv, A. (1967). *Quantum Electronics*. John Wiley, New York.
10. Basov, N. G. (1966). *International Conference on Physics of Semiconductors*, Kyoto 1966. *J. Phys. Soc. Japan*, **21**, Supplement, p. 277.
11. Laibowitz, R. L. and Sack, H. S. (1966). *phys. stat. sol.*, **17**, 353.
12. Yano, S. and Ulbrich, B. (2003). *Thin Solid Films*, **444**, 295.
13. Shionoya, S. (1998). *Photoluminescence*. In *Luminescence of Solids* (ed. D. R.Vij), p. 95. Plenum Press, New York.
14. Shionoya, S. and Yen, W. M. (eds.) (1999). *Phosphor Handbook*, p. 141. CRC Press, Boca Raton.
15. Tanabe, Y. and Sugano, S. (1954). *J. Phys. Soc. Japan*, **9**, 753.
16. Martucci, A., Guglielmi, M., Strohhöfer, C., Fick, J., Pelli, S., and Righini, G. C. (1999). *Germania based sol-gel waveguides doped with* Er^{3+}. In *Innovative Light Emitting Materials*. Advances in Science and Technology (ed. P. Vincenzini and G. C. Righini), Vol. 27, p. 197. Techna, Faenza.
17. Bresler, M. S., Gusev, O. B., Kudoyarova, V. Kh., Kuznetsov, A. N., Pak, P. E., Terukov, E. I., Yassievich, I. N., Zakharchenya, B. P., Fuhs, W., and Sturm, A. (1995). *Appl. Phys. Lett.*, **67**, 3599.
18. Kaiser, W. and Garrett, C. G. B. (1961). *Phys. Rev. Lett.*, **7**, 229.