1 Introduction

Terminological notes

By the term 'luminescence of solids' we understand a surplus of the electromagnetic (light) radiation, emitted by a solid, over its equilibrium radiation that can be described by Planck's law.[1] At the same time this radiation has to have a decay time much longer than the period of light oscillations (10^{-14}–10^{-15} s). Above all, it follows from this definition that, from the thermodynamical point of view, the luminescence is a *non-equilibrium radiation*. This means that the solid needs to be supplied in some way with extra energy (with respect to that being exchanged by the solid with its surroundings through the equilibrium electromagnetic radiation). This extra energy is transformed inside the medium into the luminescence radiation. The supplied extra energy is called *excitation energy*, and the luminescence can be classified according to the way the excitation energy is applied, as follows:

Photoluminescence is excited with light (of wavelength λ_{ex}, which is usually shorter than the emission wavelength λ_{em}. The relation $\lambda_{ex} \leq \lambda_{em}$ is called Stokes' law).

Electroluminescence originates as a consequence of the application of an electric field and the relevant electric current flow through the material (do not confuse this with the thermal radiation due to the Joule heat!).

Chemiluminescence accompanies certain types of exothermic chemical reactions—the released heat or part of it is radiated in the form of light.

Bioluminescence coexists in a similar way with certain physiological biochemical reactions.

Cathodoluminescence arises when a high-energy electron beam (10^2–10^3 eV) impinges on a luminescent screen.

Mechanoluminescence is light (usually a short flash) emitted in some cases during mechanical deformation of a solid.

Thermoluminescence occurs when a solid is first cooled to a low temperature, then illuminated (excited) using short-wavelength electromagnetic

[1] The notion of light should be understood here in a wider sense, namely, not only visible radiation, but also the near-infrared and ultraviolet regions.

radiation and finally its temperature is allowed to increase slowly, which is accompanied by emission of the luminescence radiation.

Sometimes one can also encounter terms like *X-ray luminescence* (excitation by X-rays), *sonoluminescence* (excitation by acoustic or ultrasound vibrations), *triboluminescence* (luminescence due to friction when the material is scratched or crushed), etc. In recent years, light induced through injecting low-energy electrons or holes (energy of the order of 1 eV) into a semiconductor or metal in a scanning tunnelling microscope is being investigated. In the case of metals, however, the term luminescence seems inappropriate, as luminescence in solid-state physics is—for historical reasons—connected with non-metallic solids, insulators and semiconductors. Therefore, here the term *photon emission* is used instead.

The second part of the definition of luminescence, dealing with its finite decay time, differentiates luminescence from other types of so-called *secondary radiation*: reflected light, various types of scattered light (Rayleigh, Raman and Brillouin scattering) and Cerenkov radiation. The thing is that these kinds of radiation originate as an act of very fast photon–matter interaction and there is virtually no exchange of energy between the impinging photon and the electronic system of the solid. On the contrary, during the process of luminescence the electrons are excited to higher energy states, which entails genuine absorption of the excitation energy in the material and its subsequent gradual transformation. Consequently, the succession of all involved events lasts a relatively long time. For this reason, upon cessation of the excitation the luminescence continues to decay for some time. The lower bound of this time period gets shorter with the development of time-resolved spectroscopy techniques. Nowadays it can be considered to amount to hundreds of femtoseconds. (Of course, the scale of luminescence decay times is much broader and ranges from nanoseconds up to tens of hours.)

The study of luminescence has a long history, in the course of which the relevant terminology has also been developing. Radiating rotten stumps or certain kinds of luminescent insect and fish have been known in Nature from time immemorial. Probably more than a thousand years ago the Chinese and Japanese knew about luminescent dyes. In the seventeenth century, the 'Bolognian stone' was described, emitting red light upon prior exposure to sunlight (nowadays we know that it was barium sulphide, BaS). This and similar stones were given the name *phosphors* and the relevant effect, i.e. long-term light emission after cessation of the excitation radiation, was named *phosphorescence*.[2] Later on, in order to distinguish from phosphorescence, the term *fluorescence* was introduced to designate light emission with an immeasurably short decay after stopping the excitation; this was observed for the first time in fluorite, CaF_2. The general term *luminescence*, comprising both

[2] This has nothing to do with the chemical element P, phosphorus (in the sense that the luminescing substance would have to comprise P atoms), even if historical connection with phosphorus lighting itself can be traced; see Harvey E. Newton (2005). *A History of Luminescence. From the Earliest Times until 1900*. Dover Publications, Mineola.

phosphorescence and fluorescence, was finally introduced towards the end of the nineteenth century.

In the literature even the word *phosphor* occurs nowadays. Even if its meaning is not strictly specified, it usually denotes a luminescent solid in a wider sense. Also the meaning of the notions of phosphorescence and fluorescence has shifted in time—presently, fluorescence is sometimes understood as light emission occurring in the course of the excitation event, while in speaking of phosphorescence we have in mind non-equilibrium light emission observed during its decay, provided this decay is long enough to be observable by the naked eye. However, these terms are more exactly specified in the case of luminescence of organic substances.

Luminescence of organic and inorganic materials

Luminescence occurs in most organic substances (aromatic hydrocarbons like benzene, naphthalene, anthracene, organic dyes, etc.) as well as in many inorganic solids (ionic crystals, semiconductors). Its underlying origin, however, in both cases differs substantially. In organic matter, the role of a characteristic luminescence bearer (*luminescence centre*) is played by a molecule. This means that the essential features of the luminescence radiation (spectral content, decay time) in the solid state and in a solution are very much alike. This is because organic crystals are composed of molecules, the binding between them being mediated only by weak van der Waals forces. Therefore, the molecules keep their individuality to a large extent; the weak intermolecular forces are not strong enough to modify them substantially. The same, of course, is true in the case of a solution, thus the luminescence spectrum of organic matter has virtually the same shape in solution as well as in the solid state.

The basic features of organic luminescence can thus be derived from the electron energy level scheme of large organic molecules (the Jablonski diagram). As is well known, the ground state of such a molecule is represented by a singlet state S_0, higher lying electronic states being mainly excited singlet states (S_1, S_2, S_3, ...). The higher excited singlet states S_2, S_3, ... relax very rapidly to the S_1 state and the transition $S_1 \rightarrow S_0$ is accompanied by emission of a luminescence photon. This radiative transition is spin-allowed and, consequently, is fast (the characteristic time is of the order of 10^{-9}s) and is called fluorescence. There are, however, also excited triplet states T_1, T_2, T_3, ... The transition $T_1 \rightarrow S_0$ can also be accompanied with photon emission, but is spin-forbidden and thus slow (10^{-3}s). It is called phosphorescence.

As for inorganic solids, especially semiconductors, speaking of a 'molecule' loses its meaning to a large extent (a molecule of silicon—does it exist?). Physical properties here, above all the development of the forbidden energy gap, etc., are conditioned by the existence of a minimal ensemble of at least several tens of atoms (sometimes called a 'cluster'). However, real bulk properties can only be achieved in substantially larger objects with lateral dimensions at least 1μm that contain approximately 10^{10} atoms. We shall treat these bulk semiconductor luminescence properties in Chapters 5–11. Low-dimensional semiconductor structures are examined in Chapters 12– 17.

Intrinsic and extrinsic luminescence

Solid-state luminescence can be divided into two basic types:

- *intrinsic* ('proper');
- *extrinsic* ('improper').

Intrinsic luminescence originates in an ideal, pure and defect-free crystalline lattice, while extrinsic luminescence has its origin in lattice defects or impurities. (A similar classification can be applied also to non-crystalline, disordered solids, provided we have in mind topological disorder of the amorphous network instead of the crystalline lattice.) A luminescence-active impurity (admixture) atom, ion or molecule is frequently called an *impurity luminescence centre*.

Now, a natural question arises. If—in the extrinsic case—the luminescence radiation originates in a microscopic impurity centre, what then is the role of the host solid itself, i.e. the crystalline or amorphous matrix? This matrix fulfils multiple functions: (i) First of all, it represents a host medium inside which the luminescence centres are fixed, statistically dispersed and mechanically isolated. However, this is by no means all; if we imagine these centres located in a similar way in a vacuum, they would either not emit any radiation at all or such radiation would be substantially weaker and of different spectral content. This means that (ii) the matrix also serves as an 'antenna' capturing the excitation energy and transferring it very efficiently to the luminescence centres. Next, (iii) owing to the interaction of the impurity electronic system with the matrix vibrations the electron energy levels undergo important modifications, which leads to substantial alterations of the optical spectra of the centre. Finally, (iv) in the special case of electroluminescence the matrix must ensure suitable electrical conductivity for exciting the centres.

At present we know that in the beginning of the era of the modern quantitative study of luminescence phenomena in solids (dating roughly from the first third of the nineteenth century to the 1950s), physicists and chemists were working exclusively with impurity, extrinsic luminescence. At that time, technology was not able to prepare sufficiently pure materials, while it is known today that most semiconductors owe their strong luminescence radiation to the presence of impurities. Even a negligible amount of residual impurity (1 ppm or less) can manifest itself in luminescence of enormous intensity, which masks or totally prevents intrinsic emission. This phenomenon, which will be discussed in Chapter 7, forms the basis of a sensitive qualitative and quantitative photoluminescence analysis of impurities in semiconductors. As a typical example of extrinsic emission we quote the luminescence of Cu^{+} or Mn^{++} ions in ZnS crystals, while the most characteristic example of intrinsic emission is luminescence due to the radiative decay of free excitons.

Examples of luminescence spectra

This book contains in its title the word *spectroscopy*. It follows that we shall largely deal with the interpretation of the emission spectral line shape of semiconductor luminescence. One of our goals is therefore aimed at analysing as much of the characteristic features of the experimentally acquired emission spectrum as possible, in order to assign the given spectral line to a specific microscopic luminescence centre or to some other channel of radiative recombination. Let us introduce here several examples of luminescence spectra.

Figure 1.1 shows near-infrared photoluminescence spectra of three crystalline Si samples, measured at liquid-helium temperature (4.2 K) [1]. The A sample represents extremely pure silicon without intentional doping (i.e. without introducing electrically active impurities), with a low concentration of residual unintentional impurities—in the order of $10^{12}\,\mathrm{cm}^{-3}$ only. The B sample contains impurities in the order of $\sim 10^{14}\,\mathrm{cm}^{-3}$ and the sample C has even higher impurity concentration, $> 10^{15}\,\mathrm{cm}^{-3}$. The notations of individual emission lines and bands will be clarified in Chapters 7 and 8. The spectra are rich in structure, comprising both very narrow (almost atomic) lines and relatively broad bands. It might be interesting to draw the reader's attention to the different characters of the spectral components: e.g. the lines denoted I_{LO}(FE) and I_{TO}(FE) are, as well as the band $I_{LO,TO}$(EHL), of intrinsic origin, while the lines B_{TO}(BE), P_{TO}(BE), Sb_{TO}(BE) are of extrinsic origin.

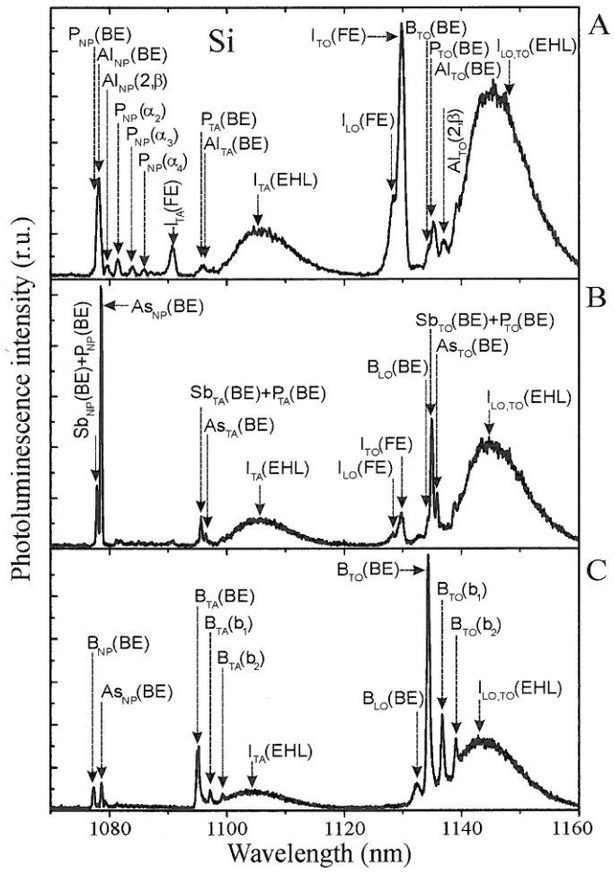


Fig. 1.1
Photoluminescence spectra of three samples of crystalline silicon close to the bandgap energy. The B and C samples were doped with (Sb + As) and (B + As), respectively; the A sample was not intentionally doped (nominally pure silicon). The samples were immersed in liquid helium ($T = 4.2\,\mathrm{K}$) and excited with a continuous-wave (cw) Ar^{+}-laser, $\lambda_{ex} = 488\,\mathrm{nm}$. After Valenta [1].

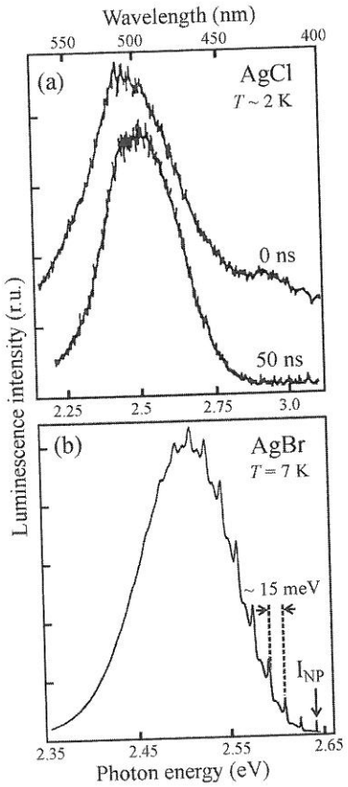


Fig. 1.2
(a) Photoluminescence spectra of pure AgCl at liquid-helium temperature under excitation with a pulsed N_2-laser ($\lambda_{ex} = 337$ nm). Numbers near the curves denote temporal delays after the excitation pulse at which the spectra were recorded. After Pelant and Hála [2]. (b) Photoluminescence spectrum of nominally pure AgBr at $T = 7$ K. The fine structure superimposed on the broad band is due to interaction with lattice vibrations—phonons. After Wassmuth *et al.* [3].

Figure 1.2(a) displays photoluminescence spectra of a pure silver chloride, AgCl, crystal, measured also at low temperature, 2 K, and under various modes of excitations [2]. Even if silver halides AgCl, AgBr are not typical semiconductors, their optical and, in particular, luminescence properties render them close to the family of wide-bandgap polar semiconductors. (Interestingly, at the same time, their mechanical properties are similar to those of metals.) In this context, they represent important model substances and throughout the book we shall often refer to their luminescence behaviour. The spectra contain a very broad emission band (full width at half maximum, FWHM, ~ 0.4 eV) at ~ 500 nm (2.5 eV) and, besides this, an indication of a weaker band at ~ 420 nm. Both bands are of intrinsic origin. The spectrum is situated in the visible region—under ultraviolet excitation and at low temperature AgCl exhibits very bright blue-green luminescence. Figure 1.2(b) shows the low-temperature (7 K) photoluminescence spectrum of nominally pure crystalline silver bromide, AgBr [3]. This material closely resembles AgCl both in chemical composition and in crystalline structure, and also the spectra displayed in Figs 1.2(a) and (b) are very much alike in their spectral position as well in their width. Despite these similarities, there is a fundamental difference in the origin of luminescence: Contrary to AgCl, the emission band at 2.5 eV in AgBr has extrinsic origin (!). (Another interesting observation is the 'fine structure' superimposed on the broad emission band in AgBr, which is completely absent in AgCl. The interpretation of these effects will be given in Chapter 7.) Finally, Fig. 1.3 demonstrates room-temperature emission spectra of zinc sulphide ZnS doped with various intentional impurities [4]. These broad extrinsic emission bands cover most of the visible region.

From these examples we can already draw several conclusions. First of all, it becomes quite obvious that—unfortunately—neither the shape nor the width of the emission spectrum can be used for judging the intrinsic or extrinsic origin of luminescence. Undoubtedly, this would be very profitable, but Nature does not reveal its mysteries so easily. Next, the reader has probably noticed

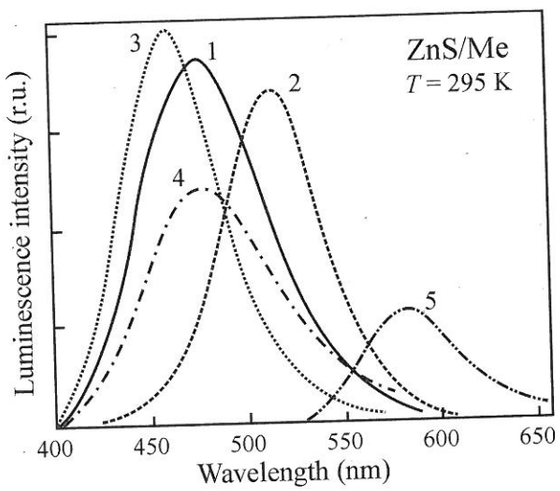


Fig. 1.3
Cathodoluminescence spectra of ZnS doped with various metals (Me), the so-called activators: 1. ZnS/Zn, 2. ZnS/Cu, 3. ZnS/Ag, 4. ZnS/Au and 5. ZnS/Mn. $T = 295$ K. After Markovskii *et al.* [4].

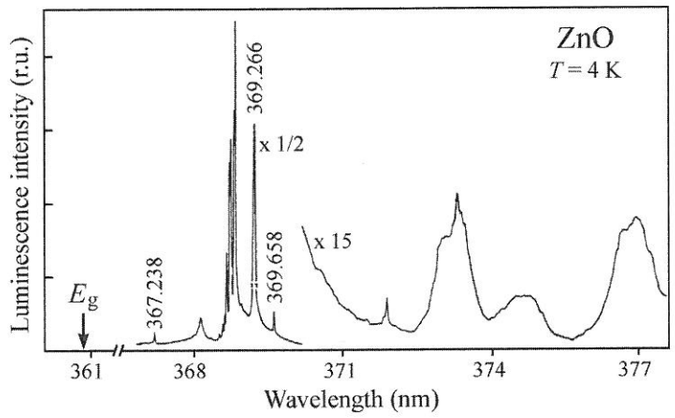


Fig. 1.4
Cathodoluminescence spectrum of edge emission in a nominally pure ZnO crystal. Excitation by bombardment with a 7 keV electron beam, $T = 4$ K. Adapted from Tomzig and Helbig [5]. The arrow labels the position of the bandgap $E_g = 3.436$ eV (~ 360.9 nm).

that most of the presented spectra are photoluminescence spectra, which anticipates the fact that the book is devoted predominantly to photoluminescence spectroscopy; the theory of electroluminescence and its applications are concentrated mainly in Chapter 11. Furthermore, the spectra displayed in Figs 1.1 and 1.2 were measured at very low temperatures. The reason for this is that the luminescence of most solids occurs only upon cooling them considerably below ordinary room temperature. And even if some materials luminesce at room temperature (Fig. 1.3), the intensity of this luminescence, as a rule, grows markedly in the course of cooling. This property contradicts flagrantly the ordinary behaviour of thermal radiation of solids (*incandescence*), described by the Planck and Stefan–Boltzmann laws, and that is why luminescence is sometimes known as *cold light*. The drop of luminescence intensity with increasing temperature is called *thermal quenching*.

Finally, in semiconductor luminescence one can encounter the term *edge emission*. This means luminescence radiation situated in the spectral range around the onset of interband optical absorption (absorption edge). More specifically, the energy of the luminescence photon $h\nu_{em}$ is only slightly less than the bandgap energy E_g. The edge emission spectrum usually consists of a large number of narrow (excitonic) lines, both intrinsic and extrinsic. An example of the low-temperature edge emission in zinc oxide, ZnO, is shown in Fig. 1.4 [5]. The bandgap energy $E_g = 3.436$ eV is marked by an arrow. The displayed wavelength range is very narrow, about 10 nm (90 meV); the individual lines are shifted with respect to E_g by 0.01–0.1 eV only. Similarly, the spectra shown in Fig 1.1 can also be specified as edge emission.

References

1. Valenta, J. (1994). *Photoluminescence characterization of selected semiconductors and insulators*. PhD Thesis, Charles University in Prague, Faculty of Mathematics and Physics, Prague.
2. Pelant, I. and Hála, J. (1991). *Solid State Com.*, **78**, 141.

3. Wassmuth, W., Stolz, H., and von der Osten, W. (1990). *J. Phys. C: Cond. Matter*, **2**, 919.
4. Markovskii, L. J., Pekerman, F. M., and Petoshina, L. N. (1966). *Phosphors* (in Russian: *Ljuminofory*). Chimija, Moskva, Leningrad.
5. Tomzig, E. and Helbig, R. (1976). *J. Luminescence*, **14**, 403.

Experimental techniques of luminescence spectroscopy

2

The aim of experimental luminescence spectroscopy is basically the same as that of other types of optical spectroscopy, namely, to disperse luminescence radiation into a spectrum, which is subsequently detected and recorded. Here, as a dispersion device, a monochromator or a polychromator is almost exclusively used; only exceptionally (in the case of Fourier luminescence spectroscopy) is an interferometer applied. Possible application of devices based on multiple-beam interference (Fabry–Perot interferometer, Lummer–Gehrke plate) brings no advantage as the luminescence spectra are usually broad and do not require extreme spectral resolution. Only exceptionally, e.g. in the case of narrow spectral lines of edge emission (Figs 1.1 and 1.4), a high-resolution monochromator or possibly a Fourier spectrometer with resolving power $R = (\lambda/\Delta\lambda)$ amounting at least to $\sim 5 \times 10^3$ ($\Delta\lambda$ denotes the minimum separation between two spectral lines around a central wavelength λ that are to be resolved) is indispensible.

Generally, what is characteristic for luminescence spectroscopy is typically an extremely low level of detected light flux rather than requirements for extreme spectral resolution. Therefore, the spectral device must have the highest possible throughput, and special care must be devoted to selecting a sensitive photodetector and, in particular, to setting up an efficient optical system to collect the luminescence signal. All these aspects will be discussed in this chapter.

2.1 Emission and excitation spectra

The purpose of monochromators in luminescence spectroscopy is, on the one hand, to observe the spectral content of luminescence emission and, on the other hand (in the case of photoluminescence) they can also be applied to select a suitable excitation wavelength from an optical excitation source. The luminescence signal can be measured either as a function of emission wavelength at a fixed excitation wavelength or vice versa. One speaks of *emission* and *excitation spectra*, respectively.

A general scheme of luminescence experiment is shown in Fig. 2.1(a). The excitation source supplies excitation energy to the sample, emitted

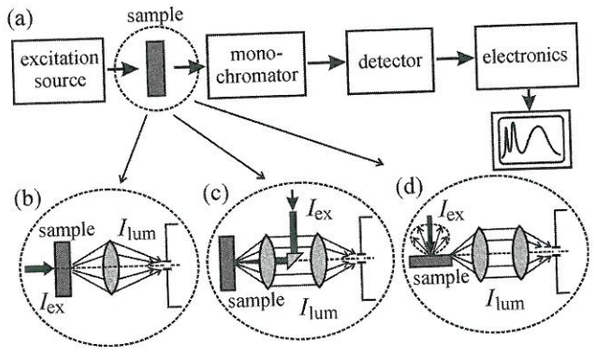


Fig. 2.1
(a) Block diagram of a luminescence experiment, (b) the 'transmission' geometry, (c) the 'back-reflection' geometry, (d) collection of luminescence from the sample edge using a waveguiding effect.

luminescence radiation is dispersed in a spectral device, a detector converts the optical signal into an electric signal that is then processed by electronic devices and finally by a control computer. The output in the form of a plot of luminescence intensity I_{lum} versus emission wavelength λ_{em} is called an *emission spectrum* (all the spectra shown in Figs 1.1–1.4 are emission spectra).

Spontaneous luminescence emission emerges from a sample practically in all directions, even if in many cases the sample might not be a perfect isotropic emitter. Various optical systems collecting the emitted light can be implemented. The simplest one is shown in Fig. 2.1(b). This is the so-called transmission geometry, where the optical axes of the excitation and luminescence rays coincide. Luminescence is thus detected from the opposite, non-excited side of the sample. This geometry is simple and easy to adjust but the luminescence signal could be significantly reduced by light scattering and especially by reabsorption in the sample. Since reabsorption may be spectrally dependent, the emission spectrum will be distorted. In addition, a significant fraction of non-absorbed excitation light could enter into the monochromator and cause problems by saturating or even damaging the detector. Therefore, the transmission geometry is seldom used.

More routinely applied is reflection geometry shown in Fig. 2.1(c). Luminescence is collected from the same spot on which the exciting radiation is focused by a lens (or an objective). The excitation beam is directed perpendicular to the optical axis of a collection system and the exciting light is sent towards the sample by means of a small totally reflecting prism (when the excitation is provided by a well-collimated narrow laser beam) or a dichroic filter (which reflects short wavelengths, i.e. excitation, and transmits longer wavelengths, i.e. luminescence). One and the same lens (objective) is employed to concentrate excitation light and to collect luminescence emission. The lens should have a very large numerical aperture in order to collect light from a wide solid angle. The collimated luminescence light is then focused by means of a second lens onto the entrance slit of a spectral device. Among merits of this geometry are the above mentioned large collection solid angle (remember that luminescence is weak secondary radiation, so the largest possible portion of the emitted power must be collected for successful spectral analysis, see Section 2.5), reduced reabsorption of luminescence radiation, as well as avoidance of stimulated emission (which will be explained later). As a

possible drawback, complicated optical adjustment required for achieving both optimum excitation power density on the sample and simultaneous maximum luminescence collection can be mentioned. Another undesirable effect may be possible reflection of a significant part of the excitation light into a detection system, but that is easily remedied by slightly rotating the sample to reflect the exciting beam out of the detection cone.

In particular cases, when the sample takes the form of thin layer(s) with high-quality surfaces, the geometry sketched in Fig. 2.1(d) (an extreme case of the reflection geometry, excitation and emission axes being perpendicular to each other) may be advantageous. Such samples may act as a waveguide into which luminescence is coupled and thereby concentrated at the output edge. In comparison with the previous case, this scheme enables the excitation and emission optical paths to be adjusted independently, and it avoids reflection of the excitation beam into the detection system. The most serious problem that might appear is the possible occurrence of stimulated emission: if population inversion is (unintentionally) achieved, then in semiconductors with high probability of radiative recombination the luminescence (spontaneous emission) guided through the excited volume may be amplified by stimulated emission. This effect can be reduced or avoided by limiting the size of the excited area, but even a small portion of stimulated emission could significantly distort the shape of the emission spectrum [1, 2]. On the other hand, this geometry is convenient especially when stimulated emission and the presence of optical gain are the subjects of investigation. Details on such experiments and measurement techniques are provided in Chapter 10.

Two particular set-ups for measurements of photoluminescence emission spectra are presented in Fig. 2.2 [3, 4]. The experimental arrangement in Fig. 2.2(a) represents a particular case of the collection geometry sketched schematically in Fig. 2.1(c). A continuous-wave (cw) Ar^+-laser with the most intense lines at 488 and 514 nm is used as an excitation source. The sample is immersed in liquid helium inside a helium-bath cryostat (the temperature of the boiling point of helium at normal pressure is 4.2 K). Luminescence is spectrally dispersed in a double-grating monochromator and detected by a photomultiplier tube (PMT) placed behind the exit slit. The PMT is cooled down by liquid nitrogen (the temperature of the boiling point at normal pressure is 77 K). The signal from the PMT is processed by a photon counter and plotted, as a function of wavelength, by a computer.

Another frequently used set-up is shown in Fig. 2.2(b). The excitation source is represented by a high-pressure mercury lamp which has one of its dominant lines at 365 nm (in the near-UV spectral region) that is almost an ideal universal excitation wavelength for most semiconductors. (Nowadays, mercury lamps have been almost completely replaced by diode lasers or (UV adapted) Ar^+-lasers.) From the point of view of geometry, this set-up is halfway between Figs 2.1(c) and 2.1(d), so-called 'reflection at 45°'. The sample is, again, placed in a cryostat whose single window is replaced by a collection lens. The diameter of the lens defines the collection solid angle. Another lens focuses the collimated luminescence beam into a point where an optical chopper (a rotating disk with regularly arranged slots) periodically chops the beam.

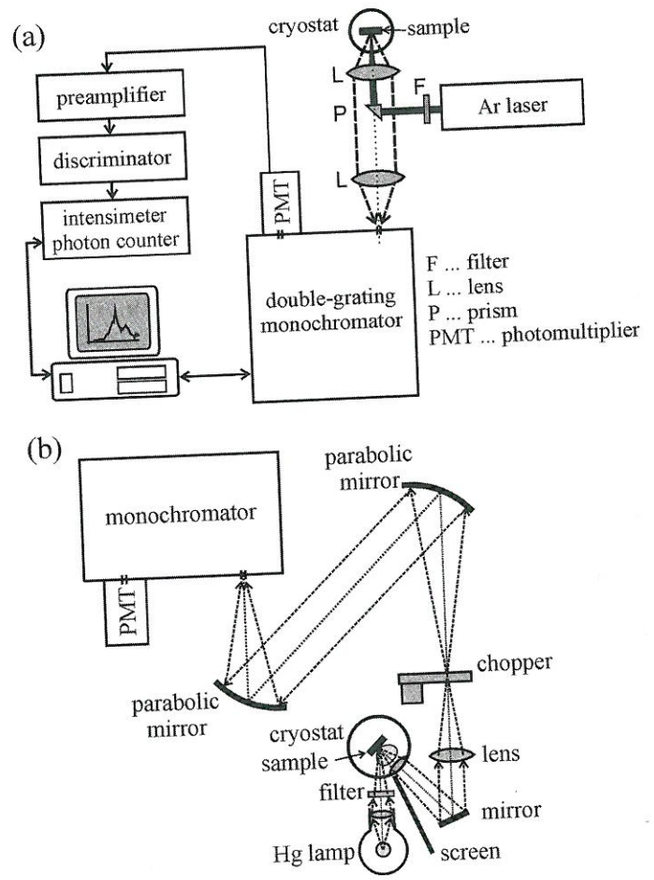


Fig. 2.2
(a) Experimental set-up for detection of photoluminescence emission spectra in a back-reflection geometry with double-lens collection system [3]. (b) Set-up for detection of emission spectra with the collection system combining lenses and two parabolic mirrors. This geometry is called 'reflection at 45°'. Adapted from Bebb and Williams [4]. The apparatus contains an optical chopper inserted into the luminescence beam, enabling application of a lock-in detection technique (Subsection 2.4.1).

The rectangular temporal trace of the luminescence signal is essential for treatment of the signal by a lock-in technique that will be discussed in Subsection 2.4.1 (in the case of chopping an unfocused luminescence beam the temporal profile of the signal would be trapezoidal and, consequently, the final signal-to-noise ratio will be degraded). Before entering the monochromator, the luminescence radiation must be refocused onto the entrance slit. This is done by a pair of parabolic mirrors; these cover a smaller area than a lens system that would play the same role, thus saving room on the optical table and, moreover, minimizing optical abberations. The detector behind the exit slit of the monochromator—again a photomultiplier tube—provides the output signal that is processed by a lock-in amplifier.

The photoluminescence *excitation spectrum* is a plot of the photoluminescence intensity I_{PL} (at a fixed emission frequency $\nu_{em} = c/\lambda_{em}$) as a function of the excitation photon energy $h\nu_{ex}$:

$$I_{PL}(\nu_{em}) = f(h\nu_{ex});\ \nu_{em} = \text{const.}$$

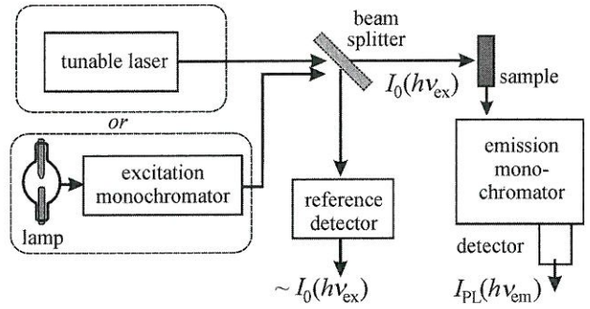


Fig. 2.3
Block diagram of an apparatu
excitation spectra measureme

A schematic sketch of an experimental set-up for recording excitation spectra is shown in Fig. 2.3. In order to understand the usefulness of the excitation spectra, let us investigate firstly the dependence of the photoluminescence intensity $I_{PL}(\nu_{em})$ on the excitation light intensity. Suppose we have a sample in the form of a plane-parallel plate with thickness d and an excitation intensity $I_0(h\nu_{ex})$ impinging on it. The transmitted excitation intensity is given by the Lambert–Beer law

$$I(h\nu_{ex}) = I_0(h\nu_{ex}) \exp[-\alpha(h\nu_{ex})d], \tag{2.1}$$

where $\alpha(h\nu_{ex})$ is the absorption coefficient. If we consider the correction for reflectivity R on the front face of the sample (for the sake of simplicity R is assumed to be independent of wavelength), the absorbed part of the excitation intensity is equal to $I_0(h\nu_{ex})\,(1-R)\,(1-\exp[-\alpha(h\nu_{ex})d])$.

The photoluminescence intensity, being proportional to the absorbed energy (possible nonlinear effects are neglected here) then reads

$$I_{PL}(\nu_{em}) \cong \eta I_0(h\,\nu_{ex})\,(1-R)\,(1-\exp[-\alpha(h\nu_{ex})d]). \tag{2.2}$$

The proportionality coefficient $\eta \leq 1$ introduced here is called the *efficiency* or *quantum yield of luminescence*. In general it depends on many variables: $\eta = \eta(h\nu_{ex}, h\nu_{em}, T, \ldots)$. Now we should distinguish two extreme cases:

1. Weakly absorbing sample $\alpha(h\nu_{ex})d \ll 1$. By applying the common approximation of the exponential function $\exp(-\alpha d) \approx (1-\alpha d)$, we obtain from (2.2)

$$I_{PL}(\nu_{em}) \cong \eta\, I_0(h\nu_{ex})\,(1-R)\,d\,\alpha(h\nu_{ex})$$

and, after dividing by a reference spectrum of the excitation source $I_0\,(h\nu_{ex})$, we find

$$\frac{I_{PL}(\nu_{em})}{I_0(h\nu_{ex})} \sim \alpha(h\nu_{ex}). \tag{2.3}$$

This means that the excitation spectrum, normalized in this way, reproduces the shape of the absorption spectrum. This can be applied to obtain (relative) absorption spectra in samples in which standard transmission spectroscopy fails, e.g. in thin semiconductor layers on an opaque (i.e. strongly absorbing or scattering) substrate, in measuring low concentrations of luminescent dopants in a strongly absorbing non-luminescent host material, in quantum wells or

superlattices, etc. This is one of the motivations for investigating photoluminescence excitation spectra. (It is worth pointing out that such measurements of optical absorption spectra can also be performed in samples of a very irregular shape, in small fragments or even in powder materials. This is a general merit of most luminescence measurements—they can be performed without paying special attention and effort to preparing a sample of a perfectly defined shape, size, surface quality, etc.)

The second motivation for measuring excitation spectra is purely practical. As mentioned above, luminescence radiation is generally very weak; hence all means of optimizing the luminescence signal should be used. In this sense the excitation spectrum, once acquired, gives a clear hint about the choice of an optimum excitation wavelength—this is a wavelength (or excitation photon energy) at which the excitation spectrum is peaked.

As an example of a photoluminescence excitation spectrum we present that of crystalline quartz, containing a pronounced structure (see Fig. 2.4(a) [5]). A narrow excitonic peak around 8.7 eV as well as onsets of indirect and direct absorption edges around 9 eV and 10.7 eV can be identified. Note that the emission spectrum is significantly red-shifted from the deep-UV excitation peaks to the near-UV and blue spectral regions (between 2 and 5 eV, see Fig. 2.4(b)). Such a red-shift between excitation and emission peaks is often called the *Stokes shift*; a more exact definition will be given in Section 4.5. A sizeable Stokes shift, of several electron-volts like in Fig. 2.4, indicates low power efficiency of luminescence, and most likely also the occurrence of a complicated mechanism of excitation energy transfer from a light-absorbing system to luminescence centres. This is additional interesting information revealed by the excitation spectra.

Let us recall that all of the above discussion is valid in the case of weak absorption of excitation light, when relation (2.3) holds true. Now we are going to discuss the opposite case.

2. Strongly absorbing sample $\alpha(h\nu_{ex})d \gg 1$, i.e. $\alpha \approx 10^3$–10^5 cm^{-1} and $d \geq 10^{-2}$ cm. In this case $\exp[-\alpha(h\nu_{ex})\, d] \ll 1$ and (2.2) is reduced to

$$I_{PL}(\nu_{em}) \approx \eta\, I_0(h\nu_{ex}). \tag{2.4}$$

This means that the excitation spectrum basically mimics the shape of the emission spectrum of the excitation source while information about the sample under study is lost. The condition $\alpha(h\nu_{ex})d \gg 1$ is valid for most semiconductors under deep band-to-band excitation, i.e. when $h\nu_{ex}$ is much larger than the bandgap energy E_g. Here, excitation spectra no longer bring useful information. In a number of experiments between the extremes 1 and 2, the usefulness of excitation spectra measurements must be assessed from case to case.

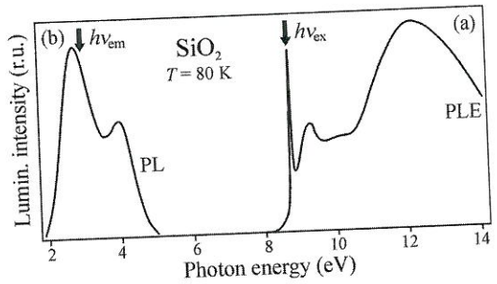


Fig. 2.4
(a) Photoluminescence excitation spectrum (PLE) of crystalline SiO_2 detected at $h\nu_{em} = 3.0$ eV and $T = 80$ K. (b) Photoluminescence emission spectrum (PL), excited by photons with an energy of $h\nu_{ex} = 8.7$ eV. Adapted from Itoh *et al.* [5].

Note: From the experimental point of view eqn (2.4) can be exploited to monitor excitation intensity variations. When the quantum efficiency η of a luminescent material (or its solution—e.g. Rhodamin B in ethylenglycol) strictly satisfies the requirement of being independent of excitation wavelength, $\eta \neq \eta(h\nu_{ex})$, then this material (the so-called quantum counter) may be applied in combination with a photomultiplier to measure the emission spectrum of an excitation source. The advantage of this mode of measurement against using detectors with a spectrally flat response (thermal detectors) is its much higher sensitivity.

2.2 Types of photodetectors

In Fig. 2.2 a photomultiplier tube (PMT) is used to detect the spectrally dispersed luminescence signal. This is the most common highly sensitive photodetector for luminescence spectroscopy, even if nowadays it is often replaced by an avalanche photodiode or by an array of small semiconductor devices—usually by a charge-coupled device (CCD) camera. We shall briefly describe these three types of detectors.

Photomultiplier tube

Detection of optical radiation in a photomultiplier tube is based on the external photoelectric effect and secondary emission of electrons. A photomultiplier is an evacuated glass (or quartz) tube containing a photocathode (from which incident photons eject electrons—with quantum efficiency $\eta_{ph} < 1$—due to the external photoelectric effect), an array of dynodes (where primary electrons are multiplied through the emission of secondary electrons) and an anode (the collector of electrons, see Fig. 2.5(a)). The outlets of all electrodes pass through a socket of the photomultiplier. A photon can eject an electron from the photocathode only when the photon energy $h\nu$ exceeds the work function ϕ of the given photocathode material, $h\nu \geq \phi$. To detect (optical) photons from the visible range ($h\nu \approx 1$–4 eV), the photocathode must be made of either an alloy containing alkali metals with low work function ϕ (e.g. $\phi = 1.8$ eV in Cs,) or a semiconductor with relatively low values of both the work function ϕ and bandgap E_g (GaAs). The applied photocathode material determines the spectral sensitivity of the detector, especially its long-wavelength cut-off, while the short-wavelength edge is determined mostly by the material of the input window.

Photoelectrons are accelerated and focused onto the first dynode, where each of these primary electrons can release δ (between 3 and 10) secondary electrons. These secondaries, upon being accelerated, hit the next dynode, and the secondary emission of electrons is repeated there so that an avalanche increase in the number of electrons is achieved: if n is the number of dynodes, the number of electrons that hit the anode will be equal to $G = \delta^n$. The quantity G is called the current amplification or photomultiplier gain and attains values

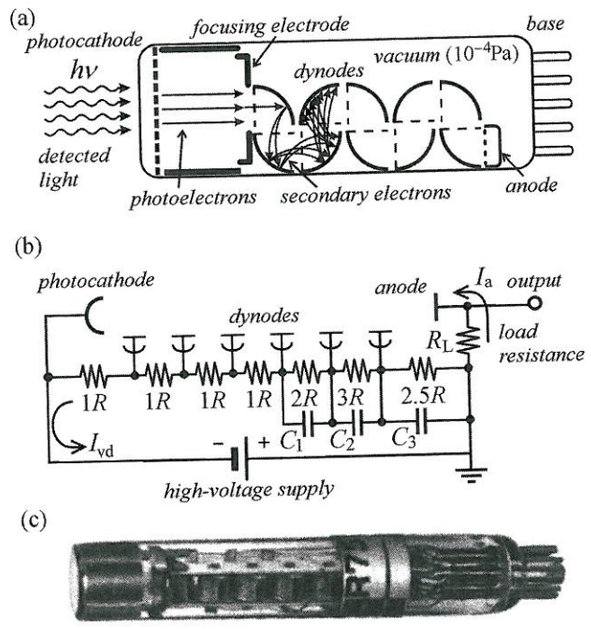


Fig. 2.5
(a) General scheme of a photomultiplier tube, (b) circuitry of the resistor voltage divider between dynodes, where I_a denotes the anode current and I_{vd} is the current flowing across the voltage divider, (c) a Hamamatsu R750 photomultiplier tube (the tube length and diameter are 88 and 20 mm, respectively). Adapted from a Hamamatsu brochure.

of 10^5–10^7. Dynodes are mostly fabricated from nickel or BeCu and are covered by a material with high coefficient δ of secondary emission (BeO, GaP, or KCsSb). The electron flux, collected by the anode, flows across a load resistance and gives an output electric signal, which is then treated by electronic circuits.

In order to accelerate the primaries as well as the secondaries, an electric field must be supplied between the relevant electrodes. This is done by applying a high voltage (usually between 1000 and 2000 V) between the photocathode and anode through a resistor voltage divider (Fig. 2.5(b)). The resistors are connected in series; the ratios of their values, defining how the applied voltage is divided between electrodes, are specified by the manufacturer and must be strictly maintained. Nowadays most manufacturers offer sockets with a built-in resistor divider for any type of photomultiplier tube. A typical value of a single resistor is of the order of 100 kΩ, determining the current through the divider to be in the order of mA (in the case of $n = 10$ dynodes and an applied voltage of 1000 V).

There are many various construction types of photomultiplier tubes. The photomultiplier sketched schematically in Fig. 2.5(a) and an overall view of which is shown in Fig. 2.5(c), has a so-called linear focus multiplier; other common types of multiplier structures are called 'Venetian blinds', boxes and grids, and circular cages. Besides, the photocathode may be deposited onto the inner surface of either the front side of the cylindrical tube (head-on type, Fig. 2.5(c)) or its side wall (side-on type). A more detailed description and comparison of different types of photomultipliers goes beyond the scope of this book and can be found in the dedicated literature, e.g. [6, 7].

We shall describe concisely the technical parameters of photomultiplier tubes. Photomultipliers owe their widespread use to, firstly, the already mentioned high gain G and also to their *spectral range of sensitivity*. This range extends between about 150 and 1100 nm; in special cases it can go up to 1400 nm. It should be stressed, however, that the entire spectral domain cannot be covered by a single type of photomultiplier. The selection of an appropriate photomultiplier type thus has to be based primarily on considerations about the spectral range of the detected light. The *anode sensitivity* $k(\lambda)$ is given as the anode current produced by 1 watt (in energy units) or 1 lumen (in photometric units) of incident radiant or luminous flux Φ, respectively:

$$k(\lambda) = \frac{\text{output signal}}{\text{incident radiant power}},$$

$$k(\lambda) = \frac{S}{\Phi} = \frac{S}{I\,(\mathrm{W/cm^2})\,\mathscr{A}\,(\mathrm{cm^2})} \quad \text{(A/W) or (A/lm)}, \qquad (2.5)$$

where $\mathscr{A}$ denotes the sensitive area of the photomultiplier. Typical values of $k(\lambda)$ are 10^3–10^5 A/W.

The above mentioned *quantum efficiency* η_{ph} of the photomultiplier tube is defined as the probability of ejecting from the photocathode one electron per incident photon (η_{ph} is typically about 0.1–20%). A very important parameter for properly operating a photomultiplier is its *maximum anode current* I_{am}, whose typical value is usually 10–100 μA and must not be exceeded (i.e. the incident radiant flux must be kept below a certain level), otherwise the excessive heating of the electrodes will induce permanent damage to the photomultiplier (due, e.g., to partial evaporation of the sensitive layer of the photocathode) which manifests itself as a decrease in sensitivity or an increase in background noise. In addition, for proper functioning of a photomultiplier tube (linearity of its response) the resistors in the voltage divider must be chosen so that the current flowing across the divider is at least an order of magnitude higher than I_{am} (Problem 2/1).

This definition of sensitivity does not take into account *noise* and the so-called *dark current* of the photomultiplier tube. The dark current I_d is a current due to the thermionic emission of electrons from the photocathode, which takes place even in darkness when no light strikes the photocathode. The thermionic current density is described by the well-known Richardson–Dushman equation [8], and thus increases exponentially with increasing temperature and decreasing work function ϕ of the photocathode. Photomultiplier tubes sensitive in the red and near-infrared spectral regions suffer from the highest dark currents, as the work function of their photocathodes must be low. Typical values of the anode dark current lie between a few nA up to hundreds of nA. Dark current at the output of a photomultiplier cannot be distinguished from the real (light-induced) signal, which is troublesome especially for very low signal levels—dark current obviously increases the limit of the lowest detectable signal (the detection limit of a photomultiplier tube). It is therefore desirable to suppress the dark current as much as possible. The strong (exponential) dependence of the dark current on temperature enables us to reduce it significantly by cooling the photomultiplier. The lowest allowed operational temperature is

given by the manufacturer and must not be exceeded because mechanical deformations at lower temperatures can induce permanent mechanical damage. Cooling can be provided by means of a thermoelectric (Peltier) element (down to about −25°C with a single-stage device) or by liquid nitrogen (boiling point −195°C = 77 K). The cooling housings for different types of photomultiplier tubes are commonly offered by manufacturers or they can be home-built (which is not an easy task, though). Sometimes, 'for less demanding applications', photomultipliers can be used even without cooling, but that depends on the photomultiplier type and even the particular photomultiplier piece of choice.

The dark current I_d can thus be suppressed to a large extent by cooling. Alternatively, because the dark current contribution to the resulting output signal has the form of a direct current (DC) component, it can be filtered out when measuring alternating current (AC) signals. A similar approach cannot, unfortunately, be applied to reduce noise in the output signal—see, e.g., Figs 1.1 and 1.2, where we reproduce luminescence spectra as recorded by a plotter after being treated by the electronic part of the detection channel. The spectra in Fig. 1.1 were detected by using a cooled (77 K) photomultiplier tube with a near-infrared sensitive photocathode while the spectra in Fig. 1.2(a) were obtained by an uncooled photomultiplier tube with a semiconductor-type photocathode. In order to characterize the noise properties of a photodetector, the quantity termed *noise equivalent power* (*NEP*) (or equivalent noise input—*ENI*) is introduced. This means the incident light power which would induce an output signal equal to the root mean square of the intrinsic (dark) noise N (Fig. 2.6). The detected noise increases, of course, with increasing transmission bandwidth Δf of the entire measurement system; it turns out that the noise amplitude scales with the bandwidth like $(\Delta f)^{1/2}$. So the detected noise N can be written down as

$$N = k(\lambda)\, NEP\sqrt{\Delta f}, \tag{2.6}$$

where $NEP = (2e\, I_d G)^{1/2}/k(\lambda)$ [7] is a parameter already independent of the measurement system, i.e. solely characterizing the photodetector. According to the definition, the signal-to-noise ratio (S/N) is equal to unity when the incident light power is equal to *NEP*. In other words, *NEP* represents the lowest signal level (detection limit) which can be detected by a given detector; weaker signals will remain buried in noise. Referring to eqn (2.6), *NEP* is expressed in $W/Hz^{1/2}$. Thus this is not directly power (in watts), but one usually supposes that $\Delta f = 1$ Hz and then $NEP(\Delta f)^{1/2}$ (W) and $NEP(W/Hz^{1/2})$ in (2.6) are numerically equal to each other. A typical value of *NEP* for photomultipliers amounts to 10^{-16}–10^{-15} W.

One can also introduce related quantities: the *detectivity* D and *normalized detectivity* D^*. The detectivity D is simply the reciprocal value of $NEP (D = 1/NEP)$, and the normalized detectivity is defined as

$$D^* = \sqrt{\mathscr{A}}/NEP = D\sqrt{\mathscr{A}}\ \ (\mathrm{cm\,Hz^{1/2}/W}). \tag{2.7}$$

A definition such as this is introduced because for most detectors the *NEP* scales with the sensitive area $\mathscr{A}$ of the detector like $\sqrt{\mathscr{A}}$. Therefore the

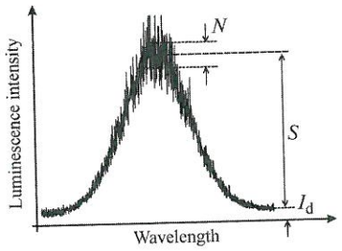


Fig. 2.6
Schematical representation of noise conditions in a DC luminescence emission spectrum. N stands for the root mean square noise value, i.e. $N = \left[\left(\int_0^{t_0} n^2(t)\,\mathrm{d}t\right)/t_0\right]^{1/2}$, where $n(t)$ is the instantaneous time-variable value of the noise related to the mean signal value S.

normalized detectivity D^* enables us to compare detectors of different sizes. A photomultiplier with a photocathode area of $\mathscr{A} \cong 5\,\mathrm{cm}^2$, $NEP = 10^{-16}\,\mathrm{W/Hz^{1/2}}$, and $\Delta f = 1$ Hz would feature $D^* \cong 2.3 \times 10^{16}\,\mathrm{cmHz^{1/2}/W}$. Let us note, however, that the detectivities D and D^* are most often used to characterize semiconductor detectors rather than photomultipliers.

A more detailed analysis of noise is beyond the scope of this book. We should only note that the total photodetector noise has several components: the shot noise of the photocathode, the noise due to the statistical character of the secondary emission from dynodes, Johnson's noise in the load resistor, the so-called $1/f$ noise, and in the case of semiconductor detectors, generation-recombination noise is added, etc. The reader is referred to [6, 9].

On the other hand, the crucial task in luminescence spectroscopy is to optimize the signal-to-noise ratio, S/N. This will be discussed in detail in Section 2.5.

Here we shall briefly deal with one type of noise only—shot noise. This noise is a direct result of the quantum (corpuscular) nature of both the photons in the input optical signal and the electric charge generated in a detector. The number of incoming photons per unit time fluctuates around a certain average value. Statistical fluctuations in the stream of photons emitted by a classical source obey Poissonian photon statistics (some special cases violating Poissonian statistics are mentioned in Chapter 17 and Appendix L). Obviously, this photon-related noise component is independent of the particular detector, and represents an inevitable property of the incoming signal and thus also the lowest possible noise level of any optical detection system. An analogous statistical description applies to the internal detector process of photoelectron emission upon absorption of a photon. The current generated by a detector thus fluctuates around an average value (i.e. the signal S) and—in analogy with similar statistical ensembles—the relevant fluctuations are equal to the square root of the signal average. This means that the mean square value of the shot noise N reads

$$N = \sqrt{S}. \tag{2.8}$$

This kind of noise, increasing at higher signal levels, is shown in Fig. 2.6 and is noticeable also in Figs 1.1 and 1.2 (the noise of the background level is of different origin).

To conclude our account of photomultipliers we give several practical hints. A photomultiplier is a sensitive, fragile, and expensive electronic tube that must be handled with care. When not in use, the photomultiplier tube should be stored in a light-tight box. Its installation into a housing must be done under dimmed light and exposure of the photocathode to daylight should be as short as possible. The operational high voltage must be applied at least half an hour prior to measurement in order to stabilize the sensitivity $k(\lambda)$, the dark current value I_d and the noise properties. Varying or even switching on and off the applied high voltage during an experiment (e.g. for fear of exceeding the maximum allowed anode current) is a quite common mistake, which could lead to erroneous results—any change in voltage may disturb the stabilized operation conditions and cause a temporary variation in photomultiplier sensitivity,

lasting up to tens of minutes (the so-called photomultiplier hysteresis). The correct approach to avoid overflow of I_{am} is decreasing the incoming light flux (using, e.g., calibrated neutral density filters). In the case of a cooled photomultiplier, the best practice is to keep cooling and permanently applying the high voltage even between measurements, the photocathode being kept in darkness. In such a way the dark current and the noise amplitude are minimized in the long term.

There might be two additional causes of irreversible noise increase (besides the above mentioned overflow of I_{am}): exceeding the maximum allowed voltage between the photocathode and anode U_{am} (this value is given by the manufacturer, as well as I_{am}) and penetration of moisture into the cooled photomultiplier housing. We recommend the maximum allowed voltage U_{am} should never be applied to a photomultiplier. First of all, the optimum operational voltage always lies below U_{am} (sometimes even by hundreds of volts) and increasing the voltage closer to U_{am} does not lead to better performance. Secondly, a possible inaccuracy of the voltmeter monitoring the applied high voltage can result in an unintentional overflow of U_{am}, entailing consequences much like the case of overflowing of I_{am}. Moisture penetrating towards a cooled photomultiplier can condense in the socket and leakage currents could cause irreversible changes of the properties of the photomultiplier. A good commercially supplied housing should be leakproof, anyhow.

Finally, let us remark that a magnetic field can be detrimental to the functioning of a photomultiplier as it alters the optimized trajectories of electrons between electrodes. The location of a photomultiplier (and of all auxilliary experimental apparatus) in a laboratory must be chosen to be free of disturbing electromagnetic fields. A special magnetic shield can possibly be used (these are offered by photomultiplier manufacturers).

Avalanche photodiode

The basic principle of an avalanche photodiode is the same as in other photovoltaic semiconductor photodetectors, namely, the *internal photoelectric effect*—the generation of an electron–hole pair through absorption of a photon. The photogenerated charge is then separated in an electric field established across the depletion layer of a p-n junction, and a photovoltage is created at the detector output, which is subsequently processed by electronics. Semiconductor photodiodes have a much wider spectral range compared to photomultipliers, especially toward the near-infrared region; their long-wavelength cut-off is determined by the bandgap width of the photosensitive semiconductor material. For example, the range of sensitivity of a Ge and InAs diode is 600–1800 nm and 1000–3200 nm, respectively. Despite their relatively low detectivities ($D^* \leq 10^{12}\,\mathrm{cm\,Hz^{1/2}/W}$), the p-n or p-i-n photodiodes[1] are

[1] An ultrasensitive germanium detector introduced by Edinburgh Instruments, Ltd., represents an exception suitable to detect low luminescence light fluxes. It is based on a p^+-i-n^+ diode made of ultrapure germanium and cooled by liquid nitrogen. The detection range is 800–1700 nm and its normalized detectivity of $5 \times 10^{14}\,\mathrm{cmHz^{1/2}/W}$ is comparable to photomultipliers. However, high sensitivity is counterbalanced by a very slow response, big size, considerable weight, and a relatively complicated operation.

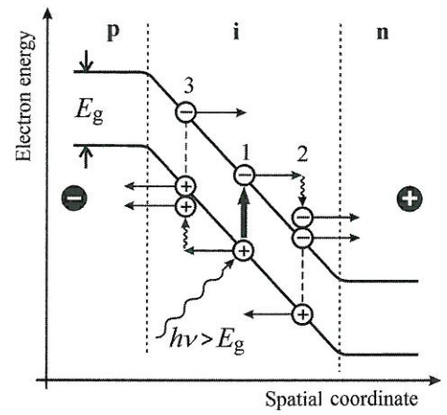


Fig. 2.7
The principle of an avalanche photodiode. Primary electron–hole pairs are created by absorption of photons $h\nu > E_g$ (1). A pair is dissociated and both the electron and hole are accelerated in opposite directions by the strong electric field in the intrinsic layer of the reverse biased p-i-n structure. After gaining enough energy the accelerated electron (hole) can create by impact ionization a new electron–hole pair (2) or (3) and the process can continue in an avalanche-like manner.

applied in luminescence spectroscopy for the detection in the infrared region beyond ~ 1000 nm, i.e. beyond the red cut-off in sensitivity of the common photomultipliers. However, current photodiodes lack the effect of photocarrier multiplication. Such multiplication is only possible in avalanche photodiodes that have internal *gain* in the range of $G_{AP} \sim 10$–500. The avalanche photodiode detectivity D^* is then increased but usually does not reach values comparable to photomultipliers.

The principal functional layout of an avalanche photodiode is shown in Fig. 2.7. Basically one deals with a p-i-n semiconductor structure under high reverse bias. Upon absorption of a photon $h\nu > E_g$ an electron–hole pair is created. This pair is dissociated in the strong electric field of the intrinsic layer (i) and both the electron and hole are accelerated in opposite directions. If the energy acquired by the accelerated carriers exceeds E_g (which is not always the case, because carriers are losing their energy in collisions with the crystal lattice), another electron–hole pair can be generated by *impact ionization*. Then the whole process may be repeated and an avalanche-like increase in number of carriers takes place.

Even if the principle of an avalanche photodiode looks very simple, there are several fundamental and technical obstacles making fabrication of an efficient and reliable device difficult. First of all, contrary to what might seem profitable at first sight, simultaneous amplification of both the electrons and holes (Fig. 2.7) is not favourable for obtaining high gain. In fact, such a situation is undesirable because this process is unstable and can result in avalanche breakdown and damage to the diode. In addition, the symmetrical amplification as a random process gives rise to a higher noise level of the detector and, moreover, prolongs the risetime of the avalanche diode, thereby reducing the frequency bandwidth of the diode. Therefore, avalanche diodes should be made of a material in which only one type of carrier is amplified. Let us characterize the probability of ionization by electrons (or holes), i.e. the number of ionization events per unit distance, by the *ionization coefficient* α_e (or α_h). Then an important avalanche photodiode parameter can be defined, namely, the ionization ratio $\kappa = \alpha_h/\alpha_e$ and it is required that κ tends either to zero $\kappa \to 0$ (electron multiplication) or to infinity $\kappa \to \infty$ (hole multiplication). The best

commercial avalanche photodiodes based on silicon have $\kappa \approx 0.006$. In the extreme case of $\kappa = 0$ the photodiode gain is simply given by $G_{AP} = \alpha_e w$, where w is the width of the multiplication region.

Another difficulty one is faced with when designing an avalanche photodiode arises from this simple reasoning: to achieve high diode sensitivity (i.e. to maximize the number of primary electron–hole pairs), the intrinsic (i) layer should be made as thick as possible. On the other hand, the inevitable presence of a high electric field across the layer requires this layer to be thin. The last requirement is of a technical rather than essential nature: the volume of the amplifying material should be limited in order to minimize the possible occurrence of material inhomogenities that could cause a local overrun of the breakdown electric field, thus creating a microplasma and eventually leading to photodiode damage. This conflict of opposite demands is solved by the special design of the avalanche photodiode—the wide photon-absorption layer is separated from the thin electron-multiplication layer ('separate-absorption-multiplication'). Evidently, fabrication of good avalanche photodiodes involves the use of both a perfect homogeneous material and advanced technology. The first avalanche photodiodes produced in the 1970s suffered from excess noise and easy breakdowns but subsequent radical development of semiconductor technologies removed those constraints.

Finally, the requirement to minimize the multiplication volume (as well as to achieve fast response) leads to a small size of the sensitive area $\mathscr{A}_{AP}$ of the avalanche photodiode (of the order of mm^2 or smaller).

The structure of an avalanche photodiode together with the profile of the internal electric field are presented in Fig. 2.8 (in this case hole multiplication occurs—photon absorption takes place in an i-InGaAs layer from which the holes drift in the weak electric field to an intrinsic InP multiplication layer) [10]. Obviously, the highest allowed bias applied to an avalanche photodiode must not be exceeded, similarly to photomultipliers.

Let us briefly compare photomultipliers with avalanche photodiodes from the point of view of luminescence spectroscopy. An advantage of avalanche photodiodes as solid-state devices over fragile vacuum tubes is their small compact size and robustness. On the other hand, noise properties of avalanche photodiodes are generally worse as expressed by the mean value of the dark noise [11]

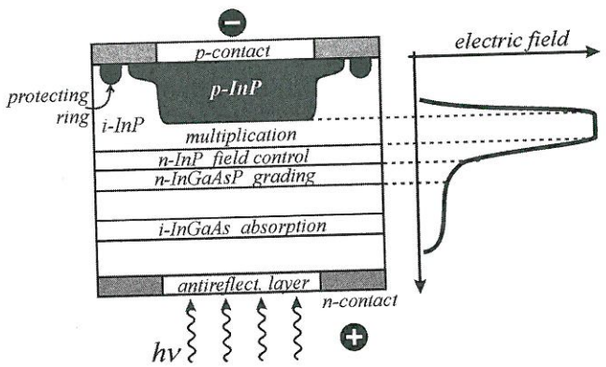


Fig. 2.8
Internal structure of an InGaAs/InP avalanche photodiode with separated absorption and multiplication layers and the corresponding profile of the electric field (right). Adapted from Ferrari *et al.* [10].

$$N_{AP} = \left(2eI_{dAP}\bar{G}_{AP}^2 F(G_{AP})\Delta f\right)^{1/2}. \tag{2.9}$$

When compared to eqn (2.6) which is valid for photomultipliers we notice firstly in (2.9) the *squared* mean value of the gain $\bar{G}_{AP}^2$ and, secondly, an additional noise factor $F(G_{AP})$. The squared factor arises as a consequence of the contribution of both types of carriers (electrons and holes) to the photodiode current (even if only one type of carrier is multiplied). The factor $F(G_{AP})$ arises from the stochastic nature of the ionization process. For multiplication of electrons this factor is given by

$$F(G_{AP}) = \kappa G_{AP} + (2 - 1/G_{AP})(1 - \kappa). \tag{2.10}$$

In the case of an Si avalanche photodiode with parameters $\kappa = 6 \times 10^{-3}$ and $\bar{G}_{AP} = 100$ we obtain $F(G_{AP}) \cong 2.6$. The *NEP* of an avalanche photodiode can be expressed, using (2.9), as

$$NEP_{AP} = \frac{\sqrt{2eI_{dAP}\,\bar{G}_{AP}^2\,F(G_{AP})}}{k_{AP}(\lambda)}$$

and for the normalized detectivity it follows from (2.7) that

$$D_{AP}^* = \frac{\sqrt{\mathscr{A}_{AP}}\,k_{AP}(\lambda)}{\sqrt{2eI_{dAP}\,\bar{G}_{AP}^2\,F(G_{AP})}}.$$

Finally, the ratio of the normalized detectivity of a typical photomultiplier D_{PM}^* (with gain $G_{PM} \approx 10^5$) to that of an avalanche photodiode D_{AP}^* is found to be

$$\frac{D_{PM}^*}{D_{AP}^*} = \sqrt{\frac{\mathscr{A}_{PM}}{\mathscr{A}_{AP}}}\sqrt{\frac{I_{dAP}}{I_{dPM}}}\,\frac{k_{PM}(\lambda)}{k_{AP}(\lambda)}\left(\frac{\bar{G}_{AP}\sqrt{F(G_{AP})}}{\sqrt{G_{PM}}}\right), \tag{2.11}$$

where I_{dAP} and I_{dPM} stand for the dark currents of the avalanche photodiode and photomultiplier, respectively, and k_{AP}, k_{PM} are the corresponding sensitivities.

If we try to compare, using eqn (2.11), two particular detectors—a Hamamatsu R316 photomultiplier with a Perkin Elmer InGaAs avalanche photodiode (the spectral sensitivities of both detectors overlap in the range 700–1100 nm), we find $D_{PM}^* \cong 50 D_{AP}^*$. However, a comparison like this must be considered with care, because all device parameters are strongly dependent on the radiation wavelength as well as on operational conditions. Sometimes even the datasheets for detectors, supplied by the manufacturers, do not contain all the necessary parameters.

Nevertheless, in general it holds true that the normalized detectivity of avalanche photodiodes does not attain the photomultiplier values. Moreover, the small size of the sensitive area of the avalanche photodiode is hardly appropriate for luminescence spectroscopy—to ensure tight focusing of luminescence radiation at the exit slit of a monochromator onto a small area of 1 mm^2 is basically impossible; a large part of the signal would thus be lost. Here, the optimum choice is indeed a photomultiplier because its photocathode of typically rectangular shape with size about $3 \times 12\,mm^2$ fits

the cross-section of the light beam emerging from the exit slit quite well. Therefore, avalanche photodiodes are applied in unconventional spectroscopic techniques only, like time-correlated single-photon counting and the determination of photon autocorrelation (see Subsection 2.4.2 and Chapter 17). They can also be used for the detection of extremely low-level luminescence signals from small sources—nanoobjects, like single molecules or nanocrystals, in a microspectroscopy set-up (Chapter 17). Single-photon counting avalanche photodiode modules are produced, e.g., by Perkin Elmer, ID Quantique and Hamamatsu.

In these special luminescence measurements, the avalanche photodiode operates in the so-called Geiger mode, the applied bias being higher than the breakdown voltage and the diode gain being very high ($G_{AP} \cong 10^5 - 10^6$). A particular diode design simultaneously ensures that no current flows across the diode except in the case when an avalanche event is initiated by a released photoelectron; after generating a current pulse the diode is closed again (in analogy with the Geiger–Müller counter of radioactive radiation). In this way individual photons may be resolved. Normalized detectivity of photon counting modules based on silicon avalanche photodiodes can be as high as 5×10^{13} cmHz$^{1/2}$/W. Obviously, in this case cooling of the avalanche photodiode is of particular importance in order to suppress the dark counts (the dark counts of the best diodes can be below 20 counts/s).

Multichannel detectors

Photomultipliers and avalanche photodiodes are single-channel photodetectors, which means that they are able to detect, at a given time, the photon flux at a single point (area) in space only. Alternatively, the so-called multichannel detectors, in the form of linear or two-dimensional arrays of tiny photosensitive solid-state devices, are able to detect the photon flux simultaneously at many spatial points. In luminescence spectroscopy, when the single-channel detector is placed behind the exit slit of a monochromator, only a single point from the spectrum is recorded. The entire spectrum is then acquired sequentially, step by step, by rotating the dispersion element (grating or prism) and thereby scanning the spectral image across the exit slit. Such a measurement is, of course, quite lengthy and the obtained spectrum can be distorted owing to, for instance, the long-term instability of both the excitation source and/or the sample (photochemical changes, generation of defects, etc.).

On the other hand, multichannel detectors, fixed in the output plane of a spectrograph, simultaneously detect the intensity (photon flux) at many spectral points within some spectral range, or even the whole spectrum without moving the dispersion element (Fig. 2.9). This allows for significant shortening of the signal acquisition time, avoiding instability effects, monitoring the entire spectrum in real time, eliminating inaccuracy of the dispersion element scanning mechanism, etc. In addition, this approach considerably simplifies those experiments that apply a pulsed mode of luminescence excitation. It may be worth evoking memories of the first type of multichannel detector used in spectroscopy, which was a photographic plate. Its replacement by photomultipliers and scanning monochromators during the 1950s and 1960s

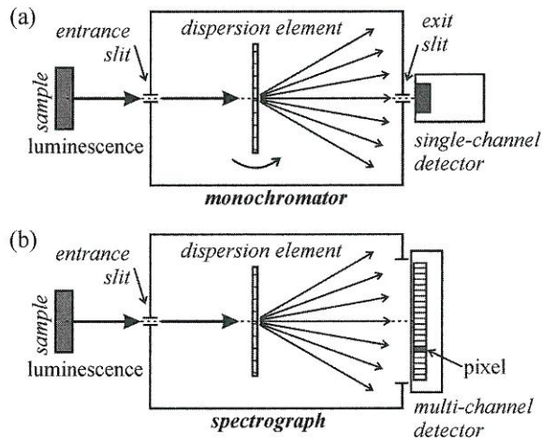


Fig. 2.9
Scheme of (a) a monochromator equipped with a rotating dispersion element (prism or grating) and a single-channel detector, (b) a spectrograph with a fixed dispersion element and a multichannel detector.

introduced a qualitative improvement to optical, and especially luminescence, spectroscopy. Even if the photomultiplier is a single-channel detector only, its many times higher sensitivity and reproducibility compared to photographic emulsions enabled researchers to investigate many new semiconductor phosphors. This finally resulted in the discoveries and inventions of, e.g., injection electroluminescence, light-emitting diodes, semiconductor lasers, and the accelerated development of optoelectronics and its applications.

Present-day multichannel detectors combine detection sensitivity—comparable to that of photomultipliers—with the advantages of a 'photographic plate' and the mechanical robustness of a solid-state detector. The better noise and detection properties of multichannel detectors arise from the so-called *Fellgett* (multiplex) *advantage*, which makes it possible to integrate a weak luminescence spectrum for an arbitrarily long time.

Let us give a more detailed explanation. Suppose we have a linear array detector containing n elements (called pixels, an abbreviation of 'picture element'), the sensitivity of each of them being the same as that of a photomultiplier. In the case of a scanning monochromator and a single-channel detector, the detector (photomultiplier) 'sees' any of the n spectral elements separately, say for one second of time, therefore the entire spectrum is recorded within n seconds, featuring a certain signal-to-noise ratio $(S/N)_0$. A multichannel detector records all of the n pixels simultaneously, therefore, supposing the integration time to be also one second, the resulting signal-to-noise ratio will be the same, i.e. $(S/N)_0$. The spectrum, nevertheless, is recorded within $1/n^{\text{th}}$ of the time required to be measured (to the same 'quality') by the photomultiplier. If we now record (integrate) the spectrum also for the time span of n seconds, we obviously obtain an n-times higher signal nS but—according to eqn (2.8)—only $\sqrt{nS}$-times higher noise level. Consequently, the spectrum recorded in this way will feature the signal-to-noise ratio

$$\left(\frac{S}{N}\right)_{\text{Fellgett}} = \frac{nS}{\sqrt{nS}} = \sqrt{n}\left(\frac{S}{N}\right)_0. \qquad (2.12)$$

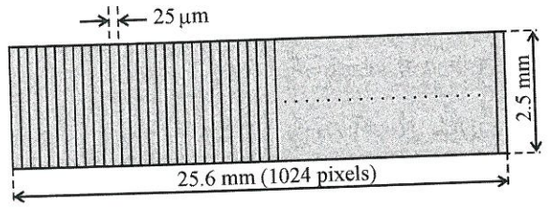


Fig. 2.10
Standard array of silicon photodiodes.

This increase in signal-to-noise ratio is known as Fellgett advantage. Let us consider $n \cong 1000$ (indeed, the number of pixels is often $1024 = 2^{10}$); then according to eqn (2.12) the increase in the signal-to-noise ratio is $\sqrt{1000} \approx 32\times$ (!). The foregoing example is valid strictly in the case of shot noise only but, as we shall see shortly, most other noise components can be successfully suppressed in multichannel detectors; in particular the thermal noise is effectively suppressed by cooling.

Of course, a further increase in S/N can in principle be achieved by further prolonging the acquisition time, as is immediately evident from eqn (2.12). With an acquisition time of m seconds ($m > n$) the signal-to-noise ratio is improved $\sqrt{m}$-times against $(S/N)_0$. It must be remembered, however, that in this way we deliberately renounce other benefits of multichannel detectors, namely, the elimination of possible sample 'fatigue' or of other unwanted variations in time of both the sample and the whole experimental set-up; on the other hand, such a measure is often beneficial because it facilitates successful recording of ('recovered from the noise') the spectra of extremely low light fluxes, undetectable by making use of photomultipliers.

Two of the most common types of multichannel detectors are *photodiode arrays* and *charge-coupled-device* (CCD) *cameras*. A linear array of photodiodes is made up of a row of tiny (silicon) photodiodes (Fig. 2.10). Every photodiode (pixel) is reverse biased and charged like a capacitor. The incident photons generate electron–hole pairs, increasing electric conductivity in the capacitor and consequently its discharging. Then, a read-out period follows, during which the diodes are charged up by a current that is proportional to the number of absorbed photons. The usual duration of the read-out period in the case of the 1024 diode array shown in Fig. 2.10 is 13 ms. The spectral range of the sensitivity is, similarly to common single Si photodiodes, about 300–1100 nm, and the spectrum acquisition time can be arbitrarily varied. The discrete structure of the photodiode array seemingly imposes certain limits to the achievable *spectral resolution*. However, luminescence spectra are usually relatively broad and so this fact does not usually pose serious problems.

The dark signal in a diode array, given by thermally generated electron–hole pairs, determines the sensitivity threshold of the array but can be substantially reduced by cooling—mostly thermoelectric cooling down to −20°C is applied.

The sensitivity of the diode array is not necessarily sufficient for the purposes of luminescence spectroscopy. Nowadays, the most sensitive photodetectors, aimed at detecting extremely weak signals and broadly exploited especially in luminescence spectroscopy, are CCD cameras. Their basic principle again consists in the generation of charge carriers by incident photons, but

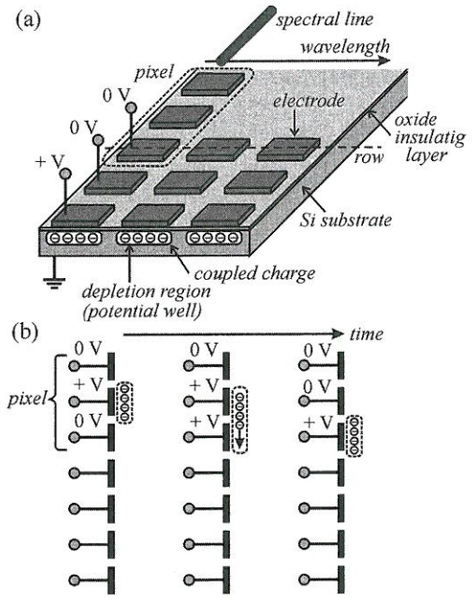


Fig. 2.11
(a) Schematics of a CCD detector. (b) Charge transfer in a CCD detector. Adopted from ISA Jobin-Yvon-Spex flyers [12].

their construction and read-out scheme differ substantially from the photodiode arrays.

One deals with a system of discrete small components, in which the individual pixel is usually defined as a set of three electrodes with a variable applied bias [12]. The electrodes are made of conductive (highly doped) polycrystalline silicon and deposited on a photosensitive silicon wafer of lower conductivity; the said materials are separated by a thin layer of insulating SiO_2 (Fig. 2.11(a)). The photoelectrons generated by incident light may, depending on the applied bias, be trapped and stored in potential wells that are created in silicon beneath the electrodes. The light-detection process involves three steps:

1. Absorption of incident photons in photosensitive silicon and generation of electron–hole pairs. The number of pairs in each pixel is proportional to the incident light intensity. When positive bias is applied to the pixel central electrode, only the electrons are trapped in the potential well while holes are drained to earth.
2. Transfer of the collected charge from a given pixel to the next one. By applying the bias to the outer electrode of the pixel, the electron cloud is shifted to the pixel edge and eventually to the next pixel in the column, i.e. along the particular spectral line or a specific wavelength. This process is repeated until the charge from each pixel of a row reaches a read-out register located along the edge of the CCD chip, i.e. in parallel with the wavelength axis. This is accomplished by gradually varying the bias on individual pixels (see Fig. 2.11(b)). It is now clear why a pixel is composed of three electrodes. A great advantage of this mode of information read-out is the extreme reduction of the read-out noise due to the use of a single

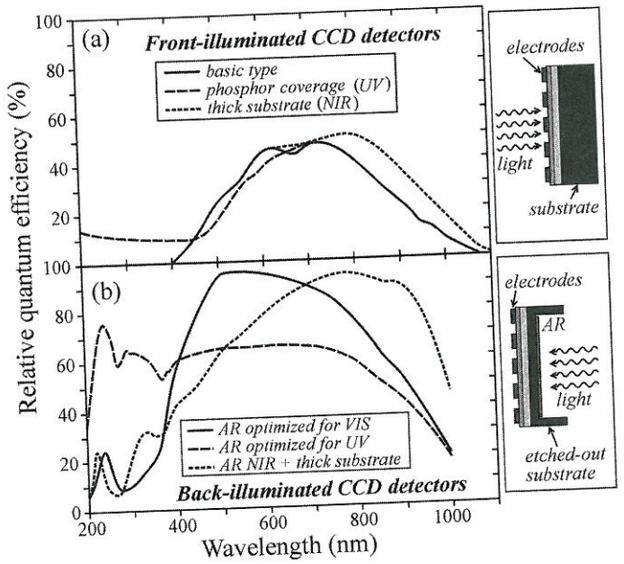


Fig. 2.12
Spectral response of a silicon CCD camera. (a) Sensitivity of a standard front-illuminated CCD detector (extension of the detection range into the IR region—short dashed line—can be achieved by using a thicker substrate that absorbs effectively longer wavelengths). (b) Spectral sensitivity of a back-illuminated chip where absorption of incident light in electrodes is avoided. Extension into the UV region can be obtained for example by depositing a phosphor layer (emitting visible light under illumination by UV light) directly onto the detector. AR means antireflection coating. Adapted from the ISA Jobin-Yvon-Spex flyer [12].

common read-out register for all of the pixels instead of employing a read-out system in every pixel.

3. Conversion of electric charge into voltage, amplification and digitization. The stored charge is transferred row by row into the read-out register and then into an output node, converted to an appropriate voltage, amplified and turned into a digital signal. Finally, the number of detected photoelectrons in every pixel of a CCD is stored, and the obtained picture (spectrum) can be visualized on a display, numerically treated, etc.

The typical area of a pixel in a spectroscopic CCD detector is about $20 \times 20\,\mu m^2$, and the full size of the sensitive area is, e.g., $26.6 \times 6.6\,mm^2$ (1024×256 pixels with a pixel size of $26 \times 26\,\mu m^2$).

The spectral sensitivity of a CCD detector is determined by its material—usually silicon. The typical range of about 400–1000 nm can be modified to some extent by adjusting the detector surface and thickness and the mode of illumination (front or back-illuminated), see Fig. 2.12. The experimental *spectral coverage* depends on the physical length of the detector and on the monochromator dispersion, and it can obviously be expressed as

$$\Delta\lambda_{cov} = (\text{reciprocal linear dispersion[nm/mm]}) \times (\text{CCD detector length [mm]}).$$

We have to stress that multichannel detectors should not be coupled to a common scanning monochromator, by merely removing its output slit and replacing it by a CCD camera or a diode array. An ordinary monochromator is able to image an aberration-free spectrum (i.e. to create a spectrally resolved projection of the entrance slit into the exit slit plane) of a very narrow wavelength range only, covering roughly the slit width (~ 1 mm). For application with multichannel detectors, special *imaging* spectrographs are available. Their optical aberrations must be corrected over the whole size (height and width)

of the detector. For example, combination of a Horiba Jobin Yvon iHR 320 imaging spectrograph having a dispersion of 2.64 nm/mm with a CCD chip of 1024×256 pixels (pixel size $26 \times 26\,\mu m^2$) shows a spectral coverage of $\Delta\lambda_{cov} = 70.3$ nm. In order to measure a wider spectral range the ultimate spectrum must be glued together from several segments with a shifted central wavelength.

The sensitivity of the CCD detector is extremely high. Thermal noise can be almost eliminated by cooling down to $-90°C$ or even $-140°C$ when the dark signal can be suppressed down to an unbelievable 1 electron/pixel/h only (!) (In practise, the recommended temperature is usually above $-110°C$ because lower temperatures would decrease the quantum efficiency of the detector; efficient cooling is provided either by liquid nitrogen or a multistage thermoelectric Peltier element.) The lowest detectable signal is then limited by the so-called read-out noise, or possibly only by the shot noise originating mostly in the signal itself.

Let us put forward a few examples:

1. Suppose the total root mean square noise value N_{CCD} of the detector is given as a combination of three components

$$N_{CCD} = \sqrt{|n_{shot}|^2 + |n_{dark}|^2 + |n_{read}|^2}, \tag{2.13}$$

with n_{shot}, n_{dark} and n_{read} standing for the shot noise, dark signal and the read-out noise, respectively. Assuming a relatively low flux of incident photons of 10 photon/pixel/s entails (in the visible range) a power flux density of $\sim 4 \times 10^{-13}\,W/cm^2$. Considering the quantum efficiency of the detector to be 50% we obtain 5 photoelectron/pixel/s. Setting the acquisition time to be 600 s (10 min), the total integrated signal is 5 photoelectron/pixel/s $\times 600\,s = 3000$ photoelectron/pixel. Shot noise of such a signal is $n_{shot} = (3000)^{1/2} = 54.77$. Suppose further this cooled detector has a dark signal of 1 electron/pixel/min, then the total dark signal accumulated during 600 s is 10 electron/pixel and the corresponding noise component is $n_{dark} = 10^{1/2} = 3.16$. Finally, the typical read-out noise of a 1024×256 pixel CCD detector with slow reading is about $n_{read} = 4$ electrons (root mean square value).
According to (2.13) the overall noise of the detection system is

$$N_{CCD} = \sqrt{(54.77)^2 + (3.16)^2 + 4^2} = 55.$$

The signal-to-noise ratio is $S/N = 3000/55 = 54.54$. Note that, were the detector completely free from both the dark current and read-out noise, then the best achievable signal-to-noise ratio would be $S/N = 3000/54.77 = 54.77$, i.e. only negligibly higher. We conclude that under an incident photon flux of 10 photon/pixel/s the intrinsic detector noise does not play any detrimental role—such measurement is referred to as 'shot noise limited'. The same conclusion would hold true even for a ten times lower incident flux of 1 photon/pixel/s!

2. Let us now consider a complelety different situation when an extremely low signal of only 1000 photon/s hits the *whole area* of a CCD detector

(1024×256 pixels$^2 = 676\,\mu$m^2). This means, at a photon energy of 2 eV, an incident power density of 1.7×10^{-17} W/cm^2. Such extremely low photon fluxes take place, for example, when detecting luminescence excited by the tip of a scanning tunnelling microscope (STM) or photoluminescence of single semiconductor quantum dots (see Chapter 17). Suppose we have at our disposal the same CCD detector as in the previous example. This means with quantum efficiency of 50%, dark signal of 1 electron/pixel/min, read-out noise of $n_{\text{read}} = 4$ electrons, and we also leave unchanged the acquisition time 600 s.

The average photon flux incident on a single pixel is 1000 (photon/s)/2.62×10^5 pixels $= 3.8 \times 10^{-3}$ photon/pixel/s; the corresponding number of photoelectrons is $0.5 \times 3.8 \times 10^{-3} = 1.9 \times 10^{-3}$ photoelectron/pixel/s, and after an integration time of 600 s we obtain the total signal as 1.9×10^{-3} photoelectron/pixel/s $\times$ 600 s $= 1.14$ photoelectron/pixel. The corresponding shot noise is $n_{\text{shot}} = \sqrt{1.14} \approx 1.07$. From the preceding example we remember that $n_{\text{dark}} = 3.16$, $n_{\text{read}} = 4$ and thus the total CCD noise reads

$$N_{\text{CCD}} = \sqrt{(1.07)^2 + (3.16)^2 + (4)^2} = 5.21.$$

The signal-to-noise ratio is thus $S/N = 1.14/5.21 \approx 0.22$. Taking $S/N = 1$ as the detection limit criterion, the light flux under consideration is not detectable and will be 'buried in noise'. Here the experiment is already clearly limited by the detector's noise itself!

Does this example imply that even sensitive CCD detectors are not able to detect photon fluxes of the order of 10^3 photon /detector/s? No, it does not, as we shall see. The solution could be to set a longer detection time and/or apply a less noisy detector.

3. Let us keep all the parameters from example 2 fixed, except the integration (acquisition) time, which we increase now from 600 s to 1 hour (3600 s). The signal will be increased to 6.84 photoelectron/pixel and the overall noise to $\sqrt{(2.61)^2 + (7.75)^2 + 4^2} = 9.1$. We obtain $S/N = 6.84/9.1 \cong 0.75 < 1$. We can see that even quite substantial prolongation of the acquisition time does not lead to $S/N = 1$, and further extension of the experimental time is usually out of the question (due to possible instabilities of sample properties or experimental conditions).
4. Let us select the best available CCD detector with a high quantum efficiency (90%) and almost zero dark signal (1 electron/pixel/h)—these are typical parameters of a back-illuminated liquid-nitrogen-cooled silicon CCD chip. Let the incident photon flux be the same as in example 2, i.e. 3.8×10^{-3} photon/pixel/s, and integration time is 600 s. Then the signal is $0.9 \times 3.8 \times 10^{-3} \times 600 = 2.05$ photoelectron/pixel, the shot noise is $\sqrt{2.05} = 1.43$, the dark signal is 1/6 electron/pixel and the corresponding noise is $\sqrt{1/6} = 0.41$. The read-out noise remains equal to 4 electrons. Now we have $N_{\text{CCD}} = \sqrt{(1.43)^2 + (0.41)^2 + 4^2} = 4.27$, $S/N = 2.05/4.27 \cong 0.48 < 1$. Even now the spectrum remains undetectable, due to the relatively high read-out noise. The last chance is to extend the acquisition time; 40 min (2400 s) is on the limit but still acceptable. The reader then easily finds that the resulting S/N ratio finally breaks the detection limit, being $S/N = 8.21/4.98 \cong 1.65 > 1$. The spectrum can thus be recorded within a 40-minute exposure time, even if quite noisy.

In conclusion, CCD detectors manage to display the whole spectrum of comparatively strong signals in real time (which, by the way, appears to be an excellent tool to adjust the experimental set-up in this case). The whole spectral range of such signals is detected almost instantaneously (no monochromator scanning is needed), all the wavelengths being recorded simultaneously. The CCD detectors have very low dark signal and their sensitivity is comparable to photomultipliers or even better. For detection of extremely low signals, the S/N ratio can be improved by increasing acquisition times or by binning signals from several neighbouring pixels. One of the drawbacks of a CCD is the relatively narrow spectral coverage, as another possible weakness may be their low read-out speed given by operating a single read-out register ('serial read-out'). CCD detectors must be handled with care, avoiding strong overexposure by light (like photomultipliers) and, in particular, discharge of static electricity into both the disconnected CCD camera and the control electronics connectors.

2.3 Monochromators and spectrographs

A *monochromator* serves to disperse the analysed light signal into a spectrum using a dispersion element (grating or prism), and then to cut out a narrow (quasi)-monochromatic part of the spectrum by the exit slit, located at the rear (output) focal plane. A monochromator can be tuned manually (e.g. for selecting the appropriate excitation wavelength) or scanned automatically by rotating the dispersion element—in this case the spectrum moves continuously across the exit slit and can be recorded using a single-channel detector.

A *spectrograph* projects the spectrum onto a fixed position in its output focal plane. The exit slit is replaced by a photographic plate or a multichannel detector. Most modern spectrographs are equipped with two output channels—a lateral output (selectable by a moving mirror), being fitted with a slit, enables application of the spectrograph in a scanning monochromator mode (Fig. 2.13).

An *imaging spectrograph* is a special type of spectrograph, equipped with purpose-built optics that reduces optical aberrations. The spectral image of the entrance slit is perfectly sharp over a large area at the output image plane (typically 1 cm high and 2 cm wide), not only in its central part (close to the optical axis) as was the case of a monochromator.

The terms *polychromator* and *spectrometer* can also be encountered in the literature. A *polychromator* has several parallel monochromatic outputs provided by multiple output slits. A *spectrometer* is a complete device designed for measurements of optical emission spectra, consisting of a monochromator (or spectrograph) coupled with a detector.[2]

[2] This nomenclature of spectral devices is not employed systematically in the scientific literature. The terms monochromator, spectrograph or spectroscope are often used arbitrarily like

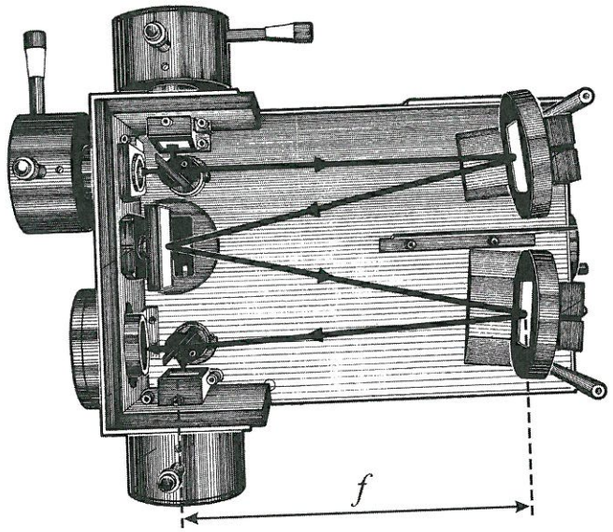


Fig. 2.13
A HR 320 monochromator–spectrograph. Adapted from Horiba Jobin Yvon flyer.

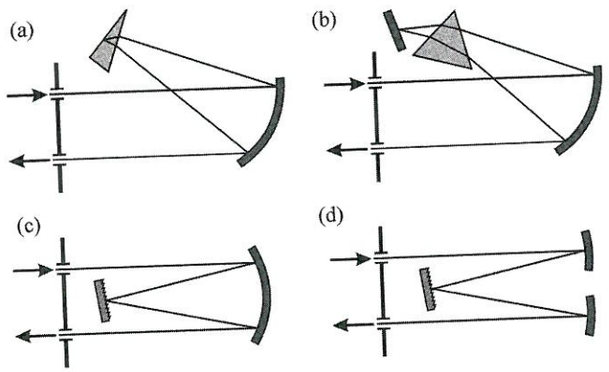


Fig. 2.14
The most common optical designs of spectral devices. (a) Littrow type with an autocollimating prism, (b) Walsch type with a prism, (c) Ebert construction with a grating, (d) Czerny–Turner type with a grating. The quality of spectral imaging increases from (a) to (d).

There are various designs of optical spectral devices with different levels of compensation for spherical aberration, astigmatism and coma, observed when displaying the spectral elements. Some of these designs are shown in Fig. 2.14. Most of the currently fabricated monochromators suitable for luminescence spectroscopy are of the Czerny–Turner type (Fig. 2.14(d))—see for example the HR 320 monochromator–spectrograph in Fig. 2.13. Other, somewhat less familiar optical schemes are those by Ebert–Fastie, Gillies, Wadsworth, the Rowland's construction with a concave grating and many others; for more details the reader is referred to the special literature [13, 14, 15].

The basic quantities describing the properties of a monochromator or spectrograph are dispersion, resolving power, and throughput (aperture). We shall

synonyms (originally a *spectroscope* denoted a device producing an optical spectrum for visual observation).

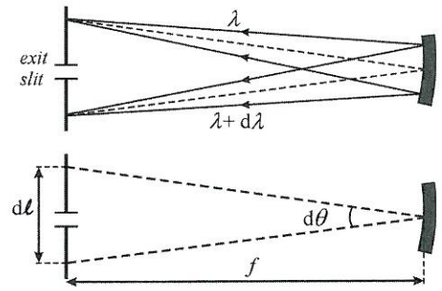


Fig. 2.15
Illustration of the definition of angular and linear dispersion.

describe them only to the limited extent necessary for the experimental foundations of luminescence spectroscopy.

2.3.1 Dispersion and resolving power

Let us consider two light beams whose wavelengths differ by $d\lambda$. They leave the dispersion element at angles differing by $d\theta$. We can define the *angular dispersion* as

$$\mathscr{D} = \frac{d\theta}{d\lambda} \text{ (rad/nm)}. \tag{2.14}$$

The spatial separation $d\ell$ between two spectral lines λ, $\lambda + d\lambda$ in the exit slit plane depends on the focal length f of the spectral device, determined by a focusing mirror or lens, see Figs 2.13 and 2.15:

$$d\ell = f\, d\theta. \tag{2.15}$$

The *linear dispersion* L is then defined as

$$L = \frac{d\ell}{d\lambda} = f\frac{d\theta}{d\lambda} = f\mathscr{D}. \tag{2.16}$$

More frequently its inverted value, the so-called *reciprocal linear dispersion*, is used:

$$L^{-1} = \frac{d\lambda}{d\ell} = \frac{1}{f}\frac{d\lambda}{d\theta} = \frac{1}{f\,\mathscr{D}} \text{ (nm/mm)}. \tag{2.17}$$

To make an order-of-magnitude estimate, let us consider a simple diffraction grating and light incident at an angle $\alpha \neq \alpha(\lambda)$. Light diffraction is described by the the grating equation $d(\sin\alpha + \sin\theta) = k\lambda$ (d is the grating constant, k is the diffraction order). By differentiating with respect to λ we obtain

$$\mathscr{D} = d\theta/d\lambda = \frac{k}{d\,\cos\theta}. \tag{2.18}$$

In order to estimate $\mathscr{D}$ we approximate $\theta \approx 0°$ (corresponding to the situation when the analysed light impinges on the grating at an angle $\alpha \neq 0$ and the spectrum is observed in a direction almost perpendicular to the grating plane). Then for a grating with 1200 grooves/mm operating in the first diffraction order we obtain an angular dispersion $\mathscr{D} = 1/d = 1.2 \times 10^3$ rad/mm $= 1.2 \times 10^{-3}$ rad/nm. A monochromator with $f = 500$ mm features a linear dispersion

$L = 500 \times 1.2 \times 10^{-3} = 0.6\,\text{mm/nm}$ and reciprocal linear dispersion $L^{-1} \approx 1.66\,\text{nm/mm}$. Therefore, in the case of luminescence spectroscopy, when—as we shall specify below—neither very high dispersion nor excessive spectral resolution is required, the spectral devices possess reciprocal linear dispersion in the range of 1–10 nm/mm. The importance of the reciprocal linear dispersion L^{-1} results from the fact that the spectral bandwidth $\Delta\lambda$ of a given device is simply determined by L^{-1} (declared by the manufacturer) and the mechanical slit width $\Delta\ell$ (set by the experimenter) via the relation

$$\Delta\lambda = L^{-1}\,\Delta\ell \tag{2.19}$$

that follows immediately from (2.17). In our example (a monochromator with $L^{-1} = 1.66\,\text{nm/mm}$) for a chosen slit width of $\Delta\ell = 0.5\,\text{mm}$ we thus find $\Delta\lambda = 0.83\,\text{nm}$. Note that eqn (2.19) is strictly valid only when both the entrance and exit slits are of the same width $\Delta\ell$. This setting is commonly used in spectroscopy for optimizing spectral brightness and facilitating spectral correction procedures (see Section 2.8).

The angular dispersion of a prism can be found by differentiating the condition of minimum deviation (see e.g. [14, 16]) which yields

$$\mathscr{D} = \frac{2\sin(\varphi/2)}{\sqrt{1 - N^2\sin^2(\varphi/2)}}\,\frac{\mathrm{d}N}{\mathrm{d}\lambda}, \tag{2.20}$$

where ϕ is the prism apex angle, and N and $\mathrm{d}N/\mathrm{d}\lambda$ are the refractive index and material dispersion of the glass (or other material) from which the prism is fabricated, respectively.

In the case of a monochromator, the measurement of the spectra is performed by rotating the dispersion element, thereby scanning the spectral image across the exit slit. Most grating monochromators are equipped with a so-called *sine-drive* scanning mechanism that ensures linear scanning of the wavelength in time, which means $\Delta\lambda = \text{const}\,\Delta t$. The principle of the sine-drive mechanism is shown in Fig. 2.16. There is a triangle ABC; one of its sides (a) is allowed to vary linearly in time t by means of a precise motor-driven screw, $a = a_0 + vt$. The length of the hypotenuse c, which is anchored by a ball-joint at point B, is constant. The second side b is carried in a sliding assembly allowing its length to be varied. Then the angle γ is a function of time, $\gamma = \gamma(t)$. By differentiating the goniometric relation $\sin\gamma(t) = (a_0 + vt)/c$ we find

$$\cos\gamma\,\mathrm{d}\gamma = (v/c)\mathrm{d}t. \tag{2.21}$$

A grating can be coupled with this mechanical system in such a way that $\gamma = \theta$. (Spectroscopic gratings have an asymmetric profile of grooves—the so-called blaze, which enables efficient diffraction of a selected spectral range into the desired angle and diffraction order.)

Then eqn (2.21) is replaced by

$$\cos\theta\,\mathrm{d}\theta = (v/c)\mathrm{d}t$$

and using (2.18) we obtain

$$(k/d)\mathrm{d}\lambda = \cos\theta\,\mathrm{d}\theta \quad \text{or} \quad (k/d)\mathrm{d}\lambda = (v/c)\mathrm{d}t,$$

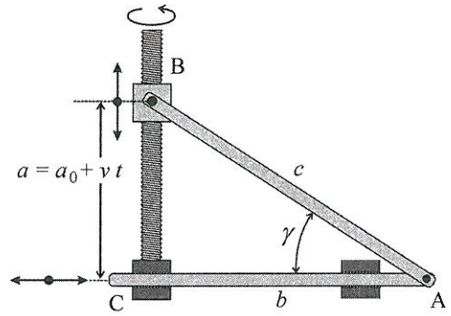


Fig. 2.16
Schematic of a sine-drive mechanism.

therefore, in the given diffraction order k we have $\Delta\lambda = \text{const}\,\Delta t$; the wavelength scan is linear in time.

A simple modification of the sine-drive (Fig. 2.16) is possible: the length of the hypotenuse c can be varied in time using the screw mechanism while a is kept constant. Then the reciprocal wavelength $\nu^* = 1/\lambda$, called the wavenumber, is a linear function of time $\Delta\nu^* = \text{const}\Delta t$. This design is referred to as a *cosecant-drive* (or just a *cosec-drive*) since here one deals with a linear temporal variation of the function $1/\sin\gamma = \text{cosec}\,\gamma$. This drive makes the spectrum to be scanned linear in the photon energy $h\nu = hc\nu^*$ instead of in wavelength, which could be advantageous for many experiments as well as their physical interpretation. Indeed, it is used in some Raman spectrometers. In the past few decades, most monochromators were provided with sine-drives but modern spectrographs are equipped with direct drives, where a motor rotates the grating directly (through some gear wheels). The number of motor steps per unit wavelength in the course of scanning is not constant and is driven by control software. The benefits of this design are a very high speed of grating rotation from one position to another, possible integration of more selectable gratings in one holder, and use of cheaper mechanical components.

Unlike in gratings, there is no simple mechanical way to ensure linear scanning of λ or ν^* in a prism monochromator. The angular dispersion of a prism as given by eqn (2.20) contains the dispersion relation $N = N(\lambda)$ which is strongly nonlinear in glasses and other optical materials. Therefore, spectra measured using a prism monochromator usually have a nonlinear wavelength scale; compare, e.g., Fig. 2.28 (a prism monochromator) with Fig. 1.1 (a grating monochromator).

Another important fact contained in relations (2.18) and (2.20) is worth pointing out: the dispersion of a grating or prism is independent of its size—e.g. two prisms made of the same material and having equal apex angles φ have equal dispersion regardless of size (Fig. 2.17). In other words, the dispersion of a spectral device is defined in the framework of geometrical optics, where diffraction due to finite beamwidth is not taken into account.

On the other hand, the *resolving power* R, defined with the aid of the smallest resolvable difference between two closely spaced spectral lines $\Delta\lambda$ as $R = \lambda/\Delta\lambda$, is limited just by diffraction of the light beam, whose diameter is driven by the lateral size of the dispersion element. A small prism or grating with a small effective area gives rise to sizeable broadening of spectral lines in

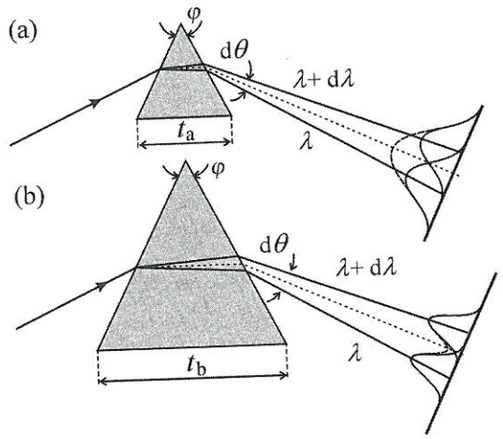


Fig. 2.17
Illustrating the distinction between dispersion and resolving power. Two prisms (a) and (b) have equal (angular and linear) dispersion but the bigger one (b) has higher resolving power, as shown by the profiles of two closely spaced spectral lines in the exit focal plane.

the exit slit plane due to light diffraction. Consequently, taking into account Rayleigh criterion for the resolution of two peaks, the resolving power is decreased (Fig. 2.17). But even having a sufficiently large dispersion element, the effective cross-section of the beam to be analysed may happen to be limited by the insufficient diameters of the lenses or mirrors used in the spectral device, and possibly also by improperly collimating the input radiation.

Along with diffraction due to the limited diameter of the light beam, another diffraction effect enters the play—diffraction at the entrance slit, which could negatively influence the output spectral image when the slit width is set too narrow, below some 'reasonable' limit.[3] Let us discuss both of these phenomena in greater detail.

To begin with, suppose we can neglect diffraction at the entrance slit. It is obvious that reducing the width of the monochromator slits leads, according to eqn (2.19), to narrowing of the transmitted spectral bandwidth $\Delta\lambda$ and, consequently, to increased spectral resolution. Let us employ the term *theoretical* or *ultimate resolving power* for $R_t = \lambda/\Delta\lambda_{opt}$, where $\Delta\lambda_{opt}$ is the (minimum) spectral bandwidth at which the input slit diffraction does not yet become evident, and thus an ultimate resolution R_t is achieved. Let us consider this situation, as occurring in a prism monochromator, displayed in Fig. 2.18. We remember that the diffraction angle on a circular aperture of diameter W (here the 'effective aperture stop of the monochromator') is expressed by the well-known relation

$$\delta\theta \approx \lambda/W, \tag{2.22}$$

if we omit the numerical factor 1.22 on the right-hand side for the sake of simplicity. Now, should two closely spaced lines be resolved ($\Delta\lambda = \Delta\lambda_{opt}$), the condition $\Delta\theta = \delta\theta$ must be satisfied according to Fig. 2.18. The angle $\Delta\theta$

[3] Note that in spectral devices equipped with a diffraction grating, the interplay between three diffraction effects is taking place—diffraction on the input slit, on the grating and diffraction due to the finite beam size.

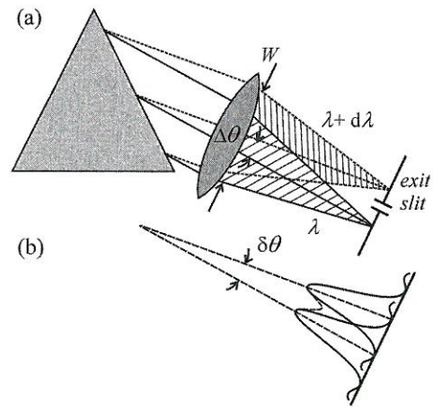


Fig. 2.18
Illustrating the resolving power of a simple prism monochromator. (a) Rays leaving the prism and propagating through an exit focusing lens. The lateral size of the beam W is basically determined by the prism dimensions. The images of two spectral lines in the plane of the exit slit have finite width due to light diffraction. (b) Profiles of two closely spaced spectral lines in the plane of the exit slit that meet Rayleigh resolution criterion (i.e. the centre of one linear diffraction pattern falls on the first minimum of the diffraction profile of the other line).

is determined by the angular dispersion of the prism; by combining eqns (2.14) and (2.22) we obtain

$$R_t = \lambda/\Delta\lambda_{opt} = \frac{\delta\theta\, W}{\Delta\lambda_{opt}} = \frac{\Delta\theta\, W}{\Delta\lambda_{opt}} \approx \frac{d\theta}{d\lambda} W = W\,\mathscr{D}. \tag{2.23}$$

Obviously, this relation is valid for any dispersion element. Importantly, it shows that there is a close relation between the angular dispersion $\mathscr{D}$ and the ultimate resolution R_t of a dispersion element, and also supplies a qualitative explanation for Fig. 2.17.

By inserting the angular dispersion into eqn (2.23) for the grating (2.18) or prism (2.20), respectively, we obtain, after some simple algebra, for the ultimate resolving power of the grating

$$R_t = m\,k \tag{2.24}$$

and for that of a prism (valid under the assumption that the whole prism entrance face is illuminated)

$$R_t = t\,\frac{dN}{d\lambda}. \tag{2.25}$$

In eqn (2.24) m stands for the number of illuminated grating grooves (it is easy to show that $m = W/d\cos\theta$), and t in eqn (2.25) denotes the width of the prism base, see Fig. 2.17. For example, a grating with 1200 grooves/mm and width of 5 cm has in the first diffraction order $R_t = 6 \times 10^4$; a prism made of a heavy flint glass ($dN/d\lambda \approx 2.7 \times 10^{-4}$/nm for blue light) with a base of $t =$ 5 cm features $R_t \approx 1.35 \times 10^4$. In reality, the ultimate resolution power may be decreased by one-third or even more due to imperfections in the mechanical and optical elements, imperfect homogeneity of the optical materials, etc.

Coming back to the problem of diffraction at the entrance slit and the determination of the optimum slit width $\Delta\ell_{opt}$, we start with the scheme in Fig. 2.19. The detected light is focused onto the entrance slit under a top angle of 2φ and adjusted so that the whole collimating element (lens) of diameter W is covered. Along with the angle φ, determined basically by the rules of geometrical optics, the deviation angle due to diffraction at the slit emerges

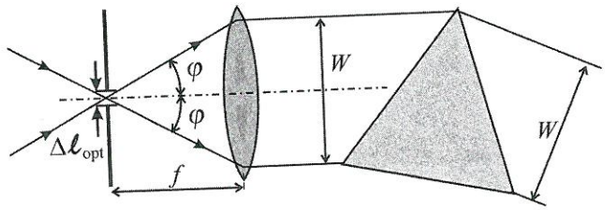


Fig. 2.19
The entrance slit of a spectral device should not be set narrower than $\Delta\ell_{opt}$, otherwise, due to diffraction, the incoming light fills up an angle larger than 2φ and part of the signal is lost.

as an independent factor. This diffraction angle $\varphi_d \sim \lambda/\Delta\ell$ increases with decreasing $\Delta\ell$. When setting $\Delta\ell$ too small (with the best of intentions to improve resolution), below some optimum value $\Delta\ell_{opt}$, the angle φ_d becomes larger than φ and part of the input light fails to meet the lens and is lost for spectral detection. Such a situation has several negative consequences: (i) first of all, the brightness of the spectrum drops which might be fatal in studying low-level luminescence signals; in any case the signal-to-noise ratio deteriorates (see Section 2.5); (ii) light passing through the lens hits the internal walls or other parts of the device, thereby producing parasitic reflections and thus increasing the background signal (the so-called stray light) which, eventually, (iii) results in futher suppression of the S/N ratio. Moreover, mechanical imperfections, e.g. microscopic damage to the slit blades, can substantially degrade the sharpness of the spectra when setting the slit very narrow.

The optimum value $\Delta\ell_{opt}$ of the slit width is thus achieved if the diffraction angle

$$\varphi_d \sim \lambda/\Delta\ell_{opt} \tag{2.26}$$

is just equal to the angle φ given by the geometrical condition

$$\varphi \sim \frac{W}{2f}, \tag{2.27}$$

imposed by Fig. 2.19. By comparing (2.26) and (2.27) we obtain

$$\Delta\ell_{opt} \approx \frac{2f\lambda}{W}. \tag{2.28}$$

The reader is, undoubtedly, aware of some oversimplification of the discussion given above; in particular relation (2.26) valid for Fraunhofer diffraction of plane waves must be considered as an approximation only. Thus, to make an order-of-magnitude estimate, we neglect also the numerical factor 2 in (2.28), we shall consider a monochromator with focal length of $f = 60\,\text{cm}$ and effective width of the dispersion element $W = 6\,\text{cm}$, we put $\lambda = 1\,\mu\text{m}$ and then we get for the optimum slit width

$$\Delta\ell_{opt} \approx f\lambda/W \approx 10\,\mu\text{m}.$$

This is a very important message for luminescence measurements: Decreasing the input slit width below approximately 10 μm makes no sense—resolution will not be improved; on the contrary the quality of the spectrum will deteriorate and the S/N ratio will decrease.

Of course, in most practical luminescence experiments much wider slits (100 μm or more) are used, partly due to low light fluxes and also because the luminescence spectra are often composed of broad bands that do not require an extremely high resolution, as already stressed above. Nevertheless, one can meet situations where the luminescence emission spectrum of a semiconductor contains a series of narrow, sometimes closely spaced lines (see, e.g., the spectrum of bulk silicon in Fig. 1.1). To resolve such lines, the width of the slits must be chosen very carefully. The optimum width is found as a trade-off between an acceptable noise level and the requirement to distinguish the closely spaced lines. Here, it appears useful to realize that by opening slits to a real value $\Delta\ell_{real} > \Delta\ell_{opt}$, the theoretical limit of the resolving power R_t is relaxed to a real value R_{real} called the 'effective-slit-width-limited resolving power': $R_{real} = \lambda/\Delta\lambda_{real} = \lambda/L^{-1}\Delta\ell_{real}$. An approximate relation between these two resolving powers may be written as

$$R_{real} \leq \left(\frac{\Delta\ell_{opt}}{\Delta\ell_{real}}\right) R_t. \tag{2.29}$$

All this is illustrated schematically in Fig. 2.20.

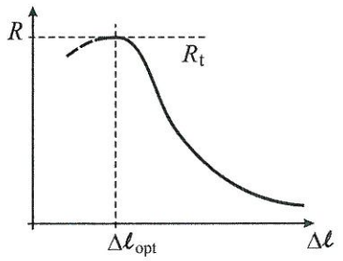


Fig. 2.20
Real resolving power of a spectral device as a function of the slit width $\Delta\ell$. The quantity R_t stands for the ultimate resolving power. When $\Delta\ell$ is large enough, diffraction effects on the slits may be neglected and the simple relation $(R_2/R_1) = (\Delta\ell_1/\Delta\ell_2)$ holds true.

To conclude our discussion of the dispersion properties we shall attempt to compare briefly prism- and grating-based spectral devices. A diffraction grating gives a linear sweep in λ or $\nu^* = 1/\lambda$, which is one of its main merits. Another advantage, as we have seen, is—as a rule—the higher resolving power compared to a prism of comparable size. On the other hand, a grating monochromator or spectrograph has at least two unpleasant properties: Firstly, different wavelengths can be sent to the same diffraction angle, in other words, spectra of different diffraction orders may partially overlap one another. An inexperienced or careless experimenter can consequently come to a completely erroneous interpretation of his observation. The problem is usually solved by inserting appropriate band-pass or edge filters in front of the input slit in order to restrict the spectral width of the detected signal. A second disadvantage of gratings is less familiar—the diffraction efficiency of a grating can often exhibit a sharp local minimum at certain wavelengths. This is known as Wood's anomaly (the physical origin of which is based presumably on the generation of surface plasmons in the metal coating, but is not known in all its details). Such a minimum is projected onto the measured spectrum and might be misinterpreted as a spectral feature. Therefore, a correction for the spectral response of the apparatus is often necessary; this will be treated in Section 2.7.

A prism spectrometer also has its pros and cons. The main advantage over a grating device is that the apparatus is absolutely free from any problems connected with the overlap of different order spectra. In general, prisms may also have a wider dispersion range than gratings; depending on the material the prism is made of (various glasses, ionic crystals, etc.), the application range can extend from the ultraviolet to the far-infrared spectral regions. On the other hand, as the main drawback of the prism, its nonlinear dispersion along with the relevant nonlinear sweep of the spectra inducing a significant decrease in resolving power at longer wavelengths—as follows from eqn (2.25)—is usually noted.

Even if none of these two types of dispersion elements is ideal, most commercial spectrometers produced during the last 20 or 30 years are based on gratings. The advanced technology of holography has led to striking improvements in the quality of holographic diffraction gratings, which are now available in a wide range of groove densities and sizes for reasonable prices. In order to overcome the disadvantage of the relatively narrow spectral range, several interchangeable gratings are mounted on a revolving table (turret) inside the spectrometer.

We shall conclude this subsection with a short note on wavelength calibration of monochromators and spectrographs. The current commercially available spectral devices have an internal control system for rotating the grating. The exit (detected) wavelength is calculated and indicated by control software. Nevertheless, the correctness of the wavelength calibration must be checked from time to time, especially after moving the device or before some high-resolution measurements. Wavelength calibration is done using a light source of well-defined wavelengths. For a fast check we can use some lines of gas lasers, e.g. the red line of a He-Ne laser at 632.8 nm. Low-pressure discharge lamps (filled with Hg, Ar, Kr, Ne, etc.), whose emission wavelengths are tabulated, are considered to be the best sources of calibration spectral lines. Calibration of a monochromator is checked by scanning the spectrum of a calibration lamp using narrow slits and comparing the result with the tabulated wavelength values. Calibration of a spectrograph is more sophisticated: in the first step the 'monochromator-like' calibration at the centre of a multichannel detector (more precisely, at the point of intersection of the optical axis of a spectrograph with the exit image plane) must be checked, and the second step calibrates the dispersion of wavelengths on the multichannel detector (off-axis points). The appropriate calibration procedures should be integrated in the control software and described in a user manual.

2.3.2 Throughput of monochromators and spectrographs

In order to define the term *throughput* we begin by recalling the definition of the *brightness* (*radiance*) of an extended emitter, Fig. 2.21(a). The radiant flux $\Delta\phi_i$, emitted by a surface element ΔS of the emitter, is proportional to the size of this element, to the cosine of the angle ϑ between the direction of radiation and the normal to ΔS, and to the magnitude of the pertinent solid angle $\Delta\omega$, i.e. $\Delta\phi_i \sim \Delta S \cos\vartheta\,\Delta\omega$. The proportionality factor in this equation is just the brightness of the emitter B:

$$\Delta\phi_i = B\Delta S \cos\vartheta\,\Delta\omega. \tag{2.30}$$

Suppose the element ΔS is represented by the uniformly illuminated (effective) area of a monochromator entrance slit. Due to the narrow field of view of the collimator optics we can take $\cos\vartheta \approx 1$ and approximate the angle $\Delta\omega$ via the relation $\Delta\omega \approx W^2/f^2$, see Fig. 2.21(b). Equation (2.30) then reads

$$\phi_i \approx B\,S\,(W^2/f^2). \tag{2.31}$$

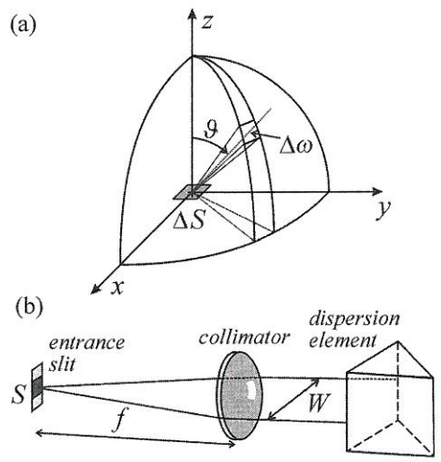


Fig. 2.21
Illustrating the definitions of (a) the brightness of an extended flat emitter and (b) the aperture ratio of a spectral device.

Let us first consider the case of a *monochromator*. Here an essential quantity is the (quasi-)monochromatic luminous or radiant flux ϕ_λ at the exit slit that induces the photoelectric response of a single-channel detector. The efficiency and/or transmittance of all optical elements inside the monochromator (mirror reflectivity, lens transmittance, efficiency of the diffraction grating, or prism transmittance) at a wavelength λ can be included in the coefficient $\kappa(\lambda) < 1$. Then, based on eqn (2.31), we can write for the output flux

$$\phi_\lambda = \kappa(\lambda)\,\phi_i = \kappa(\lambda)\,B\,S\,\left(W^2/f^2\right). \tag{2.32}$$

The *throughput* of the monochromator is then defined as

$$\sum{}_{\text{mono}} = \frac{\kappa(\lambda)\,\phi_i}{B} = \kappa(\lambda)\,S\,(W^2/f^2) \approx \kappa(\lambda)\,S(A/f^2). \tag{2.33}$$

Here A stands for the effective area of the dispersion element. Note that the throughput is proportional to the maximum solid angle that just illuminates the whole area of the collimator. It should be noted that the definition of the throughput is not standardized. Sometimes it is taken simply as (A/f^2) or even $\sqrt{A/f^2} \approx W/f$. Here, in analogy with a photographic camera, the aperture ratio is given by 1/#, where (aperture number) $\# = f/$(diameter of the entrance pupil). Producers of spectral devices usually give the throughput[4] in the form of $(f/\#)$, with $\# = f/W$ ranging from about 3.5 to 20 (the bigger this number the lower is the throughput of the device!). This simplified (commercial) definition of the throughput does not take into account the factor S in eqn (2.33), i.e. the variability of the slit width.

In luminescence spectroscopy, dealing mostly with weak signals, spectral devices with a large throughput are usually preferred. However, similar to the mutual interconnection between dispersion and resolving power, there exists a simple relation between the throughput and resolving power of a

[4] Unfortunately, the terminology is not universal. Different terms like luminosity, throughput, aperture, aperture ratio, acceptance angle, etc. are used.

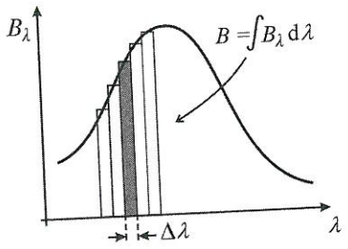


Fig. 2.22
A continuous spectrum characterized by its spectral brightness B_λ. The entrance slit is illuminated by the spectrally integrated brightness $B = \int B_\lambda d\lambda$.

device. In what follows we shall derive this relation and discuss its practical consequences.

Let w and h be the width and height of the illuminated part of the entrance slit, respectively. Then $S = wh$ and the quantity $\alpha = w/f$ stands for the angular width of the entrance slit as 'seen' from within the device. Analogously, $\beta = h/f$ means the angular height of the slit. Consequently, $\alpha\beta = S/f^2$ and instead of eqn (2.32) we write

$$\phi_\lambda = \kappa \, B \, \alpha \, \beta \, W^2. \tag{2.34}$$

Using the relations for the linear and angular dispersions (2.19) and (2.17), respectively, the transmitted bandwidth $\Delta\lambda$ determined by the slit width[5] w reads

$$\Delta\lambda = L^{-1} w = \frac{1}{f\,\mathscr{D}} w = \alpha \, \mathscr{D}^{-1}$$

and eqn (2.34) takes the form

$$\phi_\lambda = \kappa \, B \, \beta \, \Delta\lambda \, \mathscr{D} \, W^2. \tag{2.35}$$

In case of a sufficiently broad continuous spectrum (common in luminescence) we can introduce the 'spectral brightness' of the entrance slit B_λ by replacing $B \to B_\lambda \Delta\lambda$ (see Fig. 2.22). Then, the output flux is given by

$$\phi_\lambda = \kappa \, B_\lambda \, \beta \, \mathscr{D} \, W^2 (\Delta\lambda)^2, \tag{2.36}$$

which can be further modified by introducing the real resolving power R_{real}

$$\phi_\lambda = \frac{\kappa \, B_\lambda \, \beta \, \mathscr{D} \, W^2 \lambda^2}{R_{\text{real}}^2}. \tag{2.37}$$

The final relation between resolving power and throughput Σ_{mono} we are looking for is derived from (2.37) using eqns (2.32) and (2.33) and reads

$$R_{\text{real}} = \frac{\kappa \, \beta \mathscr{D} W^2 \, \lambda}{\Sigma_{\text{mono}}}. \tag{2.38}$$

This relation shows that the real resolving power and the (real) throughput of a monochromator are, in agreement with intuition, inversely proportional to each other—widening the slits boosts the output signal but reduces the spectral resolution. Again, the experimenter must find a viable compromise using his/her experience. However, more interesting conclusions can probably be drawn from eqn (2.37):

(a) The transmitted (output) light flux ϕ_λ scales with the resolution like $(1/R_{\text{real}})^2$. This means that the output signal decreases quadratically with increasing resolution. In other words, the output signal level increases with the spectral (and thus also mechanical) slit width like $(\Delta\lambda)^2$; for that matter see also eqn (2.36). This relation can sometimes be used during an experiment to estimate quantitatively the signal level. If we need (e.g. due to low signal levels at the emission band edges) to increase the 'detection sensitivity', one option is to set wider slits. Now, we know that doubling the slit width entails a fourfold increase in signal level. We have to keep in mind, however, that the above quadratic relation holds well for broad structureless spectra only (as follows from the introduction of the spectral brightness B_λ, see Fig. 2.22). Besides, there are plenty of more precise means to increase sensitivity.

(b) The output flux ϕ_λ scales linearly with β, the angular height of the illuminated slit. Therefore, it appears highly desirable to illuminate the whole height of the entrance slit, which, however, is difficult to arrange because the emitting spot is commonly of circular shape. Nevertheless, when the signal is brought by an optical cable, its termination can be shaped to fit the entire surface of the entrance slit. Also in cases of weakly absorbed or homogeneous two-photon excitation of a sufficiently voluminous sample by a laser beam (Chapter 5, Fig. 5.15(b)), the elongated luminescence spot can be oriented so that its image covers the entrance slit.

(c) The relation $\phi_\lambda \sim W^2$ implies that a high-throughput monochromator has to be equipped with a large grating or prism.

[5] We consider a symmetrical device with equal widths of both slits, consistent with the previous text.

In the case of a *spectrograph* it is the so-called *irradiance* of a multichannel detector (the radiant flux at the spectrometer output divided by the illuminated area) that has a crucial role to play, rather than the radiant or luminous flux itself. This quantity determines the exposure time and the signal-to-noise ratio. Luminescence radiation, spatially distributed by a dispersion element, projects the entrance slit area $S = wh$ characterized by the spectral brightness B_λ to its image $S' = w'h'$ (in the plane of the multichannel detector, see Fig. 2.23), the dimensions of the slit and its image being identical ($w = w'$, $h = h'$) in the case of a symmetrical device. In analogy with definition (2.33) we define the throughput of the spectrograph Σ_{spectro} using the irradiance $(\kappa(\lambda)\phi_i/S')$ on the area $S' = w'h'$ of the slit image created by quasi-monochromatic light $\lambda \pm \Delta\lambda/2$ as

$$\sum_{\text{spectro}} = \frac{\phi_\lambda}{B_\lambda \, \Delta\lambda \, S'} = \frac{\phi_i \, \kappa(\lambda)}{B_\lambda \, \Delta\lambda \, S'} = \frac{S}{S'} \kappa(\lambda) \left(\frac{W}{f}\right)^2 = \kappa(\lambda) \left(\frac{W}{f}\right)^2. \tag{2.39}$$

Now we can express the 'spectral irradiance' of the detector with respect to its full height $h_0 \geq h'$ as $E_\lambda = \phi_\lambda / h_0 w$, as follows from Fig. 2.23. This expression for E_λ can then be easily modified using eqn (2.32) and the definition of the aperture number # as

$$E_\lambda = \frac{\phi_\lambda}{h_0 \, w} = \frac{\kappa(\lambda) \, B_\lambda \, \Delta\lambda \, S \, (W/f)^2}{h_0 \, w} = \left(\frac{h}{h_0}\right) \kappa(\lambda) \, B_\lambda \left(\frac{1}{\#}\right)^2 \frac{\lambda}{R_{\text{real}}}. \tag{2.40}$$

Relation (2.40), valid for a spectrograph, is analogous to expression (2.37) relevant to a monochromator. Here, again, are some conclusions important for experiments:

(a) The irradiance E_λ decreases proportionally to $(1/R_{\text{real}})$, unlike the transmitted flux at the exit slit which scales as $(1/R_{\text{real}})^2$.

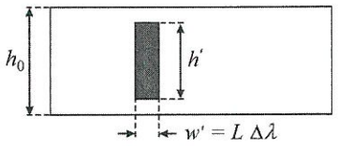


Fig. 2.23
Image of the entrance slit on a multichannel detector located in the spectrograph exit plane. h_0 is the height of the sensitive area of the detector, and h' and w' stand for the height and width of the entrance slit image, respectively.

(b) The irradiance E_λ is proportional to (h/h_0). Therefore, it is desirable to choose a combination of spectrograph, multichannel detector and a method to illuminate the entrance slit so that $h = h_0$, in order to maximize E_λ. Naturally, the case $h > h_0$ should be avoided because part of the signal flux fails to meet the detector.

(c) The irradiance E_λ scales like $(1/\#)^2$, i.e. the irradiance of the detector increases with the square of the aperture ratio. This factor is actually the most important one among those aimed at maximizing the detector irradiance.

2.4 Signal detection methods in luminescence spectroscopy

Following the discussion of spectral devices and detectors we are now going to describe briefly methods of signal treatment in spectroscopy. Essentially this means dealing with the principles of electronic devices that treat the electric signal from the photodetector. Here we concentrate on single-channel detectors because multichannel ones are quite straightforward and user friendly—the accumulated photoinduced charges, after being digitized, are simply transferred to a computer to be processed by software.

For the sake of simplicity we consider in this section only continuous-wave (cw) luminescence signals, i.e. signals excited by a continuously emitting lamp or laser. In particular, we are going to describe two methods: phase-synchronous detection and the photon-counting technique. Experimental methods of pulsed luminescence signals are left to Section 2.9.

2.4.1 Phase-synchronous detection

The simplest arrangement of optical cw signal detection is obviously direct current detection—time invariant radiant flux (at the monochromator exit slit) induces in the attached photomultiplier a constant photocurrent, which is subsequently directly measured by a voltmeter connected to a load resistance (Fig. 2.5(b)). In the case of extremely low luminescence photon fluxes a DC preamplifier could be inserted between the photomultiplier and the voltmeter. However, in reality such DC measurements are rarely applied, because low photocurrents are commonly comparable in magnitude with the photomultiplier dark current or only slightly higher. The dark current amplitude must be subtracted in this case from the detected signal, which leads to important errors like in any other indirect measurements where two comparable quantities are to be subtracted. The inevitable presence of broad-frequency noise makes the quality of the resulting output even worse because there is no electronic component available for noise suppression.

It therefore becomes obvious that weak continuous optical signals can be successfully detected only when applying some electronic device that is able to extract reliably a signal buried completely in the dark current and/or in various types of noise whose amplitude may be higher by several orders of

magnitude. This is possible, first and foremost, by employing the method of phase-synchronous detection through a device called a *lock-in amplifier*. This device is able to measure alternating voltage or currents and to provide a corresponding DC output voltage that scales linearly with the input signal level (the output signal level is usually much higher than the input level, hence the term 'amplifier'). Here, only certain frequencies of the input signal are detected while virtually all the others are suppressed (mainly all noise with a broad frequency spectrum), hence the term 'lock-in'. The ability of a synchronous detector to extract and amplify a low synchronous signal buried in noise is characterized by a quantity called the *dynamic reserve*. This is the ratio of the maximum acceptable amplitude of the asynchronous (noise) signal (i.e. the amplitude value which just starts to induce nonlinearity and saturation of input circuits) to the peak synchronous signal giving a full-scale DC output. The dynamic reserve of up-to-date lock-in amplifiers is extremely high, routinely reaching values of 100 dB (!) or even higher.

Any lock-in amplifier contains two key parts: a mixer (phase detector) and an integrator. The principle of operation of a lock-in amplifier is illustrated in Fig. 2.24. The input to the mixer (left side of Fig. 2.24) receives two electric signals—the measured sine wave signal e_1 from a photodetector and a square wave so-called reference signal e_2 (the origin of these particular forms of signals will be described later). The output of the mixer e_3 is the product of the input signals: $e_3 = e_1 e_2$. The technical solution consists in alternating usage of a direct and an inverting amplifier in the mixer input. The amplitude of the reference signal E_2 can be put equal to unity. Then Fig. 2.24(a) shows immediately that for equal frequencies of both the signals ($f_1 = f_2$) and provided their relative phase shift is equal to $0, 2\pi, 4\pi$, etc. the output mixer signal e_3 contains a pulsating voltage of positive polarity only. The measured signal is said to be synchronous with the reference. Enabling e_3 to pass through the integrator having a time constant $\tau = RC$ (right-hand side of Fig. 2.24), a DC output voltage with amplitude E_{DC} proportional to the amplitude of the measured input signal E_1 is obtained.

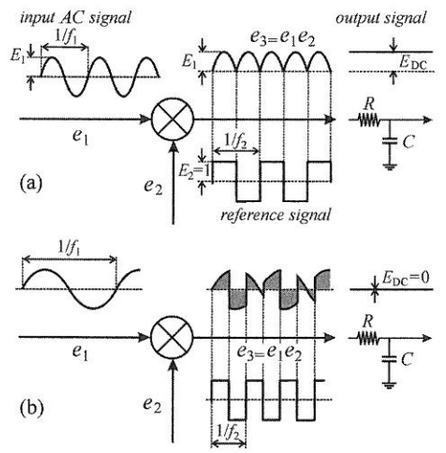


Fig. 2.24
Principle of a lock-in amplifier. The left-hand side shows a mixer and the right-hand side an integrator (low-pass filter). (a) For a synchronous input signal $f_1 = f_2$ we obtain a non-zero DC output signal $E_{DC} \neq 0$. (b) Asynchronous input signal $f_1 \neq f_2$ along with a sufficiently long time constant $\tau = RC$ result in a zero DC output level.

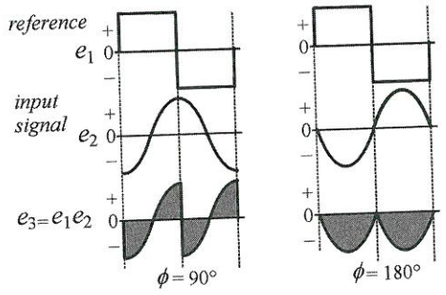


Fig. 2.25
Demonstration of the phase sensitivity of a lock-in amplifier. The input signal from a photodetector e_1 and the reference signal e_2 have equal frequencies but the DC output is strongly dependent on the phase shift between these two signals.

In the case of an asynchronous signal $f_1 \neq f_2$ (Fig. 2.24(b)) the signal e_3 contains both positive and negative components compensating each other if a long enough time constant τ is applied, and thus zero DC output results. It can be seen that the lock-in amplifier acts as an extremely selective filter processing and amplifying only the 'meaningful' signal component, i.e. that corresponding to the reference frequency; this is the succinct essence of the principle of the lock-in amplifier.

There are, however, additional useful features of synchronous detection that remain to be explained. First of all, these is the *phase sensitivity*, which strongly influences the magnitude of the output signal—for certain values of the phase shift between the signal and the reference, the output E_{DC} may be zero even if the measured and reference signals have the same frequency $f_1 = f_2$(!). Two examples ($f_1 = f_2$ while the phase shift is different from $2\pi n$, n being an integer) are demonstrated in Fig. 2.25: the phase shift $\phi = 90°$ (or $(2n + 1)\pi/2$) induces zero DC output, while under a shift of $\phi = 180°$ (or $(2n + 1)\pi$) the output signal E_{DC} is of the same amplitude as for $\phi = 0°$ but of negative polarity. This is why a lock-in amplifier is sometimes also called a phase-sensitive detector.

The phase sensitivity has several interesting consequences. But prior to entering the relevant discussion we should describe how cw luminescence radiation is converted into a sine-modulated input signal and how the reference signal is generated. Everything is based on simple mechanical chopping of the light beam (as a rule, the excitation beam is chopped, less commonly chopping of the photoluminescence signal is applied—see Fig. 2.2(b)) by a mechanical chopper, which is a rotating disk with periodically arranged slots. (Of course, mechanical chopping itself cannot create the sinusoidal signal shown in Figs 2.24 and 2.25, nevertheless, the input circuits of a lock-in amplifier are able to adapt a chopped cw signal into a sine wave.) The reference signal is generated by an optical pair (a LED and a photodiode) attached to the chopper. In this way the reference and measured signal are modulated by the same mechanical system and undoubtedly $f_1 = f_2$ holds true. However, there is certainly a non-zero phase shift ϕ between both signals due to the different response times of the photomultiplier and the reference photodiode along with different delays introduced by wiring and circuits on the path of both signals towards the mixer input.

Here, the phase sensitivity of the lock-in amplifier appears to be very useful. The phase shift ϕ between the two input signals can be continuously tuned

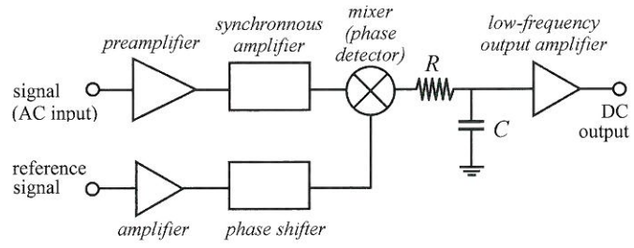


Fig. 2.26
Block diagram of a phase-sensitive detector. Both the modulated input and the reference signal are first amplified and possibly shape-adapted. The reference signal branch contains an adjustable phase shifter. After passing through the *RC* circuit (low-pass filter) the signal is amplified again.

to maximize the output signal amplitude E_{DC}. Up-to-date lock-ins can do this phase tuning automatically. Another useful feature is the possibility of alternating the polarity of the output simply by setting the phase shift to $\pm 180°$. The converse is also true: if the device, on its own, automatically changes the phase by $\pm 180°$ during measurement, this indicates that the polarity of the input signal has changed, which may carry important experimental information. (While one can hardly meet such a situation in luminescence spectroscopy, it is, however, quite common in photoelectric measurements.) Finally, the phase sensitivity proves very useful in filtering out all kinds of noise, even if their frequencies may happen to be equal $f_1 = f_2$, because they have no fixed phase relation to the reference signal, i.e. their phases fluctuate randomly.

The time constant $\tau = RC$ also deserves a short discussion. The illustrations in Figs 2.24 and 2.25 are oversimplified. It can be shown that the output of the mixer can contain not only pulsating components with frequency $f_1 = f_2$, but also with $f_1 = (2n + 1) f_2$ where n is an integer. Moreover, asynchronous input frequencies reveal themselves at the mixer output as sum and difference frequency components $(f_1 + f_2)$, $(f_1 - f_2)$ and higher harmonics. These frequencies could possibly appear as noise superimposed on the output voltage E_{DC} but they are mostly filtered out by the RC circuit, which then acts not only as an integrator but also as a low-pass filter. A block diagram of a lock-in amplifier is shown in Fig. 2.26.

Concerning the optimum choice of the time constant τ, smoothing of the DC output is evidently improved with increasing time constant (available values of τ range from microseconds to tens of seconds) but, on the other hand, very high values of τ make the overall response of the device slower. This means that the speed of scanning a spectrum across the exit slit must be adapted to the selected time constant, otherwise the acquired spectral shape may be distorted. Very long time constants can lead to intolerably long acquisition times. This problem will be discussed in more detail in Section 2.5.

We shall conclude our discussion of the synchronous detection technique by mentioning the limits of the method. One limit arises from luminescence kinetics and its relation to the chopping rate. If the mechanical chopper modulates the excitation beam, the expected square wave photoluminescence signal may be distorted owing to the finite speed of the rise and decay of luminescence (Fig. 2.27). If the relevant characteristic time constants are comparable to or even longer than the chopping period $T_2 = 1/f_2$, the lock-in amplifier input signal loses its modulated shape and can no longer be processed by the device (Fig. 2.27(d)). Possible solutions are obvious—either decrease the chopping

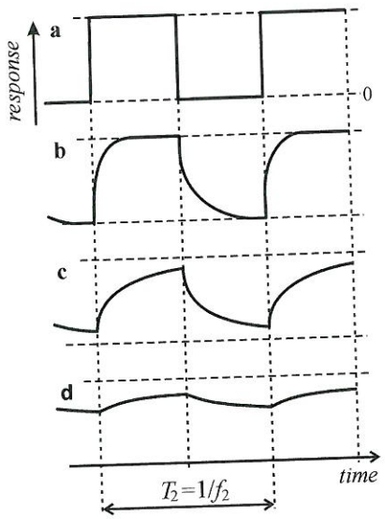


Fig. 2.27
Temporal trace of the photoluminescence response excited by a mechanically chopped beam with frequency f_2. Curve (a) pertains to luminescence with very fast kinetics, i.e. the rise and decay times are short compared to $T_2 = 1/f_2$; curves (b)–(d) illustrate luminescence with slower and slower response. A lock-in amplifier will perceive case (d) as almost an unmodulated DC signal and will not respond to it.

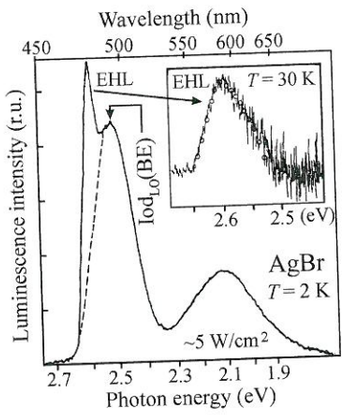


Fig. 2.28
Low-temperature photoluminescence spectrum of AgBr under excitation with a cw UV laser and using low-frequency synchronous detection ($f_2' = 500$ Hz). The spectrum contains three bands: luminescence of the electron–hole liquid (EHL) at ~2.6 eV, a band $Iod_{LO}(BE)$ due to radiative recombination of excitons bound to ions of an isoelectronic unintentional dopant (iodine) at ~2.5 eV, and a band at ~2.1 eV due to a residual impurity. The inset shows the spectrum obtained under high chopping frequency $f_2'' = 100$ kHz, which eliminates the iodine-related band $Iod_{LO}(BE)$ with a decay time of ~ 25 μs and retains only the fast band of EHL with a decay time of about 20 ns. Adapted from Pelant *et al.* [17] and Hulin *et al.* [18].

frequency f_2 (which is not always possible) or modulate the luminescence instead of the excitation beam (which is experimentally not so easy to do). On the other hand, the above drawback can sometimes be turned into an advantage and applied as a simple tool for time-resolved luminescence spectroscopy.

Suppose that the material under study shows two luminescence emission bands, partially overlapping spectrally but having substantially different kinetics. Then these spectral peaks may be cleanly separated by performing two measurements with a lock-in amplifier under two different chopping frequencies $f_2' \ll f_2''$. While the measurement at frequency f_2' reveals a spectrum comprising both the bands, the measurement at frequency f_2'' shows only the 'fast' band provided this frequency is high enough to eliminate the slow component. An example is shown in Fig. 2.28 [17, 18].

The second limit beyond which application of the lock-in technique serves no useful purpose is detection of very low light fluxes, where the discrete character of both light (photons) and detection events (photoemission from the photomultiplier photocathode) becomes apparent. This situation is nicely illustrated by Fig. 2.29. The lock-in amplifier cannot detect very weak signals since they are formed by an array of random anode pulses (Fig. 2.29(c)) rather than by a DC component processable through chopping into a form suitable for phase-sensitive treatment. Detection of such low photon fluxes requires another method called photon counting.

2.4.2 Photon counting

Every current pulse at the output of a photomultiplier, shown in Fig. 2.29(c), corresponds to the impact of one photon on the photocathode, accompanied by emission of one photoelectron. Therefore, the number of pulses N_r registered per unit time is proportional to the number of photons N_h hitting the photocathode during the same time interval, i.e. to the incident light flux or

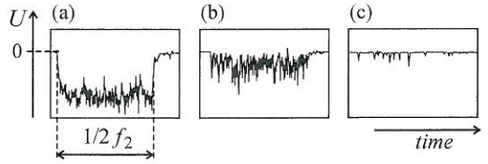


Fig. 2.29
Visualization of the output signal of a photomultiplier (chopped at frequency f_2) on an oscilloscope screen for different signal levels. (a) The light flux hitting the detector is high enough, single output voltage pulses are merged together. (b) The number of incoming photons is lower; pulses corresponding to the detection of single photons can already be resolved. (c) Very low signal photon flux. The signal from the detector is composed of discrete voltage pulses; there is almost no continuous component. Therefore the lock-in sees only incoherent noise and the DC output level E_{DC} is zero. The signal has negative polarity, which corresponds to the most common photomultiplier wiring with a grounded anode (Fig. 2.5b). Modified from Hamamatsu [7].

intensity. (In quantitative measurements, e.g. those of the quantum efficiency of luminescence, we have to keep in mind that these quantities are only proportional to each other, not equal. The number of registered pulses N_r is always lower than the number of incoming photons N_h, the proportionality coefficient being the product of the photocathode quantum efficiency η_{ph} and the collection efficiency of dynodes α_d, which means $N_r = \eta_{ph}\alpha_d N_h$.)

A block diagram of a photon counting apparatus is presented in Fig. 2.30. The output signal from the photomultiplier is composed of single voltage spikes (as already illustrated in Fig. 2.29(c)), generally superimposed on various low- and high-frequency noise and instabilities—circle A. Big pulses are due to photoelectrons emitted from the photocathode while smaller ones come from the thermionic emission in dynodes. This signal is treated by an input amplifier, which multiplies the signal approximately one thousand times and filters out the low-frequency ripple—circle B. The amplified pulses enter a discriminator (circle C), which eliminates pulses that are below a certain adjustable discrimination level (dashed line in the circle B). In this way the parasitic pulses originating in the thermionic emission of dynodes are ruled out. Each pulse passing through the discriminator generates at the output a standard TTL voltage pulse (circle C) and these pulses are counted by a common counter. The counter can send out either digital information about the number of pulses per unit time or it can produce an analogue signal (for a plotter) employing a digital-to-analogue converter.

The principle of the photon-counting method looks simple but the apparatus requires excellent temporal and noise characteristics of all of its components. The most critical is selection of a suitable photomultiplier tube meeting stringent criteria imposed on its parameters, as well as the selection of its high anode–cathode voltage and of the appropriate discrimination level of the discriminator. These issues deserve to be discussed in more detail.

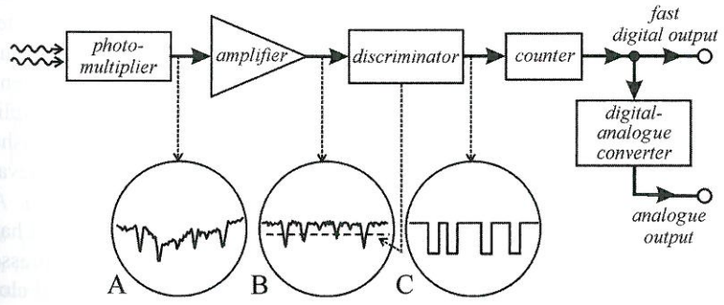


Fig. 2.30
Block diagram of a simple version of a photon counter.

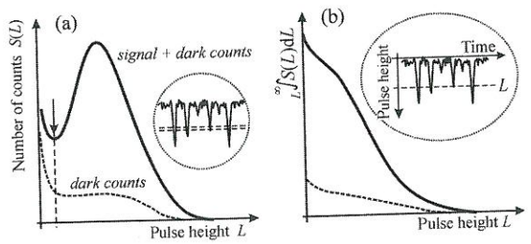


Fig. 2.31
(a) Qualitative illustration of a (differential) distribution of photomultiplier pulse heights. The dashed and solid curves are the height distributions for 'dark' pulses (photocathode not illuminated) and for 'light' pulses (signal photons hit the photocathode), respectively. Such distributions can be measured by making use of a multichannel analyser or a discriminator with adjustable upper and lower levels (see inset). (b) An integral distribution of pulse heights, which can be experimentally acquired with a single-level discriminator. Adapted after a Hamamatsu brochure [7].

By no means is every photomultiplier applicable to the photon-counting method. In selecting a suitable photomultiplier the following factors have a crucial role to play:

1. Temporal width of the output voltage pulses. This width must be sufficiently short in order to distinguish the height of pulses due to photoelectrons from that of pulses originating in thermionic emission in dynodes. Excess temporal blurring of the electron packets propagating through the photomultiplier could reduce the contrast in amplitude between the two kinds of pulses so that their separation in the discriminator would be impossible. Moreover, shorter pulses enable the photon-counting technique to be applied to higher pulse rates—the upper limit of pulse counting corresponds to the situation in which adjacent pulses become significantly overlapped and thus indistinguishable. A common value of the temporal pulse halfwidth in photomultipliers designed for photon counting is about 10–20 ns.
2. A high quantum efficiency η_{ph} of photocathode emission and a low anode dark current. It is evident from the very principle of the method that a high signal-to-noise ratio requires high yield of generation of primary photoelectrons from photons hitting the photocathode, along with minimization of the number of thermally released electrons. Quantum efficiencies of suitable photomultipliers are usually around $\eta_{ph} = 20$–30%.
3. Appropriate distribution of pulse heights. Firstly, we have to explain what we understand by the distribution of pulse heights. One can hardly expect that all the pulses or counts originating from a single photoelectron event on the photocathode (or 'dark' counts initiated by thermally released electrons) would have the same amplitude on the anode; the statistical character of secondary electron emission at the dynodes, individual variations in electron trajectories, etc. make the anode pulse heights of both the signal and 'dark' counts fluctuate. The average amplitude of signal counts is expected, naturally, to be higher compared to that of dark counts. This phenomenon is illustrated in Fig. 2.31(a) which shows a histogram $S(L)$ of the occurrence of pulses with height L; this is called the distribution of photomultiplier pulse heights. The centre of gravity of the dark-count distribution (dashed curve) is clearly left-shifted compared to the pulse distribution relevant to an illuminated photomultiplier (signal+dark counts, solid curve). An optimum pulse distribution for the photon-counting technique should have a pronounced—and as narrow as possible—peak with strongly suppressed occurrence of small pulses. The discrimination level is then adjusted close

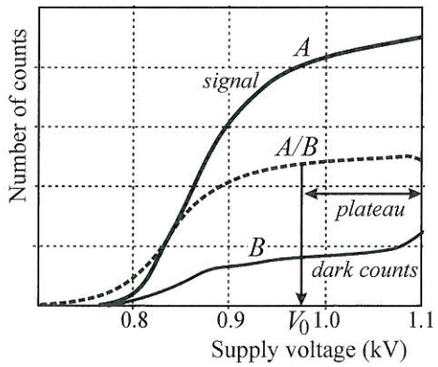


Fig. 2.32
Number of counts as a function of voltage applied to a photomultiplier in the photon-counting technique (important is the occurrence of a plateau on the curve A/B). Adapted after a Hamamatsu brochure [7].

to the local minimum of the distribution curve (see the arrow in Fig. 2.31(a)) in order to reject efficiently the dark pulses.

Here, however, the experimenter is faced with a problem. Photomultiplier manufacturers usually do not provide curves of pulse height distributions in their catalogues. Therefore, it appears advisable for the user to perform his/her own measurement of these characteristics prior to application of a particular photomultiplier[6] in photon-counting detection. A direct measurement of distributions as shown in Fig. 2.31(a) is possible with a multichannel analyser of pulse heights or using a discriminator that has two adjustable discrimination levels (such a narrow discrimination 'window' is scanned from low- to high-amplitude pulses and the occurrence of pulses belonging to this span of heights is recorded). Commonly, however, the discriminator has only one discrimination level. In this case all pulses with amplitude higher (in absolute value) than the adjusted discrimination level are recorded, resulting in the curve shown in Fig. 2.31(b). One can easily infer that this curve can be converted to the peak height distribution via numerical differentiation. Therefore, the curve (b) displayed in Fig. 2.31 is sometimes called an *integral representation*, while the curve (a) is called a *differential representation* of the pulse height distribution.

4. The choice of the photomultiplier anode–cathode high voltage and of the discrimination level. Suppose we have already acquired the differential curve of the height distribution and adjusted the discrimination level approximately to the local minimum (the arrow in Fig. 2.31(a)). We are now in a position to find a proper operational high voltage. The procedure is as follows: the applied voltage is varied in small steps (allowing the operation conditions to be stabilized adequately), and the number of pulses is recorded at each step. This is done for both non-illuminated and illuminated photomultipliers. Curves similar to those plotted in Fig. 2.32 are obtained. The number of signal pulses (upper curve) increases with increasing voltage faster than the number of dark pulses (lower curve) because the latter are

[6] There are photomultiplier types especially fabricated for use in photon-counting applications. But the manufacturer can also select an exeptionally suitable exemplar from other ranges of photomultipliers.

already substantially restricted by the discriminator. The ratio of these two curves (dashed curve in the middle) increases rapidly as the high voltage raises but then saturates, achieving a plateau. (The occurrence of a broad plateau is another prerequisite imposed upon the photomultiplier.) The correct applied high voltage is then situated in the plateau region that starts at a voltage V_0 at which the ratio of signal to dark pulses reaches saturation (Fig. 2.32) and that extends basically up to the maximum allowed voltage. The output signal is thus to a large extent independent of the anode–cathode voltage and its fluctuations, featuring an obvious benefit of the photon-counting technique.

It is important to realize that proper choice of the operational high voltage may comprise some ambiguities. The distribution of pulse heights shown in Fig. 2.31 was obtained with a certain pre-selected photomultiplier high voltage. If, later, we found that this bias was, unfortunately, outside of the plateau region, another operational voltage must be chosen, which may influence the shape of the pulse height distribution. Then, in principle, we shall have to remeasure the curves shown in Fig. 2.31, find a new discrimination level and determine a new plateau characteristic. Possibly, the operational voltage will then have to be shifted again, etc. and we risk ending up in a closed loop. The solution is a compromise based both on the experimenter's experience and on the requirements of the particular experiment. Obviously, shifting the discrimination level down (to the left side from the local minimum of the height distribution) increases the overall efficiency of pulse counting, while shifting it to higher pulse amplitudes (roughly towards the middle between the local minimum and the distribution peak) rejects more noise pulses and improves the signal-to-noise ratio.

We shall close our discussion of photon counting with a few final notes. First of all, let us remind ourselves that not only a photomultiplier but also an avalanche photodiode can be applied in the photon-counting method (Section 2.2). In fact, we have also already mentioned the natural limits of application of the technique. The high count rate limit is set by the resolution of two closely following pulses, where both the temporal width of the pulse and the dead time of the detector and discriminator are significant. A typical value of the maximum pulse rate is $1\text{–}3 \times 10^6\,\mathrm{s}^{-1}$, which corresponds to a photon flux of about 10^7 photon/s (i.e. optical power $\sim$ 3 pW) hitting a photomultiplier with a photocathode quantum efficiency of $\eta_{\mathrm{ph}} = 20\%$. The low count rate limit is determined by various noise characteristics and by the rate of dark counts. Using an up-to-date cooled photomultiplier one can easily detect signals of tens of photons per second, but if the entire spectrum is to be recorded in this case, then the overall measurement time already has a decisive role to play, as will be analysed in Section 2.5.

A troublesome peculiarity of the photon-counting technique is its excessive receptiveness to external perturbations. A photon-counting apparatus can register, for example, spurious pulses generated by a stepping motor that drives the dispersion element located inside the monochromator or high-frequency perturbations from other laboratory equipment, originating possibly even in neighbouring rooms. The only remedy is perfect shielding of all instruments

including usage of well screened cables, good earthing, and sometimes finding an optimum location of the apparatus within the laboratory. The spheres of applicability of the lock-in and photon-counting techniques overlap over a certain range of photon fluxes. If both techniques are usable, then synchronous detection is usually preferred as it is simpler and less sensitive to external perturbations.

For more information on the photon-counting technique, noise analysis, dead times, accuracy, etc. the reader is referred to, e.g., [6, 19].

2.5 Signal-to-noise ratio in a scanning monochromator

In the preceding sections we examined the basic characteristics of spectral devices, introduced parameters describing the properties of optical detectors, and analysed the two basic techniques of signal detection in luminescence spectroscopy—synchronous detection and photon counting. Now, we are ready to deduce a relatively simple expression for the signal-to-noise (S/N) ratio [20] of luminescence spectra. This is of basic importance for laboratory practice.

Consider a scanning monochromator, provided at the exit slit with a detector of sensitivity $k(\lambda)$, normalized detectivity D^* and sensitive area $\mathscr{A}$. The detector is illuminated by a monochromatic radiant flux ϕ_λ. The definition (2.7) of D^* can be modified using (2.6) to

$$D^* = \sqrt{\mathscr{A}}/NEP = \frac{\sqrt{\mathscr{A}}\sqrt{\Delta f}\,S}{\phi_\lambda N} = k(\lambda)\,\frac{\sqrt{\mathscr{A}}\sqrt{\Delta f}}{N}. \qquad (2.41)$$

To be specific, let the output from the detector be treated by a lock-in amplifier. Then the frequency bandwidth Δf of the detection path is driven by the integrator time constant τ as $\Delta f \approx 1/2\pi\tau$. This can be understood as a straightforward result of a Fourier transform, or—in a more illustrative way—we can imagine the signal as being recorded by a classical plotter (or traced in real time on an oscilloscope screen) when the time constant τ determines the speed of movement of the recorder pen. Suppose the spectrum contains very sharp features, for example an intense narrow line; then the fast upward movement of the pen is damped by the time constant τ. In the case of excessively long τ, the peak is not well reproduced, is smaller, broader and shifted along the direction of wavelength scanning. It may happen that less intense spectral peaks even disappear entirely from the record. This implies that, in order to detect a spectrum without distortion, the wavelength scanning speed must be adapted to the applied time constant τ.

The mathematical formulation of this requirement is simple. Crucial for the scanning speed are the sharpest features in the spectrum. The narrowest peak, if it is to be recorded perfectly, must be scanned for a time T' significantly longer than τ (Fig. 2.33), which means

$$\tau = \varepsilon\,T'\,(\varepsilon \ll 1),\quad \Delta f \approx 1/2\pi\tau. \qquad (2.42)$$

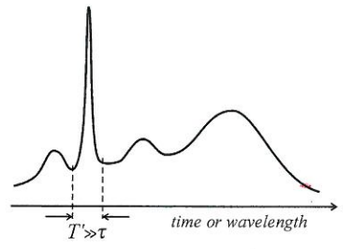


Fig. 2.33
Scanning of emission spectrum in real time. The recording time T' of the narrowest peak must be much longer than the time constant τ of the detection system.

Now, the noise level N in eqn (2.41) can be expressed using eqn (2.42) as

$$N = k(\lambda)\,\frac{\sqrt{\mathscr{A}}\,\sqrt{\Delta f}}{D^*} = \frac{k(\lambda)}{D^*}\sqrt{\frac{\mathscr{A}}{2\pi\varepsilon T'}}. \tag{2.43}$$

The magnitude of the signal S as registered by the detector for the output radiative flux ϕ_λ (2.37) is given by

$$S = k(\lambda)\,\phi_\lambda = k(\lambda)\,\frac{\kappa\, B_\lambda\, \beta\, \mathscr{D}\, W^2\lambda^2}{R^2_{\text{real}}}. \tag{2.44}$$

Finally, the signal-to-noise ratio we are looking for follows from the last two equations as

$$\frac{S}{N} = \frac{\kappa\, B_\lambda\, \beta\, \mathscr{D}\, W^2\lambda^2}{R^2_{\text{real}}}\sqrt{\frac{2\pi\varepsilon T'}{\mathscr{A}}}\, D^*. \tag{2.45}$$

This relation hold true for the narrowest spectral line. Obviously, recording any other spectral feature, like a wider peak or band, will run for even longer time. For the detection time of the whole spectrum thus the relation $T > T' \gg \tau$ holds.

Expression (2.45) for S/N contains several interesting pieces of information and is worth analysing in detail. First of all, the detector sensitivity $k(\lambda)$ does not enter eqn (2.45). This surprising fact confirms that the S/N ratio is determined predominantly by the detectivity D or D^*. Therefore, a responsible choice of a suitable detector should not be just a matter of a cursory glance at the catalogue's data on sensitivity.

Suppose now the entire experimental set-up is ready, which means the detector and monochromator parameters (detectivity D^*, transmittance κ, effective area of the dispersion element proportional to W^2, and angular dispersion $\mathscr{D}$) have been fixed. Then, for the experimenter just four remaining parameters are left to be adjusted in order to maximize S/N: the spectral brightness of the entrance slit B_λ, the angular height of the illuminated slit β, the time constant τ (or acquisition time T), and the resolving power R_{real}. In Subsection 2.3.2 we already saw that homogeneous illumination of the whole slit height is, except in special cases, hardly possible in luminescence spectroscopy. Thus, only three adjustable parameters remain, i.e. B_λ, T, and R_{real}.

First we pay attention to the resolving power and detection time T; they are closely related. An increase in resolving power R_{real} (e.g. by setting narrower slits) is accompanied, according to eqn (2.45), by a decrease in the S/N ratio. If in this situation we want to keep the S/N constant, we can attempt to increase the time constant τ and simultaneously to set a longer acquisition time T, which means a slower spectrum scan. This method is, however, very limited. Imagine we have recorded a spectrum within $T_1 = 10$ min and found that, in order to resolve better two neighbouring spectral lines, we have to increase the resolving power (i.e. to narrow down the slits), say, 3.3 times. According to eqn (2.45), the S/N ratio remains unchanged provided we prolong the acquisition time to $T_2 \approx 100T_1 = 1000$ min ≈ 16 hours (!). This is not feasible. In other words, the influence of the detection time on the S/N ratio is strongly limited. Obviously, this is because of the weak relationship $(S/N) \sim \sqrt{T}$ and, on the other hand, to the strong dependence $(S/N) \sim 1/R^2_{\text{real}}$. Let us note that a similar weak square root relation $(S/N) \sim \sqrt{T}$ holds true also for the photon-counting technique.

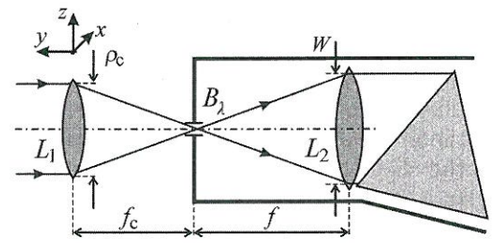


Fig. 2.34
Optical system for illuminating the spectrometer entrance slit. A collimated beam of luminescence emission comes from the left-hand side. The parameters of the focusing lens L_1 (condenser) are chosen so that the light cone inside the spectrometer just covers the whole area of the collimating element (here the lens L_2, but most of today's spectral devices utilize mirrors). At the same time this design ensures plane parallelism of the beam entering the dispersion element.

We come to the conclusion that the crucial parameter available to the experimenter is the entrance slit brightness B_λ. This can be optimized by setting up an effective optical system to collect luminescence from as wide as possible solid angle (see Figs 2.1 or 2.2, and the relevant discussion). Futhermore, the optical system focusing the luminescence signal on the entrance slit must be well designed. The basic principle is illustrated in Fig. 2.34. Luminescence emission is to be focused by a lens L_1 with minimized spherical and chromatic aberration (an achromatic doublet is preferred) exactly onto the centre of the entrance slit. Moreover, this imaging system should be designed so that the light cone entering the spectroscope just covers the whole area of the collimating lens (or mirror) L_2. Let the diameter and focal length of the lens L_1 be ρ_c and f_c, respectively; then simple geometrical similarity (Fig. 2.34) implies

$$\frac{\rho_c}{f_c} = \frac{W}{f}\left(= \frac{1}{\#}\right), \tag{2.46}$$

in other words, the aperture ratios of lenses L_1 and L_2 must be equal to one another. This equation represents a recipe for choosing the ratio ρ_c/f_c of the lens L_1 (condenser). In case relation (2.46) is not fulfilled, either part of the luminescence emission is lost for spectral analysis (if $\rho_c/f_c > 1/\#$), or the resolving power of the spectrometer is not fully exploited (if $\rho_c/f_c < 1/\#$) as the incoming light cone does not cover the whole area of the dispersion element. It is also desirable to ensure precise positioning of the properly chosen condenser L_1 with respect to the optical axis of the whole system by fixing it to a suitable $\{x, y, z\}$ translation stage equipped with micrometer adjustment screws.

Nowadays, there are many advanced numerical methods (often being part of commercial software packages) that can be applied to 'improve' noisy spectra *a posteriori*, having finished the experiment. But it is always advisable to pay attention to optimizing the luminescence signal during the experiment, especially to maximize the entrance slit brightness and simulatenously to exploit properly the throughput of the spectral device in the described way.

One way to adjust the whole optical path from the source of luminescence emission to the entrance slit is based on the principle of reversibility of light rays, a law known from geometrical optics. Here, the detector is removed

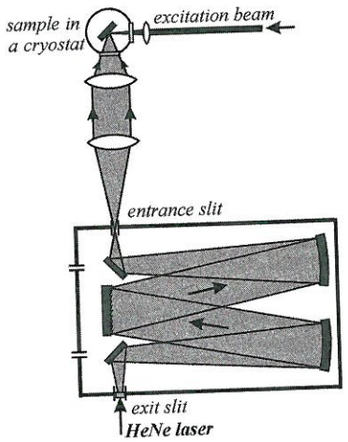


Fig. 2.35
Adjustment of the optical system for luminescence collection by reversing the direction of light propagating through a monochromator. A piece of ground glass is inserted into the exit slit.

and the exit slit is illuminated from outside, preferentially with a HeNe laser (633 nm) or other gas or diode laser emitting visible light, as shown in Fig. 2.35. A small piece of ground glass or calking paper is inserted into the plane of the exit slit in order to scatter the low-divergence laser beam into many directions, simulating the passage of the luminescence emission beam through the monochromator. The laser light, after leaving the entrance slit (the monochromator has to be set to the laser wavelength, obviously), maintains its direction going through the collection optical system and is finally focused onto the sample, where a visible red spot appears. This is the place onto which the excitation light beam is to be directed. Such an arrangement ensures correct focusing of the luminescence emission onto the entrance slit along with the optimized path through the monochromator and the right illumination of the detector. Do not forget to remove the ground glass from the exit slit before remounting the detector (!). A final signal maximization prior to recording the spectrum is done via fine adjustment of the positions of the lenses using $\{x, y, z\}$ translation stages.

Achieving maximum brightness at the entrance slit represents the experimental priority also when using a spectrograph in conjunction with a multichannel detector, because B_λ naturally figures in eqn (2.40), standing for the detector irradiance. Other experimental factors, the spectrum acquisition time and the related S/N ratio, were already described in that part of Section 2.2 dealing with multichannel detectors.

Finally, it is worth noting that, nowadays, optical waveguides (fibres) are increasingly applied to guide light signals towards spectrometers. Among numerous benefits, for example, is their capacity to bring the signal from distant and/or badly accessible places. In order to minimize signal losses, one has to choose an appropriate waveguide material with maximum transmittance over the spectral region of interest (common waveguides are made of various glasses, silica, polymers, and liquids) and then correctly couple light in and out of the waveguide.

A glass waveguide can be formed by a single fibre or a bundle of fibres. The single-fibre waveguide is mostly of cylindrical shape, the central part of which is called the core (refractive index n_1) and is surrounded by a cladding layer with lower refractive index $n_2 < n_1$ (the change in refractive index may be either step-like or gradual, hence the so-called step-index or graded-index fibres). Coupled light is guided within the core thanks to total reflection on the interface between the core and cladding, provided the entrance angle is equal to α_a or smaller (see Fig. 2.36(a)). The maximum entrance angle α_a is given by

$$\sin \alpha_a = \frac{1}{n_0}(n_1^2 - n_2^2)^{1/2}, \tag{2.47}$$

where n_0 is the relative refractive index of the ambient medium (for air we can put $n_0 = 1$). The same value of maximum angle holds for light leaving the waveguide. Typical values for glass and silica fibres are $\alpha_a = 34°$ and 12°, respectively. The angle α_a is called *the acceptance* or *aperture angle* and it determines the so-called *acceptance* (and output) *cone of the fibre*. The sine of the aperture angle is the *numerical aperture* (NA), and sometimes also the term

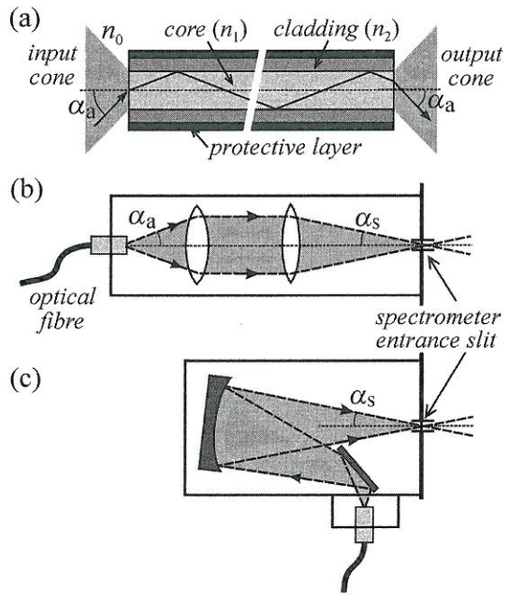


Fig. 2.36
(a) Schematic of light rays guided in an optical waveguide, (b) and (c) examples of two optical systems designed to couple correctly light from a fibre to a spectrometer using lenses (a) or mirrors (b). Based on documents by Horiba Jobin Yvon, Ltd.

aperture ratio $1/\# \approx 2\text{NA}$ is used (in analogy with photographic cameras), see Subsection 2.3.2.

The numerical apertures of a waveguide and that of a spectrometer only exceptionally happen to match one another. Therefore, a special optical system is required to adapt the output cone of a fibre to the acceptance cone of a spectrometer described by an angle α_s (Fig. 2.36(b),(c)). Such coupling adapters for various optical waveguides and spectrometers are usually designed and supplied by spectrometer manufacturers.

2.6 Fourier luminescence spectroscopy

We have shown that increasing the resolving power of a scanning monochromator leads either to a decrease in the S/N ratio or to impractically long spectral acquisition times. One possible way to solve this problem—the multichannel detectors—was also mentioned; these devices are indeed increasingly used in modern luminescence spectroscopy. However, if a need for relatively very high resolution ($\Delta\lambda \leq 0.01$ nm) arises, we shall be faced with a fundamental limitation: the ultimate resolution of the multichannel detectors is strictly limited by the finite size of their individual pixels. For example, a typical detector with pixels of 26 μm × 26 μm attached to a spectrometer with a dispersion of $L^{-1} \approx 3$ nm/mm has the resolution limit $\Delta\lambda \approx 3\,\text{nm/mm} \times 26 \times 10^{-3}\,\text{mm} \approx 0.08$ nm. Even if the need for such a very high spectral resolution is not typical for luminescence spectroscopy, it may sometimes appear, as we have already indicated several times. In this case one has no choice but to resort to Fourier luminescence spectroscopy.

A Fourier spectrometer, widely applied in infrared absorption spectroscopy, is basically a scanning Michelson interferometer. The radiation to be analysed (luminescence in our case), upon entering the spectrometer, is split into two arms and, after being reunited again, the beams interfere due to the varying length of one of the arms. The resulting temporal interference signal undergoes Fourier analysis that transforms it into a frequency (or wavelength) spectrum. There are two fundamental differences compared to standard monochromators: (i) all spectral components are analysed simultaneously (the so-called Fellgett (or multiplex) advantage which, however, takes place also in spectrographs equipped with multichannel detectors—see eqn (2.12)), and (ii) the interferometer has no slits and, consequently, its throughput/aperture is much higher than that of a traditional monochromator that shows the same resolving power (the so-called Jacquinot advantage).

It is possible to derive the following relation comparing the signal-to-noise ratio of a Fourier spectrometer $(S/N)_{\mathrm{F}}$ to that of a scanning monochromator $(S/N)_{\mathrm{scan}}$ utilizing the same detector [20]

$$\frac{(S/N)_{\mathrm{F}}}{(S/N)_{\mathrm{scan}}} \approx \frac{2\pi f}{h}\sqrt{M}, \tag{2.48}$$

where M is the number of narrow spectral components (lines) that are analysed by a Fourier spectrometer during the same time interval that a scanning monochromator needs to analyse a single component. Other variables have their usual meaning: f is the focal length of the monochromator collimator and h is the illuminated slit height (the first factor on the right side of (2.48) thus reflects the Jacquinot advantage while the second one results from the Fellgett advantage).

The right-hand side of eqn (2.48) has a typical value of 10^3. If we consider the highest real resolution of a typical scanning monochromator to be $\Delta\lambda \approx 0.1$ nm, then, keeping the same S/N ratio, a Fourier spectrometer can in principle attain a resolution of about three orders of magnitude better, i.e. $\Delta\lambda \approx 10^{-4}$ nm. However, since, as a rule, semiconductor photodiodes rather than photomultipliers are used in a Fourier luminescence spectrometer, it appears more realistic to speculate that the right-hand side of eqn (2.48) is one order of magnitude lower. Then, the practically achievable resolution of a Fourier spectrometer is about $\Delta\lambda \approx 10^{-3}$ nm; expressed in wavenumbers ν^*, which are frequently used here, this corresponds to $\Delta\nu^* \approx 10^{-2}\,\mathrm{cm}^{-1}$. We shall mention a benefit of this unique property in Subsection 7.2.2. Nevertheless, in general Fourier spectrometers are still rather expensive and their exploitation is as yet worthwhile only in very special applications (e.g. high-resolution near-infrared spectroscopy of weak luminescence signals).

2.7 Spectral corrections

Suppose we have optimized our optical collection system and our spectral apparatus consists of a high-throughput monochromator and a high-quality cooled detector. Even then, however, the experimental emission spectra of the sample under study, in spite of being minimally noisy, do not correspond—

as far as their shape is concerned—to the actual luminescence spectrum. The reason for this is due to the fact that the photodetector sensitivity $k(\lambda)$ as well as the characteristics of other optical elements—the reflectivity or transmittance of the dispersion element, mirrors and lenses—are spectrally dependent. Thus, if we denote the luminescence intensity as function of wavelength (which means the actual emission spectrum) as $I(\lambda)$, then the measured spectrum $i(\lambda)$ is equal to

$$i(\lambda) = Q(\lambda)\, I(\lambda), \tag{2.49}$$

where $Q(\lambda)$ is usually called the spectral response of the whole detection system.[7] Thus $Q(\lambda)$ includes both the transmittance of the spectral device $\kappa(\lambda)$ and the detector sensitivity $k(\lambda)$. Sometimes $Q(\lambda)$ is simply referred to as a correction function (of the detection channel). The function $Q(\lambda)$ must be determined experimentally, separately for each particular configuration of the detection system (various combinations of a dispersion element and a detector unit). A faithful spectral shape $I(\lambda)$ is then obtained, according to eqn (2.49), by dividing the measured spectrum $i(\lambda)$ by the correction function $Q(\lambda)$.

The correction function $Q(\lambda)$ is determined with the aid of a light source with well-known spectral emittance. A commonly applied calibrated source for the visible and near-infrared regions is a tungsten strip (filament) lamp, whose emission spectrum bears a close resemblance to blackbody radiation. Such calibrated lamps are supplied, together with a detailed description of their spectral energy density, either in graphical or tabular form (or simply colour temperature), valid under a specific heating current. Light is emitted by a narrow tungsten strip in a vacuum bulb; this design ensures the distribution of temperature over the entire surface of the light-emitting body to be as homogeneous as possible. (A traditional tungsten spiral in a bulb or other incandescent lamp types do not fully satisfy this condition but even special sources of this type—filament lamps—can be calibrated and applied for correction purposes.) If an experimentally acquired spectrum of the calibrated source is $w(\lambda)$ and its actual spectral emittance is $W(\lambda, T)$, the required correction function is calculated according to eqn (2.49) simply as

$$Q(\lambda) = \frac{w(\lambda)}{W(\lambda,\, T)}. \tag{2.50}$$

Then, to obtain undistorted results, any measured luminescence spectrum has to be divided by this correction curve.

The procedure for determining $Q(\lambda)$ as described above seems to be quite simple but the opposite is true. First of all, it is not easy to keep the prescribed lamp current stabilized over the whole measurement time, while even small

[7] A somewhat more exact formulation of the above debate is that the luminescence intensity means (usually) the photon flux (power) per unit interval of wavelengths, $[I(\lambda)] =$ photon/s/unit interval of λ. From the point of view of radiometry this corresponds to radiant flux $\phi_{\mathrm{L}}(\lambda)$ per unit interval of λ, $[\phi_{\mathrm{L}}(\lambda)] =$ power/unit interval of λ. This flux impinges on the entrance slit creating its irradiance $E_\lambda = \phi_{\mathrm{L}}(\lambda)/S$. Then the slit becomes a planar light source for the spectrometer, featuring brightness $B_\lambda = E_\lambda$. According to eqn (2.37), the radiant flux at the exit slit is $\phi_\lambda \sim \kappa(\lambda) B_\lambda$ and the detector signal reads $i(\lambda) \sim k(\lambda)\phi_\lambda$, which means $i(\lambda) \sim k(\lambda)\kappa(\lambda) B_\lambda \sim k(\lambda)\kappa(\lambda)\, I(\lambda)$.

deviations from the specific temperature (usually in the range of 2800–3000 K) cause significant changes of the $W(\lambda, T)$ spectrum. Too high lamp intensity may appear as another problem, because of possible detector oversaturation. Next, ageing of such lamps may be quite fast and can cause harmful alteration of the original spectrum $W(\lambda, T)$. Often, to determine $Q(\lambda)$ correctly, a diffusion reflector (made of, e.g., MgO or $BaSO_4$) is recommended, but variations in its spectral reflectivity can affect the result. In addition, problems could also arise from polarization effects—light becomes partially polarized or depolarized by reflective surfaces, by diffraction gratings, etc. In particular the last mentioned effects are often ignored even if their impact could be significant. The correction curve thus has to be determined with due care. Some difficulties in determining $Q(\lambda)$ can be eliminated by using an integration sphere fixed at the entrance of the spectrometer. This ensures correct filling of the collimator by light and also depolarizes possible partial polarization of the incoming lamp radiation.

Nowadays, performing the correction of luminescence spectra for spectral response of the detection path should be commonplace in every laboratory. Any spectrum submitted for publication in international refereed journals may be automatically supposed to have been corrected. Indeed, the difference between a corrected and an uncorrected spectrum can be considerable, as illustrated in Fig. 2.37 [21]. Naturally, the correction becomes imperative when the spectrum is going to be fitted with a theoretical curve. Actually one of the aims of this book is to teach the reader how to identify the microscopic origin of the luminescence centre, based on the corresponding spectral shape and its changes due to controlled variations of experimental parameters: temperature, excitation intensity, etc.

Nevertheless, one can admit that sometimes the correction in question is not indispensable—for example, if only a qualitative demonstration of spectral changes (e.g. one spectral band grows at the expense of another during an increase in temperature) is sufficient for a given purpose. Also in the case of very narrow emission lines the spectra can often be left uncorrected and even fitting its shape with a theoretical model is possible. The reason is that the correction curve is usually only slowly varying with wavelength and its change

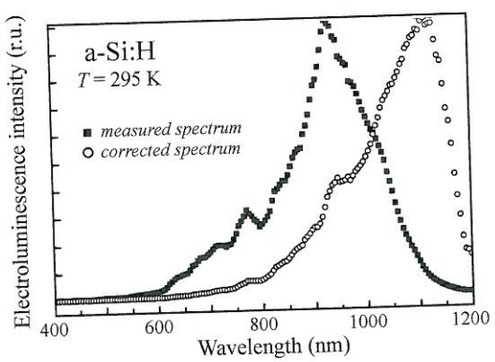


Fig. 2.37
As measured (full squares) and corrected (empty circles) electroluminescence spectra of amorphous hydrogenated silicon. The pronounced deformation of the measured spectrum at long wavelengths is mainly due to the rapid decrease in photomultiplier sensitivity in this spectral region (S1 photocathode), while the dip at ~ 790 nm is due to the so-called Wood's grating anomaly. The spectra are normalized. Adapted from Fojtik [21].

in a narrow wavelength range (a few nanometres) can be neglected. But always be vigilant!

We should stress that this correction does not serve for establishing the absolute number of emitted photons or determining the absolute luminescence quantum yield η—this is a domain of very special and quite difficult techniques, which are described for example in [22] and in Appendix K. Any spectrum corrected using (2.50) 'only' faithfully represents (in relative units) the number of photons emitted by the luminescent sample per unit interval of wavelengths or photon energies. It remains to be noted that the described correction procedure is basically the same irrespective of whether a monochromator or spectrograph is used.

In the context of corrections of optical emission spectra we must mention another experimental problem, which used to be unjustly neglected. It concerns two alternative modes of plotting the emission spectra—against either wavelengths λ or photon energy $h\nu$. Both representations are possible and widely used (sometimes even both modes are applied to the same graph, one at the bottom and the other at the upper axis or vice versa, see e.g. Fig. 1.2), conversion between wavelengths and photon energies being performed simply via the trivial relation

$$h\nu = h\frac{c}{\lambda}. \tag{2.51}$$

However, such an approach is not always fully justified. One has to take into account what type of dispersion element and wavelength scan was used during the experiment.

Let us consider firstly a grating monochromator with a linear scan in wavelengths λ (the sine-drive, see Section 2.3). The spectral measurement is carried out using constant slit widths $\Delta\lambda$, the result being either a raw spectrum $i(\lambda)$ or, after having performed the correction, the actual spectrum $I(\lambda)$. However, the *elemental step (interval) in photon energy* is not kept constant during the measurement, as is clearly seen by differentiating (2.51),

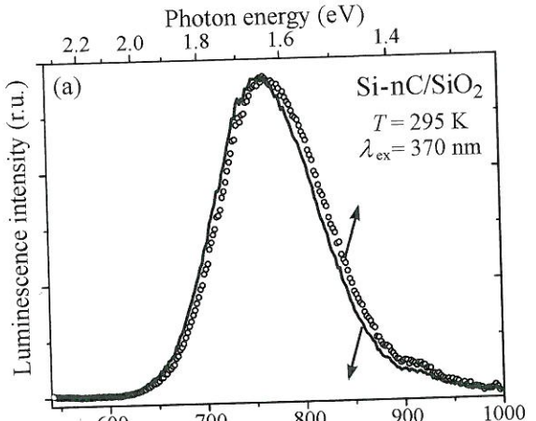
$$|\mathrm{d}(h\nu)| = \frac{hc}{\lambda^2}|\mathrm{d}\lambda|. \tag{2.52}$$

Obviously, when increasing the wavelength (keeping $\mathrm{d}\lambda = \text{const}$), the energy interval $|\mathrm{d}(h\nu)|$ drops, because it scales like λ^{-2}. Therefore, if one wishes to plot the spectrum against the photon energy scale like $I(h\nu)$, the decrease in the magnitude of the energy interval $|\mathrm{d}(h\nu)|$ should be corrected[8] by multiplying each value of $I(\lambda)$ by λ^2:

$$I(h\nu) = \lambda^2 I(\lambda). \tag{2.53}$$

This multiplication obviously boosts the long-wavelength wing of the spectrum, causing a red-shift of spectral peaks; see Fig. 2.38(a) [23]. Thus, in principle it is incorrect to convert simply λ into $h\nu$ while keeping the spectral curve unchanged. However, at the same time one can see that the correction

[8] Let us recall that $I(\lambda)$ means the photon flux per unit interval of wavelength, and by analogy $I(h\nu)$ is the photon flux per unit (therefore constant) interval of photon energies.



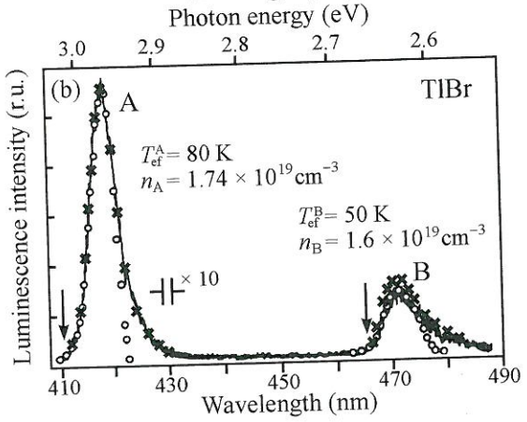


Fig. 2.38
Correction of emission spectra when transforming the wavelength-to-photon energy scale. (a) Photoluminescence spectrum of silicon nanocrystals embedded in an SiO_2 matrix. The spectrum was recorded using a grating monochromator with a linear scan of λ (solid line, the spectrum is related to the lower axis). Correction pertinent to transformation into photon energies $h\nu$ was performed by multiplying by λ^2 (open circles, the spectrum is related to the upper axis). Both curves are normalized. Adapted from Dian *et al.* [23]. (b) Emission spectra of TlBr crystals due to radiative decay of the electron–hole plasma. The solid line is the experimental spectrum measured by making use of a linear scan in λ, and the crosses show the recalculated spectrum connected with axis transformation to photon energies (circles are related to a theoretical fit, which is irrelevant to the present discussion). Both the A and B emission lines being relatively narrow, the correction as regards their spectral shape and position is of minor importance—just the relative intensities of these bands are slightly changed (the intensity of the B line increases relative to the A line). Adapted from Pelant *et al.* [24].

effect is not large and becomes significant for sufficiently broad spectra only (full width at half maximum about 100 nm or more). This is illustrated by Fig. 2.38(b) [24]: for emission lines only a few tens of nanometres in width or narrower, the distinction in shape between the recalculated curve and the spectrum as recorded becomes negligible, usually buried in noise.

In the case of a grating monochromator with a linear sweep in wavenumbers $1/\lambda$ or energy $h\nu$ (a cosec-drive, which is a much less frequent case) we obtain an experimental spectrum against photon energy like $I(h\nu)$. For transforming it into wavelengths we again use relation (2.53), but in the opposite direction: the spectrum $I(h\nu)$ will be divided by λ^2.

Monochromators equipped with a prism should in principle require two corrections. Nevertheless, the following discussion is intended to show that the situation there is, in fact, not so serious. First of all, the reciprocal linear dispersion L^{-1} is a nonlinear function of λ which, according to (2.19), causes an increase in the spectral slit width $\Delta\lambda$ (when keeping the mechanical slit width $\Delta\ell$ constant) with increasing λ.

Consequently, the shape of the measured spectrum $I(\lambda)$ does not exactly correspond to reality. In principle, for each prism device, knowing its particular $L^{-1}(\lambda)$ dependence, we should have to determine a specific correction function[9] $\xi(\lambda) = \Delta\lambda_0/\Delta\lambda < 1$ to multiply the experimental spectrum $I(\lambda)$; since $\xi(\lambda)$ is a decreasing function, the long-wavelength wing of $I(\lambda)$ would have been slightly suppressed. However, a second correction takes place during the transition from $I(\lambda)$ to $I(h\nu)$, namely, multiplication by λ^2 according to (2.53), which relatively enhances the long-wavelength side. All in all, both corrections go in opposite directions and may roughly compensate each other. This means that we can practically omit both of them within reasonable accuracy.

To conclude, this somewhat arbitrary way of interchanging between the λ and $h\nu$ axes, found frequently in the literature, is not usually at the expense of the correctness of presentation, except, perhaps, for very broad spectra. For completeness, let us also mention spectrographs, where wavelength scanning is replaced by multichannel detection. Most of these devices employ a diffraction grating whose dispersion is linear in wavelength, i.e. $\Delta\lambda = \text{const}$. This situation is analogous to that of a sine-drive monochromator—transformation from $I(\lambda)$ to $I(h\nu)$ is done using eqn (2.53), if necessary.

2.8 Influence of slit opening on the shape of emission spectra

The last important correction of emission spectra we are going to treat here concerns the effect of the finite width of the instrument slits on the spectral shape. Intuitively, we feel that setting too wide slits can bring about loss of spectral details and blurring or broadening of the spectra. In this section, the influence of the slit will be described quantitatively. Firstly, the case of a monochromator (a dispersion device with two slits) and then that of a spectrograph (having one entrance slit and a multichannel detector at the output) will be analysed.

Monochromator

In Appendix A we give an illustrative explanation of the convolution of two functions in terms of a shift of the first function over the second one along the independent variable axis, integrating simultaneously the product of the two functions (Fig. A.1). The functioning of a scanning monochromator equipped with a rotating dispersion element fits the above scheme well: the emission spectrum, created with the cooperation of the entrance slit with the dispersion element (and being affected by the spectral variations in reflectivity (transmittance) of all elements contained within the device) is projected onto the exit slit plane. Obviously, this is not the true spectrum $I(h\nu) = I(E)$ but it can be described as $\kappa(E)I(E)$, where $\kappa(E)$ comprises all the above mentioned spectral factors (see also (2.32)). If, moreover, we include now also

[9] Here $\Delta\lambda_0$ means the spectral slit width at the short-wavelength boundary λ_0 of the spectral range covered by the device, and $\Delta\lambda = L^{-1}(\lambda)\Delta\ell$ is taken for $\lambda > \lambda_0$.

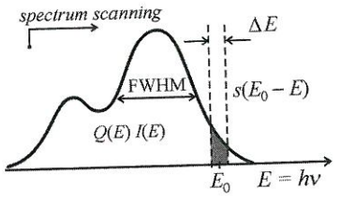


Fig. 2.39
The emission spectrum in a scanning monochromator is scanned across the exit slit (its spectral width being ΔE). The joint activity of the entrance and exit slits is described by a function $s(E_0{-}E)$. This means that the function $Q(E)I(E)$ is convoluted with $s(E_0{-}E)$; FWHM means full width at half maximum of the spectral band.

the detector spectral sensitivity $k(E)$, then the spectrum projected onto the exit slit corresponds to $\kappa(E)k(E)I(E) \equiv Q(E)I(E)$.

This spectrum is now 'swept' or scanned across the exit slit, consistently with rotation of the grating or prism (Fig. 2.39). The exit slit of spectral width ΔE transmits (and so 'integrates') slices of the spectrum towards the detector. If we denote, according to eqn (A.1) and Fig. A.1, the product $Q(E)I(E)$ as f_2, and the function describing the effect of slits as $f_1 = s(E)$, we see that the measured spectrum is given by the convolution of the functions QI and s:

$$i(E_0) = \int_{-\infty}^{+\infty} s(E)Q(E_0 - E)I(E_0 - E)\mathrm{d}E = \int_{-\infty}^{+\infty} Q(E)I(E)s(E_0 - E)\mathrm{d}E,$$

where we have applied commutativity of the operation of convolution according to eqn (A.3). Thus the measured spectrum $i(E_0)$ is, taking into account the effect of the finite slit opening, given by the convolution

$$i(E_0) = \int_{-\infty}^{+\infty} Q(E)I(E)s(E_0 - E)\mathrm{d}E. \tag{2.54}$$

The function $s(E_0{-}E)$ is usually called *the apparatus (instrumental) function* of the spectral device. It must comply with the normalization condition

$$\int_{-\infty}^{+\infty} s(E_0 - E)\mathrm{d}E = 1. \tag{2.55}$$

We are now in a position to decide under what conditions the convolution (2.54) is to be taken unavoidably into consideration when treating the experimental data; we shall discuss three typical situations.

(a) The spectral slit width ΔE is much narrower than the full width at half maximum (FWHM) of the measured spectrum, $\Delta E \ll$ FWHM. Then the function $s(E_0{-}E)$ can be approximated by the delta function $\delta(E_0{-}E)$ and (2.54) is reduced straightaway to

$$i(E_0) = Q(E_0)I(E_0). \tag{2.56}$$

Or, for a broad spectrum and narrow slits the effect of apparatus function need not be taken into account. The true emission spectrum is obtained simply by dividing the experimental result by the correction function $Q(E)$.

(b) The spectral slit width ΔE remains smaller than the FWHM of the spectrum but the apparatus function cannot be considered infinitely narrow. Then the measured spectrum and the actual spectrum are not identical, and in order to extract the actual spectrum $I(E)$ from the convolution integral (2.54), one needs to know the function $s(E_0{-}E)$.

Let us restrict our discussion to the standard case of equal entrance and exit slit widths ΔE, large enough to neglect diffraction effects. Then the geometrical image of the entrance slit in the output focal plane is represented by a homogeneously illuminated area (rectangle). Then an *ideal monochromator* (i.e. a device free of imperfections and aberrations of optical imaging) features the apparatus function of a triangle shape as described by

$$s(E_0 - E) = \begin{cases} \frac{1}{\Delta E}\left[1 - \frac{|E_0 - E|}{\Delta E}\right] & \text{for} |E_0 - E| \le \Delta E \\ 0 & \text{for} |E_0 - E| > \Delta E, \end{cases} \tag{2.57}$$

which is illustrated in Fig. 2.40(b) and whose FWHM is equal to ΔE. Figure 2.40(a) helps us to understand why this is so. When the dispersion element is turning, the monochromatic image of the entrance slit is moving across the exit slit, and the detector records a signal proportional to the overlap area (just the convolution of the entrance slit image with the exit slit). This signal increases linearly from zero to its maximum value (both areas are fully overlapping) and then drops back. As a result, the triangular shape of the function $s(E_0{-}E)$ appears.

In a *real monochromator*, however, various imperfections, inhomogeneities of all optical elements, parasitic diffraction effects, imperfect absorption of the internal blackened surfaces of the monochromator, etc. produce a weak light background that can enter the exit slit at wavelengths different from the nominally set value. This effect leads to making the triangle vertices round and the real apparatus function can be approximated by a Gaussian curve, as shown in panel (c) of Fig. 2.40. Its analytical expression takes the form

$$s(E_0 - E) = N(\Delta E') \ \exp\left[-\frac{4\ \ln 2\ (E_0 - E)^2}{(\Delta E')^2}\right], \tag{2.58}$$

where $\Delta E\prime$ is the FWHM of the curve and $N(\Delta E\prime) = 2\sqrt{\ln 2}/(\sqrt{\pi}\Delta E')$ is a normalizing factor.

The relatively simple form of the apparatus function (2.58) is advantageous for performing a deconvolution of eqn (2.54) in order to extract the real spectrum $I(E)$ from the experimental one $i(E)$. (We should not forget, however, that also the correction function $Q(E)$ must be known.) There are a large number of deconvolution methods [25] and current progress in computational techniques makes them easily available. Nevertheless, it seems that an inverse approach, the so-called iterative convolution, remains among the most widely applied methods. The principle is based on setting down a suitable model function for the actual spectrum $I(E)$ with some initial parameters and calculating the convolution (2.54). The parameters of $I(E)$ are then iterated until the best correspondence between the calculated spectrum $i(E)$ and the experimental one is achieved. Evidently, the initial selection of an appropriate microscopic model of luminescence centre (and related spectral shape) is crucial and depends strongly on both the experience and intuition of the experimenter.

Application of the apparatus function in the form of (2.58) includes one difficulty which is not evident at first glance, namely, the unknown value of $\Delta E'$. Unlike the triangular function (2.57), the halfwidth of the Gaussian peak $\Delta E'$ (2.58) is not equal to the spectral width of the entrance or exit slit ΔE. This is because of the normalization (2.55) and is illustrated in Fig. 2.40(c): the postulate of a constant area below the curve yields $\Delta E' > \Delta E$. Therefore, knowledge of the mechanical (and, consequently, also spectral) widths of slits does not enable us to determine the particular shape of the Gaussian instrumental function (2.58). How, then, can the apparatus function shape be obtained?

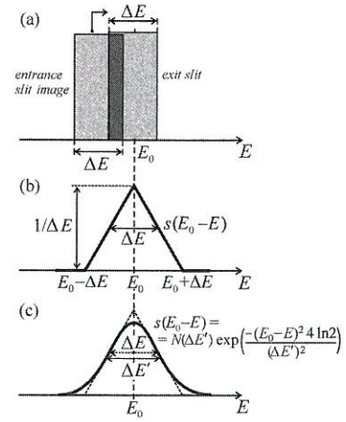


Fig. 2.40
Illustration of the apparatus function of a monochromator. (a) The monochromatic image of the entrance slit is moving to the right across the exit slit in the focal plane of an ideal device. The detector 'sees' an intensity proportional to the overlap area (dark grey rectangle) whose variation as a function of E is the apparatus function $s(E_0{-}E)$ plotted in panel (b). This function has a triangular shape with unit area, in order to fulfil the condition (2.55). Panel (c) shows the shape of this function in a real monochromator, as influenced by imperfections in the instrument and optical inhomogeneities. $N(\Delta E\prime)$ stands for a normalizing factor.

The answer to this question is related to our last situation to be discussed:

(c) The spectral slit width ΔE is much wider than the width of the measured spectrum, $\Delta E \gg$ FWHM. Then we put $I(E) \approx \delta(E - E_0')$, $Q(E) \approx$ const, and eqn (2.49) is reduced to

$$i(E_0) \approx \int_{-\infty}^{+\infty} \delta(E - E_0')\, s(E_0 - E)\, \mathrm{d}E \approx s(E_0 - E_0'). \tag{2.59}$$

Thus, paradoxically, the result of this experiment is the apparatus function rather than the spectrum (spectral peak) itself. Such a situation can be easily realized for example by measuring narrow emission lines from a low-pressure spectral lamp. If such a measurement is performed with the same slit widths as applied during the luminescence experiment, we can thereby use the obtained shape of the Gaussian apparatus function (2.58) for deconvolution (or iterative convolution) to find the actual spectral shape $I(E)$.

Let us add a final technical note: All the functions in the integral (2.54) are non-zero within relatively narrow energy intervals only; then the limits of numerical integration may be strongly restricted. In fact they are given by the narrowest function. Therefore, in the case of (b), applying the Gaussian apparatus function (2.58), reasonable integration limits are $\langle E_0 - 5\Delta E', E_0 + 5\Delta E' \rangle$.

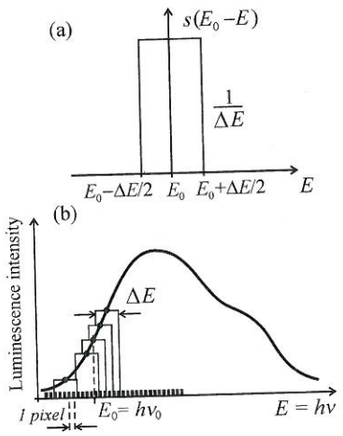


Fig. 2.41
(a) Apparatus function of a spectrograph. (b) Detection of a broad luminescence spectrum by a multichannel detector and a spectrograph with a wide entrance slit (ΔE means the spectral width of the entrance slit image).

Spectrograph

Here, assessment of the influence of the slit opening on the spectral shape is somewhat easier for two reasons. Firstly, there is only a single (entrance) slit. Secondly, the option of acquiring any spectrum for a very long time enables us to set the slit width narrow enough so as not to distort the spectral shape. Nevertheless, we are going to analyse several potential situations.

First of all, we should ask about the shape of the apparatus function $s(E)$. The answer is obvious: in any spectrograph, a uniformly illuminated monochromatic image of its entrance slit is created in the exit focal plane. Therefore, the relevant apparatus function is of a rectangular shape described by

$$s(E_0 - E) = \begin{cases} \frac{1}{\Delta E} & \text{for } |E_0 - E| \le \frac{\Delta E}{2} \\ 0 & \text{for } |E_0 - E| > \frac{\Delta E}{2}, \end{cases} \tag{2.60}$$

where ΔE is again the spectral slit width (Fig. 2.41(a)). Then the convolution (2.54) takes on the form

$$i(E_0) \approx \int_{E_0-\Delta E/2}^{E_0+\Delta E/2} Q(E) I(E) \mathrm{d}E \approx Q(E_0) \int_{E_0-\Delta E/2}^{E_0+\Delta E/2} I(E) \mathrm{d}E, \tag{2.61}$$

if we exploit the fact that the correction function $Q(E)$ is only slowly varying across the slit width. Application of relation (2.61) to deconvoluting a measured spectrum is then relatively simple.

However, a closer inspection of the preceding procedure can raise an objection: we have introduced the convolution (Appendix A) as a relative shift of two functions, which is not completely true here—the grating is not rotating, and the spectrum is not moving. Is then the proposed application of eqn (2.61)

correct? The response is yes. Although the 'monochromatic' images of the entrance slit hit the detector simultaneously, not in sequence like in a monochromator, the final effect is the same (see Fig. 2.41(b)). The mechanism of the generation of the spectral image is, however, different, as we shall indicate.

Figure 2.41(b) shows the situation when the experimenter has chosen the entrance slit width of a spectrograph[10] several times wider than the size of a single pixel. This is a standard situation; in Section 2.3 we have shown that setting the slit width narrower than about 10 μm serves no useful purpose. Therefore, a slit width of $\Delta\ell = 100$–200 μm is a reasonable choice (also in relation to an acceptable acquisition time of the spectrum), $\Delta\ell$ being really much larger than the typical CCD pixel size 25 μm. Let us select a fixed photon energy E_0. If the slit image width exceeds that of a single pixel, the signal level on the pixel at $E_0 = h\nu_0$ may 'get saturated', but this does not prevent the signal at E_0 from attaining the correct level, because 'monochromatic' images of neighbouring spectral elements join the signal. Naturally, this is accompanied by spectral broadening; everything is comprised in eqn (2.61).

We shall close Sections 2.7 and 2.8 with a brief evaluation of the effects and corrections just discussed in view of their practical importance. The reader, possibly confused and upset, might have come to the conclusion that any experimental luminescence spectrum is only a relatively useless raw product, which must undergo many mathematical operations prior to being presented and interpreted. This is not the case, though. Among the above mentioned corrections only one is really essential and almost always indispensable, namely, to divide the experimental spectrum by the spectral correction function $Q(\lambda)$ given in eqn (2.50), in order to correct the effect of spectral sensitivity of the whole detection system. The 'presentation problem' of the spectral data, i.e. care about whether to present spectra against wavelength λ or photon energy $E = h\nu$, takes place only exceptionally, in the case of very broad spectra (and if one wants to be very careful). Similarly, the correction through making use of the apparatus function $s(E_0{-}E)$ in order to remove the effect of the finite slit width is worth doing only in the case of the measurement of very narrow emission lines, like for instance those typical of radiative recombination of free excitons and biexcitons (Chapters 7 and 8). To a certain extent the corrections $Q(\lambda)$ and $s(E_0{-}E)$ may even be considered disjunctive, because the spectral correction function $Q(\lambda)$ is important for wide spectra, when the slit-width effect is negligible, while the function $s(E_0{-}E)$ is important for narrow lines when, to a good approximation, $Q(\lambda_0) \approx$ const.

2.9 Time-resolved luminescence measurements

Knowledge of the temporal behaviour of the luminescence signal, especially the shape of the luminescence decay curve after the excitation is switched off, carries valuable information contributing to the identification of the recombination process and the microscopic nature of the relevant luminescence centre (along with an analysis of the excitation and emission spectral shapes

[10] We again and again assume equal widths of the entrance slit and of its image.

and of the dependence of the pump intensity). However, the techniques of time-resolved luminescence measurements, in particular under pumping with ultrashort laser pulses, represent a largely specialized research field that cannot be fully treated within the limited space of this section. Therefore, we shall describe just the basic principles of the most common experimental methods and instruments. A comprehensive treatment of this subject can be found for example in [26] and [27].

2.9.1 Direct imaging of the luminescence response

The simplest set-up to study luminescence decay consists of a single-channel detector, usually a photomultiplier, of which the output is connected to the vertical deflection plates of an oscilloscope and swept in time by its horizontal time base. The result is either observed on a screen or transferred to a computer. Obviously, in order to record, in this way, the decay of the luminescence radiation in time, either a cw excitation beam must be chopped (then curves similar to those plotted in Fig. 2.27 are obtained) or a pulsed excitation source supplying sufficiently short pulses (a lamp or laser) must be used. Since the character of luminescence decay may be strongly dependent on the emission wavelength, the decay curves $i_\lambda(t)$ of a quasi-monochromatic luminescence signal are usually measured after inserting a monochromator or at least narrow-band interference filters in front of the detector.

The next step consists in analysing the measured curves $i_\lambda(t)$. In the simplest case, one deals with a single-exponential function $i_\lambda(t) = i_\lambda(0)\exp(-t/\tau)$ where τ is referred to as the luminescence decay time or lifetime. However, many other shapes of decay, e.g. double-exponential, hyperbolic, etc. may occur in luminescence spectroscopy. Their detailed study will be presented in Chapter 3.

Generally, the mean decay time $\bar{t}$ can be defined from the point of view of statistics as the mean value of a continuous stochastic varible by

$$\bar{t} = \frac{\int_0^\infty t\, i_\lambda(t)\, \mathrm{d}t}{\int_0^\infty i_\lambda(t)\mathrm{d}t},$$

where $i_\lambda(t)$ plays the role of the occurrence probability density of a time interval t in which an elementary emitter is found in an excited state. In the case of the single-exponential decay $i_\lambda(t) = i_\lambda(0)\exp(-t/\tau)$, simple evaluation of the integrals leads to $\bar{t} = \tau$.

For practical purposes and in analogy with the single-exponential decay, as the 'decay time' in the case of more complex decay curves $i_\lambda(t)$, the time at which the initial intensity $i_\lambda(0)$ decreases to the value $i_\lambda(0)/\mathrm{e} \approx i_\lambda(0)/2.72$, is often (somewhat arbitrarily) taken.

Although the method of direct decay imaging looks very simple, its practical implementation challenges the quality of the instrumentation and the experimenter's knowledge of the limits of applicability of the method. First of all, the whole experimental set-up (a pulsed excitation source, a photomultiplier

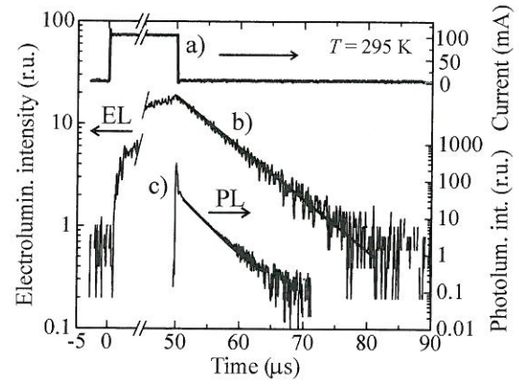


Fig. 2.42
An example of direct imaging of the luminescence response by a digital oscilloscope (the signal comes from red electroluminescence (EL) and photoluminescence (PL) of silicon nanocrystals). The upper trace (a) shows a 50 μs-long electric injection pulse, which excites electroluminescence, and curve (b) is part of the onset of EL and the subsequent decay. Curve (c) is the photoluminescence response of the same sample to excitation with 3 ns laser pulses (the third harmonics from a Nd:YAG laser). For both EL and PL, the excitation switch-off is fast enough which enables direct determination of decay times without any mathematical treatment (deconvolution). The decay curves (b), (c) are fitted by single-exponential functions with decay times of $\tau_{EL} \approx 8\,\mu s$ and $\tau_{PL} \approx 2.5\,\mu s$. After Luterová *et al.* [28].

or photodiode, and an oscilloscope) should be fast enough, highly sensitive and perfectly linearly responding to the optical signal, which requirements are difficult to be satisfied simultaneously. (For example, a highly sensitive photomultiplier must contain many dynodes, which in turn makes its time resolution worse due to statistical fluctuations of secondary emitted electrons and the long trajectory of the electron packet.) One can easily imagine a number of potential limitations of the method: (i) the insufficient response speed of the photomultiplier; if (ii) a fast p-i-n photodiode is applied, it is inherently less sensitive and a preamplifier must be used in connection with the oscilloscope, whose limited bandwidth may pose a problem; (iii) similarly, the frequency bandwidth of the oscilloscope itself may be a limiting factor, etc. The direct measurement of the luminescence response may be easily applied (without any complementary mathematical treatment) to study relatively slow decays only (under a pulsed excitation, for roughly $\tau \geq 0.1$–$1\,\mu s$). An example is shown in Fig. 2.42 [28]. If photoluminescence is excited by a chopped beam, this limit is pushed to even longer times ($\tau \geq 50\,\mu s$), considering that the typical maximum frequency of mechanical chopping is of the order of 1 kHz and the excitation pulses are of trapezoidal rather than of perfectly rectangular shape (the slope of the pulse edges is given by the speed of chopper rotation as well as by the diameter of the chopped beam).

Under the complementary (*a posteriori*) mathematical treatment, *deconvolution* of the measured decay curve is to be understood. In analogy to the influence of the monochromator slit opening on the shape of the measured emission spectra (Section 2.8), we can deduce that the experimentally determined luminescence time response $R_\lambda(t)$ is given by a convolution of the actual time response $i_\lambda(t)$ (i.e. the luminescence response to an infinitely short excitation pulse) with the temporal profile of the excitation pulse $L(t)$ measured by the same detection system as that used to obtain $R_\lambda(t)$:

$$R_\lambda(t) = \int_0^t L(t-t')i_\lambda(t')\mathrm{d}t'. \qquad (2.62)$$

Here the integration goes from zero to t. This follows from analogy with the convolution example given in Appendix A, where the actual decay shape $i_\lambda(t)$,

which has non-zero values only within an interval $\langle 0,\infty)$, substitutes now for the function defined by eqn (A.4). Convolution with the excitation pulse is defined by eqn (A.5), from which eqn (2.62) follows via simple application of the commutativity of convolution (A.3).

Often the duration of the excitation pulse $L(t)$ happens to be comparable to the decay time of the luminescence response $i_\lambda(t)$. In this case the actual luminescence response must be determined via deconvolution of relation (2.62).[11] This can also be simply formulated as follows: Due to the finite temporal width of the excitation pulse $L(t)$, the exact (sharp) initial point of decay $t = 0$ cannot be found and, consequently, the time at which the intensity drops to $i_\lambda(0)/e$ cannot be determined. Only in case the pulse width $L(t)$ is much shorter than the decay time, can we approximate $L(t-t') \to \delta(t-t')$ and we obtain $R_\lambda(t) = i_\lambda(t)$; this means that the experiment gives the actual shape of the decay curve, in complete analogy with the situation when the spectral slit width is negligible compared to the width of spectral lines and measured emission spectra do not need to be deconvoluted to recover their actual shape.

It is worth stressing that the function $L(t-t')$ appearing in (2.62) does not represent the actual shape of the excitation pulse but its 'projection' as seen by the whole detection system. Thus, it is not necessary to apply an extremely fast detection set-up that would be able to detect a distortion-free pulse shape, which might pose problems, especially in experiments with nanosecond and picosecond pulses. On the other hand, the detection system must be fast enough to distinguish reliably the shape of the excitation pulse from the shape of the relevant luminescence response, anyway. The pulse shape $L(t)$ can be most easily measured by detecting reflected or scattered excitation light.

Note that the above discussion concerns photoluminescence experiments. In the case of electroluminescence the response is determined not only by optical transitions and recombination phenomena but also by electric effects related to charge transport or capacitance of the device under study. Then, the observed electroluminescence kinetics is a combination of electric and optical effects and its analysis may be considerably more difficult.

2.9.2 Phase-shift method

The principle of this method consists in excitation of the sample with light, the intensity of which is sinusoidally modulated in time with an angular frequency ω. The luminescence response is obviously expected to be also modulated with the same frequency, but the finite value of the excited state lifetime τ induces two effects: Firstly, the luminescence response is phase shifted by an angle φ with respect to the excitation function and, secondly, its modulation depth is decreased.

[11] The relevant methods were mentioned in Section 2.8 when discussing relation (2.58). For the case of luminescence decay curves, a critical review is given by O'Connor, D. V., Ware, W. R., and Andre, J. C. (1979). *J. Phys. Chem.*, **83**, 1333.

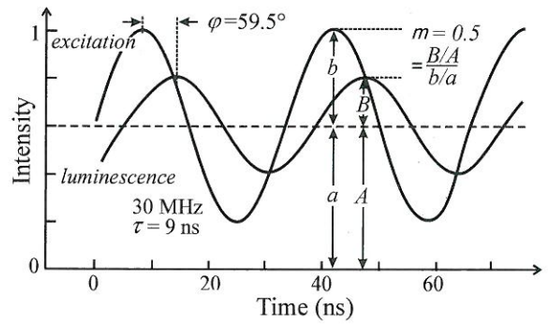


Fig. 2.43
Modulation of excitation light and luminescence response in the phase-shift method. Important parameters: modulation frequency $\omega = 2\pi \times 30\,\text{MHz}$, phase angle $\varphi = 59.5°$, demodulation coefficient $m = 0.5$ and extracted (single-exponential) decay time $\tau = 9\,\text{ns}$. Adapted from Lakowicz [27].

The overall picture is displayed in Fig. 2.43. It can be shown that [27]

$$\tan\varphi = \omega\tau, \tag{2.63a}$$

$$m = \left(1 + \omega^2\tau^2\right)^{-1/2} = \cos\varphi, \tag{2.63b}$$

where $m = (B : A)/(b : a)$ is the ratio of modulation depths of the luminescence to that of the excitation light (Fig. 2.43). As measured quantities thus φ and, possibly, m are obtained; on the basis of these the decay time τ is calculated using (2.63a). It should be stressed that the simple relations (2.63) are strictly valid for pure *single-exponential* decays only. If this is not the case, eqns (2.63a) and (2.63b) provide values of τ differing slightly from one another. They offer an order-of-magnitude estimate of the decay kinetics but their physical meaning is not evident.

To apply this method successfully, the frequency ω must be comparable to $1/\tau$. In the case of $\omega \ll 1/\tau$, the fast luminescence response does follow perfectly the modulations of the excitation light, i.e. $\varphi \to 0$, $m \to 1$, and thus analysis of the experimental curves yields no information about the decay time; if, on the other hand, $\omega \gg 1/\tau$, luminescence is not able to follow fast modulation variations, $m \to 0$, φ cannot be reliably determined and the method fails again. The applicability of the method usually lies within the modulation frequency range $\omega/2\pi \approx 1$–$30\,\text{MHz}$, which allows decay times from 10^{-1} s to $\sim 10^{-9}$ s to be measured.

The phase-modulation method has been given considerable attention for a long time. Its main benefit consists in the capacity to measure relatively fast luminescence kinetics without the need for any source of short or ultrashort pulses. Excitation can be provided by a cw lamp whose output is modulated at high frequency using, for example, light diffraction on ultrasound waves in crystals (the technique of ultrasound waves in the frequency range of 10 MHz is well mastered and a large variety of modulators are available).

Due to the development of reliable picosecond and femtosecond lasers with high repetition rate and good pulse-height stability, the phase-shift method has been put largely in the shade. Also its restriction to a simple single-exponential decay as outlined above had a role to play. A wider spectrum of information on luminescence decay can be obtained using other methods, based on high repetition rate pulsed excitation, like the time-correlated single–photon counting or sampling methods, which are the subject of the following subsections.

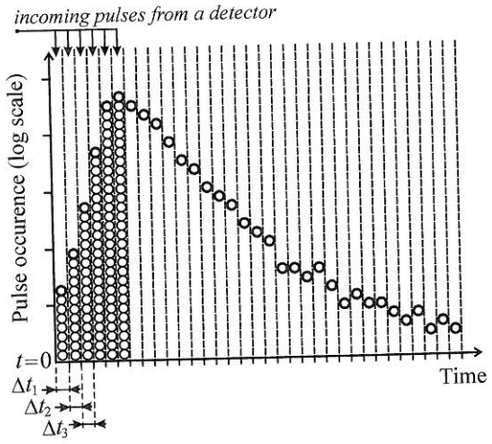


Fig. 2.44
Schematic representation of a luminescence response measured by the time-correlated photon-counting method. A narrow excitation pulse (not shown) is emitted at time $t = 0$.

2.9.3 Time-correlated photon counting

This method is based on the statistical character of luminescence photon emission. Upon excitation of a luminescent material by a very short light pulse, the luminescence emission builds up and then decays. During the decay, the number of emitted photons decreases in time. Photons are emitted randomly in all directions. Suppose we select a narrow solid angle (e.g. using a circular aperture) and record (by a photomultiplier) the arrival time of single photons within this angle, i.e. the time when a pulse appears at the photomultiplier output corresponding to a single photoelectron released from the photocathode (see Subsection 2.4.2); the moment when the excitation pulse was emitted is taken as the time origin $t = 0$. The pulse excitation of the sample is repeated many times.

The luminescence rise and decay curve we are interested in is obtained by plotting a histogram of the pulse occurrence within certain time intervals Δt_i (called time slots), as illustrated in Fig. 2.44. This is based on the fact that the probability of emitting a photon into the selected solid angle at any time is proportional to the number of luminescence centres actually being in an excited state.

One should be aware that the curve shown in Fig. 2.44 is accumulated randomly during the experiment, which means not by shifting regularly from the first time slot towards the next one in sequence Δt_1, Δt_2, $\Delta t_3, \ldots$, etc. The detection events arrive randomly into any of the time slots.

In order to perform such an experiment, two basic requirements must be satisfied:

1. The luminescence intensity must be very low to enable the arrival of single photons to be recorded.[12]

[12] In fact there are two kinds of time-correlated photon counting (TC-PC) experiments: The first one is called time-correlated single-photon counting. Here only one (or no) count is detected following each excitation event (actually, the correct proportion is one detected photon per 10 or 100 excitation pulses—if the rate is higher, the shape of the temporal response is distorted). This is used to measure fast nanosecond decays under high-repetition (> MHz) pulsed lasers (in order to accumulate enough statistics in a reasonable time). The second type of TC-PC is applied to measure slow (microsecond) decays. Here more than one count should be detected at each (low-repetition) excitation pulse, otherwise the statistics cannot be accumulated within a relatively short time period.

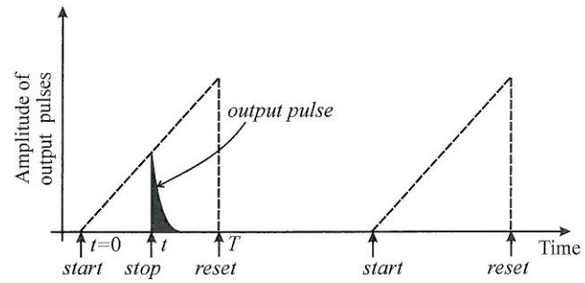


Fig. 2.45
Functioning of a time→voltage amplitude converter.

2. A high number of photons must be detected to obtain conclusive statistics. The repetition rate of special pulsed lamps emitting nanosecond pulses is of the order of 10–100 kHz; repetition rates of mode-locked pulsed lasers can be of the order of 10–100 MHz. Even for such high-repetition excitation, the accumulation of the entire rise/decay curve may take from minutes up to hours.

On the instrumental side, the essential part of this apparatus is a time → voltage converter. Its functioning is illustrated in Fig. 2.45. The converter is initiated by the excitation pulse at time $t = 0$ (start pulse) and begins to generate a DC voltage linearly increasing in time. The voltage rise is stopped by a stop pulse coming from the photomultiplier (which detects single luminescence photons). The resulting voltage amplitude is proportional to the time delay between $t = 0$ and the stop pulse. This voltage is sorted according to its amplitude and then ranked into the corresponding channel of a multichannel analyser. The immediate result of this measurement thus has the form shown in Fig. 2.44, where the horizontal axis represents channel numbers, not directly the time intervals Δt_i.

The time-correlated photon-counting method is very sensitive and accurate, but requires special electronic devices; besides, measurements are sometimes time-consuming. It is mainly applied to the detection of luminescence kinetics at the time-scale of $\sim 10^{-6}$–10^{-10} s. Obviously, the finite width of the excitation pulses $L(t)$ must also be taken into account here and the experimental result, if necessary, deconvoluted in order to extract the actual luminescence response, according to (2.62). A special deconvolution method dedicated to the case of very high repetition rates of the excitation pulses (larger than the reciprocal decay time) is described in [29].

2.9.4 Boxcar integrator

All of the above mentioned methods enable us to detect the luminescence time response at a fixed wavelength (or a range of wavelengths). Now we are going

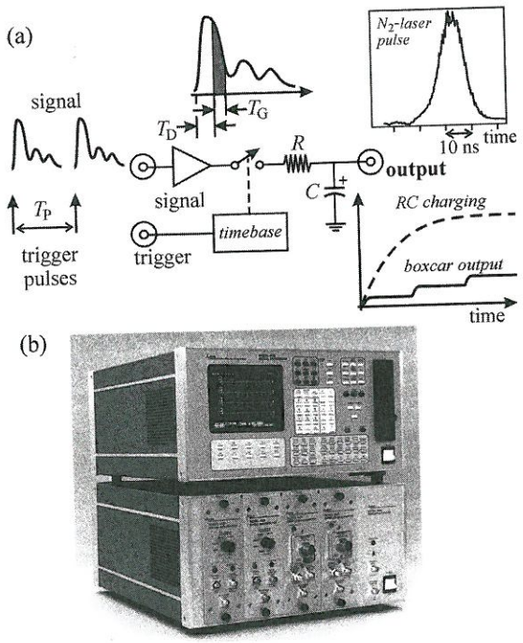


Fig. 2.46
(a) Schematic of boxcar integrator operation. The signal input is loaded by periodic short voltage pulses from a photomultiplier. The trigger input receives synchronous trigger pulses that turn on (via a time base and after a defined delay T_D) a fast sampling switch (gate) for a period T_G. The repetition rate of the whole process is $f = 1/T_P$. After Letzter [30]. The attached graph demonstrates the measured profile of N_2-laser pulses (337 nm) with a temporal FWHM of about 10 ns (detection by a photomultiplier, pulse repetition rate $f = 50$ Hz). (b) Digital EG&G PAR Model 4400 boxcar integrator. Reproduced with permission by Ametek Signal Recovery (formerly EG&G Signal Recovery).

to describe an electronic device called a boxcar integrator (or boxcar averager) which has a much greater diversity of experimental options; of practical importance for us is its capacity to record complete time-resolved luminescence spectra (i.e. emission spectra when the luminescence intensity is registered at a predefined delay after the excitation event). To begin with, we describe the principle of the device, then its application to luminescence kinetics and finally we shall address the measurement of time-resolved emission spectra.

The boxcar integrator was developed by Princeton Applied Research in the middle of the 1960s, mainly in response to demands for the detection of a wide variety of fast optical signals with high repetition rate, emerging due to the rapid development of laser techniques at that time. It can also be classified as a device aimed at detecting weak repetitive signals unprocessable by the lock-in detection technique.

A simplified scheme of boxcar functioning is shown in Fig. 2.46(a) [30]. There are two inputs: (i) the investigated signal (pulsed output from a photomultiplier or other detector) that is repeated with a period T_P, and (ii) a trigger pulse, synchronous with the signal. The signal channel contains a broadband preamplifier, a fast sampling unit, and a low-pass RC filter. With the incoming trigger pulse, the sampling unit (gate) is switched on (with a defined delay T_d) for a time interval T_G. Therefore, during this time interval a capacitor $C$is being charged and its voltage rises. The whole process is repeated with a frequency $f = 1/T_P$.

As far as the delay T_D is concerned, two modes of operation are feasible. Either T_D is kept constant, or is left to increase linearly in time. The latter mode is applied to measure the temporal shape of the incoming pulse, i.e. the

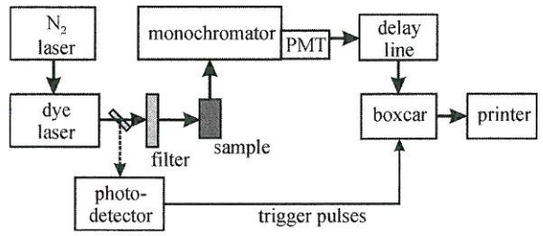


Fig. 2.47
Single-channel luminescence spectrometer with a pulsed laser excitation. Pulses from a dye laser (pumped by a nitrogen laser) periodically excite luminescence of a sample, a small fraction of the excitation pulse energy (reflected on a beam splitter) being applied to generate trigger pulses for a boxcar integrator. The delay line, inserted between the photomultiplier and the boxcar signal input, serves to compensate for internal delay in the triggering circuits of the integrator. Adapted from Letzter [30].

luminescence kinetics. Here the gate samples the temporal profile of signal pulses by shifting the delay by an increment Δt after each incoming pulse.[13]

The RC time constant is usually much larger than T_G, thus the voltage on the capacitor C increases in small steps (Fig. 2.46(a)). Due to the simultaneous slow discharging of the capacitor the output signal asymptotically approaches the average value of the input pulse amplitude within the given time interval. In other words, the latest signal acquisitions have predominant weight over the earlier ones that fade away gradually and the input signal shape is replicated at the output. This mode of signal treatment is referred to as an exponential averaging.

At the same time, an important condition is to find an appropriate relation among several experimental parameters—too fast gate scanning distorts the pulse shape while too slow scanning leads to unnecessary prolongation of the experiment and possible errors due to drift of experimental parameters. This means that an optimum time of signal scanning T_M must exist; it can be shown that this optimum scanning time is given approximately as $T_M \approx 5RCX/T_G^2 f$, where X denotes the overall sampled time interval. The inset in Fig. 2.46(a) shows an example: a N_2-laser pulse shape measured in this way ($X = 60$ ns, $T_G = 1$ ns, $T_M = 60$ s, $f = 50$ Hz).

Time-resolved emission spectra are measured in the $T_D = \text{const}$ mode (to be more specific, using several discrete values of T_D). The experiment can be performed in the set-up shown in Fig. 2.47: for any selected delay value T_D the monochromator is let to scan the spectrum. This means that this time the same portion of the temporal profile of luminescence pulse is sampled for a time T_G, while the wavelength is being scanned. On condition of a proper setting of all experimental parameters, i.e. the scanning speed in relation to RC, T_G and f, exponential averaging takes place and the luminescence spectrum is continuously recorded under a pulsed excitation. (This ability of the boxcar integrator contributed considerably to the fast development of laser spectroscopy in the 1960s and 1970s; till that time, there was practically only one way to detect emission spectra under pulsed excitation, namely, lengthy and uncertain registration on a photographic plate.) Moreover, the signal-to-noise ratio is improved when using a boxcar, because the input noise is not coherent with the trigger pulse rate and its contribution to the integrated output signal is therefore small, likewise in the lock-in technique.

[13] In principle, one deals with a stroboscopic method of measurement of periodic events, exploited also in, for example, a sampling storage osciloscope.

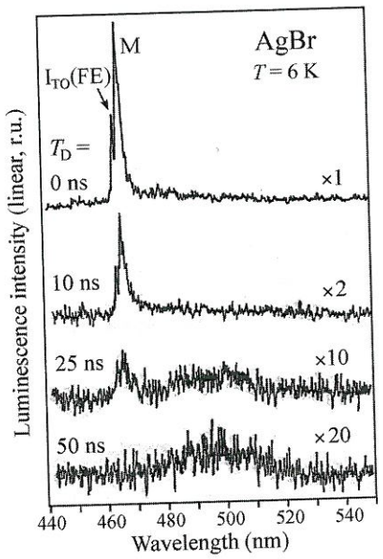


Fig. 2.48
Time-resolved luminescence spectra of AgBr at a temperature of $T = 6$ K, measured in an experimental configuration similar to that shown in Fig. 2.47 (only the dye laser was omitted, the sample was excited directly with N_2-laser pulses 337 nm, 2 MW/cm^2, 10 ns). The gate width is $T_G \sim 3$ ns and the values of the delay T_D behind the excitation pulse maximum are indicated on the left-hand side of each spectrum. With increasing delay T_D the emission lines I_{TO}(FE) and M are rapidly fading out and a slow band around 490 nm appears (this emission band is shown also in Fig. 1.2(b), but excited with cw UV light). Adapted from Baba and Masumi. [31].

Upon changing the time position of the gate, a new spectrum is registered at another delay T_D after the excitation pulse—this is the principle of time-resolved luminescence spectroscopy. An example is presented in Fig. 2.48, showing low-temperature (6 K) time-resolved emission spectra of pure AgBr under pulsed excitation with a N_2-laser (337 nm) [31]. Clearly, the spectra in different time delays (windows) behind the excitation event can differ substantially. Immediately after excitation only two narrow emission lines ('fast component'), labelled I_{TO}(FE) and M, are observed. They rapidly fade out with increasing delay T_D and a new broad emission band ('slow component') appears around ~ 490 nm. The origin of these spectral features will be discussed in Sections 7.1, 7.2, and 8.2.

2.9.5 Streak camera

A boxcar integrator is an excellent instrument to study luminescence kinetics and decay times from about 500 ps to hundreds of ns but does not serve this purpose for faster processes taking place on shorter time-scales of 10–100 ps. This constraint is due to the limited speed of the boxcar electronic circuits as well as the relatively slow response of the photomultipliers (even the special fast photomultipliers do not have the signal onset faster than ~ 500 ps). Similar limitations take place also in oscilloscopes, more specifically in their input amplifiers. For luminescence transients occurring on the time-scale of picoseconds and femtoseconds, either special methods of nonlinear optics (up-conversion) or the so-called streak camera should be used. This device enables the measurement of both time evolutions of ultrafast light signals and time-resolved emission spectra, likewise the boxcar integrator. The development of the streak camera at the beginning of the 1970s (Hamamatsu Co. Ltd.) was

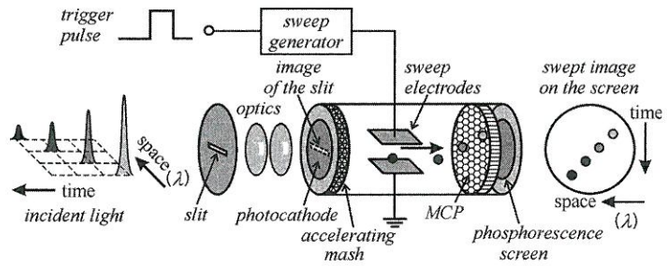


Fig. 2.49
Operation schematic of a streak camera. The light pulse passing through the entrance slit and hitting a photocathode induces emission of photoelectrons inside an evacuated tube. Photoelectrons are accelerated along the horizontal direction (i.e. the direction of the incoming light) by a voltage applied between the photocathode and a metal mesh and subsequently deflected by a high voltage in the vertical direction. Upon impinging on a phosphorescence screen, the photoelectrons produce a phosphorescent trace with spatial distribution of intensity corresponding to the temporal profile of the incoming pulse. After materials by Hamamatsu Co. Ltd.

again motivated by the fast dissemination of lasers emitting picosecond and femtosecond pulses.

The streak camera is a device which converts temporal information, originating in a light-emission event, to spatial information. Roughly speaking, the streak camera combines, in a single device, the principles of a photomultiplier with the basic components of an oscilloscope: the external photoelectric effect, deflection of an electron beam and a phosphorescence screen. The principle of a streak camera is shown in Fig. 2.49. The analysed light pulse is focused on the entrance (mechanical) slit, whose image is projected onto a photocathode deposited on the internal front face of an evacuated tube. Emitted photoelectrons (whose number is proportional to the intensity of the incident light) are accelerated in an electric field and then deflected by a high voltage applied in the direction perpendicular to the acceleration voltage. This deflection voltage is applied synchronously with the incoming light pulses, the photoelectrons thus being swept with a defined speed and known trajectory (like in an oscilloscope, but without input amplifier here) by which means the time information is converted to spatial information. Then these electrons undergo multiplication via secondary electron emission in the so-called microchannel plate, and finally they bombard a phosphorescence screen on the opposite end of the vacuum tube. This creates the phosphorescence image, rendering the temporal distribution of intensity of the analysed luminescence pulses in the form of a vertical spatial distribution of phosphorescence intensity on the screen. The phospherescence image is then recorded (e.g. by a photodiode array or a CCD detector), treated and displayed on a computer screen.

An example of a measured luminescence transient is presented in Fig. 2.50: luminescence of Hg_2Cl_2 (calomel) crystals excited by ~ 35 ps pulses of second harmonic frequency from a passively mode-locked Nd:YAG laser (532 nm) [32]. As Hg_2Cl_2 is an anisotropic uniaxial crystal, its luminescence response exhibits a slightly different temporal shape under polarization of the excitation light perpendicular to the optical axis $\mathbf{c}$ (electric field vector $\mathcal{E} \perp \mathbf{c}$) and parallel with this axis ($\mathcal{E} || \mathbf{c}$). On the left-hand side of the lower curve in Fig. 2.50 one can also see the shape of the excitation pulse; this was enabled by the appropriate setting of an optical delay line for both luminescence and the reflected fraction of the excitation light and their simultaneous focusing onto the camera entrance slit.

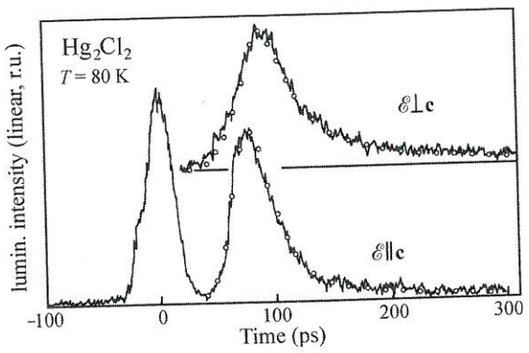


Fig. 2.50
Luminescence transients measured and displayed by a streak camera (luminescence of Hg_2Cl_2 crystals spectrally situated around ~ 396 nm, excited at $T \sim 80$ K via two-photon absorption of second harmonic pulses from a Nd:YAG laser with temporal FWHM ~ 35 ps). The lower curve shows also (on the left) the excitation pulse profile. Full lines represent experimental data while points show the calculated convolution of the excitation pulse profile with a double-exponential decay. After Svoboda *et al.* [32].

The open circles in Fig. 2.50 represent the best numerical fit using the model of double-exponential decay $i_\lambda(t) = a_1 \exp(-t/\tau_1) + a_2 \exp(-t/\tau_2)$ with parameters:

$$a_1 = 0.93,\ \tau_1 = (17 \pm 5)\,\text{ps},\ a_2 = 0.07,\ \tau_2 = (87 \pm 30)\,\text{ps}\ (\mathscr{E}\|\mathbf{c}),$$
$$a_1 = 0.82,\ \tau_1 = (25 \pm 7)\,\text{ps},\ a_2 = 0.18,\ \tau_2 = (77 \pm 26)\,\text{ps}\ (\mathscr{E}\perp\mathbf{c}). \quad (2.64)$$

Figure 2.50 indicates that the excitation pulse FWHM and the characteristic times τ_i are of the same order of magnitude, therefore to fit the experiment properly, numerical convolution of the measured system response to the excitation pulses with a presumed model shape of luminescence decay $i_\lambda(t)$ according to eqn (2.62) was applied. The numerical values given in (2.64) give an idea of the attainable accuracy in extracting the decay times τ_i by this method.

The principle of the measurement of time-resolved luminescence spectra follows from Fig. 2.49. The streak camera can be attached to the spectrograph output in such a way that the spectrum of the luminescence signal is imaged onto the entrance slit. Then the two-dimensional image displayed on the phosphorescence screen represents the temporal evolution of the pulse spectral content while the photoelectrons are vertically swept.

Various models of streak cameras are commercially available for various time regimes, with time resolution as short as ≤ 1 ps. Streak cameras can operate in a single-shot mode but also at high repetition rates of the order of 100 MHz (the so-called synchroscan camera).

2.10 Problems

2/1: *Photomultiplier nonlinearity*. When illuminating a high-gain photomultiplier by too high photon flux, the photocurrent between the last dynodes may attain such high values (without yet inducing irreversible damage to the photomultiplier) that the condition I_a (anode current) $\ll I_{vd}$ (current through the voltage divider) is no longer fulfilled—see Fig. 2.5(b). Consequently, the last stages of the voltage divider are partially 'short-circuited' by the current I_a flowing inside the photomultiplier. Then the voltage

from an external high-voltage supply is concentrated predominantly in several first stages of the divider (close behind the photocathode). Thus a positive feedback appears, the photomultiplier gain increases and the linearity of the photomultiplier response is violated. Discuss the consequences of this effect for measurements of optical gain according to the paper by Webb [33].

2/2: Show that the relation between the photomultiplier anode sensitivity $k(\lambda)$ (expressed in units of A/W) and the photocathode quantum efficiency η_{ph} has the form

$$\eta_{ph}(\%) = \frac{h\,c}{\lambda e}\frac{k}{G} \times 100 = \frac{1240}{\lambda}\frac{k}{G} \times 100,$$

where h is Planck's constant (Js), c is the speed of light (m/s), e is the electron charge (C), λ denotes the wavelength of the incident light, and G is the photomultiplier gain.

2/3: You are in a position to investigate the emission spectrum of a semiconductor with low luminescence yield. This spectrum is expected to be situated between 600 and 800 nm and composed of spectral bands and lines, the FWHM of the narrowest ones being about 1 nm and that are separated from one another by 2.5 nm. Decide which of the following monochromators is more convenient for this task and justify your selection:

1. Focal length $f = 1000$ mm, aperture $f/7.5$, holographic grating $120 \times 140\,\text{mm}^2$ (blaze at 630 nm), reciprocal linear dispersion $L^{-1} = 0.8$ nm/mm, resolving power $R = 7 \times 10^4$ (at the slit width of 10 μm).
2. Focal length $f = 320$ mm, aperture $f/4.2$, holographic grating $120 \times 140\,\text{mm}^2$ (blaze at 750 nm), reciprocal linear dispersion $L^{-1} = 2.4$ nm/mm, resolving power $R = 6 \times 10^3$ (at the slit width of 10 μm).

What is the maximum slit width of the selected monochromator you may apply in the experiment?

2/4: A scanning monochromator is used to measure one of the emission lines of the material mentioned in the previous problem. The signal from a photomultiplier is treated by a lock-in amplifier. What is the maximum time constant you can apply (in order to obtain the spectrum as smooth as possible but without distortions), if the wavelength scan speed is 10 nm/min? What is the total acquisition time of the spectrum? Note: The values of the time constant can usually be set in steps 1–3–10, e.g. 1 ms–3 ms–10 ms.

2/5: (a) Show that the FWHM of a bell-shape curve $f(x) = x \exp(-x/a)$ is FWHM $\sim 2.45\,a$.

(b) Suppose you excite photoluminescence with optical pulses of temporal FWHM = 5 ns and measure the corresponding luminescence decay. Let the decay curve $i(t)$ be of single-exponential type with the characteristic time constant $\tau = 15$ ns. Estimate the uncertainty in determining τ if no deconvolution of the experimental curve is applied. For the sake of simplicity suppose that the excitation pulse is described by a function $L(t) = t \exp(-t/a)$ and use the result of part (a).

References

1. Voos, M., Leheny, R. F., and Shah, J. (1980). *Radiative recombination*. In: *Handbook on Semiconductors* (ed. T. S. Moss), Vol. 2 (ed. M. Balkanski), p. 329. North Holland, Amsterdam.
2. Kulakovskii, V. D. and Timofeev, V. B. (1983). *Thermodynamics of electron–hole liquid in semiconductors*. In: *Modern Problems in Condensed Matter Sciences* (general ed. V. M. Agranovich and A. A. Maradudin), Vol. 6, *Electron–Hole Droplets in Semiconductors* (ed. C. D. Jeffries and L. V. Keldysh), p. 95. North Holland, Amsterdam.
3. Valenta, J. (1994). *Photoluminescence characterization of selected semiconductors and insulators*. PhD. thesis, Charles University in Prague, Faculty of Mathematics & Physics, Prague.
4. 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.
5. Itoh, C., Tanimura, K., Itoh, N., and Itoh, M. (1989). *Phys. Rev. B*, **39**, 11183.
6. Engstrom, R. W. (1980). *Photomultiplier Handbook*. RCA/Burle, USA.
7. *Photomultiplier Tube. Principle to Application*. Hamamatsu Photonics K. K. 1994.
8. Rieke, G. H. (2003). *Detection of Light. From Ultraviolet to Submillimeter*, 2nd edn, p. 194. Cambridge University Press, Cambridge.
9. Saleh, B. E. A. and Teich, M. C. (2007). *Fundamentals of Photonics*, 2nd edn. Wiley-Interscience, New York.
10. Ferrari, M., Pavesi, L., and Righini, G. C. (eds) (2002). *Micro-Optoelectronics: Materials, Devices and Integration*, p. 26. Italian Society of Optics and Photonics, Centro Editoriale Toscano.
11. Mac Intyre, R. J. (1966). *IEEE Trans. Elect. Dev.*, **13**, 164.
12. *Spectrum One, Spectroscopic CCD Detectors*. ISA Jobin Yvon-Spex, Horiba.
13. Michel, P. (1953). *La spectroscopie d'emission et ses applications*. Collection Armand Colin, Paris.
14. Sawyer, R. A. (1951). *Experimental Spectroscopy*. Prentice Hall. New York.
15. James, J. and James, J. F. (2007). *Spectrograph Design Fundamentals*. Cambridge University Press, Cambridge.
16. Demtröder, W. (2003). *Laser Spectroscopy: Basic Concepts and Instrumentation*, p. 109. Springer, Heidelberg.
17. Pelant, I. , Hála, J., Parma, L., and Vacek, K. (1980). *Solid State Comm.*, **36**, 729.
18. Hulin, D., Mysyrowicz, A., Combescot, M., Pelant, I., and Benoit à la Guillaume, C. (1977). *Phys. Rev. Lett.*, **39**, 1169.
19. *Photon Counting using Photomultiplier Tubes* (1998). Technical information, Hamamatsu Photonics K. K.
20. Bagdanskis, N. I., Bukreev, V. S., Žižin, G. N., and Popova, M. N. (1982). High resolution infrared spectrometers (in Russian: *Infrakrasnyje spektrometry vysokogo razrešenija*). In *Current Trends in Spectroscopy Techniques* (in Russian: *Sovremennyje tendencii v technike spektroskopii*, ed. S. G. Rautman), p. 153. Nauka, Novosibirsk.
21. Fojtík, P. (1999). *Selected photoelectrical properties of amorphous and microcrystalline silicon with respect to lateral resolution* (in Czech: *Vybrané fotoelektrické vlastnosti amorfního a mikrokrystalického křemíku s přihlédnutím k laterálnímu rozlišení*). Diploma thesis, Charles University in Prague, Faculty of Mathematics & Physics, Prague.
22. Shionoya, S. and Yen, W. M. (eds) (1999). *Phosphor Handbook*, Chap. 14. CRC Press, Boca Raton.
23. Dian, J., Valenta, J., Luterová, K., Pelant, I., Nikl, M., Muller, D., Grob, J. J., Rehspringer, J.-L., and Hönerlage, B. (2000). *Mater. Sci. Engin. B*, **69–70**, 564.
24. Pelant, I., Kohlová, V., Hála, J., Ambrož, M., and Vacek, K. (1987). *Solid State Com.*, **64**, 939.

25. Bertero, M. and Boccacci, P. (1998). *Introduction to Inverse Problems in Imaging*. Institute of Physics Publishing, Bristol.
26. Shah, J. (1999). *Ultrafast Processes of Semiconductors and Semiconductor Nanostructures*. Springer Series in Solid-State Sciences, Vol. 115. Springer, Berlin.
27. Lakowitz, J. R. (1983): *Principles of Fluorescence Spectroscopy*, Chap. 3. Plenum Press, New York and London.
28. Luterová, K., Pelant, I., Valenta, J., Rehspringer, J.-L., Muller, D., Grob, J. J., Dian, J., and Hönerlage, B. (2000). *Appl. Phys. Lett.*, **77**, 2952.
29. Sakai, Y. and Hirayama, S. (1988). *J. Luminescence*, **39**, 145.
30. Letzter, S. (1983). *Laser Focus*, December 1983.
31. Baba, T. and Masumi, T. (1977). *Nuovo Cimento B*, **39**, 609.
32. Svoboda, A., Večeř, J., Barta, Č., and Pelant, I. (1990). *phys. stat. sol. (b)*, **157**, K69.
33. Webb, C. (2003). *Optics & Photonics News*, May 2003, 14.