Chapter 10
Recombination

Les hommes discutent, la nature agit.
Men argue, nature acts.

Voltaire

Abstract The various mechanisms and statistics of carrier recombination in semi-conductors including band-band, excitonic, band-impurity (Shockley–Read–Hall kinetics) and Auger recombination are explained. Also recombination at extended defects and surfaces is treated. Using the diffusion-recombination theory, the one-dimensional carrier profiles for typical situations in experiments and devices are derived.

10.1 Introduction

In thermodynamic nonequilibrium excess charges can be present in the semiconductor. They can be created by carrier injection through contacts, an electron beam or the absorption of light with wavelength smaller than the band gap. After the external excitation is turned off, the semiconductor will return to the equilibrium state. The relaxation of carriers into energetically lower states (and energy release) is called recombination. The term stems from the electron recombining with the hole created after absorption of a photon. However, there are other recombination mechanisms. A dedicated textbook is [820].

In the simplest picture an excitation generates carriers with a rate G (carriers per unit volume and unit time). In the steady state (after all turn-on effects) a constant excess charge n carrier density is present. Then the generation exactly compensates the recombination processes. The principle of detailed balance even says that each microscopic process is balanced by its reverse process. If the time constant of the latter is τ, n is given by $n = G\,\tau$. This follows from the steady-state solution $\dot{n} = 0$ of

$$\frac{\mathrm{d}n}{\mathrm{d}t} = G - \frac{n}{\tau}. \qquad (10.1)$$

In the literature two limiting cases have been discussed, the *relaxation* and the *lifetime* semiconductor, depending on the relation of two time constants. The one time constant τ_0 is the relaxation time constant due to recombination as discussed in the following. The smaller τ_0 is, the faster excited electrons and holes recombine and 'disappear'. Fast lifetimes are typically present in direct semiconductors (compared to indirect ones), semiconductors with high defect density and amorphous semiconductors. The other time constant is $\tau_D = \epsilon/\sigma$, the dielectric relaxation time; it describes the transport of carriers due to mobility (and diffusion). Large dielectric relaxation times are present in semiconductors with high mobility (low defect density, small carrier mass), small τ_D typically for hopping conduction. The relaxation case is given for $\tau_0 \ll \tau_D$; carriers will recombine quickly and it is hard to build up non-equilibrium carriers and separate them with an applied electric field. In the recombination case $\tau_D \ll \tau_0$, non-equilibrium carriers can assume non-uniform distributions and an applied electrical field generates separate quasi-Fermi levels for electrons and holes[1] (cmp. Sect. 7.6).

10.2 Band–Band Recombination

The band–band recombination is the relaxation from an electron in the conduction band into the valence (the empty state there is the hole). In a direct semiconductor, electrons can make an optical transition between the bottom of the conduction band to the top of the valence band. In an indirect semiconductor, this process is only possible with the assistance of a phonon and is thus much less probable.

10.2.1 Spontaneous Emission

We consider the spontaneous recombination of an electron of energy E_e and a hole of energy E_h (Fig. 10.1a). $C(E_e, E_h)$ is a constant proportional to the matrix element of the optical transition (cf. Sect. 9.5). The spontaneous recombination rate r_{sp} at photon energy $E \geq E_C - E_V = E_g$ is (assuming energy conservation, i.e. $E = E_e - E_h$, but without k-conservation in a dense plasma [821]),

$$r_{sp}(E) = \int_{E_C}^{\infty} dE_e \int_{-\infty}^{E_V} dE_h \, C(E_e, E_h)$$
$$\times D_e(E_e) \, f_e(E_e) \, D_h(E_h) \, f_h(E_h) \, \delta(E - E_e + E_h) \qquad (10.2)$$

---

[1]In the relaxation case, the separation of quasi-Fermi levels is $\ll kT$.
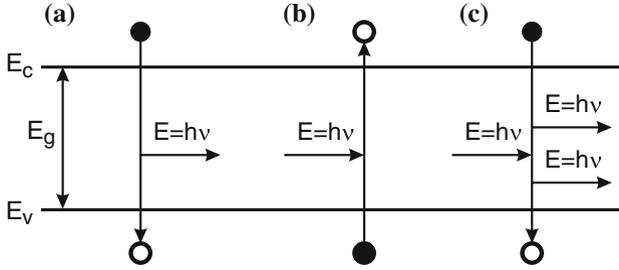
Fig. 10.1 Processes of band–band recombination: (**a**) spontaneous emission, (**b**) absorption and (**c**) stimulated emission. A *full* (*empty*) *circle* represents an occupied (unoccupied) electron state

$$= \int_{E_C}^{E+E_V} dE_e \, C(E_e, E_e - E)$$
$$\times D_e(E_e) \, f_e(E_e) \, D_h(E_e - E) \, f_h(E_e - E),$$

where f_h denotes the hole occupation $f_h = 1 - f_e$.

The lineshape of the band–band recombination with k-conservation[2] is proportional to the joint density of states (9.30) and the Fermi distribution function. At small excitation and at low doping it can be approximated by the Boltzmann distribution function and the lineshape is given as

$$I(E) \propto \sqrt{E - E_g} \, \exp\left(-\frac{E}{kT}\right). \tag{10.3}$$

An experimental spectrum is shown in Fig. 10.2 together with a fit according to (10.3). The expected FWHM of the peak is $1.7954 \, kT$, which is about 46 meV at $T = 300$ K. At low sample temperature, the temperature of the carrier gas is typically higher than the lattice temperature, depending on the cooling mechanisms (carrier–carrier scattering, optical phonon emission, acoustic phonon emission, recombination, ...) and the excitation rate. The carrier temperature in GaAs, determined from the Boltzmann tail of spontaneous emission (photoluminescence) is depicted in Fig. 10.3 as a function of excitation density; clearly it increases with increasing excitation.

The recombination rate in indirect semiconductors is small since the transition is phonon-assisted. For silicon, an internal quantum efficiency in the 10^{-6}-range has been reported [824]. For germanium, the direct transition is energetically fairly close to the fundamental, indirect L–Γ band edge transition (Fig. 9.12).The energy difference can be reduced from its bulk value of 136 meV by tensile strain. Additionally, the direct transition can be favored by heavily n-doping and filling the L conduction band minimum states (see Sect. 9.8.2). In this case, direct recombination from

---

[2]Excitonic effects are neglected here, e.g. for temperatures $kT \gg E_X^b$. Such effects are discussed in Sect. 10.3.
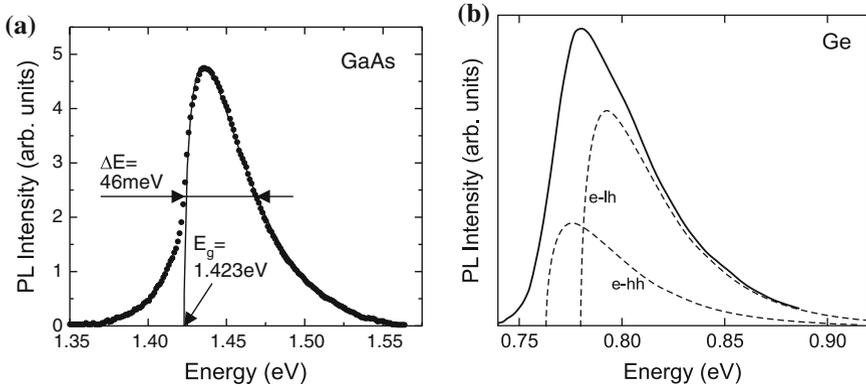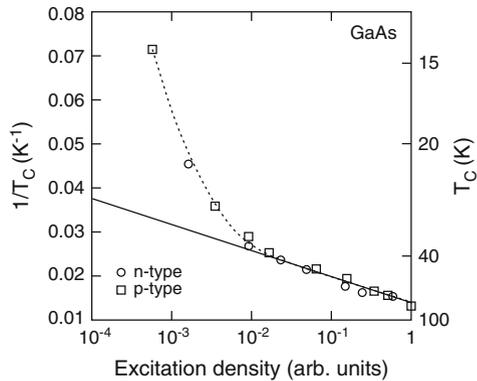
Fig. 10.2 (**a**) Photoluminescence spectrum of an undoped LPE-grown epitaxial GaAs layer at room temperature and low cw ($\lambda = 647$ nm) excitation density (10 W/cm²). The *solid line* is a lineshape fit with (10.3) and $E_g = 1.423$ eV and $T = 293$ K. (**b**) Room temperature, direct (e_Γ–h_Γ) recombination from heavily n-doped (10^{19} cm^{-3}) germanium (1 µm thick Ge layer on silicon (001)) with biaxial (thermal) tensile strain. The strain-split valence band edge (Fig. 6.43) causes the e–hh and e–lh transitions (individual contributions with lineshape according to (10.3) shown as *dashed lines*) to occur at different energies. Adapted from [822]

Fig. 10.3 Carrier temperature T_C in GaAs as a function of excitation density at a lattice temperature of 1.6 K. The *dashed line* is guide to the eyes, the *solid line* corresponds to an activation energy of 33 meV, similar to the GaAs optical phonon energy. Adapted from [823]

the conduction band Γ-minimum can be observed [822] and the effective energy difference has been lowered to about 100 meV.

10.2.2 *Absorption*

A similar consideration is made for the absorption process (Fig. 10.1b). An electron is transferred upon light absorption from a valence-band state (occupied) to a conduction-band state that must be empty. The coefficient is B_1. Also, the process

is proportional to the light intensity, represented by the density of occupied photon states $N_{ph}(E)$,

$$r_{abs}(E) = \int_{E_C}^{E+E_V} dE_e \; B_1(E_e, E_e - E)$$
$$\times D_e(E_e) \, (1 - f_e(E_e)) \, D_h(E_e - E) \, (1 - f_h(E_e - E)) \, N_{ph}(E). \quad (10.4)$$

10.2.3   Stimulated Emission

In this case, an incoming photon 'triggers' the transition of an electron in the conduction band into an empty state in the valence band. The emitted photon is in phase with the initial photon (Fig. 10.1c). The rate is (with coefficient B_2):

$$r_{st}(E) = \int_{E_C}^{E+E_V} dE_e \; B_2(E_e, E_e - E)$$
$$\times D_e(E_e) \, f_n(E_e) \, D_h(E_e - E) \, f_h(E_e - E) \, N_{ph}(E). \quad (10.5)$$

The photon density N_{ph} at a given energy is given by Planck's law and the Bose–Einstein distribution (Appendix E)

$$N_{ph}(E) = N_0 \, \frac{1}{\exp(E/kT) - 1}. \quad (10.6)$$

The pre-factor is the density of states of the electromagnetic field[3] $N_0(E) = 8\pi E^2 (n_r/hc)^3$.

10.2.4   Net Recombination Rate

In thermodynamical equilibrium the rates fulfill

$$r_{sp}(E) + r_{st}(E) = r_{abs}(E). \quad (10.7)$$

The population functions are Fermi functions with quasi-Fermi levels F_n and F_p. Since for absorption and stimulated emission the same quantum-mechanical matrix element is responsible, $B_1 = B_2$. The detailed balance (10.7) yields

$$C(E_1, E_2) = B_1(E_1, E_2) \, N_{ph} \left[\exp\left(\frac{E - (F_n - F_p)}{kT} \right) - 1 \right]. \quad (10.8)$$

---

[3]The total number of photon states in vacuum between the frequencies zero and ν is $N(\nu) = 8\pi\nu^3/(3c^3)$. With $\nu = E/h$ and $N_0 = dN(E)/dE$ and considering $c \to c/n_r$ we obtain the given value for N_0.

In thermodynamic equilibrium, i.e. $F_n = F_p$,

$$C(E_1, E_2) = N_0 B_1(E_1, E_2) = B. \tag{10.9}$$

If the constant B, the bimolecular recombination coefficient, is independent of the energy E, the integration for the net bimolecular recombination rate r_B can be executed analytically and we find

$$r_B = \int_{E_g}^{\infty} \left[r_{sp}(E) + r_{st}(E) - r_{abs}(E) \right] \mathrm{d}E$$
$$= B\, n\, p \left[1 - \exp\left(-\frac{F_n - F_p}{kT} \right) \right]. \tag{10.10}$$

In thermodynamic equilibrium, of course, $r_B = 0$. The recombination rate Bnp is then equal to the thermal generation rate G_{th}

$$G_{th} = B\, n_0\, p_0. \tag{10.11}$$

The bimolecular recombination rate typically used in Shockley–Read–Hall (SRH) [825, 826] kinetics is

$$r_B = B\, (n\, p - n_0\, p_0). \tag{10.12}$$

Values for the coefficient B are given in Table 10.1. In the case of carrier injection, np is larger than in thermodynamical equilibrium, i.e. $n\, p > n_0\, p_0$, and the recombination rate is positive, i.e. light is emitted. If the carrier density is smaller than in thermodynamical equilibrium, e.g. in a depletion region, absorption is larger than emission. This effect is also known as 'negative luminescence' [827] and plays a role particularly at elevated temperatures and in the infrared spectral region.

Table 10.1 Bimolecular recombination coefficient at room temperature for a number of semiconductors

Material	B (cm^3/s)
GaN	1.1×10^{-8}
GaAs	1.0×10^{-10}
AlAs	7.5×10^{-11}
InP	6.0×10^{-11}
InAs	2.1×10^{-11}
4H-SiC	1.5×10^{-12}
Si	1.1×10^{-14}
GaP	3.0×10^{-15}

Data for GaN from [828], Si from [829], SiC from [830], other values from [831]

10.2.5 Recombination Dynamics

The carrier densities n and p, are decomposed into the densities n_0 and p_0 in thermodynamic equilibrium and the excess-carrier densities δn and δp, respectively

$$n = n_0 + \delta n \tag{10.13a}$$

$$p = p_0 + \delta p. \tag{10.13b}$$

Here, only neutral excitations are considered, i.e. $\delta n = \delta p$. Obviously the time derivative fulfills $\frac{\partial n}{\partial t} = \frac{\partial \delta n}{\partial t}$, and correspondingly for the hole density. The equation for the dynamics

$$\dot{n} = \dot{p} = -Bnp + G_{\text{th}} = -B\,(n\,p - n_0\,p_0) = -B\,(n\,p - n_{\text{i}}^2) \tag{10.14}$$

can be written as

$$\frac{\partial \,\delta p}{\partial t} = -B\,(n_0\,\delta p + p_0\,\delta n + \delta n\,\delta p). \tag{10.15}$$

The general solution of (10.15) is given by

$$\delta p(t) = \frac{(n_0 + p_0)\,\delta p(0)}{[n_0 + p_0 + \delta p(0)]\,\exp\left[B\,t\,(n_0 + p_0)\right] - \delta p(0)}. \tag{10.16}$$

In the following, we discuss some approximate solutions of (10.15). First, we treat the case of a small (neutral) excitation, i.e. $\delta n = \delta p \ll n_0,\,p_0$. The dynamic equation is in this case

$$\frac{\partial \,\delta p}{\partial t} = -B\,(n_0 + p_0)\,\delta p. \tag{10.17}$$

Then the decay of the excess-carrier density is exponential with a time constant (lifetime) τ given by

$$\tau = \frac{1}{B\,(n_0 + p_0)}. \tag{10.18}$$

In an n-type semiconductor additionally $n_0 \gg p_0$, and thus the minority carrier lifetime τ_p is

$$\tau_p = \frac{1}{B \, n_0}. \tag{10.19}$$

If the nonequilibrium carrier densities are large, i.e. $n \approx p \gg n_0, p_0$, e.g. for strong injection, the kinetics obeys

$$\frac{\partial \delta p}{\partial t} = -B \, (\delta p)^2, \tag{10.20}$$

and the transient has the form

$$\delta p(t) = \frac{\delta p(0)}{1 + B \, t \, \delta p(0)}, \tag{10.21}$$

where $\delta p(0)$ is the excess hole density at time $t = 0$. Such a decay is called *hyperbolic* and the recombination is bimolecular. The exponential decay time is formally $\tau^{-1} = B \delta p(t)$ and is thus time and density dependent. A detailed discussion of minority carrier lifetime is given in [832].

10.2.6 Lasing

The net rate for stimulated emission and absorption is

$$r_{st}(E) - r_{abs}(E) = \left[1 - \exp \left(\frac{E - (F_n - F_p)}{kT} \right) \right]$$
$$\times \int_{E_C}^{E + E_V} dE_e \, B \, D_e(E_e) \, f_e(E_e) \, D_h(E_e - E) \, f_h(E_e - E) \, N_{ph}(E). \tag{10.22}$$

The net rate at photon energy $E = \hbar\omega$ is only larger than zero (i.e. dominating stimulated emission) when

$$F_n - F_p > E \geq E_g. \tag{10.23}$$

When the difference of the quasi-Fermi levels is larger than the band gap, the carrier population is inverted, i.e. close to the band edges the conduction-band states are more strongly populated with electrons than the valence-band states, as shown in Fig. 10.4. An incoming optical wave of energy E will then be net amplified by stimulated emission. Equation (10.23) is also called the thermodynamic laser condition. We note that lasing requires further conditions as discussed in Sect. 23.4.

Fig. 10.4 Charge-carrier distribution during inversion, necessary for lasing. *Shaded areas* are populated with electrons. A stimulated transition between an electron and a hole is indicated
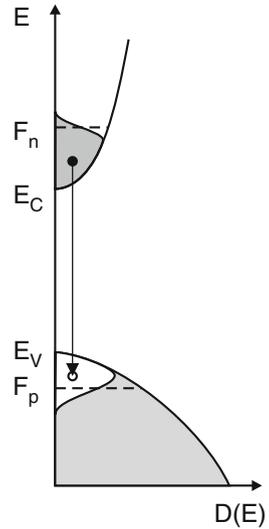
10.3   Exciton Recombination

10.3.1   Free Excitons

The observation of free-excitons is limited for semiconductors with a small exciton binding energies (such as in GaAs) to low temperatures. However, for large exciton binding energy, recombination from free-excitons is observed even at room temperature, as shown in Fig. 10.5 for ZnO.

A low temperature recombination spectrum of silicon is shown in Fig. 10.6. In pure silicon, phonon-assisted exciton recombination (cmp. Sect. 10.4) is observed involving acoustic (I^{TA}) and optical (I^{TO}) phonons. The weakly observed no-phonon line (I^0) is forbidden in perfect Si.

10.3.2   Bound Excitons

Excitons can localize at impurities, defects or other potential fluctuations and subsequently recombine [834, 835]. Excitons can be bound to neutral or ionized donors and acceptors impurities [836]. Also they can be bound to isoelectronic impurities, the most prominent example being N in GaP [837] (cmp. Sect. 9.6.9). or isoelectronic clusters [838]. The recombination of excitons localized in quantum wells (Sect. 12.4) and quantum dots (Sect. 14.3.4) is discussed later.
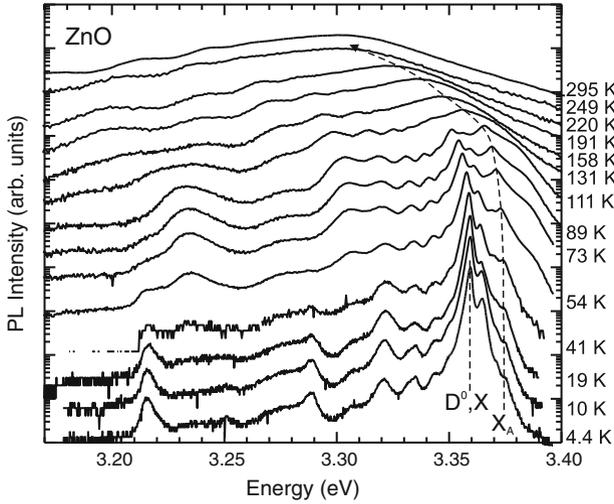
Fig. 10.5 Temperature-dependent luminescence spectra of a ZnO thin film (on sapphire). At low temperatures, the spectra are dominated by donor-bound exciton transitions (Al^0, X)). The *vertical dashed line* indicates the low-temperature position of the donor-bound exciton transition (D^0, X). The *curved dashed line* visualizes the energy position of the free-exciton transition (X_A) that becomes dominant at room temperature

Fig. 10.6 Low temperature recombination spectra from silicon with low (*solid lines*) and sizeable (*dashed line*) phosphorus dopant concentration. Spectrum for $N_P = 2 \times 10^{14}$ cm^{-3} ($N_P = 8 \times 10^{16}$ cm^{-3}) taken at 26 K (15 K). Transitions in pure Si are label with 'I', transitions involving P donors are labeled with 'P'. Q indicates the dissociation energy of the bound exciton. Adapted from [833]
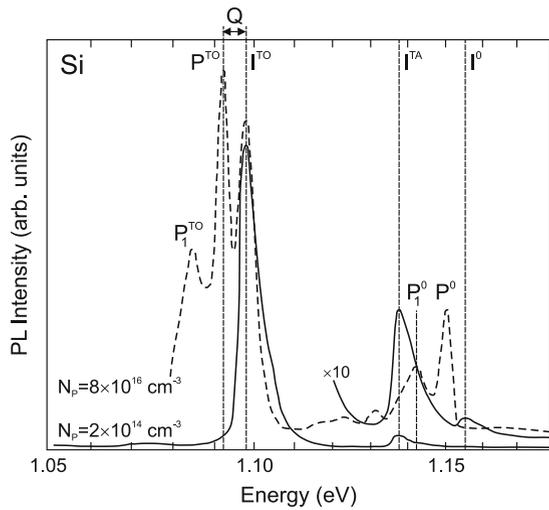
Table 10.2 Localization energy Q (Q^*) of excitons on selected impurities (ionized impurities, D^+ or A^-, respectively) in various semiconductors

Host	Donor	Q (meV)	Q^* (meV)	Q^*/Q	σ	Reference
GaAs	EMD	0.88	1.8	2.0	0.28	[840]
	Zn	8.1	31.1	3.8		[841]
GaN	EMD	6.8	11.2	1.6	0.36	[842]
	Mg	20				[843]
AlN	Si	16				[843]
	Mg	40				
CdS	EMD	6.6	3.8	0.6	0.17	[844]
ZnSe	Al	4.9	5.4	1.1	0.27	[845, 846]
	Ga	5.1	6.6	1.3		
	In	5.4	7.5	1.4		
ZnO	Al	15.5	3.4	0.21	0.3	[847]
	Ga	16.1	4.1	0.25		
	In	19.2	8.5	0.44		

σ is the ratio of effective electron and hole (polaron) masses. EMD: effective mass donor

The transition energy $\hbar\omega$ of an exciton bound to a neutral impurity is

$$\hbar\omega = E_\mathrm{g} - E_\mathrm{X}{}^\mathrm{b} - Q, \tag{10.24}$$

where Q is the binding (or localization) energy of the exciton to the impurity. The binding energy of an exciton to an ionized impurity is denoted with Q^*. A transition involving an exciton bound to a neutral donor is denoted (D^0, X); correspondingly (D^+, X), also denoted as (h, D^0), and (A^0, X). Values for donor-bound excitons in various semiconductors are listed in Table 10.2. The (D^0, X) complex is stable for $0 < \sigma = m_\mathrm{e}^*/m_\mathrm{h}^* < 0.43$ according to [839]. The (D^+, X) peak can occur on the low- or high-energy side of the (D^0, X) recombination. Whether $Q^* < Q$ or $Q^* > Q$ depends on σ being smaller or larger than 0.2, respectively [839], and is fulfilled for many semiconductors, e.g. GaAs, GaN, CdS, and ZnSe.

Recombination in silicon due to excitons involving phosphorus donors is depicted in Fig. 10.6. The (D^0, X) transition in Si:P is labeled 'P^0' ($Q = 6$ meV). Other P-related transitions are discussed in [833]. In Si, the binding energy to the impurity is about one tenth of the binding energy of the impurity (Haynes's rule [834, 848]), i.e. $Q/E_\mathrm{D}^\mathrm{b}$ and $Q/E_\mathrm{A}^\mathrm{b} \approx 0.1$ (Fig. 10.7a). In GaP the approximate relations $Q = 0.26E_\mathrm{D}^\mathrm{b} - 7$ meV and $Q = 0.056E_\mathrm{A}^\mathrm{b} + 3$ meV have been found [837]. For donors in ZnO, the relation $Q = 0.365E_\mathrm{D}^\mathrm{b} - 3.8$ meV holds (Fig. 10.7b) [849]. In Fig. 10.8, the recombination spectrum of GaAs:C is shown that exhibits recombination from excitons bound to the acceptor (carbon) and shallow donors. The exciton is more strongly bound to an ionized donor (D^+) than to a neutral donor.

Varying the concentration of a specific impurity and observing the corresponding change in the intensity of the (D^0, X) transition allows to identify the chemical species to which the exciton is bound. This can be achieved via the comparison of different samples or more elegantly by introducing radioactive isotopes. This is shown in
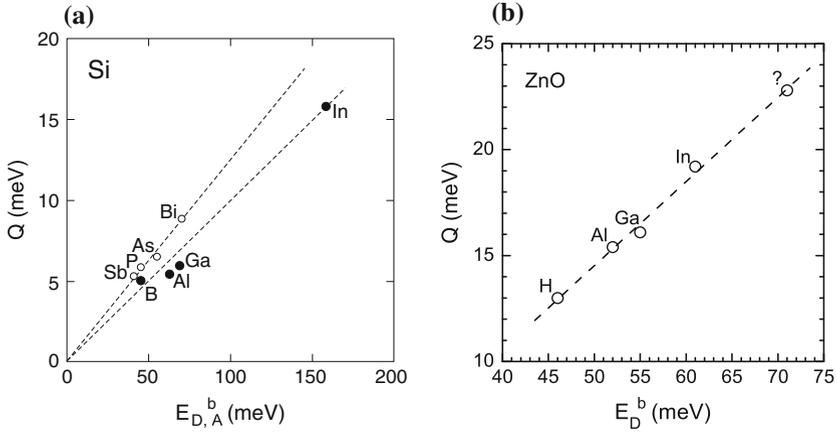
Fig. 10.7 Energy Q required to remove an exciton from a neutral impurity (10.24) as a function of the ionization energy E_D^b (*open circles*) or E_A^b (*solid circles*) of the involved impurity in (**a**) silicon (experimental data from [848]) and (**b**) ZnO (experimental data from [849]
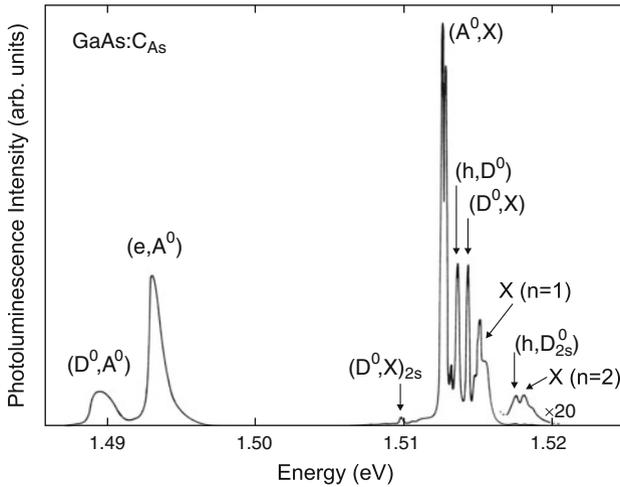

Fig. 10.8 Photoluminescence spectrum ($T = 2\,\mathrm{K}$, $D = 10\,\mathrm{mW\,cm^{-2}}$) of GaAs:$C_{As}$ ($N_A = 10^{14}\,\mathrm{cm^{-3}}$) with donor- and acceptor-related bound-exciton recombination around 1.512 eV, (e, A^0), (h, D^0) and (D^0, A^0) pair and free-exciton recombination. Adapted from [840]

Fig. 10.9 for In in ZnO; the $(^{111}In^0, X)$ transition disappears with the characteristic time constant close to that (97 h) of the nuclear decay of ^{111}In into stable ^{111}Cd. However, in such experiments it should be considered that the decay product and accompanying high-energy radiation can create new electronic and structural defects, respectively.
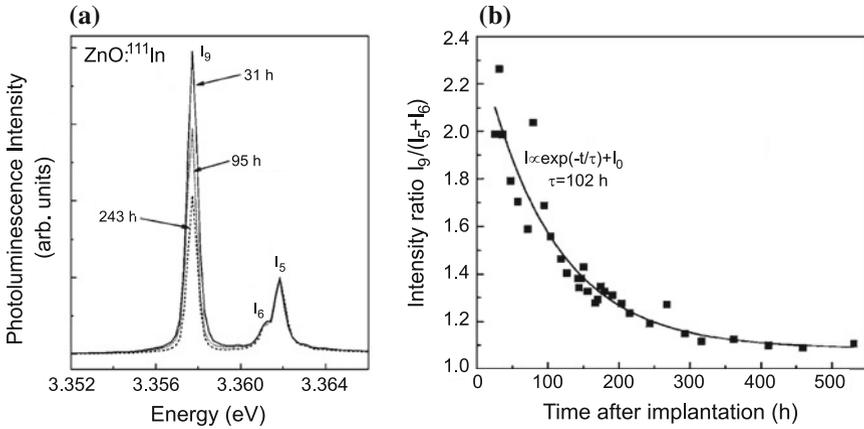
Fig. 10.9 (**a**) Low-temperature photoluminescence spectrum of ZnO implanted with ^{111}In featuring the so-called I$_9$-line. Spectra are recorded at various times after implantation as labeled. (**b**) Intensity of I$_9$-line as a function of time. Adapted from [850]
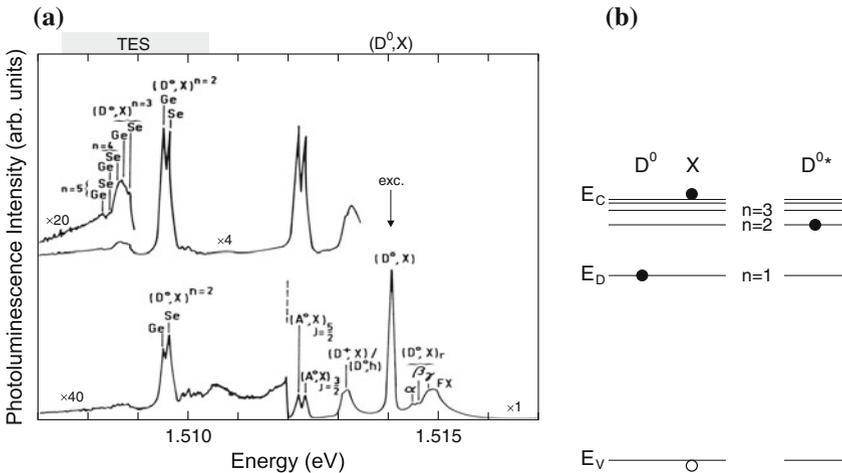

Fig. 10.10 (**a**) Photoluminescence spectrum ($T = 1.5$ K, $D = 50$ mW cm^{-2}) of high-purity GaAs with two donors (Ge and Se/Sn). The *lower* spectrum has been excited 6 meV above the band gap, the *upper* spectrum has been resonantly excited with the laser set to the (D^0, X) transition and exhibits $n = 2, 3, 4$, and 5 TES transitions. α, β, γ denote excited (hole rotational) states of the (D^0, X) complex. Adapted from [852]. (**b**) Schematic representation of the $n = 2$ TES process, *left* initial, *right* final state

The peak labeled (D^0,X)$_{2s}$ in Fig. 10.8 is called a two-electron satellite (TES) [851]. High-resolution spectra of the TES in GaAs [506, 852] are shown in Fig. 10.10a. The TES recombination is a (D^0,X) recombination that leaves the donor in an excited

state as schematically shown in Fig. 10.10b. Therefore a hydrogen-like series with $n = 2, 3, \ldots$ is observed with energies

$$E_{\text{TES}}^n = E_{(\text{D}^0,\text{X})} - E_{\text{D}}^{\text{b}} \left(1 - \frac{1}{n^2} \right). \tag{10.25}$$

The effect of isotope disorder on the sharpness and splitting of impurity states has been investigated in [853, 854]. The recombination of excitons bound to Al, Ga and In in natural silicon (92.23 % ^{28}Si, 4.67 % ^{29}Si, 3.10 % ^{30}Si) is split into three lines due to the valley-orbit splitting [855] of electron states at the band minimum (Fig. 10.11). Each of these (A^0, X) lines is split by $0.01\,\text{cm}^{-1}$ for Si:Al due to a symmetry reduction of the 4-fold degenerate A^0 ground state, as observed in the presence of applied axial strain or an electric field. The comparison to spectra from enriched ^{28}Si shows that the observed splitting without external perturbation is due to isotope disorder that causes random strains and splits the A^0 ground state into two doublets [854] (Fig. 10.11). Similarly, the (unsplit) phosphorus-induced (D^0, X) transition in enriched Si is found to be much sharper ($<40\,\mu\text{eV}$) than in natural Si ($330\,\mu\text{eV}$) [853]. At higher resolution, a hyperfine splitting of $485\,\text{neV}$ due to the ^{31}P nuclear spin $I = 1/2$ ($2 \times 10^{12}\,\text{cm}^{-3}$) in isotopically pure (99.991 %) ^{28}Si ($I = 0$) is observed for the (P^0, X) recombination [856]. In a magnetic field, the Zeeman-split lines have a FWHM of about $150\,\text{neV}$.

In Fig. 10.12 the recombination of excitons bound to the N isoelectronic impurity in lowly doped GaP is shown. The efficient recombination of nitrogen-bound electrons with holes at the Γ point is due to the wave-function component of the localized electron at $k = 0$ [612] (Fig. 7.41). The decay time of the A exciton is about $40\,\text{ns}$ [857] and thus larger than the typical lifetime of excitons in direct semiconductors (ns-range). The forbidden B exciton has a much longer lifetime of $4\,\mu\text{s}$ [857].

Fig. 10.11 High-resolution photoluminescence (PL) spectra of (A^0, X) recombination in natural and ^{28}Si-enriched silicon doped with aluminum ($T = 1.8\,\text{K}$). The ^{28}Si PL spectrum is shifted up in energy by $0.114\,\text{meV}$, as indicated by the *arrow*, to compensate for the shift in band gap. The *inset* shows a level scheme for the recombination in natural silicon. Adapted from [854], reprinted with permission, ©2002 APS
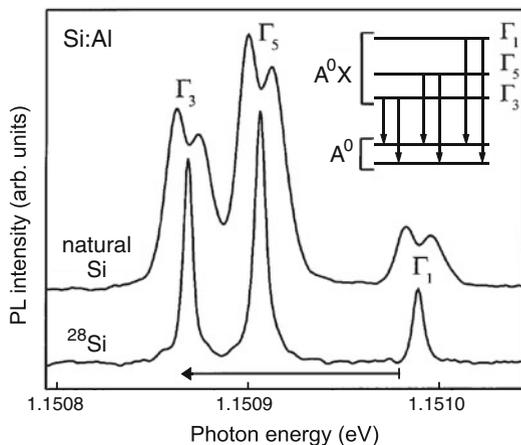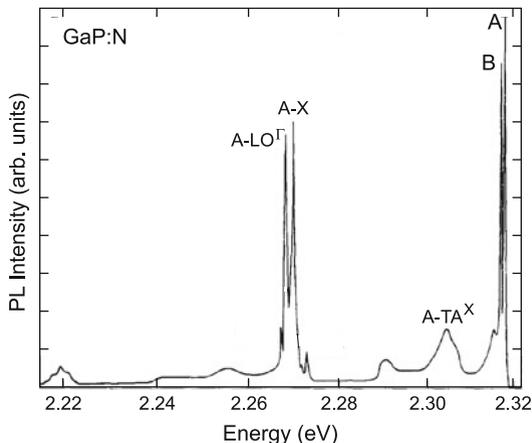
Fig. 10.12 Photoluminescence spectrum ($T = 4.2$ K) of GaP:N ($N_N \approx 5 \times 10^{16}$ cm^{-3}). The A exciton is bound to an isolated nitrogen impurity, cmp. to Fig. 9.25. Adapted from [612]

Fig. 10.13 Spectral position of neutral donor- and acceptor-bound exciton photoluminescence transition ($T = 2$ K) in GaAs doped with different amounts of indium relative to the donor-bound exciton luminescence in pure GaAs (1.5146 eV). Adapted from [858]
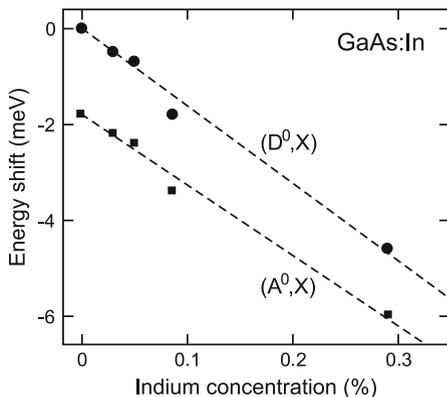
In the case of In in GaAs it has been found that down to the regime of $N_{In} <$ 10^{19} cm^{-3} the indium does not act as a substitutional isoelectronic impurity but still fully participates in the composition of a pseudo-binary system (Sect. 6.5). Recombination from excitons bound to single indium atoms or In–In pairs could not be found. The energy shift of donor- and acceptor-bound excitons in the dilute limit (Fig. 10.13) follows the band-gap dependence established for larger indium concentrations. The non-occurrence of localization effects is attributed to the small effective electron mass in InAs [472].

The luminescence intensity $I(T)$ of bound exciton lines is quenched with increasing temperature due to ionization of the excitons from the impurities. The temperature dependence can be modeled using the relation [859]

Fig. 10.14 Temperature dependent PL intensity of (D^0, X) in GaN and $(A0, X)$ in AlN:Mg recombination. *Solid lines* are fits with (10.26). Data from [843, 861]
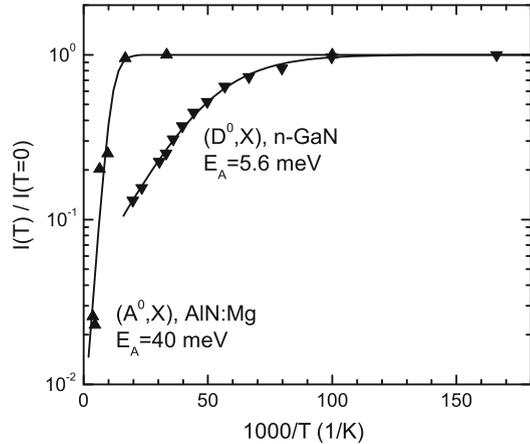
$$\frac{I(T)}{I(T=0)} = \frac{1}{1 + C \exp(-E_A/kT)}, \tag{10.26}$$

E_A being the thermal activation energy and C a pre-factor. Often the activation energy is found equal to the localization energy, $E_A = Q$ (Fig. 10.14, cmp. Table 10.2). If several processes contribute, additional exponential terms can be added with further activation energies. For acceptor-bound excitons in GaAs two processes are found to contribute, the ionization from the impurity into a free exciton ($E_A^1 \approx Q$) and into an electron-hole pair ($E_A^2 \approx Q + E_X^b$) [859]. In [860] the model is refined by considering the temperature dependence of the parameter C due to the ionization of the impurity itself.

So far *single* excitons bound to a center have been discussed. Also bound exciton *complexes* [862] containing up to six excitons have been observed at sufficient excitation density, e.g. for substitutional boron [863] or phosphorus [864] and interstitial Li [865] in silicon. In a multi-valley semiconductor several electrons are available to form bound excitons which follow approximately a shell model and exhibit further fine structure.

10.3.3   Alloy Broadening

The bound-exciton recombination peak in a binary compound is spectrally fairly sharp (Sect. 10.3.2), even in the presence of isotope disorder (Fig. 10.11). In an alloy (see Sect. 3.7), the random distribution of atoms (with different atomic order number Z) causes a significant broadening effect of the luminescence (and absorption) line, the so-called *alloy broadening* [866, 867]. As an example, we treat $Al_xGa_{1-x}As$. The

exciton samples, at different positions of the lattice, different coordinations of Ga and Al atoms. If the experiment averages over these configurations, an inhomogeneously broadened line is observed.

The cation concentration c_c for the zincblende lattice is given as $c_c = 4/a_0^3$, for the wurtzite lattice as $c_c = 4/(\sqrt{3}\,a^2c)$. For example, $c_c = 2.2 \times 10^{22}\,\mathrm{cm}^{-3}$ for $Al_xGa_{1-x}As$ in the entire composition range $0 \leq x \leq 1$ since the lattice constant does not vary significantly, and $c_c = 4.2 \times 10^{22}\,\mathrm{cm}^{-3}$ for wurtzite $Mg_xZn_{1-x}O$ [868]. In a random alloy, the probability $p(N)$ to find exactly N Ga atoms in a given volume V (with a total of $c_c V$ cations) is given by the binomial distribution

$$p(N) = \binom{c_c V}{N} x^N (1-x)^{c_c V - N}. \qquad (10.27)$$

The sampling volume for a luminescence event is the exciton volume (cf. (9.39)) that is given for the free-exciton (in 1s hydrogen state) as [866, 869]

$$V_{ex} = 10\,\pi\,a_X^3 = 10\,\pi \left(\frac{m_0}{m_r^*} \epsilon_s a_B \right)^3. \qquad (10.28)$$

One should note that due to the variation of the involved material parameters V_{ex} depends itself on x. In GaAs there are about 1.2×10^6 cations in the exciton volume. In $Al_xGa_{1-x}As$, there are on average $x c_c V_{ex}$ Al atoms in the exciton volume. The fluctuation is given by the standard deviation of the binomial distribution [869]

$$\sigma_x^2 = \frac{x\,(1-x)}{c_c\,V_{ex}}. \qquad (10.29)$$

The corresponding energetic broadening (full width at half-maximum) of the spectral line is given by $\Delta_E = 2.36\,\sigma$ with

$$\sigma = \frac{\partial E_g}{\partial x}\,\sigma_x = \frac{\partial E_g}{\partial x} \sqrt{\frac{x\,(1-x)}{c_c\,V_{ex}}}. \qquad (10.30)$$

We note that instead of the quantum mechanically correct factor $10\,\pi$ [866, 869], often the factor $4\pi/3$ [867] is used, resulting in larger theoretical broadening.

Experimental data for Cd_xSe_{1-x} in Fig. 10.15a are consistent with (10.30). The theoretical dependence (10.30) is shown in Fig. 10.15b also for $Al_xGa_{1-x}As$ together with experimental data and found to disagree [870]. Since the exciton volume is much smaller (cf. 9.6.6) than in $Al_xGa_{1-x}As$, alloy broadening in $Mg_xZn_{1-x}O$ is much larger for a given for given x.

The spectral broadening due to alloy disorder masks the fine structure of recombination lines near the band edge present for binary semiconductors. Often for all temperatures only a single recombination line appears for alloys. Spectra for three different $Mg_xZn_{1-x}O$ alloys are shown in Fig. 10.16a. The increasing inhomogeneous
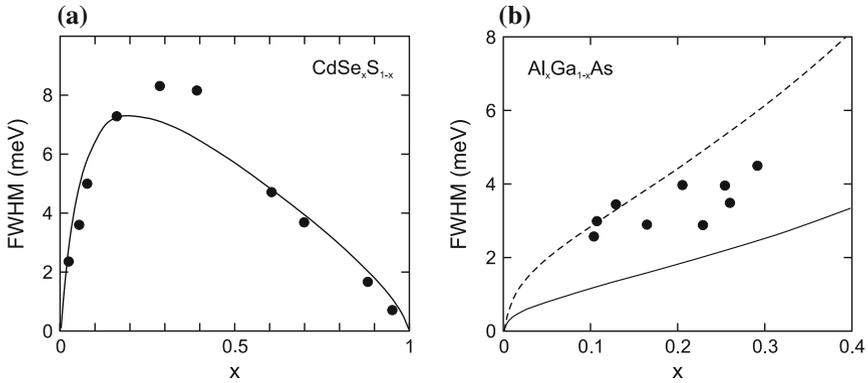
Fig. 10.15 (**a**) Spectral width of the photoluminescence from CdS_xSe_{1-x} alloys. *Solid line* is theory according to (10.30). Adapted from [866]. (**b**) Spectral width of the bound exciton recombination in $Al_xGa_{1-x}As$ with various Al content within the direct-bandgap regime. *Solid line* is (10.30) with (10.28), *dashed line* with pre-factor $4\pi/3$ instead of $10\,\pi$. Adapted from [870]
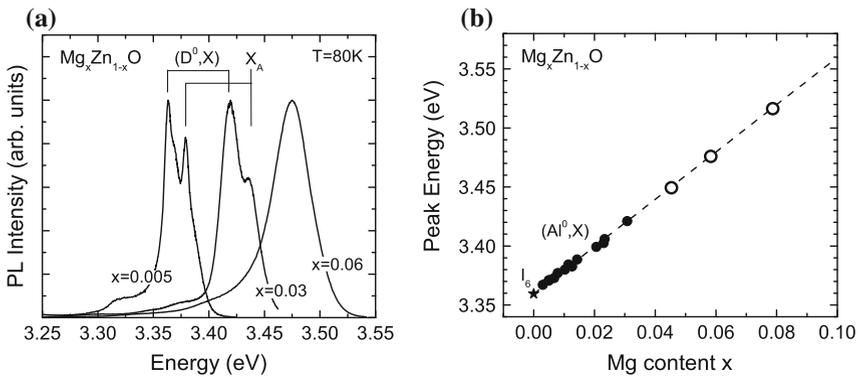

Fig. 10.16 (**a**) Photoluminescence spectra ($T = 80\,\text{K}$, scaled) of three $Mg_xZn_{1-x}O$ alloy layers on sapphire with three different Mg-contents, $x = 0.005$, $x = 0.03$, and $x = 0.06$ as labeled. The energy positions of (D^0, X) and X_A peaks are marked. Adapted from [860]. (**b**) Peak energy of the photoluminescence spectrum ($T = 2\,\text{K}$) of ZnO (I_6-line, *star*) and various $Mg_xZn_{1-x}O$ alloys (*circles*). For $x \leq 0.03$ (*filled circles*) the (D^0, X) recombination peak (Al donor) can be spectrally separated from the free exciton (X_A) recombination. For the samples with higher Mg content (*empty circles*) a single recombination peak is present at all temperatures. The *dashed line* is a linear least square fit for the alloys with $0 \leq x \leq 0.03$, showing that also for $x > 0.03$ the low temperature recombination peak is due to donor-bound excitons. Adapted from [871]

broadening is obvious, causing a single peak for $x > 0.03$. The temperature dependence of the peak positions is shown in Fig. 10.17 for the same samples. For $x = 0.005$ the bound exciton (Al-donor) (D^0, X) and free exciton (X_A) recombination lines can still be resolved despite the inhomogeneous broadening of $\sigma = 2.6\,\text{meV}$. At
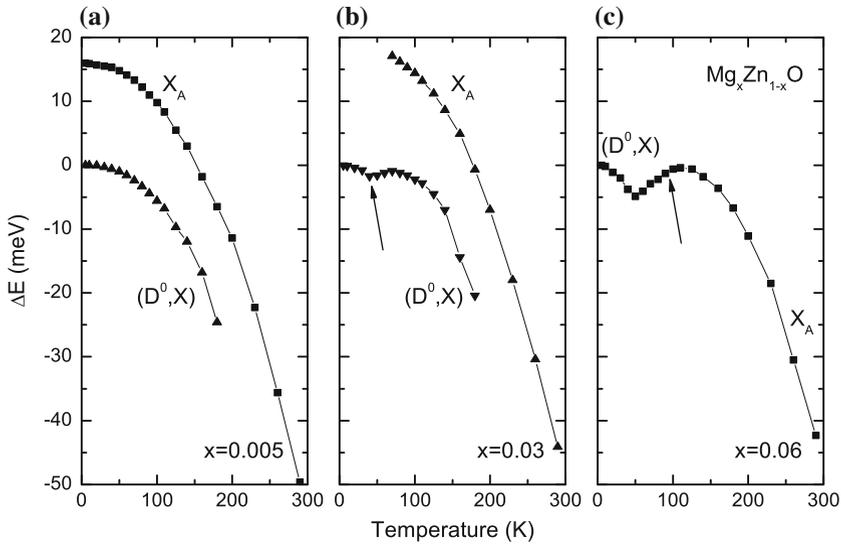
Fig. 10.17 Temperature dependence of the shift of energy position of (D^0, X) and X_A photoluminescence peak in $Mg_x Zn_{1-x}O$ alloys with three different Mg-contents, (**a**) $x = 0.005$, (**b**) $x = 0.03$, (**c**) $x = 0.06$. The energy positions are given relative to the low temperature position of the respective (D^0, X) peaks. Adapted from [860]

low temperature the luminescence intensity is dominated by (D^0, X) recombination, at room temperature by free exciton (X_A) recombination. Both peaks are present at low temperatures and exhibit a red-shift with increasing temperatures due to the shrinking of the band gap (Fig. 10.17a). The (D^0, X) peak vanishes at about $180\,K$ due to ionization of the excitons from the donors ($Q \approx 15\,meV$, similar as in pure ZnO).

For larger Mg-content of $x = 0.03$ the two peaks can still mostly be separated ($\sigma = 6.0\,meV$). The (D^0, X) energy position shows a small dip (about $2\,meV$) due to exciton localization in the alloy disorder potential (arrow in Fig. 10.17b). At low temperatures excitons are frozen in local potential minima and have a non-thermal (non-Boltzmann) population. With increasing temperature they can overcome energy barriers and thermalize, leading to a shift of the recombination peak to lower energies. Further increase of temperature populates higher levels and leads to a shift of the recombination peak to higher energies. Superimposed is the red-shift due to the band gap shrinkage. This so-called "S"-shape effect of $E(T)$ is discussed in Sect. 12.4 in detail with regard to exciton localization in a quantum well disorder potential.

For $x = 0.06$ only a single photoluminescence peak is observed for the alloy ($\sigma = 8.5\,meV$). The (D^0, X) peak is the dominant for the $Mg_x Zn_{1-x}O$ alloys at low temperatures even in the presence of large alloy broadening (Fig. 10.16b). The peak changes its nature from (D^0, X) at low temperatures to X_A at room temperature. In between, first exciton thermalization (red-shift) in the disorder potential and

subsequently exciton ionization from the donors (blue-shift, arrow in Fig. 10.17c) are observed [860]. Such exciton ionization from impurities has also been observed for AlGaN:Si [549, 872].
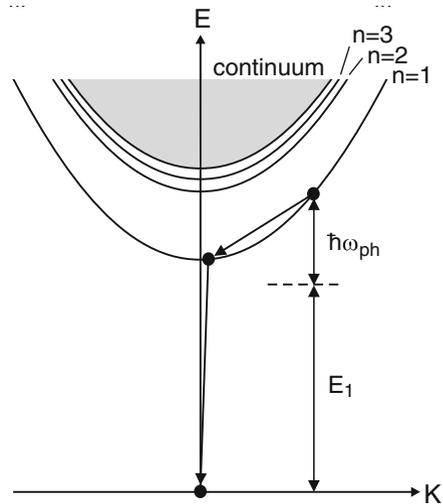
10.4 Phonon Replica

The momentum selection rule for free-exciton recombination allows only excitons with $\mathbf{K} \approx 0$ (for $\mathbf{K}$, cf. (9.37)) to recombine. The fine structure of this recombination is connected to polariton effects (cf. Sect. 9.6.8). Excitons with large $\mathbf{K}$ can recombine if a phonon or several phonons are involved [873] that provide the necessary momentum $\mathbf{q} = \mathbf{K}_1 - \mathbf{K}_2$, with $\mathbf{K}_1$ ($\mathbf{K}_2$) being the wave vector of the initial (intermediate) exciton state (Fig. 10.18). A so-called *zero-phonon* line at energy E_0 is then accompanied by *phonon replica* below E_0 at integer multiples (at low temperature) of the (LO) phonon energy $\hbar\omega_{\text{ph}}$

$$E_n = E_0 - n\,\hbar\omega_{\text{ph}}. \tag{10.31}$$

Phonon replicas have been observed in many polar semiconductors such as CdS [874] and ZnSe [875]. A sequence of such phonon replica, as observed in GaN [876], is depicted in Fig. 10.19a.

The lineshape of the nth phonon-assisted line is proportional to the exciton population at a given excess energy, which is proportional to the density of states and the Boltzmann distribution function [877]

Fig. 10.18 Schematic representation of 1 LO exciton scattering of an exciton at $\mathbf{K} \neq 0$ to an intermediate state with $\mathbf{K} \approx 0$ and subsequent radiative decay. $\hbar\omega$ represents the phonon energy and E_1 the energy of the emitted photon
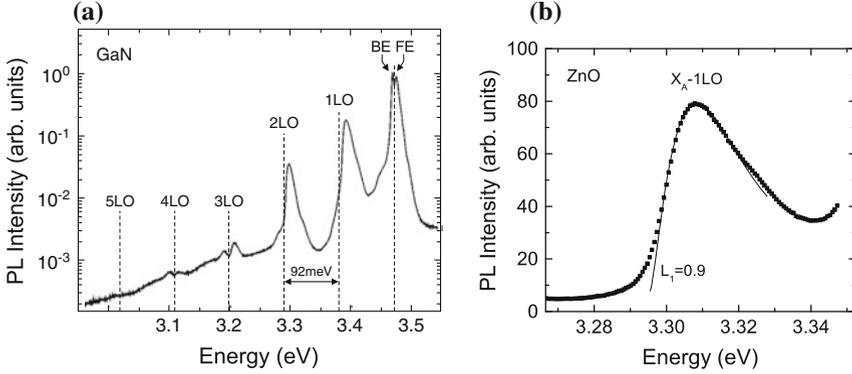
Fig. 10.19 (**a**) Photoluminescence spectrum of GaN (grown on SiC substrate) at $T = 50$ K. In addition to emission from free (FE) and bound (BE) excitons several phonon replica (labeled as 1LO–5LO) are observed. *Vertical dashed lines* indicate energy positions of multiple LO-phonon energies ($\hbar\omega_{LO} = 92$ meV) below the FE peak. Adapted from [876]. (**b**) Photoluminescence spectrum of 1LO phonon-assisted recombination peak at $T = 103$ K (from the data of Fig. 10.5). Data points (*dots*) and lineshape fit (*solid line*) according to (10.32) with the parameters $L_1 = 0.9$ and $E_1 = 3.2955$ eV (and background)

$$I_n(E_{\mathrm{ex}}) \propto \sqrt{E_{\mathrm{ex}}}\ \exp\left(-\frac{E_{\mathrm{ex}}}{kT}\right) w_n(E_{\mathrm{ex}}). \tag{10.32}$$

Here, E_{ex} represents the exciton kinetic energy. The factor $w_n(E_{\mathrm{ex}})$ accounts for the **q**-dependence of the matrix element. It is typically expressed as

$$w_n(E_{\mathrm{ex}}) \propto E_{\mathrm{ex}}^{L_n}. \tag{10.33}$$

Accordingly, as temperature dependent refinement of (10.31), the energy separation ΔE_n of the energy of the peak maximum of phonon replica from E_0 is given by

$$\Delta E_n = E_n - E_0 = -n\,\hbar\omega_{\mathrm{ph}} + \left(L_n + \frac{1}{2}\right)kT. \tag{10.34}$$

It is found theoretically that $L_1 = 1$ and $L_2 = 0$ [877]. These relations are approximately fulfilled for GaN [878]. A lineshape fit for the 1 LO phonon-assisted transition in ZnO is shown in Fig. 10.19b.

In Fig. 10.20a the 'green band' emission of ZnO is shown as presented in [879]. This band is mostly attributed to a Cu impurity; recently, evidence has grown from isotope decay and annealing studies that it is related to the zinc vacancy [880] (Fig. 10.20b). The zero phonon line is followed by many replica with a maximum at about 6 LO phonons. The intensity I_N of the N-th replica is given by [881, 882]

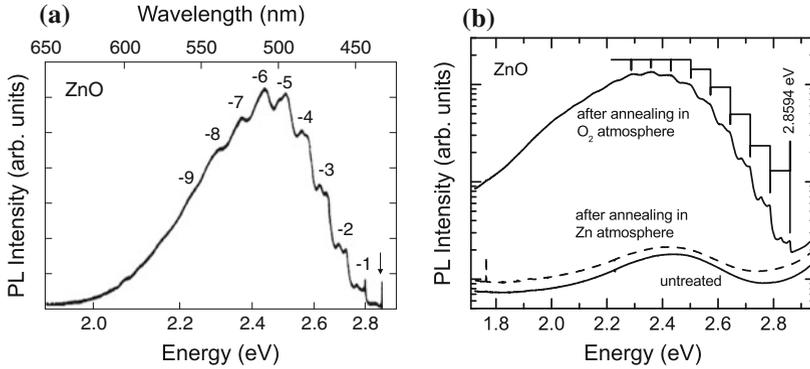$$I_N \propto \exp(-S)\,\frac{S^N}{N!}, \tag{10.35}$$

Fig. 10.20 (**a**) Luminescence spectrum of ZnO in the visible. The *arrow* denotes the zero-phonon line at 2.8590 eV. The numbers of the phonon replica are labeled. Adapted from [879]. (**b**) Luminescence spectra (*solid lines*) of a ZnO bulk crystal before ('untreated') and after annealing in O_2 atmosphere at $T = 1073$ K. After annealing in Zn atmosphere at the same temperature, the green band disappears again (*dashed line*). From [880]

where S is the so-called Huang–Rhys parameter. In [880], a coupling parameter of $S = 6.9$ has been determined.

Equation (10.35) is obtained from the consideration of transitions in the configuration diagram [881, 883] (Fig. 10.21), using the Born–Oppenheimer approximation. Here the electronic wavefunctions are separated from the vibrational wavefunctions, leading to the Franck–Condon principle, that optical transitions occur with the positions of the nuclei fixed and thus vertical in the configuration diagram Fig. 10.21. Assuming low temperatures, only the lowest state is (partially) occupied. The Huang–Rhys parameter, the average number of phonons involved in the transition, is related to the displacement $\delta q = q_1 - q_0$ of the two configurations

$$S = \frac{C\,\delta q^2}{2\,\hbar\omega_{\mathrm{ph}}}, \tag{10.36}$$

where C is the 'spring constant' of the parabola, $C = \mathrm{d}^2 E/\mathrm{d}q^2$.

For small $S \ll 1$, we are in the weak coupling regime and the zero-phonon line is the strongest. In the strong coupling regime, $S > 1$, the maximum is (red-) shifted from the zero-phonon line. We note that in absorption, phonon replica occur on the high-energy side of the zero-phonon absorption. For large S the peak intensities are close to a Gaussian. The correspondence of emission and absorption is nicely seen for excitons on isoelectronic oxygen traps in ZnTe [882]. The oxygen is on substitutional Te site. Up to seven phonon replica are visible in Fig. 10.22 around the zero-phonon or A-line with a separation of about 26 meV, the optical phonon energy in ZnTe. The Huang–Rhys parameter is about 3–4. Other peaks are due to acoustic phonons.
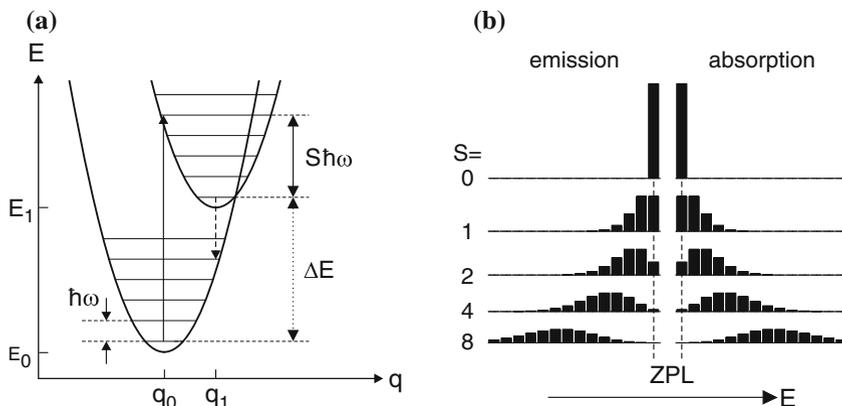
Fig. 10.21 (**a**) Configuration diagram of two states that differ in their configuration coordinate by $\delta q = q_1 - q_0$. Both are coupled to phonons of energy $\hbar\omega$. The absorption maximum (*solid vertical line*) and emission maximum (*dashed vertical line*) are shifted with respect to the zero-phonon line (*dotted vertical line*) with energy $E_1 - E_0$. The Huang–Rhys parameter is $S \sim 4$. (**b**) Intensity of zero-phonon line ('ZPL') and phonon replica (10.35) for emission and absorption processes with different values of the Huang–Rhys parameter S as labeled
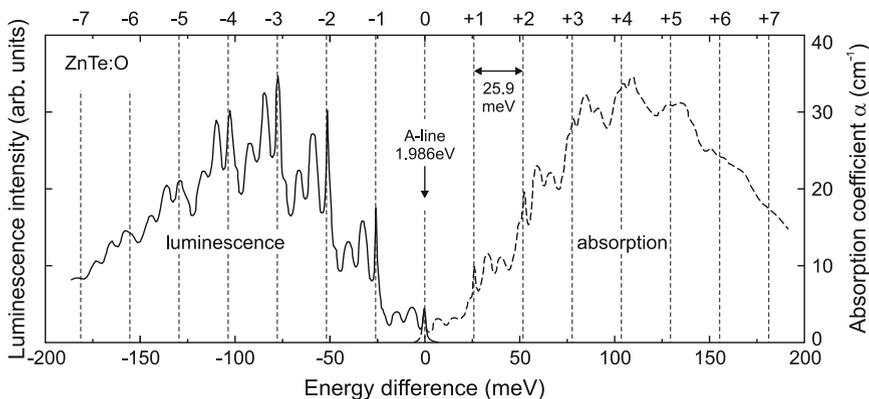

Fig. 10.22 Photoluminescence (*solid line*) and absorption (*dashed line*) spectra of excitons bound to substitutional oxygen in ZnTe at $T = 20$ K. The energy position is relative to the A-line at 1.9860 eV. The *vertical dashed lines* have a separation of 25.9 meV. Adapted from [884]

10.5 Self-Absorption

Luminescence that is emitted within the semiconductor can be (re-)absorbed before it may reach the surface and can leave the crystal. This effect is called self-absorption. It is particularly strong for radiation with an energy where the absorption $\alpha(\hbar\omega)$ is high, i.e. above the band gap of a direct semiconductor. Similarly to the penetration
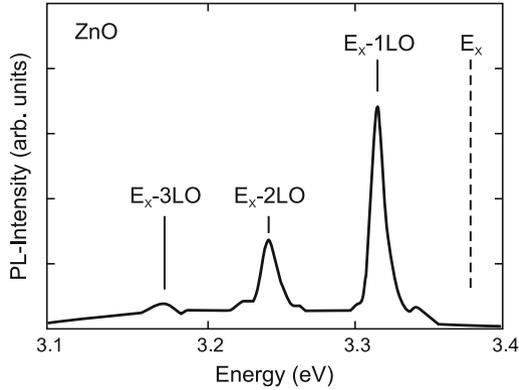
Fig. 10.23 Photoluminescence spectrum (at $T = 55$ K) from bulk ZnO excited homogeneously via two-photon excitation by a Q-switched ruby laser (pulse width 40 ns). Adapted from [885]

depth $1/\alpha$ for radiation entering the crystal, emission approximately occurs only from a layer of such thickness. For typical values of α in the range of 10^5 cm^{-1}, the 'skin' of the semiconductor that emits light with energy above the band gap is 100 nm. For light at the low energy side of the band gap or with energy within the band gap (deep levels), the emission depth can be much larger.

After re-absorption, the energy has another chance to relax non-radiatively, thus reducing the quantum efficiency. Alternatively it can be reemitted, either at the same energy or at a lower energy. Possibly several re-absorption processes occur before a photon eventually leaves the semiconductor ('photon recycling'). Such processes are important in LED structures where photon extraction has to be optimized (Sect. 23.3.4). Emission on phonon replica (Sect. 10.4) is red-shifted from the energy range of strong absorption and thus suffers no (or only little) self-absorption. This can be seen from the spectrum of a thick ZnO crystal excited homogeneously (via two-photon absorption with a red Ruby laser), Fig. 10.23. The zero phonon line (at E_X), originating from the $\approx$100 nm skin of the samples and being by far the strongest in thin films (Fig. 10.5), is practically absent and emission on the phonon replica collected from the entire volume dominates the spectrum.

10.6   Donor–Acceptor Pair Transitions

Optical transitions can occur between neutral donors and acceptors. The (spatially indirect) donor–acceptor pair (DAP) recombination is present in (partially) compensated semiconductors and follows the scheme $D^0 A^0 \rightarrow D^+ A^- eh \rightarrow D^+ A^- + \gamma$, where γ is a photon with the energy $\hbar\omega$. The energy of the emitted photon is given by

$$\hbar\omega = E_g - E_D^b - E_A^b + \frac{1}{4\pi\epsilon_0} \frac{e^2}{\epsilon_r R}, \tag{10.37}$$
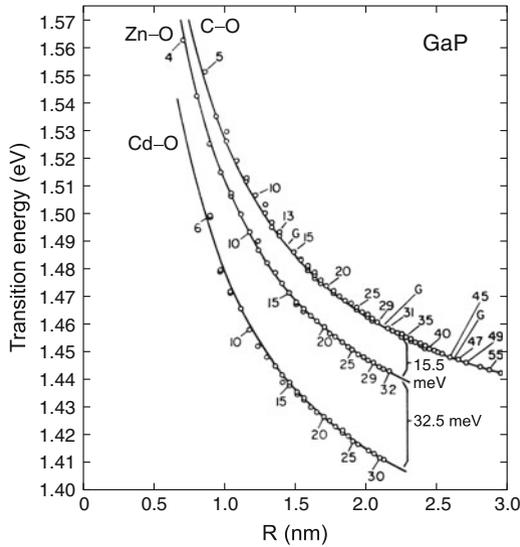
Fig. 10.24 Transition energies in GaP ($T = 1.6$ K) of the donor–acceptor recombination involving the deep oxygen donor and C, Zn, and Cd acceptors, respectively. The lines follow (10.37) for $E_g^{GaP} = 2.339$ eV, $\epsilon_r = 11.1$ and $(E_D^b)_O = 893$ meV, $(E_A^b)_C = 48.5$ meV, $(E_A^b)_{Zn} = 64$ meV, and $(E_A^b)_{Cd} = 96.5$ meV. Predicted missing modes for GaP:C, O are labeled with 'G'. Adapted from [886]

where R is the distance between the donor and the acceptor for a specific pair. Since R is discrete, the DAP recombination spectrum consists of several discrete lines. If the donor and acceptor occupy the same sublattice, e.g. O and C both substituting P sites in GaP, the spatial distance of the donor and acceptor is $R(n) = a_0\sqrt{n/2}$, where a_0 is the lattice constant and n is an integer. However, for certain 'magic' numbers $n = 14, 30, 46, \ldots$ no lattice points exist and therefore the corresponding lines are missing (labeled 'G' in Fig. 10.24). No such gaps exist in DA spectra where donors and acceptors occupy different sublattices, e.g. GaP:O, Zn (see also Fig. 10.24). In this case, the spatial separation is given by $R(n) = a_0\sqrt{n/2 - 5/16}$. If significant broadening is present, the lines are washed out and a donor–acceptor pair band forms.

10.7   Inner-Impurity Recombination

The transitions of electrons between different states of an impurity level can be nonradiative or radiative. As an example, the radiative transition of electrons in the Fe^{2+} state in InP $^5T_2 \rightarrow{}^5E$ (Fig. 10.25) and its fine structure were observed first in [888] at around 0.35 eV.
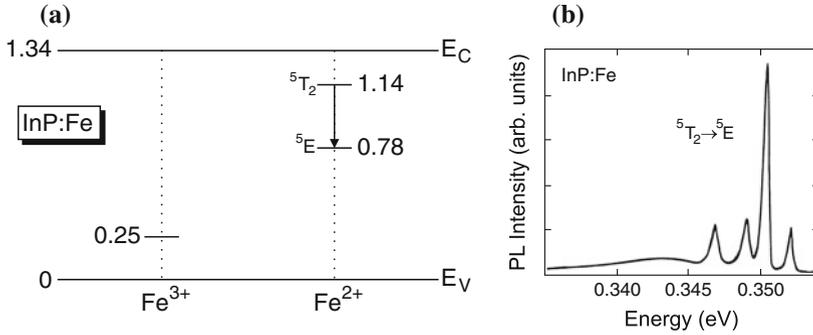
Fig. 10.25 (**a**) Schematic band diagram of InP with levels of Fe impurities in the 3+ and 2+ charge states at low temperature. All energies are given in eV. The *arrow* denotes the optical transition from an excited Fe^{2+} state to the Fe^{2+} ground state. (**b**) Photoluminescence spectrum (at $T = 4.2$ K) of InP:Fe sample with $[Fe] = 5 \times 10^{16}$ cm^{-3}. Part (**b**) adapted from [887]

10.8 Auger Recombination

In competition with the radiative, bimolecular recombination is the Auger recombination (Fig. 10.26). In the Auger process, the energy that is released during the recombination of an electron and hole is not emitted with a photon but, instead, transferred to a third particle. This can be an electron (eeh, Fig. 10.26a) or a hole (hhe, Fig. 10.26b). The energy is eventually transferred nonradiatively from the hot third carrier via phonon emission to the lattice. The probability for such process is $\propto n^2 p$ if two electrons are involved and $\propto np^2$ if two holes are involved. The Auger process is a three-particle process and becomes likely for high carrier density, either through doping, in the presence of many excess carriers, or in semiconductors with small band gap. Auger recombination is the inverse of the impact ionization (cf. Sect. 8.4.4). Phonon-assisted Auger recombination relaxes the momentum conservation rule for the involved charge carriers at the cost of an additional particle being involved in the scattering process. It has been pointed out that this process is dominating in bulk material [889, 890] and quantum wells [891].

In thermodynamic equilibrium the rates for Auger recombination and thermal Auger generation must be equal, thus

$$G_{\text{th}} = C_n \, n_0^2 \, p_0 + C_p \, n_0 \, p_0^2, \tag{10.38}$$

where C_n and C_p denote the Auger recombination coefficients. The equation for the dynamics in the presence of excess carriers (if solely Auger recombination is present) is given as

$$\frac{\partial \, \delta n}{\partial t} = G_{\text{th}} - R = -C_n \, (n^2 \, p - n_0^2 \, p_0) - C_p \, (n \, p^2 - n_0 \, p_0^2). \tag{10.39}$$
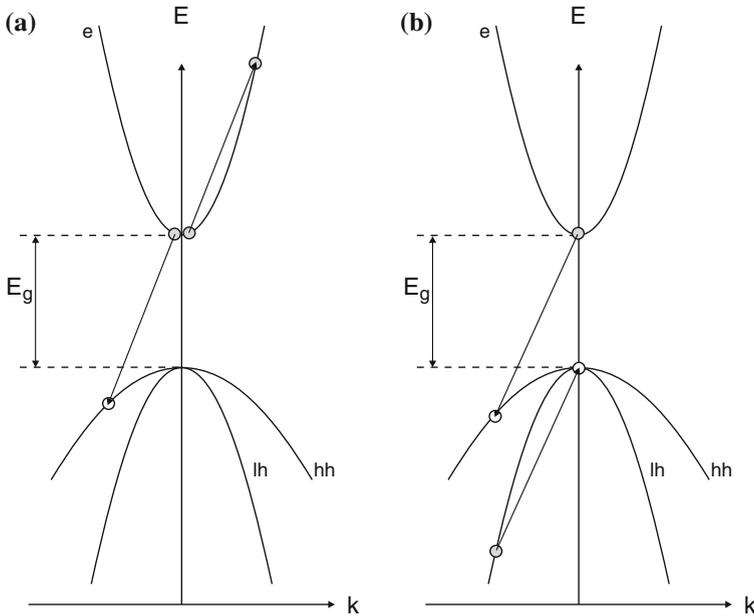
Fig. 10.26 Schematic representation of Auger recombination. An electron recombines with a hole and transfers the energy to (**a**) another electron in the conduction band, (**b**) another electron in the valence band

The Auger recombination rate typically used in SRH kinetics is

$$r_{\text{Auger}} = (C_n\,n + C_p\,p)\,(np - n_0 p_0)\,. \tag{10.40}$$

Typical values for the Auger recombination coefficients are given in Table 10.3.

In Fig. 10.27a the electron lifetime in heavily p-doped InGaAs (lattice matched to InP) is shown [893]. It follows $\tau_n^{-1} = C_p\,N_A^2$ as expected from (10.39) for p-type material. The Auger process in silicon has been discussed in detail [894]. In Fig. 10.27b experimental data for n-type and p-type Si are summarized. Auger theory can predict the lifetimes in n-type material. The predicted rate in p-type material is too small, thus a phonon-assisted process is evoked [894].

Material	C_n (cm^6/s)	C_p (cm^6/s)
4H-SiC	5×10^{-31}	2×10^{-31}
Si, Ge	2.8×10^{-31}	9.9×10^{-32}
GaAs, InP	5.0×10^{-30}	3.0×10^{-30}
InSb	1.2×10^{-26}	

Table 10.3 Auger recombination coefficients for some semiconductors. Data for InSb from [892], SiC from [830], others from [831]
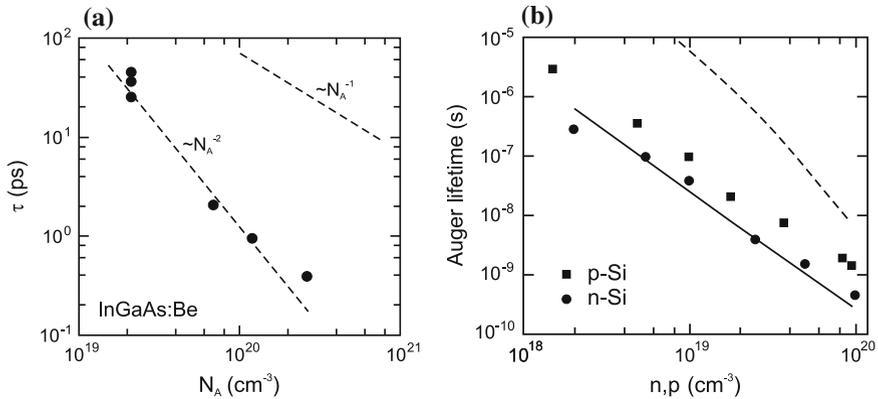
Fig. 10.27 (**a**) Experimental values of the electron lifetime in heavily p-doped InGaAs on InP at room temperature. The *dashed lines* show dependencies of Auger ($\propto N_A^{-2}$, $C_p = 8.1 \times 10^{-29}\,\mathrm{cm^{-6}\,s^{-1}}$) and band-band recombination ($\propto N_A^{-1}$, $B = 1.43 \times 10^{-10}\,\mathrm{cm^{-3}\,s^{-1}}$). Adapted from [893]. (**b**) Experimental Auger lifetimes in p-type (*squares*) and n-type (*circles*) silicon at 300 K. The *dashed* (*solid*) *line* is theory for p-type (n-type) material. Adapted from [894]

10.9   Band–Impurity Recombination

A very important recombination process is the capture of carriers by impurities. This process is in competition with all other recombination processes, e.g. the radiative recombination and the Auger mechanism. The band–impurity recombination is the inverse process to the carrier release from impurities and intimately related to carriers statistics (Chap. 7). It is particularly important at low carrier densities, for high dopant concentration and in indirect semiconductors since for these the bimolecular recombination is slow. This process is generally considered to be non-radiative since no photons close to the band edge are emitted.[4]

10.9.1   Shockley–Read–Hall Kinetics

The theory of capture on and recombination involving impurities is called Shockley–Read–Hall (SRH) kinetics [825]. An example of radiative band–impurity recombination (of the type shown in Fig. 10.28a) is shown in Fig. 10.8 (e, A^0) recombination at the carbon acceptor in GaAs.

---

[4]Depending on the energetic depth of the trap, mid or far infrared photons can be emitted.
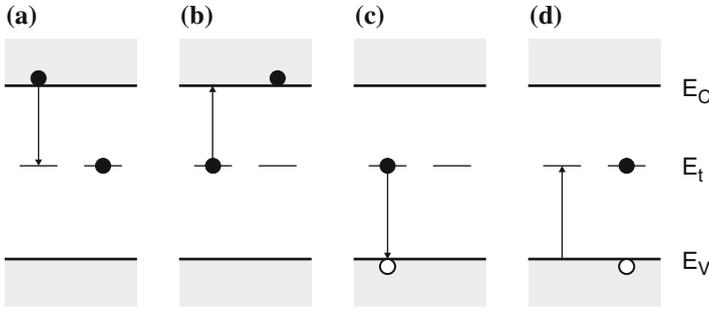
Fig. 10.28 Band-to-impurity processes at an impurity with one level (*left* initial, *right* final state in each part): (**a**) electron capture (from conduction band), (**b**) electron emission (into conduction band), (**c**) hole capture (from valence band), (**d**) hole emission (into valence band). The *arrows* indicate the transition of the electron

We consider electron traps [895] (see Fig. 10.28) with a concentration N_t with an energy level E_t. In thermodynamic equilibrium they have an electron population

$$f_t^0 = \frac{1}{\exp\left(\frac{E_t - E_F}{kT}\right) + 1}, \tag{10.41}$$

where f_t is the nonequilibrium population of the trap. Then the capture rate r_c is proportional to the unoccupied traps and the electron concentration, $r_c \propto n N_t (1 - f_t)$. The proportionality factor has the form $v_{th} \sigma_n$, where v_{th} is the thermal velocity $v_{th} = \sqrt{3kT/m^*} \approx 10^7$ cm/s and σ_n is the capture cross section that is of atomic scale, typically $\sim 10^{-15}$ cm^2. The capture cross section can be related to the optical absorption cross section [513, 514].

In order to make the following calculation more transparent, we put the effective-mass ratio $\sqrt{m_0/m^*}$ into σ in the following and thus have the same thermal velocity $v_{th} = \sqrt{3kT/m_0}$ for electrons and holes. The capture rate of electrons is

$$r_c = v_{th}\, \sigma_n\, n\, N_t\, (1 - f_t). \tag{10.42}$$

The emission rate from filled traps is

$$g_c = e_n\, N_t\, f_t, \tag{10.43}$$

where e_n denotes the emission probability. In a similar way, the emission and capture rates for holes can be written:

$$r_v = v_{th}\, \sigma_p\, p\, N_t\, f_t \tag{10.44}$$

$$g_v = e_p\, N_t\, (1 - f_t). \tag{10.45}$$

In thermodynamical equilibrium, capture and generation rates are equal, i.e. $r_c = g_c$ and $r_v = g_v$. Thus, the emission probability is

$$e_n = v_{th}\, \sigma_n\, n_0\, \frac{1 - f_t^0}{f_t^0}. \tag{10.46}$$

Using $\frac{1-f_t^0}{f_t^0} = \exp\left(\frac{E_t - E_F}{kT}\right)$, (7.10) and (7.11) the emission probabilities can be written as

$$e_n = v_{th}\, \sigma_n\, n_t \tag{10.47}$$
$$e_p = v_{th}\, \sigma_p\, p_t, \tag{10.48}$$

with

$$n_t = N_C \exp\left(\frac{E_t - E_C}{kT}\right) \tag{10.49}$$

$$p_t = N_V \exp\left(-\frac{E_t - E_V}{kT}\right). \tag{10.50}$$

We note that $n_t\, p_t = n_0\, p_0$ (cf. (7.15)).

The temperature dependence of the thermal velocity is $\propto T^{1/2}$, the temperature dependence of the band-edge density of states is $\propto T^{3/2}$ (7.8) and (7.9). Thus, the temperature dependence of the emission rate e_n is (apart from the exponential term) $\propto T^2$ if σ is temperature independent.

Charge conservation requires in nonequilibrium (and of course in equilibrium) $r_c - r_v = g_c - g_v$. From this we obtain the population of the trap in nonequilibrium:

$$f_t = \frac{\sigma_n\, n + \sigma_p\, p_t}{\sigma_n\, (n + n_t) + \sigma_p\, (p + p_t)}. \tag{10.51}$$

The recombination rate r_{b-i} of the band–impurity recombination is then

$$\begin{aligned}
r_{b-i} &= -\frac{\partial \delta n}{\partial t} \\
&= r_c - g_c \\
&= \frac{\sigma_n\, \sigma_p\, v_{th}\, N_t}{\sigma_n\, (n + n_t) + \sigma_p\, (p + p_t)}\, (n\, p - n_0\, p_0).
\end{aligned} \tag{10.52}$$

Using the 'lifetimes'

$$\tau_{n_0} = (\sigma_n\, v_{th}\, N_t)^{-1} \tag{10.53}$$
$$\tau_{p_0} = (\sigma_p\, v_{th}\, N_t)^{-1}, \tag{10.54}$$

this is typically written as

$$r_{b-i} = \frac{1}{\tau_{p_0} (n + n_t) + \tau_{n_0} (p + p_t)} (n\, p - n_0\, p_0)\,. \tag{10.55}$$

For an n-type semiconductor the Fermi level is above E_t and the traps are mostly full. Thus hole capture is the dominating process. The equation for the dynamics simplifies to

$$\frac{\partial \delta p}{\partial t} = -\frac{p - p_0}{\tau_{p_0}}\,. \tag{10.56}$$

Thus, an exponential decay with minority-carrier lifetime τ_{p_0} (or τ_{n_0} for p-type material) occurs.

A recombination center is most effective when it is close to the middle of the band gap (midgap level). The condition $\partial r_{b-i}/\partial E_t = 0$ leads to the trap energy E_t^{max} with the maximum recombination rate being located at

$$E_t^{max} = \frac{E_C + E_V}{2} - kT\,\ln\left(\frac{\sigma_n\, N_C}{\sigma_p\, N_V}\right). \tag{10.57}$$

The curvature $\partial^2 r_{b-i}/\partial E_t^2$ at E_t^{max} is proportional to $-(np - n_0 p_0)$ and thus indeed is negative in the presence of excess carriers. However, the maximum can be fairly broad.

The SRH kinetic presented here is valid for low densities of recombination centers. A more detailed discussion and a more general model can be found in [896].

A typical example for a recombination center is gold in silicon. The minority carrier lifetime decreases from 2×10^{-7} s to 2×10^{-10} s upon increase of the Au concentration from 10^{14} to 10^{17} cm^{-3}. The incorporation of recombination centers is an important measure for the design of high-frequency devices [897]. Due to importance in silicon technology the recombination properties of many metals in silicon have been investigated, in particular Fe-contamination and the role of FeB-complexes [898–900].

A reduction in minority-carrier lifetime can also be achieved by irradiation with high-energy particles and the subsequent generation of point defects with energy levels at midgap.

In Fig. 10.29 various data on minority carrier lifetime in silicon are compiled. Over some doping range, a dependence of the lifetime $\propto N^{-1}$ as in (10.53) prevails. For doping beyond the 10^{19} cm^{-3} range, Auger recombination (Sect. 10.8) with $\tau \propto N^{-2}$ is dominant. A more detailed discussion can be found in [901, 902]. Generally the lifetimes are temperature dependent [903] as expected from (10.52).
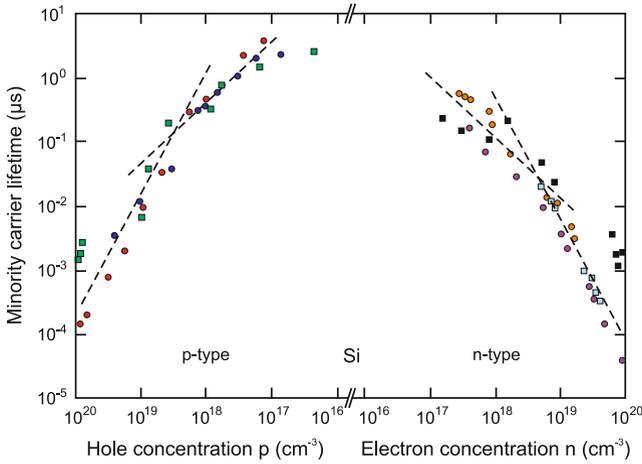
Fig. 10.29 Minority carrier lifetime at room temperature as a function of majority carrier concentration in n-type and p-type silicon. The *dashed lines* have the slopes N^{-1} and N^{-2}. Data from [902]

10.9.2 Multilevel Traps

Traps with multiple levels in the band gap have generally similar but more complicated dynamics as compared to single-level traps. Lifetimes are an average over negatively and positively charged states of the trap.

10.10 ABC Model

Summarizing the results on band-impurity recombination (Sect. 10.9), bimolecular recombination (Sect. 10.2) and Auger recombination (Sect. 10.8), the total recombination rate R can be written simplified as

$$R = A\,n + B\,n^2 + C\,n^3, \tag{10.58}$$

where A is the coefficient for the band-impurity recombination, B the bimolecular recombination coefficient and C the Auger recombination coefficient; n denotes the carrier density. This model is known as the 'ABC' model. It can be refined separating effects of electrons and holes and including higher terms. Often such model is used to investigate recombination in devices as a function of injection, e.g. [904, 905].

The internal radiative quantum efficiency η_{int} is given by ratio of the radiative recombination rate and the total recombination rate,

$$\eta_{int} = \frac{B\,n^2}{A\,n + B\,n^2 + C\,n^3} = \frac{B\,n}{A + B\,n + C\,n^2}. \tag{10.59}$$

10.11 Field Effect

The emission of electrons from a trap is thermally activated with an ionization energy $E_i = E_C - E_t$. If the trap is in a strong electric field E, the emission probability can change. An acceptor-like trap after removal of the electron is neutral and its potential is short range. A donor has a long-range Coulomb potential after ionization. In an electric field, these potentials are modified as visualized in Fig. 10.30. Various additional processes can now occur.

10.11.1 Thermally Activated Emission

For the δ-like potential the ionization energy remains unchanged. For the Coulomb potential the barrier in the field direction is lowered by

$$\Delta E_i = e \sqrt{\frac{e}{\pi \, \epsilon_0 \, \epsilon_r}} \, \sqrt{E} \,. \tag{10.60}$$

The emission rate e_n is increased in the field by $\exp(\Delta E_i / kT)$. This effect is called the Poole–Frenkel effect [906] and can be quite important. For silicon and $E = 2 \times 10^5$ V/cm and $\Delta E_i = 100$ meV a 50-fold increase of the emission rate at room temperature is expected. As an example the Poole–Frenkel effect for the electron emission from (neutral) interstitial boron in silicon ($B_i^0 \rightarrow B_i^+ + e^-$) is
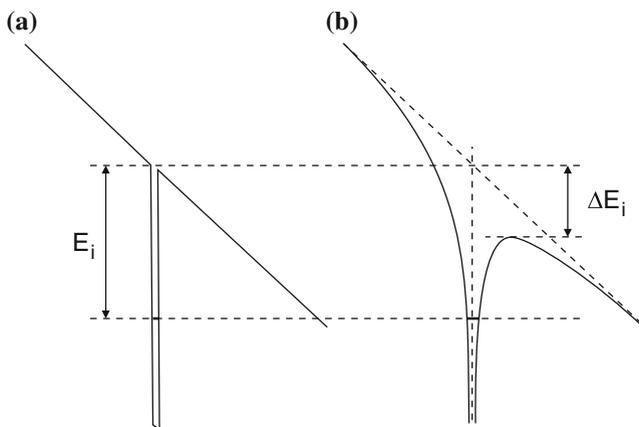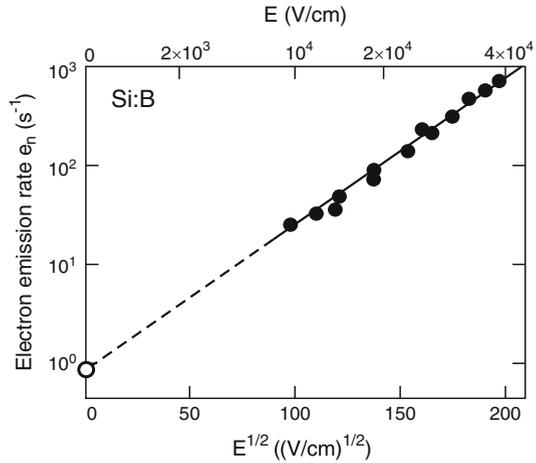

Fig. 10.30 Field effect at (**a**) a δ-like potential and (**b**) a Coulomb potential

Fig. 10.31 Field effect for electron emission from interstitial boron in silicon ($T = 65$ K). The *filled circles* represent experimental data from DLTS, the data point marked with an *empty circle* is from EPR (zero field). The *line* is a linear fit and extrapolation. Adapted from [907]

shown in Fig. 10.31, following the enhancement of $e_n \propto \exp(\sqrt{E})$. The extrapolation to $E = 0$ agrees with the EPR result [241, 907].[5]

10.11.2 Direct Tunneling

Carriers can tunnel from the trap level through the barrier in the field direction into the conduction band. This process is temperature independent. The transmission factor of a barrier is (in WKB approximation) proportional to $\exp[-(2/\hbar)$ $\int \sqrt{2m\,[V(x) - E]}\,\mathrm{d}E]$. The emission probability for a triangular barrier is then

$$e_n = \frac{e\,E}{4\sqrt{2m^* E_i}} \exp\left(-\frac{4\sqrt{2m^*}\,E_i^{3/2}}{3\,e\,\hbar\,E} \right). \tag{10.61}$$

In the case of a Coulomb-like potential the argument of the exponent in (10.61) needs to be multiplied by a factor $1 - (\Delta E_i/E_i)^{5/3}$ with ΔE_i from (10.60).

10.11.3 Assisted Tunneling

In a thermally assisted tunneling process the electron on the trap level is first excited to a virtual level $E_t + E_{ph}$ by phonon absorption and then tunnels out of the trap

---

[5]The slope of the line in Fig. 10.31 is slightly smaller than expected from (10.60).

(photon-assisted tunneling). From the energetically higher level the tunneling rate is higher. The probability is proportional to $\exp\left(E_{ph}/kT\right)$. The additional energy can also be supplied by a photon (photon-assisted tunneling).

## 10.12   Recombination at Extended Defects

10.12.1   Surfaces

A surface (cmp. Chap. 11) is typically a source of recombination, e.g. through midgap levels induced by the break of crystal symmetry. The recombination at surfaces is modeled as a recombination current

$$j_s = -e\,S\,(n_s - n_0), \tag{10.62}$$

where n_s is the carrier density at the surface and S is the so-called *surface recombination velocity*.

The surface recombination velocity for GaAs is shown in Fig. 10.32. For InP, if the surface Fermi level is pinned close to midgap, the surface recombination velocity increases from $\sim 5 \times 10^3$ cm/s for a doping level of $n \sim 3 \times 10^{15}$ cm^{-3} to $\sim 10^6$ cm/s for a doping level of $n \sim 3 \times 10^{18}$ cm^{-3} [908]. For Si, the surface recombination rate depends on the treatment of the surface and lies in the range between 1–10^4 cm/s [911]. The Si-SiO$_2$ interface can exhibit $S \leq 0.5$ cm/s. Time-resolved measurements and detailed modeling for Si have been reported in [911].

Fig. 10.32 Surface recombination velocity for GaAs as a function of n-type doping concentration. Different experimental points correspond to different surface treatment methods. *Dashed line* is a guide to the eye. Experimental data from [912]
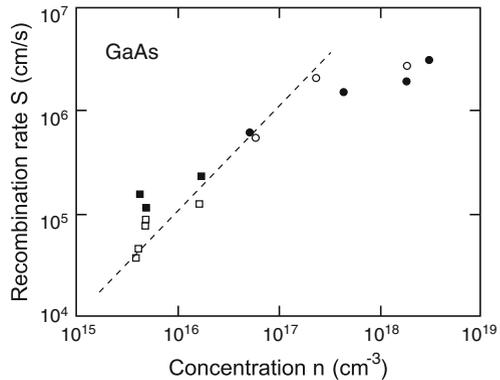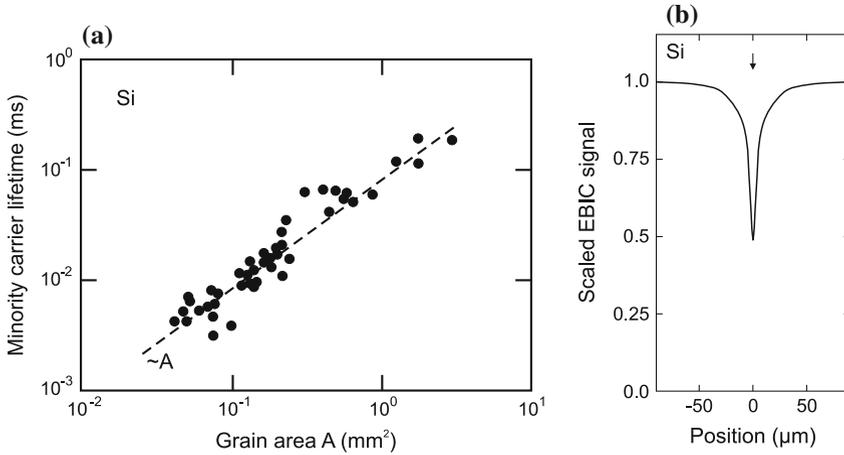
Fig. 10.33 (**a**) Minority carrier lifetime in (p-type) silicon as a function of grain boundary size. The *dashed line* has the slope $\propto A$. Data from [916]. (**b**) Linescan of the electron beam induced current (EBIC) perpendicular to a single grain boundary in silicon. The *arrow* denotes the position of the grain boundary. Compiled from [917]

10.12.2   Grain Boundaries

Grain boundaries can be a source of non-radiative recombination. This is technologically important for solar cells made from polycrystalline silicon (cf. Sect. 22.4.6). The grain boundary can be understood as an inner surface in the crystal. Modelling of recombination at a grain boundary can be done using an interface recombination velocity [913, 914] or considering deep traps [915]. The minority carrier lifetime decreases with decreasing grain boundary area A (Fig. 10.33a). The carrier loss at a grain boundary can be imaged directly via the efficiency of the collection of an electron beam induced current (EBIC) as shown in Fig. 10.33b. The minority carrier lifetime is only unaffected when the average distance to a grain boundary is much larger than the minority carrier diffusion length, $\sqrt{A} \gg L_{\mathrm{D}}$, otherwise the entire grain volume is subject to non-radiative recombination.

10.12.3   Dislocations

Also dislocations typically act as recombination centers, sometimes called carrier sinks. In Fig. 10.34 it can be seen that the minority carrier lifetime depends on the dislocation density n_{d} and follows a $\tau^{-1} \propto n_{\mathrm{d}}$ law, as if each dislocation is a recombination center [918]. The non-radiative recombination makes dislocations appear as 'dark line defects' in luminescence imaging [919]. In [920] also the decrease of carrier lifetime around (misfit) dislocations has been imaged. The effect of dislocations on the radiative recombination efficiency depends on the diffusion length [920].

Fig. 10.34 Inverse minority carrier lifetime in n-type silicon (40 Ω cm), low resistivity Ge (3–5 Ω cm) and high resistivity Ge (30–40 Ω cm). Data from [918]
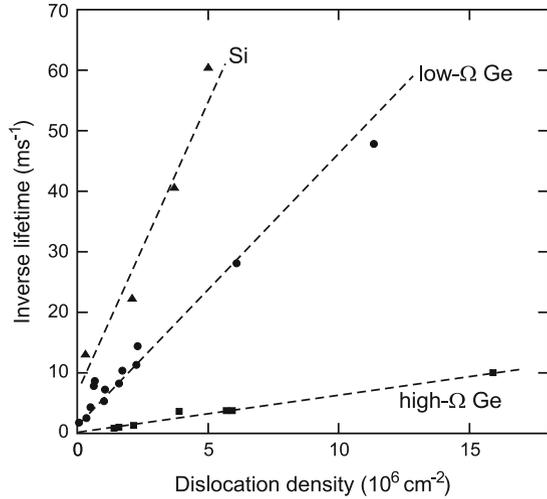
10.13 Excess-Carrier Profiles

In this section, some typical excess-carrier profiles (in one-dimensional geometry) are discussed that arise from certain excitation conditions. The excess-carrier density Δp (here holes in an n-type semiconductor, i.e. $\Delta p = p_{\mathrm{n}} - p_{\mathrm{n}_0}$) is determined by the diffusion equation (cf. (8.64a))

$$D_{\mathrm{p}} \frac{\partial^2 \Delta p}{\partial x^2} = -G(x) + \frac{\Delta p}{\tau_{\mathrm{p}}} .$$
(10.63)

10.13.1 Generation at Surface

First, the generation of excess carriers in a semi-infinite piece of semiconductor shall occur only at the surface at $x = 0$. The generation is zero everywhere else and the excitation is incorporated via the boundary condition $\Delta p(x = 0) = \Delta p_0$. The general solution for the homogeneous equation (10.63), i.e. $G = 0$, is

$$\Delta p(x) = C_1 \exp\left(-\frac{x}{L_{\mathrm{p}}}\right) + C_2 \exp\left(\frac{x}{L_{\mathrm{p}}}\right),$$
(10.64)

with the diffusion length $L_{\mathrm{p}} = \sqrt{D_{\mathrm{p}}\tau_{\mathrm{p}}}$. Taking the boundary condition $\Delta p(x \to \infty) = 0$ the solution is

$$\Delta p(x) = \Delta p_0 \exp\left(-\frac{x}{L_{\mathrm{p}}}\right).$$
(10.65)

In order to connect Δp_0 with the total generation rate per unit area G_{tot}, we calculate

$$G_{\text{tot}} = \int_0^\infty \frac{\Delta p(x)}{\tau_{\text{p}}}\,\mathrm{d}x = \frac{\Delta p_0\,L_{\text{p}}}{\tau_{\text{p}}} = \Delta p_0 \sqrt{\frac{D_{\text{p}}}{\tau_{\text{p}}}}. \qquad (10.66)$$

10.13.2   Generation in the Bulk

Now, a generation rate following (9.16), realistic for photodiodes and solar cells, is considered,

$$G(x) = G_0 \exp\left(-\alpha\, x\right), \qquad (10.67)$$

i.e. due to light absorption with the (wavelength-dependent) absorption coefficient α. The total generation rate is

$$G_{\text{tot}} = \int_0^\infty G(x)\,\mathrm{d}x = \frac{G_0}{\alpha}. \qquad (10.68)$$

The total generation rate is equal to the number of photons per second Φ_0 that enter the semiconductor.

The solution of (10.63) is the sum of the homogeneous solution (10.64) and a particular solution that is given by

$$\Delta p(x) = C\, \exp\left(-\alpha\, x\right). \qquad (10.69)$$

The constant C is determined to be

$$C = \frac{G_0\, \tau_{\text{p}}}{1 - \alpha^2\, L_{\text{p}}^2}. \qquad (10.70)$$

Therefore, the solution is

$$\Delta p(x) = C_1 \exp\left(-\frac{x}{L_{\text{p}}}\right) + C_2 \exp\left(\frac{x}{L_{\text{p}}}\right) + \frac{G_0\, \tau_{\text{p}}}{1 - \alpha^2\, L_{\text{p}}^2}\, \exp\left(-\alpha\, x\right). \qquad (10.71)$$

Using again $\Delta p(x \to \infty) = 0$ (leading to $C_2 = 0$) and a recombination velocity S at the front surface, i.e.

$$- e\, S\, \Delta p_0 = -e\, D_{\text{p}}\, \frac{\partial \Delta p}{\partial x}\bigg|_{x=0}. \qquad (10.72)$$
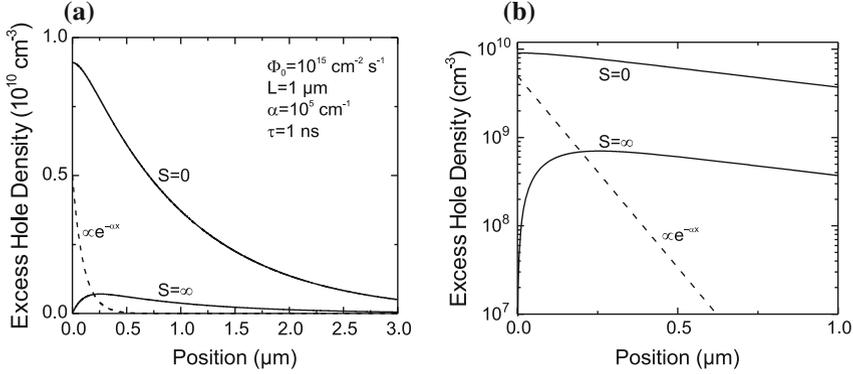
Fig. 10.35 Excess carrier density profile (10.73) in (**a**) linear and (**b**) semi-logarithmic plot for $S = 0$ and $S = \infty$. Other parameters are given in panel (**a**)

The solution is given as

$$\Delta p(x) = \frac{G_0\, \tau_p}{1 - \alpha^2\, L_p^2} \left[\exp\left(-\alpha\, x\right) - \frac{S + \alpha\, D_p}{S + D_p/L_p}\, \exp\left(-\frac{x}{L_p}\right) \right]. \qquad (10.73)$$

For vanishing surface recombination, $S = 0$, the solution is (Fig. 10.35)

$$\Delta p(x) = \frac{G_0\, \tau_p}{1 - \alpha^2\, L_p^2} \left[\exp\left(-\alpha\, x\right) - \alpha\, L_p\, \exp\left(-\frac{x}{L_p}\right) \right]. \qquad (10.74)$$

For $\alpha L_p \gg 1$, (10.65) is recovered. This dependence is the excess-carrier profile if the absorption is strong, which is a tendency for short wavelengths. The current at the surface, $j(x = 0) \propto \nabla \Delta p$, is zero.

In the case of very strong surface recombination, $S \to \infty$, (10.73) becomes

$$\Delta p(x) = \frac{G_0\, \tau_p}{1 - \alpha^2\, L_p^2} \left[\exp\left(-\alpha\, x\right) - \exp\left(-\frac{x}{L_p}\right) \right], \qquad (10.75)$$

with $\Delta p(0) = 0$ (Fig. 10.35). The current at the surface is ($D_p\, \tau_p = L_p^2$)

$$j(x = 0) = -e\, D\, \left. \frac{\partial \Delta p}{\partial x} \right|_{x=0} = -e\, \frac{G_0\, L_p}{1 + \alpha\, L_p} = -e\, \Phi_0\, \frac{\alpha\, L_p}{1 + \alpha\, L_p}. \qquad (10.76)$$