Chapter 1
Signal, Distortion, and Noise

Our usual aim is to measure the brightness of an astronomical object – the amount of energy per second arriving at the Earth from that object, usually in the form of light waves but sometimes in the form of particles or of gravitational waves. The arriving light could be characterised in several ways – its time of arrival, its arrival direction, the frequency of the light, and its polarisation state. More generally, we might measure the brightness of light in various different directions, thus making a *map* of the sky brightness; or we might measure how the brightness of the light from an object is distributed in frequency, thus measuring the *spectrum* of that object. Considering individual photons, each one is characterised by exactly four quantities – time, direction, frequency, and polarisation.

However ... when we wish to measure the brightness of light from some astronomical object, we cannot determine it directly. Instead, we rely on the arriving light causing some physical change in our detection system, so that what we actually measure may be a current, or a voltage. The detected effect is known as the *signal*. Furthermore, between the arrival of the light at the top of the atmosphere and the detected signal, much of the incoming light may be lost by a variety of processes in the atmosphere, the telescope, the camera optics, and so on. How to convert from the detected signal to the desired astronomical quantity is the problem of *calibration*. The pattern of light can also be smeared out by a variety of effects in the atmosphere, optics and detector, limiting our *resolution*, that is, the finest scale of information we can distinguish. Even worse, the signal may also be accompanied by *noise*. i.e. random fake signals caused by background light, detector fluctuations and so on, so that it can be hard to be sure when we have actually seen a signal at all. Before looking at how light is lost or distorted in specific situations, we need to consider the general principles of such processes.

In the opening sections of this introductory chapter, we will review how the basic quantities – frequency, direction, time, and polarisation – are expressed in an astronomical context. Much of this material will be well known to likely readers of this book, but it is worth reviewing carefully. Following this, we will look in turn at signal loss and calibration, at distortion and smearing, at noise, and how the presence of noise affects sensitivity and resolution.

A. Lawrence, *Astronomical Measurement*, Springer Praxis Books,
DOI 10.1007/978-3-642-39835-3_1,

1.1 Basic Quantities

1.1.1 Frequency, Wavelength and Energy

Astronomers tend to jump promiscuously between frequency, wavelength, and photon energy, depending on the circumstances, and use different units for different wavelength regimes.

In radio and mm-wave astronomy it is most normal to characterise light by its *frequency*, usually given the symbol ν and almost always expressed in Hz (i.e. cycles per second) or derivatives, such as MHz (10^6 Hz) or GHz (10^9 Hz). In other areas of physics one occasionally sees angular frequency, i.e. $\omega = 2\pi\nu$, in units of radians per second, but this is almost never used in astronomy.

In UV, optical and infra-red astronomy, and sometimes in radio astronomy, it is more normal to characterise light by its *wavelength*. This is of course related to the frequency by

$$\lambda = c/\nu,$$

where c is the speed of light. If we use SI units with ν in Hz and $c = 2.998 \times 10^8\,\mathrm{m\,s^{-1}}$, then the wavelength is in metres. However, the units of wavelength used in practice depends on the type of astronomy – m or cm in radio astronomy, μm in infra-red astronomy, and nm or Å in optical and UV astronomy, where $1\,\text{Å} = 10^{-10}$ m. Note also that the speed of light in a physical medium such as air or glass is slower than the speed of light in a vacuum; as a light wave changes from one medium to another the frequency (rate of arrival of wavefronts) stays the same, but the wavelength changes. For very accurate work therefore, you have to be careful to note whether wavelengths quoted are in vacuum, in air, or whatever. Occasionally in infrared astronomy you might see light characterised by its *spatial frequency*, that is, the number of waves per unit length, often referred to as the "wavenumber" and quoted in units of $\mathrm{m^{-1}}$ or $\mu\mathrm{m^{-1}}$. This of course is simply $1/\lambda$.

In X-ray and gamma-ray astronomy, frequency and wavelength are rarely used, with spectral dependence almost always being expressed in terms of photon energy. Photon energy E is related to the light frequency by the Planck formula

$$E = h\nu = hc/\lambda,$$

where h is Planck's constant. In SI units, with ν in Hz and $h = 6.626 \times 10^{-34}\,\mathrm{m^2\,kg\,s^{-1}}$, energy E would be in Joules. For energy in general, many astronomers still use the cgs unit of ergs (where $1\,\mathrm{J} = 10^7$ ergs). For photon energy however (and also for the energy of other kinds of particle) the normal unit is the electron volt (eV). This is defined as the energy an electron will gain when accelerated through a potential gap of 1 V. Because the volt is an SI unit, this means

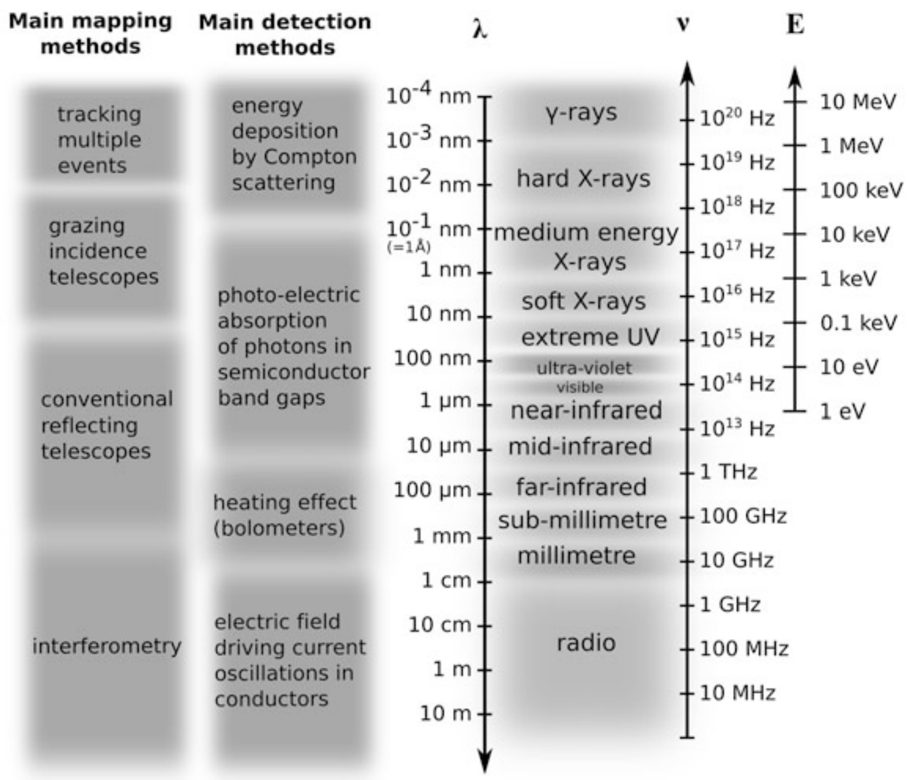

Fig. 1.1 A graphic illustration of the relation between wavelength, frequency, and photon energy, and the location of customary wavebands. The figure shows how the main techniques used for mapping and detection vary across the EM spectrum

that the energy gained is $E = 1 \times e$ where e is the electron charge in Coulombs, 1.602×10^{-19} C. The conversion between eV and J is therefore

$$1 \text{ eV} = 1.602 \times 10^{-19} \text{ J}.$$

Note that optical photons are of the order eV, X-rays of the order keV, and gamma-rays MeV to GeV.

Figure 1.1 shows graphically how the three ways of characterising light – by frequency, wavelength and photon energy – relate to each other. The figure also shows the location of the standard *wavebands* used in astronomy – radio, X-ray, ultraviolet etc. The names reflect history of course, and the precise boundaries are somewhat arbitrary. However the existence of such differing regions of the spectrum is quite meaningful in several ways. From the point of view of astrophysics, different kinds of light tend to originate via different emission mechanisms – blackbody radiation, synchrotron radiation, atomic transitions, nuclear transitions and so on. More importantly for the purposes of this book, the typical *detection methods* are different in different regimes – for example gamma-rays are usually detected via Compton scattering, optical photons by absorption in semi-conductors, and radio waves by the excitation of oscillating currents in conductors. Likewise, the different wavelength regimes differ in how the angular distribution of brightness on the sky is typically measured – gamma-rays by shadowing or particle tracking, optical light by focusing and imaging, and radio waves by interferometry. The major differences are indicated in Fig. 1.1, and the physical and technological issues concerned are taken up in more detail in Chaps. 3 and 4.

1.1.2 Position

Much of the time astronomers are concerned simply with *relative positions* on the sky, with angular distances $\Delta\theta$ measured in degrees, arcminutes, or arcseconds. It is also quite common to see angles quoted in decimal degrees, where $1'' = 0.0000278°$. Many astrophysical calculations will of course require the angle to be in standard units of radians; to convert it is simplest to just recall that $180° = \pi$ rad. Often we want to specify a solid angle on sky; the standard unit is the steradian (sr), such that there are 4π sr over the whole sky, but astronomers often measure areas on the sky in square degrees or square arsceconds etc. To convert, it is easiest to recall that $1\,\mathrm{sr} = (180/\pi)^2$ square degrees. For quick mental calculation it can also be handy to remember that there are $\sim 41{,}253$ square degrees on the whole sky.

The *local position* of an object in the sky, as seen from a specific location on Earth, can be specified by two co-ordinates – azimuth and elevation. The zero point of azimuth is normally given by the local direction of North. Elevation is angle upwards from the horizon, normally specified in degrees. Quite often however, astronomers work in terms of the *zenith distance* of an object, which is the angle downwards from the vertical towards the horizon. The zenith distance determines the path length of atmosphere through which the object is being observed. As we will discuss in Chap. 2, three effects change with elevation/zenith distance – the brightness of the sky background; the absorption (extinction) of light; and the bending (refraction) of light. Note that because of the rotation of the Earth, the local position (θ, ϕ) of an object will change with time.

Absolute positions on the sky are given in two co-ordinates equivalent to longitude and latitude, but these need to be tied to something which is the same for everybody, as opposed to being tied to local directions. There are several different ways of doing this. These absolute co-ordinate systems are not important for this book, but we describe them briefly here for completeness. (See any book on observational astrophysics for more detail.) The standard system is that of *celestial co-ordinates*. Here the longitude co-ordinate, known as *Right Ascension* or RA, is defined by the projection of the Earth's equator out into space, with the zero point of longitude being where the equator is intersected by the ecliptic, that is, the annual path of the Sun around the sky. The latitude co-ordinate, known as *Declination* or Dec, is the elevation upwards from the celestial equator. Declination is normally specified in degrees – either decimal degrees, or degrees, minutes and seconds. Right Ascension can also be specified in degrees, but traditionally is specified in hours, minutes and seconds, with 24 h representing 360°. This has its origin in the fact that, because of the rotation of the Earth, the pattern of stars in the sky appears to move, so that any given star, at a specific (RA, Dec) co-ordinate, re-appears at the same local position once every $\sim$24 h. An absolute sky position given in hours, minutes and seconds for RA, and degrees, minutes and seconds for Dec, is said to be a *sexagesimal* co-ordinate. Although the pattern of stars remains fixed, the movement of the Earth around the Sun means that the position of the Sun with respect to the stars changes, so that the range of RAs available at night-time gradually shifts around the year. The motion of the Earth around the Sun also produces a very slight annual modulation of the position of very nearby stars with respect to more distant ones – the effect of parallax.

Another way to specify an absolute position on the sky is to tie the longitude to the observed plane of the Milky Way, with the zero of longitude defined by the Galactic Centre, giving Galactic Co-ordinates (l, b). How to tie such reference frames down precisely is a significant issue in astronomy, but will not concern us further in this book.

1.1.3 Time

Accurate time measurement crops up in four ways in astronomy. First, because of the rotation of the Earth, we need to know the time in order to convert the RA and Dec position of an object into its current local position in the sky. Secondly, we sometimes need to record the *absolute time* of an observation in some standard system, so that we can compare observations made with two different telescopes or spacecraft. For example, did a gamma-ray burst seen with the Fermi spacecraft occur before or after an optical flash seen with the Gemini telescope in Hawaii? Another important example will be seen in Chap. 4, when we discuss very long baseline interferometry, where radio observations from observatories thousands of miles apart have to be tied accurately together. The convention in astronomy is to record the Universal Time (UT) of an event, which requires observatory clocks to be kept synchronised with international standards. (Note that UT is the same as Greenwich Mean Time, GMT.) The UT time has to be expressed within a given calendar date, which in turn is normally expressed as a Julian Date (JD), which is formally defined as starting at Greenwich noon on Jan 1st 4713 B.C. on the Julian Calendar. More usefully, noon U.T. on March 1st 2000 was JD 2,451,605.0 Some astronomers quote the Modified Julian Date (MJD) which is just JD-2,400,000.5 and which is equivalent to starting the system on Nov 17, 1858.

The third use of accurate time measurement is in recording the *relative timing* of events within an observation. For example, in X-ray astronomy, photons are generally detected individually, and each one tagged with the time of arrival. This needs an accurate internal clock. However, it also imposes constraints on the ideal design of detectors – for example, we do not want a "dead period" after detection of a photon during which we miss other photons.

The fourth use of accurate time measurement is in recording the *temporal fluctuations* of an incoming signal. To perform interferometry (see Chap. 3) we need to preserve the phase information of incoming waves, so that they can be mathematically combined later in a computer. To perform Fourier Transform spectroscopy, we need to calculate the autocorrelation function of the fluctuating signal, as this will be the Fourier Transform of the spectrum of the signal. This is performed by specialised fast electronics, as explained in Chap. 5.

1.1.4 Polarisation

The polarisation state of the light emitted by astronomical objects can hold important physical clues. Fully characterising the possible state of received light waves involves five numbers – the frequency, the intensity, and up to three numbers characterising the polarisation state. Consider first a simple plane EM wave. In addition to the frequency and intensity we need only one number – the *position angle* of the E-vector on the sky (or equivalently the B-vector of course). However, very few objects emit straightforward plane-wave radiation. Most natural light is made up of short bursts ($\sim 10^{-8}$ s) of such plane waves, with random orientations. Such light is *unpolarised* and therefore characterised only by its frequency and intensity. A common situation is that light is composed of a mixture of unpolarised light and plane waves. To characterise such light we need four numbers – frequency, total intensity, the *position angle* of the polarised component, and the *percentage polarisation*, that is the percentage of the total intensity which is in the polarised component.

More generally, and occasionally of importance in astronomy, we have to allow for the fact that the polarised component may show *elliptical polarisation* rather than simple plane polarisation. If two synchronised orthogonal plane waves are added, the result is a plane wave at 45° to either of the originals. If there is a 90° phase difference, the result is a kind of corkscrew wave – seen in the plane of the sky, the tip of the E-vector rotates in a circle. More generally, for various phase differences and relative component sizes, the result will be an E-vector that traces an ellipse. (See any good optics textbook for a more careful explanation.) Such more general polarised light needs two numbers – not just position angle, but also the *eccentricity* e of the ellipse, or some such equivalent number, with plane waves corresponding to $e = 1$.

The complete theoretical description of light of a given frequency therefore involves four numbers: intensity, percentage polarisation, position angle, and elliptical eccentricity. The measurement of polarisation normally involves measuring the intensity of light after passing through a variety of polarising filters; a standard set of such operations produces four numbers known as the *Stokes parameters*, from which the theoretical description can be derived. Astronomical papers will quite often simply report the Stokes parameters.

1.2 Rate of Arrival: Fluxes

In observational astronomy, we are usually aiming to measure some kind of *flux*: the flow of something per second through a given surface. For example we might wish to know the total amount of energy per square metre per sec arriving at the Earth that has been emitted by some astronomical object:

$$F \quad [Wm^{-2}].$$

Note that sometimes the term "flux" is used to mean the flow of energy through a particular surface, with *flux density* used specifically to mean "flux per unit area". However, astronomers often use the term "flux" to mean "flux per unit area', and use "flux density" to mean the amount of flux per unit spectral frequency, as we will discuss below. It should usually be clear from context and/or units how these terms are being used.

It is not always light that we are trying to measure. We might want to know how many cosmic ray protons are arriving at the Earth, in which case the flux would be measured in units of particles per square metre per second. High energy photons (X-rays and gamma-rays), like particles, are detected in discrete events, so in that case we might wish to measure photons per square metre per second. Astronomers tend to use all sorts of hybrid units, so one simply has to get used to switching between them. For example, an X-ray flux, rather than being quoted in $\mathrm{W\,m^{-2}}$, may be quoted in units of $\mathrm{keV\,cm^{-2}\,s^{-1}}$.

In this book, we will assume that we are talking about the flux of lightwave energy unless otherwise stated.

1.2.1 Surface Brightness

Often we are concerned with making maps of the sky – either by imaging, or by some other technique such as scanning or interferometry (see Chap. 3). Although we typically end up with a linear image in (x,y) co-ordinates, what we are really doing is measuring the brightness of the sky in different directions (θ,ϕ). For single objects, such as stars, our aim would be to collect light from a wide enough range of angles that we "catch" all the light from that object; but for an extended object, such as a nebula or a galaxy, we want a map of the *surface brightness* as a function of position on the sky. We can define the surface brightness $I(\nu,\theta,\phi)$ such that $I d\nu d\theta d\phi$ is the energy per sq. m. per sec arriving at the surface of the Earth coming from frequency ν and position $(\theta,\ \phi)$ over the small range $d\nu$, $d\theta$, $d\phi$. The standard unit for surface brightness would be $\mathrm{W\,m^{-2}\,Hz^{-1}\,sr^{-1}}$, but it is often quoted within a standard passband in magnitudes per square arcsecond. (See below for an explanation of magnitudes.)

Note that this definition of "surface brightness" looks essentially the same as the quantity described as "Intensity" in books on radiative transfer or optics. In that case, one is considering various possible surfaces within an object or within an optical system, and describing the flow of energy through these surfaces, specifying how much energy is coming from each different direction (θ, ϕ) as seen from the surface, per unit solid angle.

1.2.2 Spectral Flux

The discussion above assumes that we measure all the radiation from our object, regardless of wavelength/frequency. This is known as the *bolometric flux*. However, to get useful astrophysical information, we often want to know how the incoming energy varies with frequency. We can specify this by the *spectral flux density*, or *monochromatic flux* F_ν, defined such that $F_\nu d\nu$ is the flux in the infinitesimal range ν to $\nu + d\nu$, so that F_ν has units of $\mathrm{W\,m^{-2}\,Hz^{-1}}$. (Often you will see such measurements quoted in Jansky units, where $1\,\mathrm{Jy} = 10^{-26}\,\mathrm{W\,m^{-2}\,Hz^{-1}}$.) We could likewise define F_λ in units of $\mathrm{W\,m^{-2}\ m^{-1}}$, or $\mathrm{W\,m^{-2}\,\mu m^{-1}}$ and so on. Because $c = \nu\lambda$ we can swap easily between them by noting that, ignoring signs,

$$F_\lambda = F_\nu \frac{d\nu}{d\lambda} = F_\nu \frac{c}{\lambda^2} = F_\nu \frac{\nu^2}{c}.$$

In practice we estimate the spectral flux density by measuring the flux F through some restricted range of frequencies $\Delta\nu$ centred on ν and so estimating $F_\nu = F/\Delta\nu$. If this is a narrow range, for example as is the case for individual pixels in a dispersive spectrograph, the estimate will be accurate. More often, what we measure in practice is neither a bolometric flux nor a monochromatic flux, but the signal through a *broad bandpass*, set by the range of frequencies or photon energies your detector is sensitive to, or the range of frequencies allowed through by some kind of filter in the system. Interpreting broad-band fluxes can be a subtle problem, requiring for example a knowledge of how the sensitivity of the specific measurement system varies with frequency. Sometimes therefore the practice is to leave the measurements in a "native" form – counts per second when observed with that specific setup – provide data on the instrument performance, and leave the astronomer to interpret the measurements. Alternatively, the observer can try to "correct" their measurement as if it came from a *standardised passband*. For example the (blue) B-band flux F_B is that measured through a filter that, approximately, lets through light centred at $\lambda_B = 440\,\mathrm{nm}$, within a range of $\Delta\lambda \sim 95\,\mathrm{nm}$. (See Chap. 5 for more detail.) Likewise the (visual) V-band flux F_V is centred at 550 nm and the (red) R-band flux F_R at 640 nm, in similarly broad wavelength ranges. One can then use these broad band fluxes to calculate rough estimates of monochromatic fluxes, e.g. $F_{\lambda_B} \sim F_B/\Delta\lambda_B$. Alternatively, they can be left as fluxes on the B-band system etc, for more careful interpretation. They are then often quoted as *magnitudes*, as explained in the next section. The use of broad passbands for estimating the spectral properties of astronomical objects is examined a little more closely in Chap. 5.

1.2.3 Magnitudes

Magnitudes are a relative measure of brightness, expressed logarithmically. A *logarithmic* scale makes sense because of the extremely large range of fluxes seen, from the Sun to the most distant high redshift galaxy. Furthermore, many effects, such as the absorption of light through the atmosphere, which we will examine in Chap. 2, have an exponential dependence, which makes working in logarithmic terms mathematically simpler. A *relative* scale makes sense, because as we shall discuss later in this chapter, it is often very hard to get reliable absolute measurements, while the ratio of the brightnesses of two objects may be accurately measured. Furthermore, if we are making broad-band measurements, as discussed in the previous section, converting from the measured signal to absolute units is non-trivial.

One could work in *dex*, so that 1 dex = a factor 10, 2 dex = a factor 100 and so on. However, historically, stars were assigned "magnitudes" such that the brightest stars were "first magnitude" and the faintest stars visible to the naked eye were "sixth magnitude". Now that we have physical measurements of the stars, we know that this corresponds to approximately a factor of 100 in brightness. The modern convention is therefore to *define* a factor of 100 as a difference of 5 magnitudes. In other words, if we have measured the fluxes of two objects A and B as F_A and F_B, then their *magnitude difference* is

$$\Delta m = -2.5 \times \log_{10}(F_A/F_B)\,.$$

Note that the minus sign means that a larger magnitude is a fainter object, in accordance with tradition. The idea of magnitude scales can be applied to many "magnitude systems" so that for instance if we have measured fluxes through the standard B-band, we get the difference in B-magnitudes, ΔB.

To turn this scale of differences into an absolute system, we need a *zeropoint*. The tradition in optical astronomy is to take the star Vega as having zero magnitude in all the standard bands, i.e. B = 0, V = 0, R = 0 etc. If an astronomer measures a star to be 12.4 magnitudes fainter than Vega in the standard B-passband, then we say that it has B = 12.4 in the "Vega system". How do we turn this into a statement in ordinary physics units? As described in the previous section, this is actually a somewhat ambiguous problem, but we can at least turn standard broad-band magnitudes into rough estimates of monochromatic flux densities. If for example we measure the monochromatic flux of Vega at the centre of the B-band, we get $F_{440\,\mathrm{nm}} = 4630\,\mathrm{Jy}$, and we can then use this as the *zeropoint flux* F_{B_0}. This *calibration* then allows us to calculate approximate monochromatic fluxes for objects from their magnitudes in the Vega system:

$$F_B\,(\text{estimated}) = F_{B_0} \times 10^{-B/2.5},$$

where F_{B_0} is the zeropoint, F_B is the (monochromatic) flux of the object in the B-band, and B is the B-magnitude of the object. Note that we have taken the fundamental thing as being the B-magnitude, tied to Vega, so that the zeropoints are different for each passband, and the monochromatic fluxes are seen as estimates. An alternative approach is to take the monochromatic fluxes as fundamental, to pick a universal zeropoint F_0, and then to *define* a magnitude system which can be applied at any frequency such that

$$m_{AB} = 2.5 \times \log_{10}\,(F_\nu/F_0)\,.$$

For historical reasons, this is known as the "AB magnitude system", and has F_0 defined to be 3,631 Jy (Oke (1974)). The AB system is really just a way of expressing normal physical units in magnitude terms. In optical astronomy, it is gradually taking over from traditional magnitudes – when reading astronomical papers, you have to be careful to spot whether magnitudes are being quoted in the AB system or the Vega system.

Notably, magnitudes are commonly used in optical and IR astronomy, and almost never used in radio and X-ray astronomy. This is partly because the former are the "old fashioned" bits of astronomy, but it is also because these are the wavelength regions where we have the most problem with the atmosphere, making absolute calibration particularly hard (See Chap. 2).

1.3 Loss of Signal

In any astronomical measurement, we always lose some signal, typically in a series of steps – through the atmosphere, through reflections, through inefficiencies of the detector, and so on. We will first step through an example in some detail, and then consider how we can recover from those losses to arrive at an estimate of the original signal.

1.3.1 The Chain of Losses

Consider a star with true spectral flux density F_ν being measured by a CCD camera on a ground-based optical telescope. (We will discuss how CCDs work in Chap. 4). Some of the light is absorbed while travelling through the atmosphere. If the frequency dependent atmospheric transmission is T_ν, and the collecting area of the telescope is A, then the total flux arriving at the telescope aperture is $AT_\nu F_\nu$. The light bounces off various mirrors until it arrives at the entrance window of the CCD camera. Each reflection absorbs some of the light, mirror imperfections scatter light out of the beam, and various telescope supporting structures cause shadowing, losing more light. These effects could all be summed up in a telescope efficiency E_ν. The camera will also have its own optics with similar losses, which we could express as camera efficiency C_ν.

To make a standard astronomical measurement through a particular passband we would typically pass the light through a filter. This filter would have a transmission f_ν which may be close to unity near the central frequency, and fall off smoothly either side. Likewise the CCD itself will have a detection efficiency D_ν which depends on frequency. The signal finally seen by the detector is then a broad-band signal, after integrating over those frequency dependent losses:

$$K = A \times \int F_\nu T_\nu E_\nu C_\nu f_\nu D_\nu d\nu.$$

The CCD collects charge over time t (see Chap. 4) but when we read out the detector what we actually measure is a voltage V which depends on some conversion factor G (usually known as the "gain") which is different for different devices. Finally then what we really get is

$$V = G \times A \times t \times \int F_\nu T_\nu E_\nu C_\nu f_\nu D_\nu d\nu.$$

For different types of astronomical observation – an X-ray image taken in space, a radio interferometry observation made with phase-sensitive receivers etc – the details will be different, but the general principle will be the same. Between the light arriving at Earth and a final signal measurement taking place, there will be a substantial chain of physical effects in which light will be lost in a frequency dependent manner.

1.3.2 Calibrating the Signal

In principle we could try to understand the physics of all those effects, calculate the losses caused, arrive at values for all the factors E_ν, D_ν etc, and then correct for them. We could make this job a little easier if we make simplifying assumptions. For example we could assume that all the frequency dependent items are constant over the small range of frequency that the filter lets through, and approximate the filter passband itself as a top hat with height f and width $\Delta\nu$. Then

$$V = F_\nu \times \Delta\nu \times f \times A \times t \times T \times E \times C \times D \times G.$$

In practice, we know the loss factors rather imperfectly, and furthermore they could change with time – the transmission of the atmosphere may be different on different nights; the scattering loss from the telescope mirrors may get worse if they become dusty; and the detector may gradually degrade. Astronomers therefore tend to rely on empirical calibration. This comes in three stages.

(i) In the lab. The components of a system – the mirrors, the filters, the detector – can be measured in the lab before the whole system is taken to the mountain observatory, or the launchpad. For example, a light source of known properties can be shone onto the system. However, many things are only measurable once the whole system is integrated and working in situ.

(ii) At the start of mission. Before any scientific observations start, there is normally an extended period of testing, adjusting, and observing a wide range of cosmic objects of known properties, or internal calibration sources. In principle, this should tell us for example the end-to-end efficiency of the system. Typically this needs to be done for several different operational modes. These calibration factors can then be applied to scientific observations. Some things might slowly change – for example detectors may degrade due to cosmic ray bombardment – so periodical re-calibration campaigns will be undertaken.

(iii) Every night. For ground based telescopes, the atmosphere can change significantly from night to night and indeed from hour to hour. Photometric calibration is therefore achieved by observing a *standard star* of known flux $F_{\rm std}$. If this is observed close in time and close on the sky to our object, then our object flux is

$$F_{\rm obj} = F_{\rm std} \times \frac{V_{\rm obj}}{V_{\rm std}},$$

where V is the actual measured voltage (or similar) in each case. If the standard star is observed in a different part of the sky, then we would have to take into account how, on that particular night, the atmospheric absorption varies with elevation, as explained further in Chap. 2.

1.4 Distortion and Smearing

As well as losing some of the signal, astronomical measurement systems *distort* it. This can happen in two ways. It could be a *simple distortion*, which can be calibrated, so that we can recover the original signal. Alternatively, it could involve *smearing*, which is a kind of probabilistic scrambling of the signal. In this case we cannot reliably recover the original signal. We will look at these two kinds of effect in turn.

1.4.1 Simple Distortion

Collecting and detecting signals from the sky always involves some kind of *mapping* from a structure on the sky to a structure on the detector. For example, the pattern of stars at angular positions on the sky θ, ϕ, after being focused and bounced around various mirrors, ends up making a pattern of signal as a function of x, y position on our detector. What is the net mapping function $(\theta, \phi) \rightarrow (x, y)$? Likewise, the spectrum of a star as a function of wavelength λ, after passing through the spectrograph, maps into position on the detector, $\lambda \rightarrow x$. One more example is the way that the sensitivity of a detector may vary across its surface, producing apparent structure that is not there in the sky, so that we measure $V(x, y) = R(x, y) \times V_{expected}(x.y)$ where $R(x, y)$ describes how the relative sensitivity varies from spot to spot.

If we are lucky, the distortion mapping may be linear, and so described by a single scaling factor. In our imaging example, this would mean that the distance between any two points on the image, Δx, is related to the distance $\Delta\theta$ between the corresponding points on the sky by

$$\Delta x = p\Delta\theta.$$

Here the scaling constant p is known for historical reasons as the *plate scale*, and is normally expressed in units of arcsec/mm. However, in practice the mapping is not linear, especially for wide field systems. To calibrate this, one needs some kind of simplified model of the distortion. For example, in some optical systems the distortion is largely radial, and can be reasonably modelled as a low order polynomial, so that distance from the centre of the detector is given by

$$r = p_1\theta + p_2\theta^2,$$

and so needs two constants to describe the distortion. A positive value of p_2 gives the classic "pincushion" distortion, whereas a negative value of p_2 gives the "barrel" distortion. (See Fig. 1.2.) Similarly, the mapping $x(\lambda)$ in a spectrograph is normally modelled as a polynomial.

1.4.2 Calibrating Distortions

In general astronomers refer to the distortions caused by the various components of a measurement system as the *instrumental signature*. The aim of calibration is remove the instrumental signature and so solve for the true properties of the sky. Just as with photometric calibration, the instrumental signature may be largely determined in advance, in the lab, or in a calibration campaign, but for ground based telescopes nightly calibrations are also wise. This is because ground based telescopes are

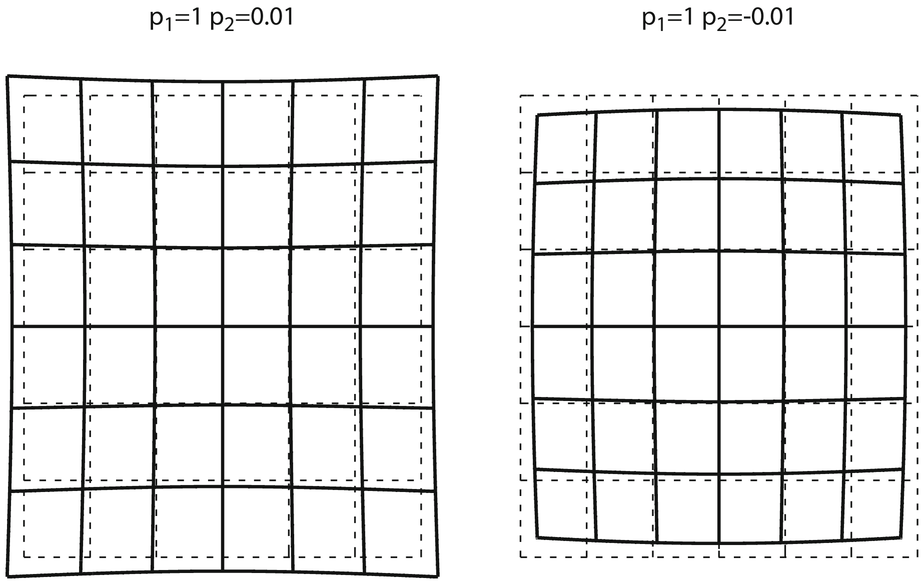

Fig. 1.2 Illustrating what optical distortions do to a regular grid. On the left is an example of a pincushion distortion, where the distortion is purely radial, and the second component p_2 is positive; on the right is barrel distortion, where p_2 is negative

very large, and operating in a gravitational field. There are therefore significant problems of mechanical flexure, including hysteresis, which means that it is highly desirable to measure distortions on the spot, with exactly the same configuration as the scientific observation.

Measuring the positions of stars is known as *astrometry*, as opposed to measuring their brightness, which is *photometry*. Just as with photometric calibration, astrometric calibration involves standard stars which have accurately known positions in advance. The constants p_1 and p_2 (or any other constants in a more complicated model of the distortion) are found by taking an image containing a network of such astrometric standards, measuring their observed x, y positions, and fitting the model function to the data. Calibrating a spectrograph involves shining light of known wavelength through the same optical system; this normally involves the use of an "arc lamp" which produces many spectral lines at known wavelengths. Spatial sensitivity variations in a measurement system are calibrated by obtaining an image of a source of light known to be spatially uniform, such as the twilight sky, or a white patch on the inside of the telescope dome, normalising this to unity, and dividing observed images of the night sky by this normalised image. This is known as *flat-fielding*.

1.4.3 Smearing

In some cases it is not possibly to uniquely remove the instrumental signature. This is because the detection process is intrinsically probabilistic in nature, smearing out the original signal. A familiar example is image convolution. The true map of brightness is usually spatially smeared – for example by seeing effects in the atmosphere, by the telescope optics, or by the antenna pattern of a radio dish. Stars are so far away that they should produce tiny pinpricks of light. A star with the same radius as the Sun ($\sim 7 \times 10^8$ m) at distance of 10 light years ($\sim 10^{17}$ m) should have apparent diameter 0.003″, but the smearing produces a blurry image three orders of magnitude larger than this. Another example is the smearing that occurs in an optical spectrograph due to diffraction as the light passes through the slit, so that even an infinitely sharp spectral line produces a measured line with finite width. A third example is the way that X-ray spectra are measured. We will look at this example carefully in order to understand the principles of how smearing occurs.

As we shall see later, X-ray detectors are capable of estimating the energy of individual photons that they detect, which enables one to measure the "energy spectrum" of a source. Each photon makes a pulse of some size, and the pulses are binned into channels. However a photon of given energy E could end up in many different possible energy channels C, centred on the "correct" channel, but with a broad spread. Figure 1.3 illustrates how photons of slightly different energies E_A and E_B then produce overlapping spreads; the distribution of photon energies becomes scrambled in the distribution of observed channel counts.

For a given channel C let us write the probability that a photon of energy E will end up in that channel as $P(C, E)$. This 2D function is known as the "detector response matrix". Suppose that the true energy spectrum – the number of photons per second per unit photon energy – arriving at the detector is $F(E)$ so that the number of photons in energy range E to $E + dE$ is $F(E)dE$. These detected photons are allocated to channels with probability $P(C, E)$. The observed distribution of counts in the various channels, known as the "count spectrum" is then

$$N(C) = \int_{E=0}^{E=\infty} F(E)P(C, E)dE,$$

i.e. the observed count spectrum is a *convolution* of the true spectrum with the detector response. The problem is that even if we know the detector response matrix $P(C, E)$ very well, we cannot uniquely invert such an integral to recover the true spectrum. (It is possible sometimes to derive a statistically reasonable possible inversion – see Sect. 1.4.5 and Appendix A.)

1.4.4 Calibrating Smearing: Resolution

Even if we cannot remove the smearing, we can at least characterise and quantify it. For an X-ray spectrum, we want the detector response matrix $P(C, E)$. Likewise image smearing is characterised by the *point spread function (PSF)* $P(\theta, \phi)$, i.e. the pattern of light that a point source would make. As usual with calibration issues, we could try to use our knowledge of the measurement system to predict these

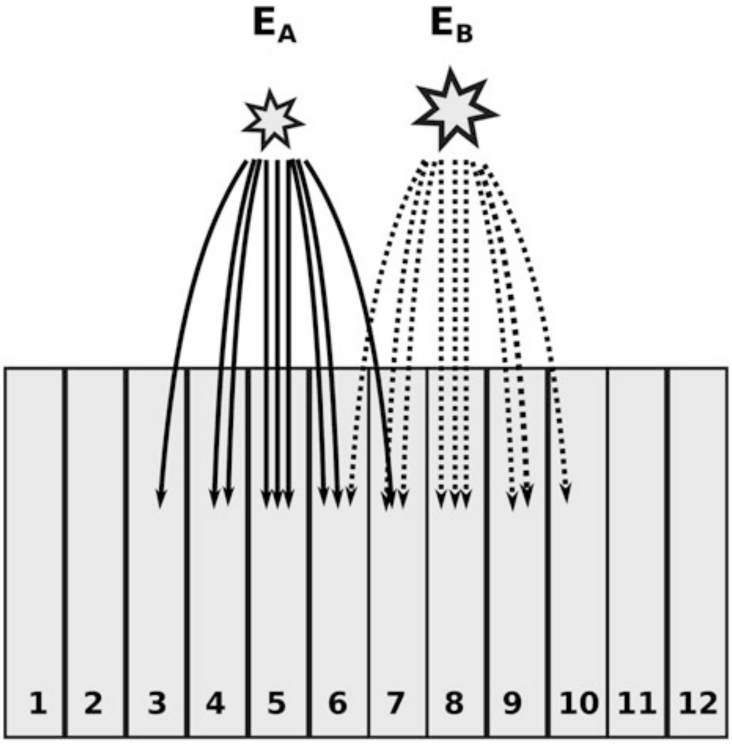

Fig. 1.3 Illustration of spectral convolution in an X-ray detector. Two streams of photons with different energies land on the detector. Those with energy E_A on average cause a pulse in channel 5, but with some chance of causing a pulse 4 or 3 etc. Likewise photons with energy E_B on average cause a pulse in channel 8, but with a probabilistic spread into other channels. For any given pulse in say channel 7, we cannot tell whether it was caused by a photon of type A or type B. If we know how many photons of the two types A and B are arriving, we can predict the distribution of counts in the channels; but we cannot go the other way; i.e. if we know the counts in the various channels, we cannot unambiguously deduce the number of photons of types A and B

smearing functions, but more often we use an empirical calibration – for example a bright star in our image will give us the PSF accurately. Often but not always the PSF is more or less azimuthally symmetric so it is well approximated by a one-dimensional function, $P(\theta)$. Very often, but not always, because smearing functions are probabilistic in nature, they are well approximated by a Gaussian distribution

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right],$$

where μ is the mean of the distribution and σ is the standard deviation. (See Appendix A for a brief review of statistics and statistical functions.)

Very often we simply want to quantify how bad the smearing is, by characterising the width of the smearing function in some way. We will then know that scales smaller than this width do not hold useful information – they have been completely scrambled – whereas larger scales are reliable. Such a characteristic smearing width is referred to as the *resolution* of the image or spectrum of whatever. The most

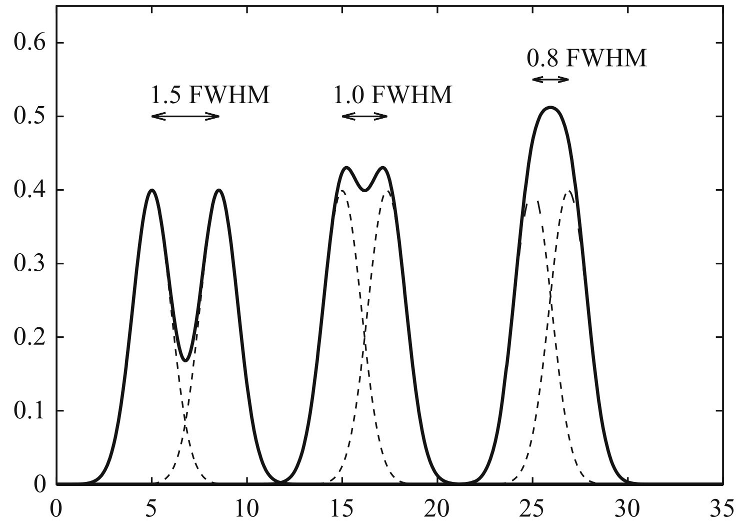

Fig. 1.4 This figure illustrates the dependence of resolution on separation compared to the smearing function width. In this example, all the curves are Gaussians with a standard deviation $\sigma = 1$. Note that for a Gaussian, FWHM = 2.354σ

usual way to do this is to quote the Full Width at Half Maximum (FWHM) of the smearing function. For the case of a Gaussian smearing function, FWHM = 2.354σ. Note however that smearing functions are not *always* Gaussian.

For a given quantity X, as well as specifying the resolution ΔX, we could specify the *resolving power* $R = X/\Delta X$ where X is taken as a characteristic value. This is most commonly used in spectroscopy, where $R = \lambda/\Delta\lambda$, as it allows us to compare how good our discrimination is at quite different wavelengths. For example, as we shall discuss in Chap. 5, simple broad band filters in the optical, at $\lambda \sim 500$ nm, have a width of $\Delta\lambda \sim 100$ nm, giving a resolving power of $R \sim 5$. Dispersive spectrographs using gratings on the other hand achieve a resolving power of thousands. An energy-resolving X-ray detector, working at $E \sim 1$ keV, which corresponds to $\lambda \sim 1.24$ nm might have an effective $\Delta E \sim 50$ eV. In wavelength terms, this corresponds to $\Delta\lambda \sim 0.06$ nm, much smaller than the width of a typical optical filter, $\Delta\lambda \sim 100$ nm. The *resolution* of the X-ray detector is therefore much smaller than the optical filter, but its *resolving power* is about the same, as $E/\Delta E = \lambda/\Delta\lambda \sim 20$.

Figure 1.4 illustrates graphically why FWHM is a good guide to resolvability. Consider two stars in an image that very are close together – can we separate them? The figure shows what happens if we add two Gaussian shapes. When the separation is greater than the FWHM, the two stars are clearly resolved. For separations less than the FWHM, we cannot distinguish two separate peaks. The resulting image is a little broader than the PSF, but you wouldn't know whether the broadening was because the underlying object was extended, or because it was the sum of two close objects, or the sum of three etc.

1.4.5 Calibrating Smearing: Modelling

Although we cannot go backwards from the smeared signal to the original signal, we can go forwards to test whether a possible signal is consistent with the measurements. Taking the X-ray spectrum example, one can hypothesise a spectral shape, calculate the count spectrum that would result if such a spectrum were to be observed, and see if this agrees with the observed data. For example, we could hypothesise that the spectrum is of the form:

$$F(E) = KE^{-\gamma}.$$

The usual procedure is to try many different values of the two parameters K and γ, and in each case to calculate the predicted count spectrum, and then calculate the goodness of fit of this prediction when compared to the data and its errors, for example using the χ^2 statistic (see Appendix A). Then one searches for the values of K and γ which give the best fit.

Is this laborious process really necessary? Suppose we have estimated the FWHM of the energy smearing function, ΔE. The *spectral resolving power* is then $R = E/\Delta E$. If the resolving power is large (i.e. the relative energy-smearing is small), say, $R > 100$, then the count spectrum is a reasonable approximation to the true spectrum, and the above model fitting approach is unnecessary. However many X-ray detectors have low resolving power, $R \sim 5$, and so the model fitting approach is crucial. Likewise, in an image smeared by atmospheric seeing (see Chap. 3) of size $\Delta\theta$, if we are considering structure on scales $\gg \Delta\theta$ then the smearing doesn't matter. However, two close stars in an image may be just on the edge of resolution, or one may not be sure whether there are two objects or three. In such circumstances, one can fit a model, blurring with the PSF, to decide what is reasonable to believe. This is always a statistical matter; one can only attach confidence levels to various hypotheses. (See Appendix A.)

Suppose we postulate a model of the sky which is not some mathematical function, but just a collection of gridded brightness values, on a grid scale smaller than the smearing scale? We could in principle run through huge numbers of such gridded images, perhaps applying some generic smoothness constraint, and although many different image-grids would be statistically consistent with the measured data, we could find the one that makes the observed data most likely to have happened. This is the idea behind *image sharpening*. This technique can give interesting results, but ones that have to be treated with great caution.

1.5 Noise and Uncertainty

Every measurement has an uncertainty, arising for many different reasons. We will be concerned here with uncertainty arising in the measurement process, which may be thought of as the process of combining signal with noise. If the true signal is S and we attempt to measure it many times, then we will find many different values $M = S + P$ where P is a random variable, usually (but not always) with mean zero, and often (but not always) with a Gaussian distribution. Then we can characterise the noise added to the signal as the standard deviation σ of the random variable P.

Careful treatment of noise is very important in designing instruments, in planning observations, and in understanding the results of our measurements. It is useful to distinguish two main types of noise.

(a) *Intrinsic noise* is inescapable because light comes in discrete packets – photons – and so is subject to counting statistics. (See Appendix A for a reminder.) For example, if the flux arriving at the detector is F, with a mean frequency ν, then in integration time t, the number of photons collected will be $N = Ft/h\nu$. This number follows a Poisson distribution, so repeat experiments would give a range of values with standard deviation $\sigma = \sqrt{N}$.

(b) *Extrinsic noise* refers to random fluctuations imposed by the environment or by the measurement process, as opposed to being inherent in the signal. For example, in ground-based IR astronomy the light from astronomical objects sits on top of a very bright sky background, but this background is constantly changing, so that repeat measurements would give different values. Similarly in radio astronomy we measure an alternating current in an antenna induced by the incoming radio waves; but thermal processes in the electronics produce a randomly fluctuating current even in the absence of any signal. Another important example is how the sensitivity or gain of the detector may vary from spot to spot, producing a kind of graininess or *pattern noise*.

A key quantity is the *signal to noise ratio*, $r = S/\sigma$ where S is the net signal – counts, voltage, whatever – and σ is the standard deviation of the noise process. This determines not just the error on our measurement, but whether we have managed to measure a signal at all. If the true signal is zero, every so often we will see a large value just by chance. For example, if the noise distribution is Gaussian with standard deviation σ then the probability in one experiment of getting $r > 2$, i.e. a fake signal with $S = 2\sigma$, is 1/20. The probability of $r > 3$ is 1/370. So if we make a measurement and find $r = 1$ we certainly would not believe it to be real, but at $r = 3$ the signal is very likely real. However, one has to be very careful about exactly what question is being asked. Consider an X-ray image with 1,000 × 1,000 pixels, containing a uniform background with a pixel-to-pixel noise level σ. Although for any specific pixel the chance of a 3σ high spot is small, there will be $N = 1{,}000 \times 1{,}000/370 = 2{,}700$ such fake high pixels spread over the image. The question "is the signal at this spot real?" is quite different from the question "how many real sources are there in this image?".

1.5.1 Sensitivity in the Presence of Noise

What is the faintest object we can see? How fast can we detect a given object? To answer these related questions, one has to think about both signal and noise. Given a flux density F_ν, how does one increase the final signal? Recall this equation from Sect. 1.3.2:

$$S = F_\nu \times \Delta\nu \times f \times A \times t \times T \times E \times C \times D \times G,$$

where S is the signal measured. The first thing to do is to increase the collecting area – e.g. make a bigger telescope. The second thing is to increase the efficiency of the system – to reduce light losses in the telescope optics, camera, and detector. The third thing is to integrate for longer. However detection is also limited by noise; if the noise is too large compared to the signal, we will not be sure whether we have seen a real signal or not. If we pick a given signal to noise ratio r as our threshold of believability, we can ask how long an integration it takes to achieve this value of r, and how this scales with system parameters. This turns out to depend strongly on what kind of noise is dominating our measurement.

(i) *Photon limited case.* Simplifying somewhat, for a source with flux F the signal measured in counts is $N_S = FAEkt$ where A is collecting area, E represents all the efficiency factors, t is the integration time, and k is a factor which converts to photon rate, which will depend on the details of the system. If the source is much brighter than the background, and any extrinsic noise is comparatively small, then we have $\sigma \sim \sqrt{N_S}$. Note that noise increases with time, but more slowly than the signal. The time taken to reach signal to noise r is

$$t = \frac{1}{kAE}\frac{r^2}{F}.$$

We see that the dependence on collecting area and system efficiency is linear, as is the dependence on source flux F. To detect a source that is half as bright takes twice as long. However, the dependence on signal-to-noise ratio r is quadratic. If we wish say to achieve 10σ rather than 5σ, it will take four times as long.

(ii) *Background limited case.* If the source flux F sits on top of a background B, then the noise is given by the total counts $N = kAEt(F + B)$. If $B \gg F$ then $\sigma \sim \sqrt{kAEtB}$. To actually estimate F we have to make two measurements – one at a location containing the source, and another containing only background – and subtract. Each measurement has error σ, so following standard transmission of errors (see Appendix A), the difference estimate has $\sigma_{net}^2 = 2\sigma^2$. The signal is $N_S = kAEtF$ and the noise is $\sigma_{net} = \sqrt{2kAEtB}$ and so the time taken to reach signal to noise r is

$$t = \frac{1}{kAE}\frac{2r^2B}{F^2},$$

The dependences on signal-to-noise, collecting area, and efficiency are the same, but we see that it also pays to keep the external background as low as possible. We also see that time taken varies quadratically with F. In contrast to the photon limited case, if you want to detect something at half the flux density, it takes 4 times as long, and detecting things 10 times fainter takes a 100 times as long. When the sky background is brighter than the things you are looking for, detecting fainter things is *very* hard.

(iii) *Readout noise limited case.* For CCDs and related detectors, signal accumulates as charge during an integration (see Chap. 4), so that $S = aFt$ where a is some constant depending on the system. The process of reading out the charge at the end adds noise σ_R of a *fixed size*. Then the time taken to achieve signal to noise r is

$$t = \frac{1}{a}\frac{r\sigma_R}{F},$$

An aim of detector design will be to make σ_R as small as possible. Note that because the noise is of fixed size, taking many short exposures rather than one long one would be foolish. This contrasts with the photon limited case, where the net noise is the same for a single long exposure or many short exposures added together. A sensible aim might be to integrate long enough that the photon noise from the signal becomes bigger than the readout noise, so that the readout noise becomes unimportant.

(iv) *Fluctuation limited case.* This is often referred to as the detector noise limited case, but also applies to external fluctuations, such as erratic variations in sky background. Rather than integrating, we sample the signal repeatedly over time and average our estimates. (Some types of external fluctuation, loosely referred to as "1/f" noise are not statistically stationary, and so this averaging over time does not behave as you might expect.) The amplitude of the fluctuation noise will depend on the sampling timescale; suppose that for sampling time t_0 the noise is σ_D. Then if we integrate for time t we will make $N = t/t_0$ samples, so that when we average these the error on the mean is $\sigma^2 = \sigma_D^2 t_0/t$. So the time to reach signal to noise r is then

$$t = t_0 \frac{1}{b^2}\frac{r^2\sigma_D^2}{F^2}.$$

This is like the background limited case – the dependence on flux is quadratic, so that detecting fainter sources in the presence of detector noise is extremely hard. The background limited case and the fluctuation limited case differ markedly however in their behaviour with respect to spectral multiplexing, which will be explained in Chap. 5.

Overall, an aim of experimental design is to make readout noise, fluctuation noise, and background all as small as possible.

1.5.2 Resolution in the Presence of Noise

As well as determining your sensitivity limit, noise will determine your effective resolution. Figure 1.5 shows two marginally resolved Gaussians, with separation of the peaks equal to the FWHM of the smearing function, as in the middle example

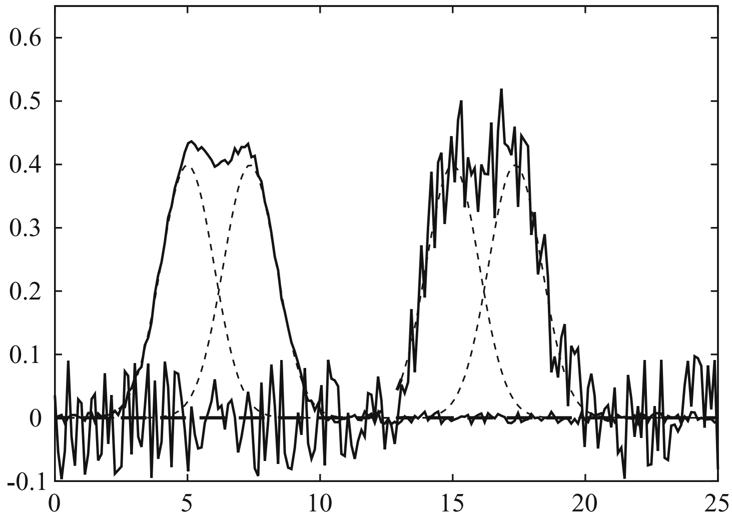

Fig. 1.5 Illustrating the effect of noise on resolution. Both examples show the sum of two Gaussians with $\sigma = 1$, separated by FWHM $= 2.354\sigma$, but with differing amounts of random noise added

of Fig. 1.4. In this case some random noise is added. When the noise is small, the two peaks can still be clearly seen in the sum of the two curves; however when the noise is increased information is lost and the peaks are not resolved. The effective resolution is significantly larger. As ever, improved information could be obtained by model fitting, but this is also limited by noise.

1.6 Qualities of Astronomical Measurement Systems

We have seen how generic problems occur repeatedly in different areas of astronomy – loss of light, unknown calibration factors, distortion of images and spectra, smearing of signals, and the presence of noise which masks the information we want. We will see more examples as we step through different examples of the measurement process. This gives us a way to characterise the strengths and weaknesses of different systems. There are five key questions.

(1) What is the effective collecting area of this system?
(2) What is its overall efficiency?
(3) What are the distortions introduced, and can we correct for them?
(4) What is the degree of statistical blurring introduced by this system?
(5) What are the sources of noise in this system?

1.7 Further Reading

Some of the issues covered briefly in this chapter, such as positional co-ordinate systems, radiation quantities, and photometric systems and magnitudes, are covered in a little more detail in basic astronomy textbooks such as Carroll and Ostlie (2006), and Karttunen et al. (2007). (Works cited here are detailed in the References section.) A good introduction to measurement quantities is in Tomey (2010). Tomey's book covers some of the same material as this book, but more heavily concentrated on optical astronomy.

For brushing up on optics, good general textbooks are by Hecht (2003), and Lipson et al. (2010). A slightly more advanced treatment also covering photonics is given by Graham Smith et al. (2007).

More advanced treatments of co-ordinate systems and astrometry are given in the review article by Johnston and de Vegt (1999), and the book by Perryman (2012). Photometric systems and magnitudes are covered in detail in the review articles by Bessell (2005) and Fukugita et al. (1996).

1.8 Exercises

1.1. A large single dish radio telescope has spatial resolution with FWHM$\sim$1 arcminute. How many separate radio sources could be distinguished over the whole sky?

1.2. The Hydrogen $H\alpha$ emission line has a wavelength of $\lambda = 656.3$ nm. In energy terms, estimate roughly how many $H\alpha$ photons equate to a golf ball in motion.

1.3. Show that the quantity νF_ν is proportional to the flux per decade of frequency. How does νF_ν compare to λF_λ? Suppose you plot on paper the spectral energy distribution of an astronomical source, using (i) F_ν vs ν, (ii) F_ν vs $\log \nu$, and (iii) νF_ν vs $\log \nu$. In which of these cases do equal areas on the piece of paper represent equal amounts of radiant energy?

1.4. The zero point of the B-band magnitude system is 4,260 Jy. Its central wavelength is 440 nm, and the width of a typical B-band filter is 97 nm. Roughly how many photons per sq.m. per second would we get from a star with B = 17.0? How many photons per second would be detected on a CCD on a 2 m diameter telescope, if the telescope optics have an efficiency of 80 %, the camera optics has efficiency 50 %, and the detector has efficiency 70 %? Ignoring any background or detector noise, what signal-to-noise ratio would be achieved in a 2 min integration?

1.5. An X-ray detector has resolving power $R = 10$ near photon energies with $E = 2$ keV, and the energy response function is Gaussian to a good approximation. If an incoming photon has true energy $E_t = 2.1$ keV, what is the probability that the detector will record it as having apparent energy $E_a > 2.5$ keV?

1.6. On a particular telescope, with a specific detector, a star with magnitude $B = 22.3$ takes 150 s to be detected at signal-to-noise of $r = 10$. The measurement is against a bright sky, such that the measurement is background limited. If a better detector is used, with twice the efficiency, how long would it take to detect an object with $B = 24.0$?

1.7. A star with flux F is being observed with a CCD camera on a large telescope. In order for the signal-to-noise of the measurement to be dominated by photon-counting rather than the CCD readout noise, the exposure has to be 5 s or more. With a better camera we could also use 5 s exposures on objects that were 10 times fainter. What options would we have on improving the CCD camera in order to achieve this aim?

References

Bessell, M.S.: Standard photometric systems. Ann. Rev. Astron. Astrophys. **43**, 293–336 (2005)
Carroll, B.W., Ostlie, D.A.: An Introduction to Modern Astrophysics, 2nd edn. Addison Wesley (2006)
Graham Smith, F., King, T.A., Wilkins, D.: Optics and Photonics, 2nd edn. Wiley (2007)
Hecht, E.: Optics, 4th edn. Addison Wesley (2003)
Johnston, K., de Vegt, C.: Reference frames in astronomy. Ann. Rev. Astron. Astrophys. **37**, 97–125 (1999)
Karttunen, H., Kröger, P., Oja, H., Poutanen, M., Donner, K.J.: Fundamental Astronomy, 5th edn. Springer (2007)
Lipson, A., Lipson, S.G., and Lipson, H.: Optical Physics, 4th edn. Cambridge University Press (2010)
Oke, J.B.: Absolute spectral energy distributions for white dwarfs. ApJSupp. **27**, 21 (1974)
Perryman, M.: Astronomical Applications of Astrometry – Ten Years of Exploitation of the Hipparcos Satellite Data. Cambridge University Press (2012)
Fukugita, M., Ichikawa, T., Gunn, J.E., Doi, M., Shimasaku, K., and Schneider, D.P.: The sloan digital sky survey photometric system. Astron. J. **111**, 1748–1756 (1996)
Tomey, F.R.: To Measure the Sky: An Introduction to Observational Astronomy. Cambridge University Press (2010)