Chapter 7
Matter and light

> Because atomic behavior is so unlike ordinary experience, it is very difficult to get used to, and it appears peculiar and mysterious to everyone – both to the novice and to the experienced physicist. Even experts do not understand it the way they would like to, and it is perfectly reasonable that they should not, because all of direct human experience and of human intuition applies to large objects.
>
> – Richard Feynman, *The Feynman Lectures on Physics*, 1965

Chapter 1 introduced the situations that produce line and continuous spectra as summarized by Kirchhoff's laws of spectrum analysis. This chapter descends to the microscopic level to examine the interaction between photons and atoms. We show how the quantum mechanical view accounts for Kirchhoff's laws, and how atomic and molecular structure determines the line spectra of gasses.

To understand modern astronomical detectors, we also turn to a quantum mechanical account – this time of the interaction between light and matter in the solid state. The discussion assumes you have had an introduction to quantum mechanics in a beginning college physics course. We will pay particular attention to some simple configurations of solids: the metal oxide semiconductor (MOS) capacitor, the p–n junction, the photo-emissive surface, and the superconducting Josephson junction. Each of these is the physical basis for a distinct class of astronomical detector.

7.1 Isolated atoms

7.1.1 Atomic energy levels

A low-density gas produces a line spectrum, either in absorption or emission, depending upon how the gas is illuminated (review Figure 1.7). The formation of lines is easiest to understand in a gas composed of single-atom molecules, like helium or atomic hydrogen. Consider the interaction between a single atom and a single photon. In either the absorption or the emission of a photon, the atom usually changes the state of one of its outermost electrons, which are therefore termed the ***optical electrons***. The same electrons are also called the

valence electrons, since they largely influence an atom's chemical properties by participating in covalent and ionic bonds with other atoms.

Observations of atomic spectra and the theory of quantum mechanics both demonstrate that the energy states available to any bound electron are ***quantized***. That is, an electron can only exist in certain permitted energy and angular-momentum states. In theory, these permitted states arise because an electron (or any other particle) is completely described by a ***wave function***. In the situation in which the electron is bound in the potential well created by the positive charge of an atomic nucleus, the electron's wave function undergoes constructive interference at particular energies, and destructive interference at all others. Since the square of the wave function gives the "probability density" of the electron existing at a certain location and time, the electron cannot have energies that cause the wave function to interfere destructively with itself and go to zero. Physicists call these ***forbidden*** states. In the isolated atom, most energies are forbidden, and the energies of the rare ***permitted*** states are sharply defined.

Figure 7.1a illustrates the permitted energy levels for a fictitious atom, which appear as horizontal lines. In the figure, energy, in units of electron-volts, increases vertically. (One electron-volt (eV) is the energy gained by an electron accelerated by a potential difference of one volt. 1 eV = 1.6022×10^{-19} J.) There are seven ***bound states***, labeled a – g, . . . , in this particular atom. (Real atoms have an infinite number of discrete states – see below.) These different energy levels correspond to different configurations or interactions of the outer electrons. The idea here is that the atom must exist in one of these permitted energy states. The lowest energy state, the one assigned the most negative energy, is called the ***ground state*** (level a in the figure). This is the configuration in which the electrons are most tightly bound to the nucleus,

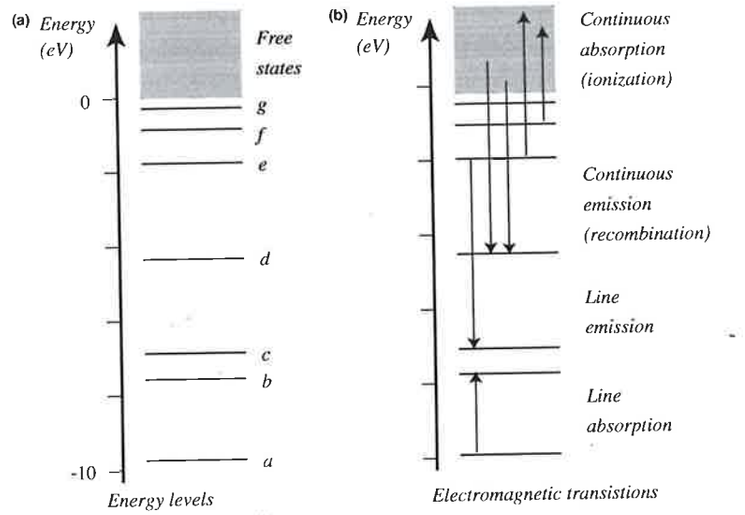


Fig. 7.1 (a) Permitted energy levels for an electron in a hypothetical atom that has seven bound states (a–g). The most tightly bound states (lowest energy) correspond to an electron location closer to the nucleus. (b) Absorption or emission of photons. The probabilities of different transitions can be vastly different from one another.

and would be the state of an undisturbed atom at zero temperature. Above the ground state are all other permitted ***excited*** states, up to the ***ionization level***. The ionization level, conventionally assigned zero energy, corresponds to an atom that has so much internal energy that an electron is just able to escape. In that situation, the free electron is no longer part of the atom, and the remaining positive ion will have internal energy states described by a completely different diagram. Because the wave function of the free electron in a vacuum is not restricted by variations in potential energy, the energy of the free electron is not quantized.

You can think of bound states with higher energies as situations in which the optical electrons are *on average* physically further away from the nucleus. Be aware, though, that the vision of electrons orbiting the nucleus like planets in the Solar System (i.e. the early Bohr theory) is limited in its usefulness. The best answer to the question "where is this electron?" is a function that says certain locations are more likely than others, but, unlike the energy situation, a rather broad and sometimes complicated range of positions is possible for each bound state.

7.1.2 Absorption of light by atoms

Even though we can't see the positions of an atom's electrons, we *can* measure their energies when light interacts with atoms. Remember that a photon carries a specific amount of energy that is directly proportional to its frequency, ν:

$$E = h\nu = \frac{hc}{\lambda} \tag{7.1}$$

The atom can make a transition from one bound state to another by either absorbing (the process is called ***photo-excitation***) or emitting a photon of the correct frequency or wavelength, as illustrated in Figure 7.1b. In the process of photo-excitation, the photon is truly absorbed, and ceases to exist. The figure shows a photo-excitation transition from the ground state (level a, which has energy E_a) to the first excited state (level b, which has energy E_b). The photon responsible for this transition must have wavelength

$$\lambda_{ab} = \frac{hc}{\Delta E_{ab}}$$

where

$$\Delta E_{ab} = E_b - E_a$$

This explains why a beam of light with a continuous spectrum that passes through an atomic gas will emerge exhibiting an absorption line spectrum. Since only photons with energies corresponding to the energy difference between bound electron states, ΔE_{ij}, can be absorbed, only lines with the corresponding wavelengths of λ_{ij} will be present as absorption features in the spectrum that emerges.

As Figure 7.1 illustrates, photons capable of *ionizing* the atom can have any wavelength, so long as they are energetic enough to move an electron from a bound to a free state. This is observed in the spectrum as a feature called an ***absorption edge*** – a drop in the intensity of the transmitted continuum at wavelengths shorter than the ionization wavelength.

7.1.3 Emission of light by atoms

An isolated hot gas produces an emission line spectrum. Again, you can understand why by considering the quantized energy levels. In Figure 7.1, for example, an atom changing from state e to state c must lose energy. It can do so by creating a photon with energy ΔE_{ec}. This process of ***de-excitation*** by photoemission can occur spontaneously, or it can be stimulated to occur by an incoming photon of exactly the transition energy. This latter process is the equivalent of negative absorption: one photon collides with the atom and two identical photons emerge. Stimulated emission is the basis for the operation of lasers and masers.

If there is a significant number of free electrons in a hot gas, then the gas will emit continuous radiation along with the usual emission lines. As illustrated in Figure 7.1, a photon is emitted if a free electron loses energy and recombines with a positive ion, forming the bound state of the neutral atom. The resulting radiation will be continuous, since the energy of the free electron is not quantized. Transitions from one free state to another are also possible, and will also contribute to a continuous spectrum.

7.1.4 Collisions and thermal excitation

Atoms prefer to exist in the lowest possible energy state, the ground state. An isolated atom in any excited state will spontaneously decay to a lower state. The length of time an atom can expect to remain in a particular excited state depends on the rules of quantum mechanics, but if there is a quantum-mechanically "permitted" transition to a lower state, the half-life of the excited state usually is on the order of 10^{-8} seconds. How do atoms get into an excited state in the first place? One way, of course, is by absorbing electromagnetic radiation of the proper wavelength. A second path is via collisions with other atoms or particles. Atom-on-atom collisions can convert kinetic energy into internal energy in the form of optical electrons in excited states. In the very eventful environment of a hot gas, atoms that want to stay in the ground state have little chance of doing so for long, because they are kicked up into higher states by collisions. A hot gas glows because the resulting excited atoms will decay back to lower energy levels, emitting photons in the process.

Collisions can transfer energy out of an atom as well as into it. With many collisions, at constant temperature, the population and de-population rates for

one level due to all processes are equal, and the expected number of atoms in a particular bound state is well defined. The ***Boltzmann distribution*** describes the number of atoms in each energy state in such a situation of ***thermodynamic equilibrium***. Consider any two bound states, i and j, having energies E_i and E_j. The Boltzmann equation gives the ratio of the number of atoms in these two states as

$$\frac{n_i}{n_j} = \frac{g_i}{g_j}\exp\left\{\frac{E_j - E_i}{kT}\right\} \tag{7.2}$$

Here g_i and g_j are the ***statistical weights*** of each level (g is the number of distinct quantum mechanical states at the specified energy – see the next section). Boltzmann's constant, k, has the value 1.381×10^{-23} J K^{-1} $= 8.62 \times 10^{-5}$ eV K^{-1}.

7.1.5. Specification of energy levels

In the terminology of quantum mechanics, the state of every bound electron is specified by four ***quantum numbers***:

n, the principal quantum number, can take on all positive integer values 1, 2, 3,.... This number is associated with the radial distribution of the probability density of the electron as well as with its energy, and in the terminology used by chemists, specifies the ***shell***.

l, the azimuthal quantum number, can take on values $0, 1, \ldots, (n-1)$. It can be associated with the angular distribution of the probability density, and can have a secondary effect on the energy of the state.

m, the magnetic quantum number, can take on values $0, \pm 1, \ldots, \pm l$. It describes the possible interaction between the electron and an imposed magnetic field. It can have an effect on the energy of the electron only if a field is present.

s, the electron spin quantum number can have only two values, $+1/2$ or $-1/2$. It can affect the electron energy by interacting with the angular momenta of other parts of the atom.

In particle physics, a ***fermion*** is a particle like the electron, proton or neutron, whose spin quantum number has a half-integer value like $\pm 1/2$, $\pm 3/2$, etc. Any particle's intrinsic angular momentum has the value $(h/2\pi)\sqrt{s(s+1)}$, where h is Planck's constant and s is the spin quantum number. Particles with integer spin (0, ± 1, etc.) are called ***bosons***.

The ***Pauli exclusion principle*** states that no two identical fermions may occupy the same quantum state. This demands that ***no two electrons bound in an atom may have the same four quantum numbers (n, l, m, s)***. Table 7.1 lists all possible values of the four quantum numbers for electrons in the first few levels. Each of the states listed must be either empty or occupied by a single electron. The ground state of an atom with atomic number Z will have the lowest-energy configurations occupied, up to the Zth available state, and all other states empty. The actual energy of a particular state depends not only on the atomic number and the values of the four quantum numbers for the occupied states, but also on other details like the atomic weight, and magnetic interactions between the electron, nucleus, and electrons in other states.

Table 7.1. *Quantum numbers of the first 30 bound atomic states (up to the ground state of zinc). In the periodic table, the 4s states are usually filled before the 3d states, 5s before 4d, etc. See Figure 7.2*

Quantum numbers					
n	l	m	s	Name of configuration	Number of states
1	0	0	±1/2	1s	2
2	0	0	±1/2	2s	2
2	1	−1	±1/2		
		0	±1/2	2p	6
		+1	±1/2		
3	0	0	±1/2	3s	2
3	1	−1	±1/2		
		0	±1/2	3p	6
		+1	±1/2		
3	2	−2	±1/2		
		−1	±1/2		
		0	±1/2	3d	10
		+1	±1/2		
		+2	±1/2		
4	0	0	±1/2	4s	2

The energy of an electron will depend most strongly upon both n and l quantum numbers. The ***configuration*** of electrons in an atom is therefore usually described by giving these two numbers plus the number of electrons in that n, l level. The ***spectroscopic notation*** for a configuration has the form:

$$ny^x$$

where

n is the principle quantum number,

x is the number of electrons in the level – many electrons can have the same n, l so long as they have different m and/or s values, and

y codes the value of the l quantum number according to the following scheme:

l	0	1	2	3	4	5	6	7, 8, ...
Designation	s	p	d	f	g	h	i	k, l, etc.

Lithium, atomic number 3, for example, has the ground-state configuration $1s^2\ 2s^1$; that is, two electrons in the $n = 1$ state, one with quantum numbers (1, 0, 0, –1/2), the other with (1, 0, 0, 1/2). The third lithium electron (this is the valence electron) is in the $n = 2$ level with quantum numbers either (2, 0, 0, –1/2) or (2, 0, 0, 1/2). Table 7.2 gives some further examples of electron configurations.

Figure 7.2 is a schematic energy-level diagram that shows the relative energies of the electron configurations in atoms. As one moves from element to element in order of increasing atomic number, electrons are added from the bottom up in the order suggested by Figure 7.2. (There are minor exceptions.)

The ***periodic table***, one of the triumphs of human learning, summarizes our knowledge of the chemical properties of the elements, and recognizes that chemical behavior is periodic in atomic number. The table is organized according to similarities in optical electron configurations. Each row or period contains elements with identical values of n for outer electrons. In chemical terminology, the valence electrons of atoms in the same period are all in the same ***shell***. The

Table 7.2. *Examples of a few electron configurations*

Element	Atomic number	Electron configuration of the ground state
Hydrogen	1	$1s^1$
Helium	2	$1s^2$
Boron	5	$1s^2\ 2s^2\ 2p^1$
Neon	10	$1s^2\ 2s^2\ 2p^6$
Silicon	14	$1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^2$
Argon	18	$1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6$ = [Ar]
Potassium	19	$1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 4s^1$ = [Ar] $4s^1$
Scandium	21	[Ar] $3d^1\ 4s^2$
Germanium	32	[Ar]$3d^{10}\ 4s^2\ 4p^2$
Krypton	36	[Ar]$3d^{10}\ 4s^2\ 4p^6$ = [Kr]
Rubidium	37	[Kr] $5s^1$

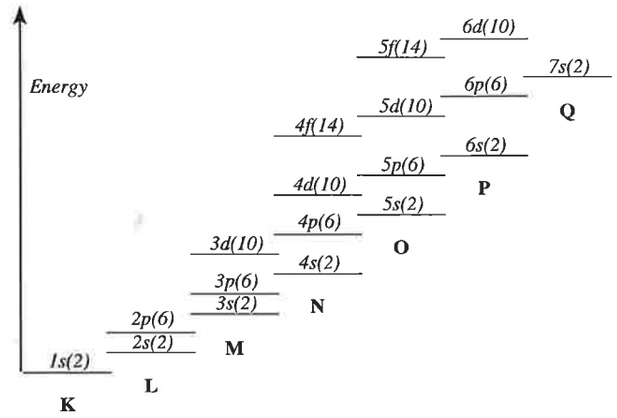


Fig. 7.2 Schematic energy levels of electronic configurations. Each level is labeled with the spectroscopic designation, including the number of electrons needed to fill the configuration in parentheses. Chemical shell designations (K, L, etc.) are at the bottom of each column. The diagram indicates, for example, that the two 5s states will fill before the ten 4d states. Energy levels are illustrative only of the general order in which configurations or sub-shells are filled and are not to scale. There are several exceptions to this overall scheme.

atomic properties of elements, like electro-negativity, ionic radius, and ionization energy all trend generally in one direction along the row. Chemical behavior likewise trends from one extreme to the other along a row. Period three, for example, ranges from the reactive metal sodium, through the less reactive metals magnesium and aluminum, the semi-metal silicon, the increasingly reactive non-metals phosphorus, sulfur, and chlorine, and the inert gas, argon. Elements in the same *column* of the table, in contrast, have the same electron configuration in their outer shells, and therefore all have very similar chemical properties. The noble gases, for example, (helium, neon, argon, krypton, xenon, and radon – column 18 or group VIIIA) all exhibit chemically inert behavior, and all have a filled outer shell with eight electrons in the s^2p^6 configuration. Similarly, the halogens in column 17, all highly reactive non-metals like fluorine and chlorine, have outer shells with the s^2p^5 configuration. There is also a secondary trend in properties moving down a column: the chemical reactivity of the halogens, for example, decreases steadily from fluorine, the lightest, to astatine, the heaviest.

Because of the order in which configurations are filled (see Figure 7.2) many elements have identical valence electron configurations and differ only in their inner electron shells. For example, the rare-earth elements, or lanthanoids – cerium ($Z = 58$) through ytterbium ($Z = 70$) – are all in period 6 of the table, but have chemical properties that are almost indistinguishable from one another. This is because they have identical outer shells ($6s^2$), and differ only in the configurations of their inner (primarily 4f) sub-shells.

For atoms with multiple valence electrons, the energy level of an excited configuration may depend not only on the quantum numbers of the electrons, but upon the interactions between the electron spins and angular momenta. For example, the excited state of helium that has configuration $1s^1\ 2p^1$ has *four* possible energies spread over about 0.2 eV. States differ because of different relative orientations of the two electron spins and the $l = 1$ angular momentum of

the p electron (directions are quantized and thus limited to four possibilities). The details of how multiple electrons interact are beyond the scope of this book, but for now, it is sufficient to recognize that such interactions can cause the energy level of a configuration to split into multiple values.

7.2 Isolated molecules

The outermost electrons of a molecule see a more complex binding potential due to the presence of two or more positively charged nuclei. Generally, this results in a greater number of electronic energy states. Each electronic state is still characterized by four quantum numbers, but in the molecule, the value of the m quantum number has an important effect on the energy level. More importantly, the molecule itself has internal degrees of freedom due to its ability to rotate around its center of mass, as well as its ability to vibrate by oscillating chemical bond lengths and angles. These internal rotational and vibrational modes are quantized as well, and they vastly increase the number of energy states permitted to the molecule.

Quantum mechanical theory approximates the total internal energy of a molecule as the sum of three independent terms:

$$E = E_{\text{electron}} + E_{\text{vibration}} + E_{\text{rotation}}$$

In addition to the quantum numbers specifying the electronic state, a diatomic molecule like CO or TiO will have one quantum number, J, to specify the rotational state, and one, v, for the vibrational state. Specification of the vibrational mode for molecules with more than two atoms becomes quite complex.

Figure 7.3 is a schematic energy-level diagram for a fictitious diatomic molecule. The energy levels in the figure are not scaled precisely. Transitions between the ground state and the first excited electronic state are usually in the

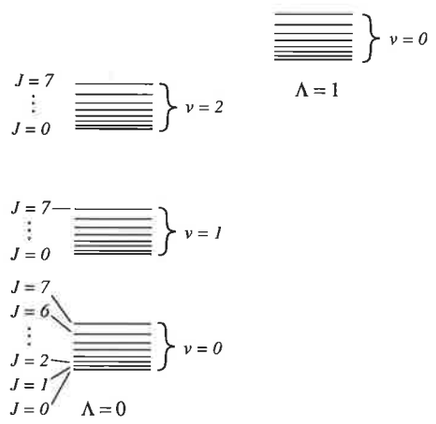


Fig. 7.3 Energy levels in a simple molecule. Right and left columns are different electronic states, as indicated by the quantum number Λ. Quantum numbers J and v specify the rotational and vibrational states, respectively. We show only three rotation states and seven vibration states in the lower electronic level.

range 0.5 to 100 eV. Transitions between adjacent vibrational states are about 100 times smaller than this, and between adjacent rotational states, about 10^6 times smaller yet.

The spacings between the rotational levels at different electronic and vibrational states are similar. As a result, the spectra of even simple diatomic molecules show a complicated pattern of lines consisting of extensive *bands*, with each band composed of many closely packed lines.

7.3 Solid-state crystals

7.3.1 Bonds and bands in silicon

A crystal is a mega-molecule in which the pattern of atoms and bonds repeats periodically with location. Many of the detectors we discuss in the next chapter are made of crystalline solids, so we now describe in detail the electronic structure of silicon, the most important of these materials. The silicon atom, located in column IVa of the periodic table, has 14 electrons, four of which are in the outer shell, with configuration $3s^2\,3p^2$. The outer shell will be filled when it contains eight electrons, not four. According to the theory of chemical valence, the component atoms of a molecule try to attain the electron structure of an inert gas (eight outer-shell electrons) by an appropriate sharing or transfer of electrons. Shared or transferred electrons produce, respectively, covalent or ionic bonds between atoms.

Consider the formation of a silicon crystal. Figure 7.4a shows what happens to the energy levels of an isolated silicon atom when a second silicon atom is brought closer and closer to it. As the electron wave functions begin to overlap, the levels split into two, outermost first. The nearer the neighbor, the greater is its influence, and the greater the splitting of levels. The outer electrons of both atoms can enter those levels since their wave functions overlap.

If we construct a crystal atom by atom, new energy states appear with each addition. For five atoms in a row, we expect something like Figure 7.4b. As crystal construction continues, more and more electron states become available as more and more atoms are added to the structure. Since even a tiny crystal

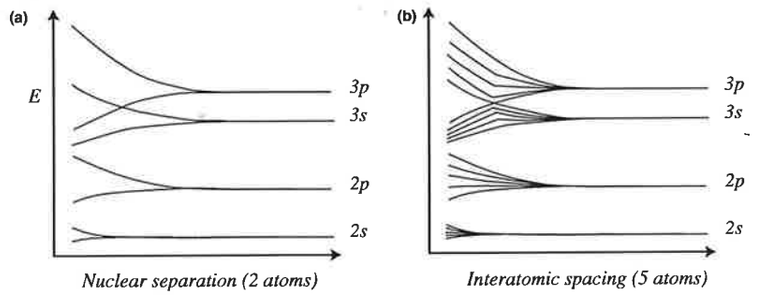


Fig. 7.4 (a) Changes in the electron energy levels in a silicon atom as a second atom is brought into close proximity. (b) The same diagram for the case of five atoms in a linear matrix.

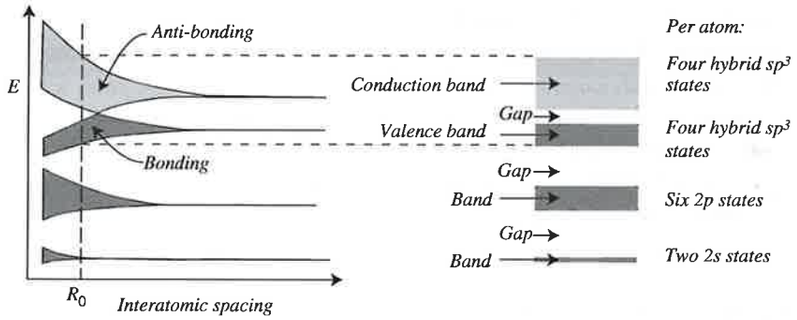


Fig. 7.5 Schematic diagram of the bands in silicon crystals. The diagram at right shows the bands formed at the preferred inter-atomic spacing. Dark-gray bands are occupied, light-gray are empty but permitted. Energies in the band gaps (white) are forbidden to electrons.

contains on the order of 10^{20} atoms, each causing a split in the energy levels, the spacing between levels must be on the order of 10^{-20} eV. These levels are so closely spaced that for practical purposes we treated them as a continuous ***band*** of available energies. If bands do not overlap, they will be separated by energy ***band gaps***. An electron anywhere in the crystal lattice is permitted an energy anywhere in a band, and is forbidden an energy anywhere in a gap.

Figure 7.5 shows the energy situation in crystalline silicon. The preferred inter-atomic spacing between nearest neighbors is R_0 (0.235 nm at room temperature). Note that at this spacing, the 3p and the 3s energy levels overlap. The result is called a crossover degeneracy, and energies in the crossover region are forbidden by quantum mechanics. The permitted states in the quantum mechanical view are certain linear combinations of s and p states, not the separate s and p states of the isolated atom. The periodic potential pattern of the regularly spaced nuclei in the crystal causes electron wave functions to interfere constructively at particular locations, and produces a set of states called ***sp^3-hybrid orbitals***. Each silicon atom contributes eight such states to the bands. Four of the sp^3 hybrid orbitals, the ones with lowest energy, correspond to an electron having its most probable location midway between the atom and one of its nearest neighbors. The nearest neighbors are at the four vertices of a tetrahedron centered on the nucleus. These four sp^3 hybrid orbitals all have energies that lie in the ***valence band*** and constitute the ***bonding states***. Four other sp^3 hybrid states have energies in the ***conduction band*** and locations away from the bonding locations. These are the ***anti-bonding states***.

From now on, we will use band diagrams, like the right side of Figure 7.5, to account for all the electrons in the entire crystal. At zero temperature all the anti-bonding states are empty and make up the ***conduction band***. The difference between the energy of the top of the valence band, E_v, and the bottom of the conduction band, E_c, is called the ***band gap energy***:

$$E_G = E_c - E_v$$

In silicon, the band gap is 1.12 eV at room temperature.

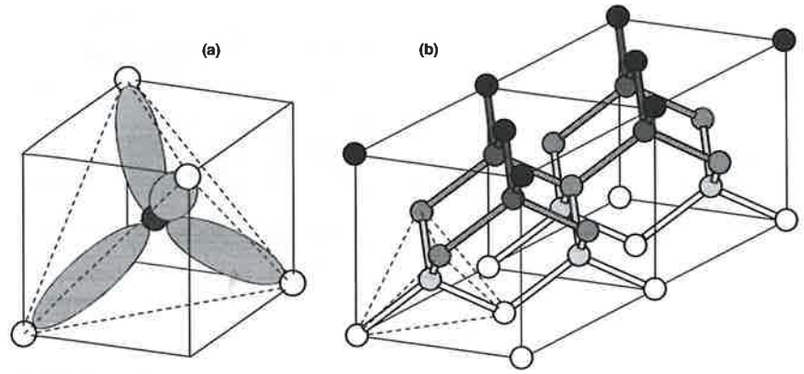


Fig. 7.6 (a) Tetrahedral covalent bonds for sp^3 hybridized orbitals for one atom (black) pictured at the center of a cube (solid lines). Its nearest bond-forming neighbors are at the four corners of the cube. These define the vertices of a tetrahedron (dashed lines). The electron bonding states are shown as light-gray ellipsoids – the regions where there is the highest probability of finding a valence electron. (b) A stick-and-ball model of the diamond lattice. Darker grays represent atoms higher up in the vertical direction. Each cube outlines a unit cell of the crystal, and a complete crystal is built by assembling many identical adjoining unit cells in three dimensions.

The most probable physical location of the valence electrons is on the line joining neighboring nuclei. In a silicon crystal, one would find two electrons in each of the light-gray region of Figure 7.6a – one from each atom – and this pair of shared electrons constitutes a covalent bond. Each atom forms four such bonds, symmetrically placed, and each atom therefore "sees" eight outer electrons – a complete shell. The bonds arrange themselves in a tetrahedral structure about each nucleus. The symmetry here reflects the transformation of the two s state and two p state electrons of the isolated silicon atom into the four hybrid sp^3 state electrons of the silicon crystal. X-Ray diffraction studies confirm that this tetrahedral structure repeats throughout the crystal in a three-dimensional pattern called a ***diamond lattice***, as sketched in Figure 7.6b.

7.3.2 Conductors, semiconductors, and insulators

It is instructive, although overly simple, to explain the differences between electrical conductors, semiconductors, and insulators as arising from differences in the size of the band gap and in electron populations within the bands. The important principle is that a material will be a good conductor of electricity or heat if its electrons can accelerate (i.e. change quantum state) easily in response to an applied electric field.

An analogy may help you understand the effect of band structure on conductivity. Imagine that you are standing in an almost empty room. You are free, in this environment, to respond to a whim to run across the room at top speed. On the other hand, if the same room is packed shoulder-to-shoulder with people, running is out of the question, no matter how strong your desire. Indeed, in a sufficiently dense crowd, moving at all is impossible.

Similarly, an electron in relative isolation can help conduct electricity or heat because it can (and must) accelerate when a global electric field is imposed, or when a strong local field arises during a collision with another particle. In a

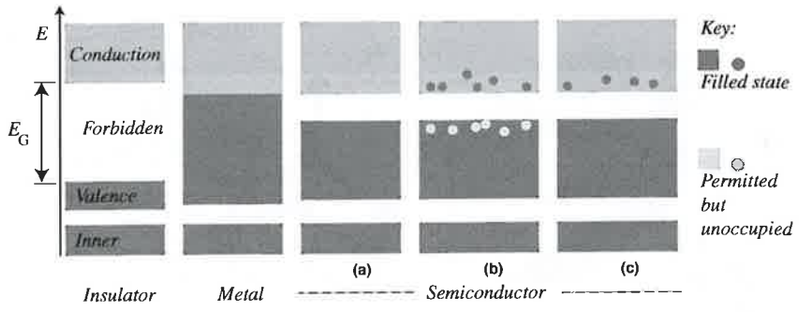


Fig. 7.7 Band structure of insulators, conductors, and semiconductors: (a) an intrinsic semiconductor at zero temperature; (b) the same material at a higher temperature; and (c) an extrinsic semiconductor.

crystalline solid, however, options are more restricted. The Pauli exclusion principle explains that an electron can only accelerate (i.e. change quantum states) if it can move into a new state that is (a) permitted and (b) not occupied by another electron. In a perfect silicon crystal at zero temperature, these two conditions are difficult to satisfy: every electron is in the valence band and every electron part of a covalent bond. Every permitted state in the band is occupied. The electrons, in effect, are packed shoulder-to-shoulder. Although we have ignored movement of the nuclei (which can oscillate around their mean positions), as well as surface effects, the basic conclusion is: electron crowding makes cold silicon a poor conductor of electricity, heat, and sound.

There *are* available states at much higher energies – the anti-bonding states in the conduction band. If an electron can acquire at least enough energy to jump the band gap, then it finds itself in the relatively empty conduction band where it is able to move around. In the crowded-room analogy, you might have enough energy to climb up a rope through a trap door to the empty room on the next storey, and there you are free to run.

Silicon conductivity will improve at higher temperatures, because an electron in a hot crystal might gain enough energy from a thermal collision to reach a state in the conduction band.

Figure 7.7 shows simplified diagrams of the band structures typical of metals, insulators, and semiconductors. In a typical insulator, the valence band is completely filled. The band gap is large compared to both the thermal energy, kT (at room temperature (300 K), $kT = 0.026$ eV), and any other energy sources. Because of the large gap, valence electrons cannot reach any permitted states in the conduction band. The exclusion principle forbids any electron to move into an already occupied state, so electrons cannot move at all – the material is a non-conductor.

A metallic conductor, in the second panel of the figure, has unoccupied permitted states immediately adjacent to the occupied valence states. If an electron near the top of the valence band absorbs even a tiny amount of energy, it may move into the conduction band, and from there to virtually anywhere in the material. The horizontal coordinate in these diagrams represents position in the material.

The figure shows three different views of materials called ***semiconductors***. The first, (a) an ***intrinsic semiconductor***, looks like an insulator, except it has a small band gap. This is similar to the band diagram of silicon at zero temperature. A valence electron can jump the gap into the conduction band by absorbing a modest amount of energy, either from thermal excitation or from some other energy source. Illustration (b), for example, shows the material in (a) at a high temperature. A few electrons have absorbed sufficient thermal energy to rise to the conduction band. This material will conduct, but the size of the current is limited because only these few conduction band electrons can easily change states. More electrons, of course, will rise to the conduction band to improve the conductivity if the temperature is increased further, and materials of this kind, in fact, can be used to make temperature gauges (thermistors).

The other thing to notice in illustration (b) is that whenever an electron is boosted into the conduction band, it must leave behind an empty state in the valence band. Another valence electron can shift into this vacated state and create a new empty state in the location it vacates. Since yet another electron can now move from a third location to fill this second empty state, it is clear that valence electrons can move through the crystal by occupying and creating empty states. It is easier to concentrate on the motion of the empty states, and to think of these ***holes*** as the entities that are moving. Holes thus behave like mobile positive charges in the valence band, and will contribute to the overall electrical conductivity. In intrinsic semiconductors, holes are usually less mobile than conduction-band electrons.

The third semiconductor (c) also has a few electrons in the conduction band, but without any corresponding holes in the valence band. Materials of this kind, called ***extrinsic semiconductors***, are extremely important in the construction of most electronic devices. A second class of extrinsic semiconductors has valence-band holes without corresponding conduction-band electrons.

7.3.3 Intrinsic semiconductors

Semiconductor crystals

A pure silicon crystal forms by linking all atoms with the tetrahedral covalent bond structure pictured in Figure 7.6. Part (b) of that figure shows the arrangement of a few dozen silicon atoms and bonds, and although not all bonds have been drawn, you can assure yourself that, as the pattern repeats, each atom will end up with four bonds. This geometry, called the ***diamond lattice***, insures that each atom shares eight electrons, completely filling the outer shell and producing a chemically stable structure. Indeed, the regularity of the diamond-lattice structure is tightly enforced, even if impurities are present in the silicon.

Table 7.3. *Periodic table of the elements near column IVA*

IIB s^2	IIIA s^2p^1	IVA s^2p^2	VA s^2p^3	VIA s^2p^4
	B	^{6}C	N	O
	Al	^{14}Si	P	S
Zn	Ga	^{32}Ge	As	Se
Cd	In	^{50}Sn	Sb	Te
Hg	Tl	^{82}Pb	Bi	Po

Elements with similar outer-electron configurations form similar diamond lattice crystals. These are in column IVA (also called column 14) of the periodic table, and include carbon, germanium, and tin.[1] Similar bonds also form in binary compounds of elements symmetrically placed in the table on either side of column IVA. For binary compounds, the crystal structure is called the "zinc blend" structure, which resembles Figure 7.6 except for alternation of the chemical identity of the nuclei on either end of each bond. Most useful semiconductors exhibit the diamond or zinc blend crystal structure. (Exceptions include lead sulfide and zinc oxide.) Table 7.3 shows part of the periodic table containing elements that combine to form important semiconductors.

Examples of binary-compound semiconductors are gallium arsenide (GaAs, a III–V compound) and cadmium teluride (CdTe, a II–VI compound). Some ternary compounds, notably $(Hg_xCd_{1-x})Te$, and quarternary compounds like $In_xGa_{1-x}As_yP_{1-y}$, also form useful semiconductors. Commercially, silicon is by far the most commonly used semiconductor. Germanium and gallium arsenide also find important commercial applications.

Semiconductor materials generally have a resistivity in the range 10^{-2} to 10^{9} ohm cm, midway between that of a good conductor (10^{-6} ohm cm) and a good insulator ($>10^{14}$ ohm cm). As we have already seen, resistivity depends critically on both temperature and the size of the band gap. Table 7.4 lists the band-gap energies for several semiconductors. Note that since the lattice spacing in a crystal is likely to change with temperature, so too will the band gap. Carbon in the diamond allotrope is an insulator because its band gap is so large that very few electrons can be thermally excited to the conduction band at room temperature; other carbon allotropes (graphite, carbon nanostructures) are conductors.

[1] The most common allotrope of tin, white tin, the familiar metal, has a tetragonal crystal structure. Gray tin, a less common allotrope, crystallizes in the diamond lattice. Lead, the final member of column IVA, crystallizes in a face-centered-cubic lattice.

Table 7.4. *Some common semiconductors. Forbidden band-gap energies and cutoff wavelengths at room temperature. A more complete table is in Appendix H1. Data from Sect. 20 of Anderson (1989)*

Material		Band gap (eV)	λ_c (μm)
IV			
Diamond	C	5.48	0.23
Silicon	Si	1.12	1.11
Germanium	Ge	0.67	1.85
Gray tin	Sn	0.0	
Silicon carbide	SiC	2.86	0.43
III–V			
Gallium arsenide	GaAs	1.35	0.92
Indium antimonide	InSb	0.18	6.89
II–VI			
Cadmium sulfide	CdS	2.4	0.52
Cadmium selenide	CdSe	1.8	0.69
Mercury cadmium teluride	$Hg_xCd_{1-x}Te$	0.1–0.5 (x = 0.8–0.5)	12.4–2.5
IV–VI			
Lead sulphide	PbS	0.42	2.95

Conductivity and temperature

At zero temperature, all the materials in Table 7.4 are non-conductors. As temperature increases, thermal agitation causes ionizations: electrons are promoted to the conduction band, free of any one atom; similarly mobile holes are created in the valence band. The material thus becomes a better conductor with increasing temperature. At equilibrium, we expect the rate of electron-hole recombinations to exactly equal the rate of thermal ionizations.

How, exactly, does an electron in a bonding state receive enough energy to jump the band gap? Optical electrons can collide with one another, of course, but it is important to note also that the lattice itself is an oversized molecule that can vibrate by oscillating bond length or angle. Just as with molecules, lattice vibration states are quantized with respect to energy. Solid-state theory often associates each discrete lattice vibration energy with a particle, called a ***phonon***, an entity analogous to the photon. Changes in electron state may thus involve the absorption or emission of a phonon. An electron can jump the band gap because it absorbs a phonon of the correct energy, and can lose energy and momentum by creation of, or collision with, a phonon.

At a particular temperature, the density of electrons at any energy within the bands will depend upon the product of two functions, (a) the probability,

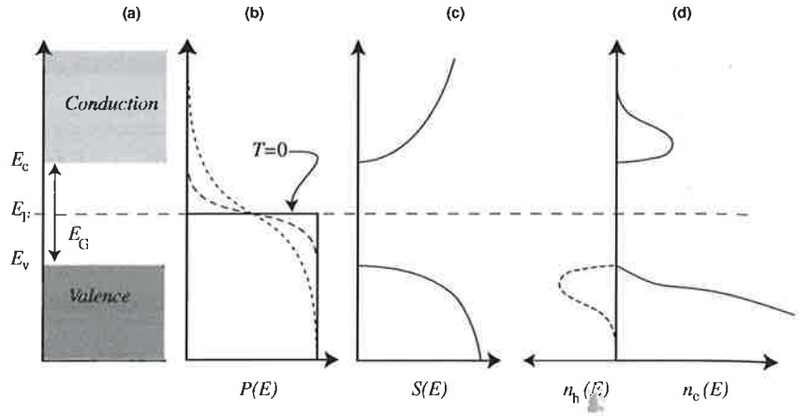


Fig. 7.8 Electron and hole density in an intrinsic semiconductor. (a) The locations of the band edges and the Fermi level midway between them. (b) The probability of finding an electron in a permitted state as a function of energy, $P(E)$. The solid line shows P at zero temperature, and the broken lines at two higher temperatures. (c) The density of permitted states as a function of energy, $S(E)$. (d) The density of electrons (the product of (b) and the highest temperature curve in (c)) and holes as functions of energy. The horizontal scale of plot (c) has been expanded to show detail.

$P(T, E)$, of an electron having that energy, and (b) the number density of available states at each energy, S:

$$n_c(T,E) = P(T,E)S(E)$$

With respect to the probability of a fermion having a particular energy, recall that the exclusion principle causes important restrictions on occupancy. This is certainly the case for the electrons in the bands of a semiconductor, where most of the valence states are fully occupied. In such a situation of ***electron degeneracy*** the probability per unit energy that an electron has energy, E, is given by the ***Fermi–Dirac*** distribution:

$$P(T,E) = \frac{1}{[1 + \exp\{(E - E_F)/kT\}]} \tag{7.3}$$

This expression reduces to the Boltzmann distribution, Equation (7.2), at high temperatures. At the limit of zero temperature, the Fermi–Dirac distribution requires that all of the lowest energy states be occupied, and all of the higher states (those with energies above E_F) be empty. That is, at $T = 0$,

$$P(E) = \begin{cases} 1, E < E_F \\ 0, E > E_F \end{cases} \tag{7.4}$$

The parameter E_F is called the ***Fermi energy***, and might be defined as that energy at which the probability for finding an electron in a permitted state is exactly one half. According to this definition, the Fermi energy will itself be a function of temperature for some systems at high temperature. However, for all cases we are concerned with, the Fermi energy can be treated as a constant equal to the energy of the highest permitted state at $T = 0$.

Figure 7.8a shows the energy bands for silicon at absolute zero, where electrons will fill all available states in the permitted bands up to the Fermi level, which falls midway between the valence and conduction bands. Figure 7.8b plots Equation (7.3) at three different values of temperature.

Figure 7.8c shows a schematic representation of the ***density of permitted states***, $S(E)$, for the valence and conduction bands of silicon. Note that $S(E)$, which gives the number of quantum states that are available at a particular energy (per unit energy and volume), is approximately quadratic near the permitted band edges and vanishes in the band gap.

The product $P(E)S(E)$ gives $n_e(E)$, **the *number density of electrons*** at energy E. The ***number density of holes*** at energy E in the valence band is just

$$n_h = [1 - P(E)]S(E)$$

Figure 7.8d shows these two functions, n_e and n_h, for a non-zero temperature. The total number densities of charge carriers of each kind (negative or positive) are given by the integrals of these functions in the appropriate band – conduction electrons and valence holes:

$$\begin{aligned} n_N &= \int_{E_F}^{\infty} n_e dE \\ n_P &= \int_{-\infty}^{E_F} n_h dE \end{aligned} \tag{7.5}$$

In intrinsic semiconductors, the number density of these two kinds of charge carriers in equilibrium must be equal, so $n_P = n_N$. The temperature dependence in Equation (7.5) follows from the Fermi distribution, and has the form

$$n_P = n_N = AT^3 e^{-\frac{E_G}{kT}} \tag{7.6}$$

The number density of charge carriers in intrinsic semiconductors, and therefore properties like resistivity and conductivity, should vary approximately exponentially with temperature.

7.3.4 Intrinsic photoabsorbers

An electron can leave a covalent bond if it is given sufficient energy (at least the value of E_G) to jump the band gap into an anti-bonding state in the conduction band. The required energy could be supplied by a photon, if the photon has a wavelength less than the ***cutoff wavelength*** for the material:

$$\lambda_c = \frac{hc}{E_G} = \frac{1.24\,\mu m}{E_G[eV]} \tag{7.7}$$

The band gap for silicon, for example, corresponds to a cutoff wavelength λ_c of 1.1 μm.

Photo-absorption changes a material's properties. For example, Figure 7.9a shows a simple device that utilizes such a change to measure light intensity. Photons absorbed by a block of semiconductor material produce ionization

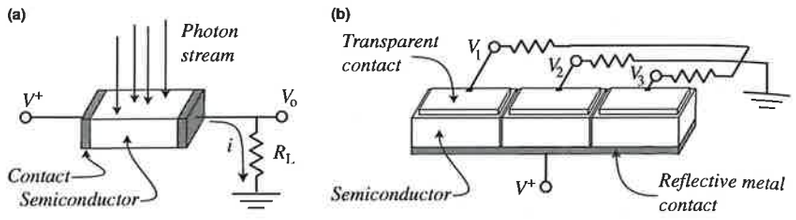


Fig. 7.9 Simple photoconductors. In (a), light strikes the exposed surface of a semiconductor linked to a simple circuit by two metal contacts. Photo-ionization produces charge carriers that reduce semiconductor resistance. Current through the device will increase with increasing illumination, and output is the voltage across a load resistor. In (b), a three-pixel device registers three different voltages in response to local illumination. Here photons pass through upper (transparent) contacts. The lower contact is reflective so that photons passing through the device are redirected for a second pass.

events – electrons in the valence band promoted to the conduction band, leaving an equal number of holes. The greater the flux of incoming photons, the greater the equilibrium concentration of charge carriers, and the greater the conductivity of the detector. If you maintain a constant voltage across the semiconductor, as in the figure, then the electrical current through the circuit should increase with the number of photons absorbed per second. A measurement of the voltage at V_0 thus monitors light intensity. Figure 7.9b shows an alternative illumination strategy that facilitates a close-packed array of detectors.

Notice that this ***photoconductor*** responds to the *number* of photons per second absorbed, not, strictly, to the rate of *energy* absorbed. Of course, if you know their spectral distribution, it is an easy matter to compute the energy flux carried by a given number of photons.

There are at least three reasons why a photon incident on the top of the device in Figure 7.9a will fail to generate an electron–hole pair. First, we know that those with frequencies below the band-gap frequency, E_G/h, cannot move an electron from the valence to conduction band, and thus cannot be detected.

A second failure is due to reflection of photons from the top surface of the device. As we saw in Chapter 5, minimal reflection occurs at normal incidence, and depends on the refractive index of the material:

$$R = \frac{(n_1 - n_2)^2}{(n_1 + n_2)^2}$$

The refractive index (and thus reflectivity) for silicon and most other semiconductors is very high in the ultraviolet, decreases through visible wavelengths, and is low (3.5 to 4) in the red and infrared. Reflectivity is also low in the X-ray band. Anti-reflection coatings can considerably reduce reflectivity for a particular wavelength.

A third reason for detection failure is that photons above the band-gap frequency might pass completely through the device without interaction. We now examine this transmission phenomenon in greater detail.

Once entering a semiconductor, the distance a photon can travel before being absorbed depends very strongly on its wavelength, as well as the quantum mechanical details of available electron states in the material. If a beam of photons enters material in the z-direction, its intensity at depth z will be

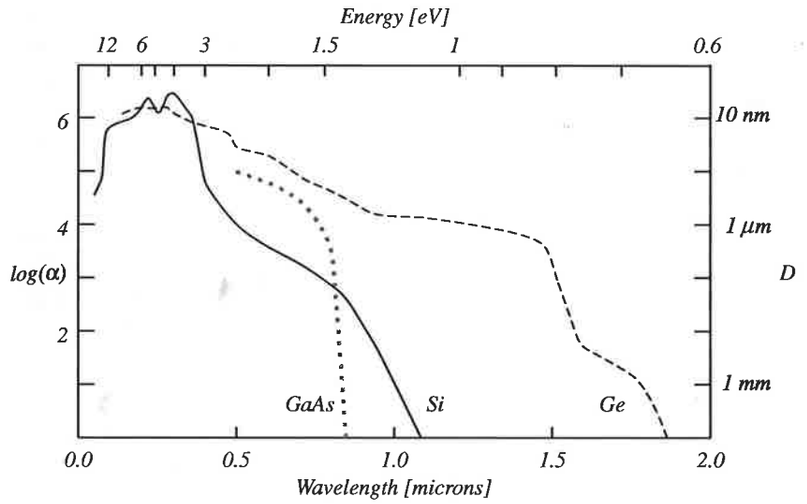


Fig. 7.10 The absorption coefficient, α, measured in m^{-1}, as a function of photon wavelength or energy. The absorption depth, D = 1/α, is on the right axis. Two indirect transition materials, silicon and germanium, show a much more gradual change with energy than does gallium arsenide, a direct transition material.

$$I(z) = I_0 e^{-\alpha z}$$

where I_0 is the intensity at $z = 0$, and α is the ***absorption coefficient***. A large absorption coefficient means light will not travel far before being absorbed. Figure 7.10 shows the absorption coefficient as a function of wavelength for silicon, germanium, and gallium arsenide. These curves illustrate an important divergence in behaviors caused by the details in the available energy states in two classes of materials. Notice that GaAs absorbs strongly right up to the cutoff wavelength, whereas Si and Ge very gradually become more and more transparent approaching that wavelength. Materials with an abrupt cutoff, like GaAs and InSb, are called ***direct transition*** semiconductors. Materials of the second kind, like Si and Ge, are called ***indirect transition*** semiconductors.

For both direct and indirect materials, when photons of the proper energy are absorbed, they almost always produce electron–hole pairs. The exceptions are usually due to flaws in the material. In some cases, a photon can interact with the lattice (particularly defects in the lattice) and deposit its energy as a phonon, not as a photo-ionization. For this reason, light-detecting devices require a semiconductor material that has been crystallized with strict controls to assure chemical purity and lattice integrity.

7.3.5 Extrinsic semiconductors

In practical devices, crystals inevitably have some chemical impurities and mechanical imperfections. These alter the energies and momenta of the states available near the sites of the defects, usually in undesirable ways. Curiously, though, some of the most useful semiconductor devices are made by intentionally introducing impurity atoms into the lattice.

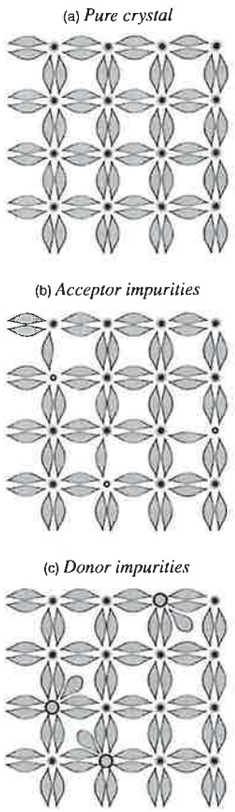


Fig. 7.11 A schematic of the bond structure in intrinsic and extrinsic semiconductors. In an actual crystal, the bond arrangement is three-dimensional – see Figure 7.5: (a) a pure intrinsic material; (b) three p-type impurity atoms in an extrinsic material; and (c) three n-type impurity atoms.

Figure 7.11a shows a flattened schematic of the positions of the atoms and outer electrons in an intrinsic semiconductor like silicon. Each atom shares eight valence electrons, forming four complete bonds. All atoms and bonds in the lattice are identical. Diatomic semiconductors like GaAs have a similar structure, except the chemical identity of the atoms alternates along rows and columns.

Now we intentionally introduce an impurity into the lattice, as in Figure 7.11b, where a few of the silicon atoms have been replaced by atoms that have only three valence electrons, like boron, gallium, or indium. Each impurity creates a vacancy in the electron structure of the lattice – a "missing" electron in the pattern.

The crystal, in fact, will try to fill in this "missing" electron. The impurity creates what is called an ***acceptor state***. It requires relatively little energy (on the order of the room temperature thermal energy, kT) to move a valence electron from a silicon–silicon bond elsewhere in the lattice into this gap at the impurity site. This creates a hole at the site that donates the electron. Such a hole behaves just like a mobile hole in an intrinsic semiconductor – a positive charge carrier that increases the conductivity of the material. Semiconductors in which impurities have been added to create positive charge carriers are termed ***p-type extrinsic*** semiconductors.

Figure 7.12a is an energy-band diagram for a p-type semiconductor. At zero temperature, a small number of (unoccupied) acceptor energy states exist within the band gap of the basic material. The energy difference, E_i, between the top of the valence band and the acceptor states is typically on the order of 0.05 eV in silicon (see Table 7.5). At a finite temperature, excitation of electrons from the valence band into these intermediate states creates valence-band holes (Figure 7.11b). Because the electrons in the intermediate states are localized at the impurity sites, they are immobile and cannot contribute to the conductivity. The mobile holes in the valence band, of course, can contribute, and are termed the ***majority charge carriers***. In contrast with intrinsic semiconductors, $n_P > n_N$ in p-type materials. Adding impurities to create an extrinsic semiconductor is called ***doping***, and the more heavily doped the material is, the higher is its conductivity. The transparent conductors used as contacts in Figure 7.9b, for example, are often made of highly doped silicon.

There is a second kind of doping. Figure 7.11c illustrates intrinsic material doped with atoms that have five valence electrons, like arsenic or antimony. The result is an ***n-type extrinsic*** semiconductor. Here, the "extra" electrons from the ***donor*** impurities are easily ionized into the conduction band. This ionization restores the bond structure to that of a diamond crystal (only eight shared outer-shell electrons, not nine) and consequently produces conduction-band electrons that constitute the majority carriers. Figure 7.12c shows the band structure of an n-type material, with occupied impurity states at energy E_i below the bottom of the conduction band at zero temperature. At higher temperatures, as shown in

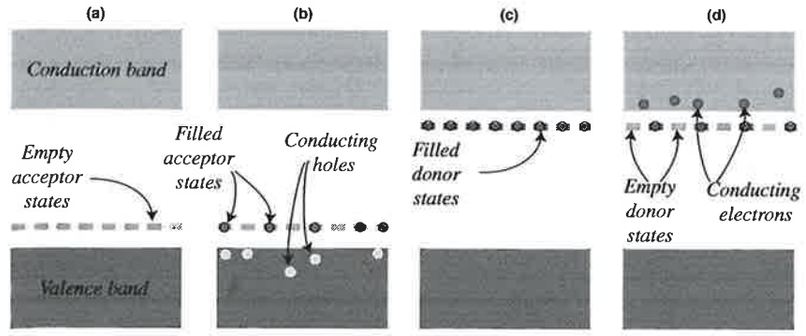


Fig. 7.12 Band structure of extrinsic semiconductors. (a) A p-type material at zero temperature. The energy difference between the top of the valance band and the acceptor states is typically on the order of 0.05 eV. (b) The same material at a higher temperature. Electrons excited into the acceptor states have created valence-band holes. (c) An n-type material at zero temperature; (d) shows the same material at a higher temperature, where electrons from the donor states have been ionized into the conduction band.

Figure 7.9d, some of these donor states are ionized to add electrons to the conduction band.

Extrinsic semiconductors respond to light in nearly the same way as intrinsic material. In fact, because the concentration of impurity atoms is always quite small (typically one part in 10^3 or 10^4), the presence of dopants does not appreciably modify intrinsic photo-absorption *above* the band-gap energy. The important difference occurs with photons whose energies lie *below* the intrinsic band-gap energy but above the dopant ionization energy, E_i.

Suppose, for example, a sample of boron-doped silicon (usually symbolized as Si:B), a p-type material, is kept so cold that the acceptor states, which lie 0.045 eV above the top of the valence band, are mostly empty. Intrinsic absorption in silicon cuts off at wavelengths longer than 1.12 μm. Short-ward of this cutoff wavelength, our sample absorbs as if it were intrinsic silicon. However, because photons with wavelengths shorter than $\lambda_i = hc/E_i = 26$ μm can ionize electrons from the valence band into the acceptor states, the number of majority carriers in the extrinsic material will increase in proportion to the number of photons short-ward of 26 μm. In effect, extrinsic absorption moves the cutoff to the longer wavelength. The implication for the construction of detectors for infrared light is obvious.

The absorption coefficient for extrinsic operation depends upon the dopant concentration. An important difference, then, between intrinsic and extrinsic photo-absorption is that the coefficient for extrinsic absorption can be adjusted in the manufacturing process. However, there are limits to the amount of impurity that can be added, so the absorption coefficient for extrinsic operation will always be low. One set of limits arises because, at high concentrations, the dopant atoms are so close together that their electron wave functions overlap, producing an ***impurity band***. If the states in this band are partially occupied, then the material will be conducting – with charge carriers "hopping" from one impurity state to another, effectively short-circuiting any photoconductive

Table 7.5. *Ionization energies, in eV, for different impurity states in silicon and germanium. Data from Kittel (2005) and Rieke (1994)*

Acceptors	Si	Ge
B	0.045	0.0104
Al	0.057	0.0102
Ga	0.065	0.0108
In	0.16	0.0112
Tl	0.26	0.01
Be	0.146	0.023
Cu	0.23	0.039
Donors		
P	0.045	0.0120
As	0.049	0.0127
Sb	0.039	0.0096
Bi	0.069	

effect. Extrinsic detectors therefore tend to be rather thick (1 mm) to provide adequate depth for photo-absorption.[2]

7.4 Photoconductors

Both intrinsic and extrinsic semiconductors, employed in a circuit like the one illustrated in Figure 7.9, in principle make excellent light detectors. If the voltage across the semiconductor is maintained as a constant (i.e. if $R_L = 0$), then the current will be directly proportional to the number of charge carriers in the material, which (at a sufficiently low temperature) will be directly proportional only to the rate at which it absorbs photons. More precisely, the electric current due to photons will be something like

$$I_{photo} = nqP_aP_c$$

where n is the number of photons incident per second and q is the electron charge. The quantity P_a is the probability that an incident photon will generate a pair of charge carriers in the detector; P_a depends upon the factors discussed above: surface reflectivity, the absorption coefficient, and the thickness of the sensitive layer. (The fraction of the photons entering a layer of

[2] There are some important exceptions, e.g. the ***blocked impurity band (BIB) detector***, also called the ***impurity band conduction (IBC) detector***, in which a thin layer of highly doped material is bonded to a layer of intrinsic material, so that the intrinsic material breaks the continuity (and conductivity) of the impurity band. See McLean (2008).

thickness z that are absorbed is just $1 - e^{-\alpha z}$.) The absorption coefficient, in turn, will depend upon wavelength and (for extrinsic materials) impurity concentration.

The quantity P_c is the probability that a charge carrier, once created, will actually move from the semiconductor to the appropriate electrode. It depends on a number of factors: electric field strength, thickness of the material, charge carrier mobility, and the presence of flaws in the crystal that might promote recombination.

Although is desirable to make I_{photo} as large as possible, there are limits. For example, increasing the voltage across the electrodes increases P_c, but at large potential differences electrons will gain considerable kinetic energy. At high energies, electrons will ionize atoms by collision to create new charge carriers, and these secondaries will in turn be accelerated to produce additional carriers. This avalanche of charge produces a spike in the output current. At high enough voltages, in a condition called ***breakdown***, the avalanche becomes constant, destroying the resistance of the material and making it useless as a detector. Similarly, increasing the thickness, z, of the material increases the probability that a photon will be absorbed. At the same time, however, increasing z in Figure 7.9b reduces the probability that a charge carrier will be able to move to an electrode before it recombines.

7.5 The MOS capacitor

The metal-oxide-semiconductor (MOS) capacitor is the basic element of an important class of astronomical detectors. Figure 7.13 illustrates the essentials of the device, which is a three-layer sandwich. In the figure, the left-hand layer is a block of p-type semiconductor. The left-hand face of this block is connected to electrical ground. A thin layer of insulator forms the middle of the sandwich. The semiconductor is usually doped silicon, and the insulator layer is usually silicon dioxide. The right-hand layer is a thin coating of metal, which is held at a positive voltage. If the insulating layer is not made of SiO_2 (silicon nitride,

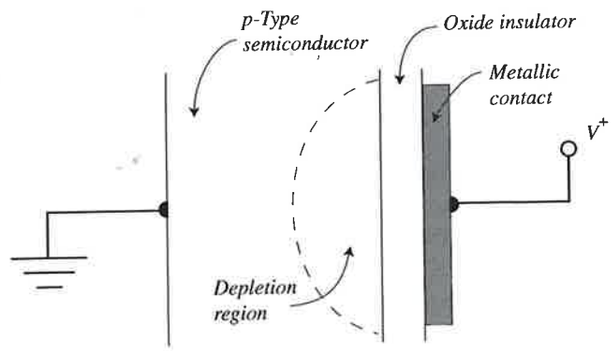


Fig. 7.13 Cross-sectional view of the physical structure of a MOS capacitor. Positive voltage (usually a few volts) applied to the metal layer creates a depletion region in the semiconductor.

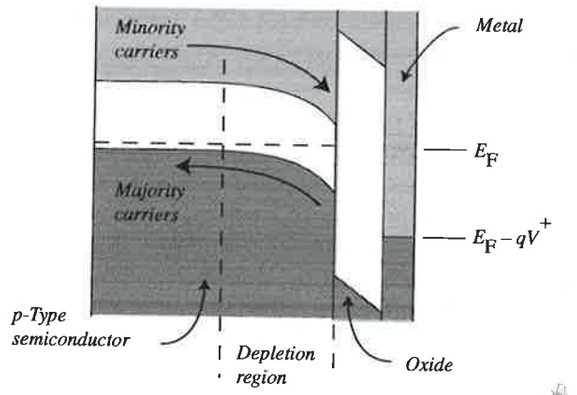


Fig. 7.14 An energy-band; diagram for the MOS capacitor. Majority carriers are swept out of the depletion region. Minority carriers are swept towards the boundary with the insulator.

Si_3N_4, is the usual alternative), then the device is called an ***MIS (metal-insulator-semiconductor) capacitor***.

Figure 7.14 shows the band structure of the device. The positive voltage of the metal layer distorts the energies of the bottom and top of the semiconductor forbidden gap. The tilt of the band reflects the strength of the electric field. In the diagram, the electric field forces electrons to move to the right and holes to the left. If a hole loses energy, it moves upwards in the diagram.

The large band gap in the insulator prevents minority electrons from crossing into the oxide layer. The flow of majority holes to ground in the valence band, in contrast, is not impeded. The result is that, in equilibrium, a ***depletion region*** devoid of the majority charge carriers develops in the semiconductor adjacent to the insulator. The minority carriers here are immobile – trapped in the potential well formed by the bottom of the semiconductor valence band and the band gap of the insulator.

The MOS capacitor is especially useful because it will *store* electrons that are generated by ionization. Referring to Figure 7.15a, it is clear that if an electron–hole pair is created in the depletion region, the pair will be swept apart before they can recombine: the electron goes into the well, and the hole leaves the material. Electrons in the well remain there indefinitely, since they sit in a region depleted of holes. Ionizations outside the depletion zone are less likely to produce stored electrons since charges there move by diffusion, and the longer it takes for the electron to reach the depletion zone, the greater are its chances of encountering a hole and recombining.

For ionizations in the depletion zone, however, charge storage can be nearly 100% efficient. Eventually, if enough electrons accumulate in the zone, they will neutralize the effect of the positive voltage and remove the potential well for newly generated electrons. Figure 7.15b illustrates this saturated situation. Saturation destroys the depletion zone and allows generated charge carriers to move only by diffusion, eventually recombining in equilibrium just as in an

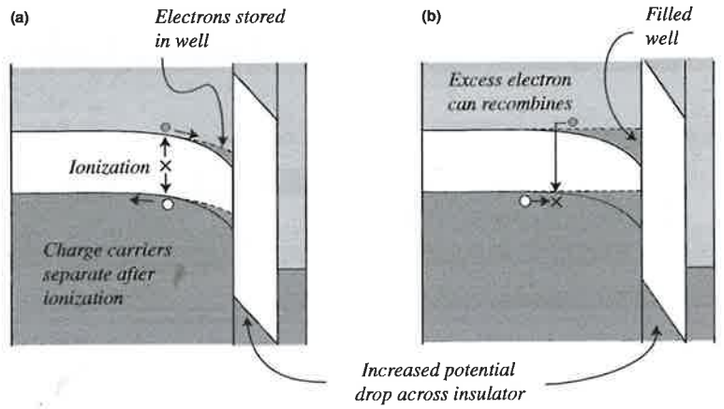


Fig. 7.15 (a) The movement of charge carriers created by ionization in the semiconductor layer of an MOS capacitor. Conduction-band electrons will move into the potential well, while valence-band holes move out of the material to ground. There is a net increase in the negative charge in the semiconductor layer. (b) In a saturated device, there is no longer a potential gradient in the semiconductor, so recombination and ionization will be in equilibrium, and there will be no further gain in net negative charge.

ordinary semiconductor. Newly created electrons are no longer stored. The capacitor has exceeded its ***full-well capacity***.

Short of saturation, the MOS capacitor is a conceptually simple detector of light. For every photon below the cutoff wavelength absorbed in the semiconductor layer, the device stores something like one electron. Making a photometric measurement then consists of simply counting these electrons. For an astronomer, this is a wonderful characteristic. It means that a very weak source can be detected by simply exposing the capacitor to light from the source for a time long enough to accumulate a significant number of electrons.

7.6 The p–n junction

A very significant situation arises if a p-type material and an n-type material are brought into contact. Junctions of this sort are the basis for many solid-state electronic devices and for some astronomical detectors. Figure 7.16 illustrates the behavior of charge carriers at a p–n junction. In the figure, we imagine that a block of n-type material has just been brought into contact with a block of p-type material.

Figure 7.16a shows the non-equilibrium situation immediately after contact. The majority charge carriers start to flow across the junction. Electrons in the n-side conduction band will move across the junction and drop down in energy to fill the available acceptor states on the p side (broken lines). Likewise, mobile holes in the valence band of the p-type material will move across the junction to neutralize any electrons in donor states in the n-type material. Opposite charges build up in the doping sites on either side of the junction – excess negative charge on the p side, excess positive charge on the n side. Electrostatic repulsion eventually halts further transfer of carriers across the junction.

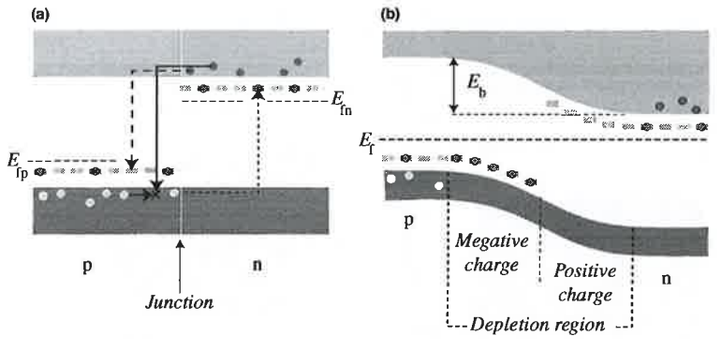


Fig. 7.16 The p–n junction; (a) shows the flow of charge carriers immediately after contact between the two regions. Majority carriers (n-side electrons and p-side holes) recombine, fill acceptor sites, and ionize donor sites. The band structure in equilibrium is shown in (b). The accumulation of charges near the junction creates a built-in field, which alters the energy levels of available states so that the Fermi energy is the same everywhere in the crystal.

7.6.1 Generation and recombination

Figure 7.16b shows the situation once equilibrium is established. As in the MOS capacitor, a ***depletion region***, constantly swept clear of mobile charge carriers, has formed in the volume surrounding the junction. The lack of charge carriers means this region should have very high electrical resistance. In equilibrium, the energy of the bottom of the conduction band and the top of the valence band changes across the depletion region – it requires work to move an electron from the n region to the p region against the electrostatic force. The potential difference across the depletion zone, E_b, is just sufficient to bring the Fermi energy to the same level throughout the crystal. In equilibrium, charges do move through the depletion region, but the two electric currents here cancel:

$$I_r = -I_g$$

The first current, the ***recombination current***, I_r, is due to the majority carriers that are able to overcome the potential barrier, cross the depletion region, and undergo recombination. This I_r is a positive current that flows from p to n; it has two components: one caused by n-side electrons, the other by p-side holes. Figure 7.17 illustrates the flow of the recombination current, whose magnitude *will depend on the size of the barrier* and on the temperature.

The second current, I_g, the ***generation current***, is due to minority carriers and flows in the opposite direction (from n to p). The minority carriers are thermally ionized conduction-band electrons on the p side and valence-band holes on the n side, which diffuse away from their creation sites. If such a carrier reaches the depletion region, it will be swept across. Diffusion speed outside the depletion region depends on the temperature and the impurity concentration, but is (to first order) *independent* of E_b. Thus, I_g depends on temperature, but in contrast to I_r is virtually independent of the size of E_b.

7.6.2 p–n Junction diodes

The different behaviors of the two currents mean that the p–n junction can function as a ***diode***: it will carry (positive) current in the direction

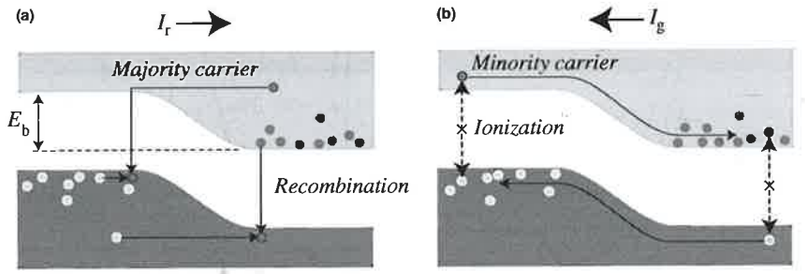


Fig. 7.17 Recombination (a) and generation (b) currents at a p–n junction.

p to n, but not in the reverse direction. Figure 7.18 illustrates the basic process.

In the condition known as ***forward bias***, a positive voltage connected to the p side of the junction reduces the size of the potential barrier E_b. The recombination current, I_r, will flow more strongly. (That is, more electrons will have energies greater than the barrier, and can move from n to p.) The size of this current will depend in a non-linear fashion on the size of the applied voltage, V_{ext}. The applied voltage, however, does not affect the generation current in the opposite direction, I_g, due to minority carriers. The relatively poor conductivity of the depletion region guarantees that almost all of the potential drop will occur here, and the applied voltage will have little influence on the diffusion rate outside the depletion region. Thus, in the forward bias case, $I_r >> -I_g$, and current flows from p to n.

A negative voltage connected to the p side of the junction – a condition known as ***reverse bias*** – increases the size of the potential barrier E_b. This chokes off the flow of majority carriers and lowers I_r from its equilibrium value. Again, the minority carrier current, I_g, remains little changed, so the result of the reverse bias circuit is a very small current in the direction n to p. Boltzmann's law and the above arguments suggests a ***diode equation*** that gives the voltage–current relationship for an "ideal" diode:

$$I_{TOTAL} = I_r + I_g = I_s\left(e^{\frac{qV_{ext}}{kT}} - 1\right) \tag{7.8}$$

Here, q is the electron charge, and current and voltage are assumed to be positive in the p to n direction. You can verify that this formula corresponds to the behavior seen in an actual diode illustrated in Figure 7.19. The formula does not describe the phenomenon of diode breakdown at large reverse biases.

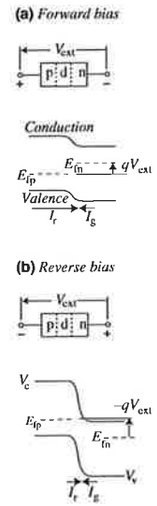


Fig. 7.18 Biased diodes: (a) forward bias reduces the size of the barrier, so the recombination current increases; (b) reverse bias increases the barrier and decreases the recombination current. In both cases, the generation current remains unchanged.

7.6.3 Light detection in diodes

Figure 7.20 shows the result of photo-absorption in a p–n diode. Each absorption of a photon causes an ionization and the creation of a conduction electron and

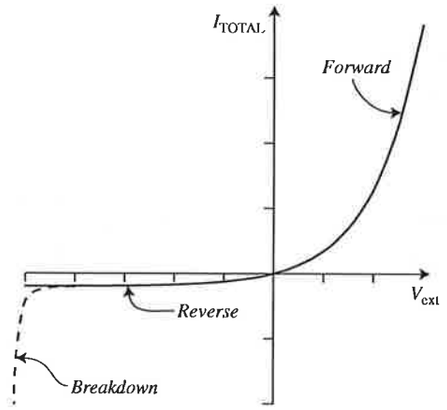


Fig. 7.19 The current–voltage relation for an ideal p–n diode. The solid line is the relation given by Equation (6.8). The dotted line shows the phenomena of breakdown in real diodes, which become conducting at very negative external voltages.

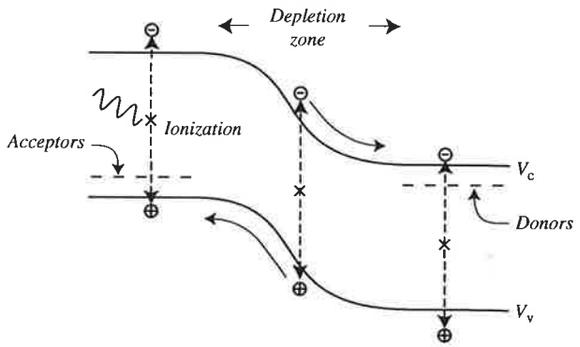


Fig. 7.20 Photo-absorption in a p–n junction diode.

valence hole. This adds a new contribution to the generation current, this one dependant on ϕ, the number of photons that enter the detector per second. The inclusion of a photocurrent modifies Equation (7.8):

$$I_{TOTAL} = I_{ph} + I_r + I_g = -q\phi\eta + I_s\left(e^{\frac{qV_{ext}}{kT}} - 1\right) \qquad (7.9)$$

Here η is a factor that depends on the fraction of incident photons absorbed as well as the probability that a generated charge carrier will cross the junction before recombining. Note that charge pairs created in the n or p material must move by diffusion to the junction, as discussed above, and have a finite probability of recombining before crossing the junction. Electron–hole pairs created in the depletion zone, on the other hand, are immediately swept apart by the strong electric field there, and have little chance of recombining. Majority

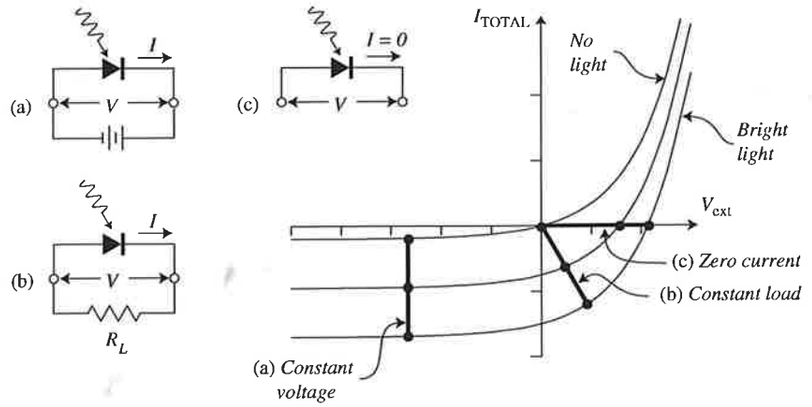


Fig. 7.21 Current–voltage relations for a photodiode at three different levels of incident photon flux. Heavy lines show electrical properties as a function of photon intensity for three different modes of operation.

carriers will thus tend to accumulate on either side of the depletion zone, and the junction will behave like a charge storage capacitor if an external circuit does not remove the carriers.

There are different strategies for employing the light sensitivity of a photodiode. Figure 7.21 is a plot of Equation (7.9) for three different light levels, as well as three different modes of operation: (a) In the ***photo-conductor*** mode, a battery holds the external voltage to a constant value, and the current is a linear function of the incident photon flux. (b) In the ***power-cell*** mode, the diode is connected to a constant-load resistance, and the power output depends on the incident photon flux. This is the principle of operation for solar power cells. (c) In the ***photovoltaic*** mode, current from the diode is held at zero (making it a storage capacitor by connecting it to a very high impedance voltmeter, for example), and the voltage across it is a non-linear function of the photon flux.

7.6.4 Variations on the junction diode

Some modifications of the simple p–n junction can improve the device's response to light. Several are important in astronomy.

The ***PIN diode*** sandwiches a thin layer of intrinsic (undoped) silicon between the p material and the n material of the junction. This increases the physical size of the depletion zone, and the resulting p–intrinsic–n (PIN) diode has larger photosensitive volume, higher breakdown voltage, lower capacitance, and better time response than the simple p–n device.

The ***avalanche photodiode*** is both a physical modification and a mode of operation. Consider a photodiode (a modified PIN type) that is strongly back-biased at close to its breakdown voltage. Because of the large voltage drop across the intrinsic region, charge carriers created by photo-absorption will accelerate to high kinetic energies – high enough to produce impact ionization of additional charge carriers. These secondaries will in turn accelerate to

produce further ionizations. The resulting avalanche of carriers constitutes a current pulse that is easy to detect. At low levels of illumination, counting the pulses is equivalent to counting photons. At higher illuminations, the pulses are too close together to count, but the resulting current, although noisy, is very large and therefore easy to detect.

The ***Schottky photodiode*** is especially useful as a detector in the near- and mid-infrared. Figure 7.22 shows a junction between a metal and p-type silicon as well as the corresponding electron energy bands. At the junction, electrons will spontaneously flow from the metal to neutralize majority holes in the semiconductor until the Fermi levels in the two materials match. The resulting charge transfer sets up a potential barrier at the junction, as well as a depleted zone on the semiconductor side. Light detection occurs because holes created by photo-absorption in the metal layer, if energetic enough, will move across the potential barrier into the depleted region, where they will be swept out into the semiconductor (remember, our band diagrams show energies for electrons, so holes will move upwards in the diagram). The barrier height for holes, E_{ms}, determines the long-wavelength cutoff of the diode; E_{ms} depends on the metal used. Common choices are all metallic silicides, which form easily when a very thin metallic layer is deposited on silicon by evaporation. The most useful is PtSi, with $E_{ms} = 0.22$ eV, and $\lambda_c = 5.6$.

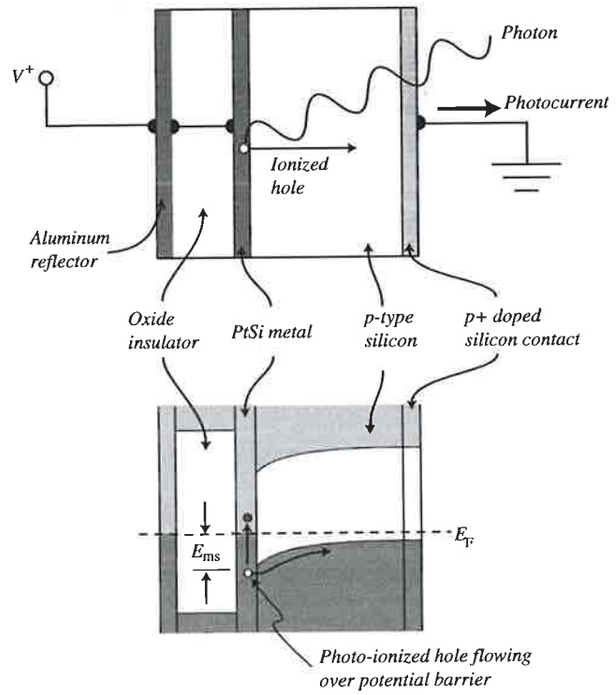


Fig. 7.22 A Schottky photodiode. The upper diagram shows a cross-section of the material structure, and the lower shows the energy bands. Photo-ionized holes produced in the metal move upward in the diagram, and will be detected if they overcome the potential barrier between metal and semiconductor. The reflector sends photons back to the PtSi layer for a second pass.

Schottky diodes have relatively low efficiencies, but they also have several virtues: they are very easy to manufacture and mate to read-out electronics. They also tend to have uniform responses and therefore are a good choice for the elements in an array. Their sensitivities extend into the infrared, where intrinsic silicon is useless.

7.7 The vacuum photoelectric effect

The vacuum photoelectric effect depends on the ejection of electrons from the surface of a solid, and has important applications in astronomical detectors. Figure 7.23 illustrates the effect, which is simplest in metals. A thin slab of the metal cesium occupies the left side of the figure, which shows the band structure. We use cesium as an example because it has a relatively loose hold on its surface electrons. The surface of the metal runs vertically down the center. If the potential energy of an electron at rest well away from the metal is zero, then the ***work function, W***, of the material is the difference between this free electron energy and the Fermi energy of the solid. In the case of cesium, the work function is 2.13 eV.

We would like to use the energy of one photon to move one electron from the metal to the vacuum. This operation has two requirements: the electron must be given a positive energy, and it must be located at the surface. In general, the absorption of a photon with energy $h\nu \geq W$ will take place in the interior of the metal, and will promote an electron there into the conduction band. This electron *might* have positive energy. If, after diffusing to the surface, the electron still has both positive energy and an outward-directed momentum (*case* B in the figure) it can move into the vacuum.

A simple device called a ***photocell*** (or more properly, a ***vacuum photodiode***), illustrated in Figure 7.24, uses this effect to measure the intensity of light. In the

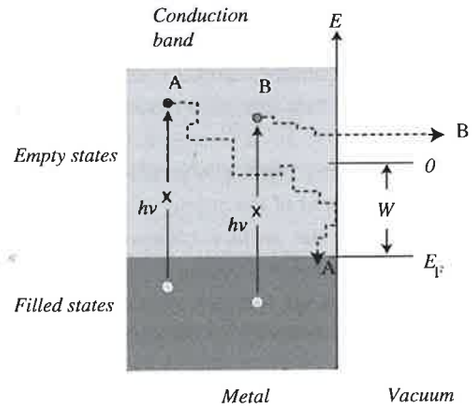


Fig. 7.23 The vacuum photoelectric effect in a metal. Photoelectron B reaches the vacuum with positive energy, while photoelectron A does not. Both photoelectrons make collisions with the lattice, and execute a random walk to the surface. Photoelectrons gradually become thermalized – if the metal is cold, they tend to lose energy on each lattice collision.

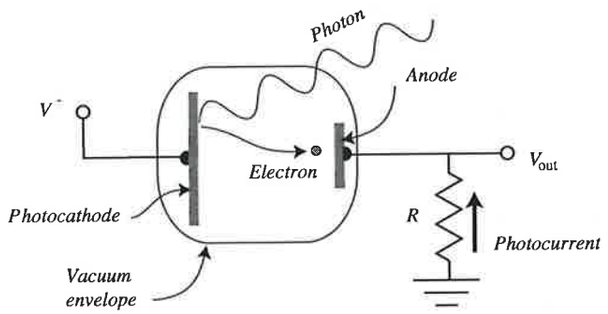


Fig. 7.24 A vacuum photodiode. A photon with sufficient energy to eject an electron strikes the photocathode. The photoelectron then accelerates to the anode and flows through the load resistance to ground.

diagram, two conductors are sealed in an evacuated cell with a transparent window. One conductor, the ***photocathode***, is made from some material (e.g. cesium) that exhibits an efficient vacuum photoelectric effect. The photocathode is held at a negative voltage. The other conductor, the anode, is connected through a load resistor to the ground as illustrated. Illumination of the photocathode ejects electrons into the vacuum. These accelerate to the anode. The result is an output current and voltage across the resistor that is proportional to the photon arrival rate at the cathode.

Metals actually make rather poor photocathodes. For one thing, they exhibit large work functions. (Cesium, the metal with one of the smallest values for W, will only detect photons with wavelengths shorter than 580 nm.) A second, even more serious, disadvantage is that metals are highly reflective. Semiconductors usually make better photocathodes since they are much less reflective. The photoelectric effect is slightly more complex in a semiconductor, as illustrated by the band diagram in Figure 7.25a. The zero of energy and the work function are defined as in a metal, and a new variable, the ***electron affinity***, χ, is defined as the difference between the zero point and the energy at the bottom of the conduction band. For a simple semiconductor, as in Figure 7.25a, the electron affinity is a positive number. Since there are no electrons at the Fermi level in a semiconductor, the energy required to eject an electron is

$$h\nu \geq E_G + \chi$$

This can be relaxed by creating a p–n junction near the emitting surface. In Figure 7.25b, the displacement of the energy of the bottom of the conduction band in the n material near the surface means that a photon with energy greater than just E_G can cause electron emission if it is absorbed in the p-type material. In this case, the effective electron affinity of the p-type material is a negative number. The n-type layer is so thin and transparent that its more stringent energy requirements are not a serious detriment to the cathode's sensitivity to long wavelengths. Materials of this type, termed ***NEA photocathodes*** (negative electron affinity), are usually fabricated with a III–V semiconductor as

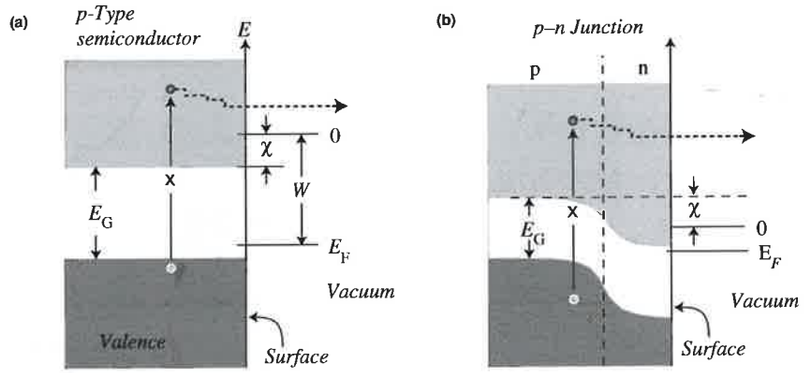


Fig. 7.25 The vacuum photoelectric effect in semiconductors

the p-type material and oxidized cesium as the n-type material. For example, an NEA photocathode made from p-doped gallium arsenide (E_G = 1.4 eV) with a surface layer of n-doped Cs_2O (E_G = 2.0 eV, χ = 0.6 eV) is sensitive out to 880 nm and has been important for some astronomical applications.

We have assumed that emitted photoelectrons will leave from the surface that is illuminated. This need not be the case, and many photocathodes are ***semi-transparent***: photons enter on one side and electrons emerge from the opposite side.

7.8 Superconductivity

Superconducting material has an electrical conductivity that falls to zero at and below a critical transition temperature, T_c. The simplest superconductors, and the first investigated, are all metallic elements with very low critical temperatures ($T_c < 10$ K). These are called type I superconductors. More complex materials (alloys, ceramics, and various exotic compounds) may have higher transition temperatures. Currently (2010), the highest claimed T_c is 242 K.

Type I superconductors are the basis of some potentially important light detectors in astronomy, so we briefly describe their behavior here. The website superconductors.org or the modern physics text by Harris (1998) gives a more complete introduction, and chapter 10 of Kittel (2005) provides a more advanced treatment, as does Blundell (2009).

7.8.1 The superconductor band gap

Above the critical temperature in a superconducting metal like lead, the Fermi–Dirac formula describes the energy distribution of the valence electrons in the conduction band. At T_c (7.19 K for lead) a lattice-mediated force between electrons makes new energy states available below the Fermi level – two spatially separated electrons can form a ***Cooper pair*** of exactly canceling

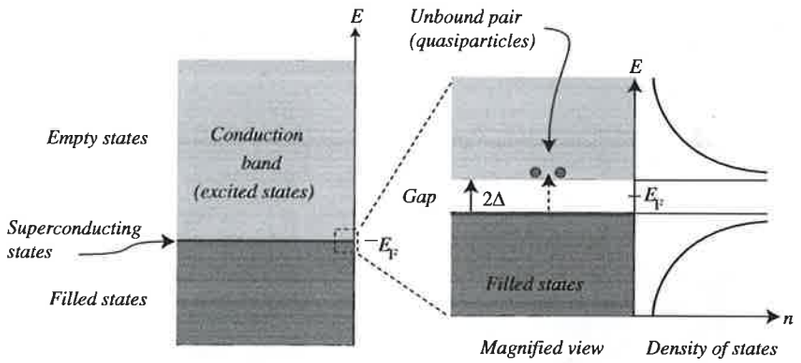


Fig. 7.26 Energy bands in a superconductor. The band-gap energy is the energy required to break apart two electrons bound in a Cooper pair, placing them in an excited quasiparticle state (dotted arrow). The density of states just below and just above the band gap is very high, although there are no states in the gap itself.

momenta and spins. Each pair has a binding energy well below the thermal energy of the lattice and, with zero spin, behaves in many ways like a boson – the Pauli exclusion principle does not apply to these states, and all pairs have the same momentum (zero, when there is no current). It is the Cooper pair states that are responsible for superconductivity and many resultant behaviors – including perpetual electric currents and magnetic levitation.

Our concern, however, is the manner in which a superconductor interacts with light. Figure 7.26 shows the special energy-band diagram for a superconductor. At temperatures below T_c, an unlimited number of superconducting states exist at an energy Δ below the Fermi level. Single electrons will therefore occupy only states of energy $(E_F - \Delta)$ or lower. The value of Δ is a strong function of temperature, rising from zero at T_c to a maximum value of Δ_m at temperatures below about $0.3\,T_c$. The value for Δ_m, which measures the binding energy per electron of a Cooper pair, is tiny, 1.4×10^{-3} eV for lead, which is typical.

Consider what must happen for a superconductor to absorb a photon: only if the photon has energy larger than 2Δ can it break apart a Cooper pair and promote the two electrons to higher energies. Lower energy photons will not be absorbed: the material has an effective band gap of magnitude 2Δ, as shown in Figure 7.26. The electrons promoted to the excited states in the "conduction" band in the superconductor have quantum characteristics that differ from energetic electrons in an ordinary metal, and are therefore termed ***quasiparticles***. For example, the number of states available to quasiparticles at energies just above the gap is very large. Table 7.6 lists the gap energies and transition temperatures of a few superconductors that have been useful in astronomical detectors.

7.8.2 Light detection in an SIS junction

Two superconductors separated by a thin layer of insulator (SIS = superconductor–insulator–superconductor) constitute a ***Josephson junction*** if the insulator is thin enough (around 1 nm) to permit quantum-mechanical tunneling.

Table 7.6. *Some type I superconductor characteristics. Data from Kittel (2005)*

Atom	T_c, K	Band gap $2\Delta_m \times 10^4$ eV
Al^{13}	1.14	3.4
Nb^{41}	9.5	30.5
Hf^{72}	0.12	0.4
Ta^{73}	4.48	14

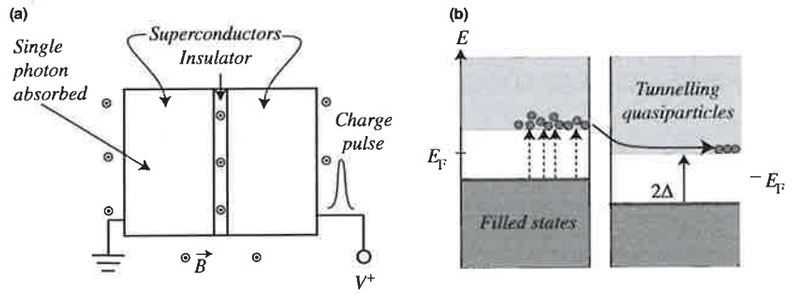


Fig. 7.27 An STJ diode. (a) A cross-section of the physical device. In most practical detectors the three layers and their contacts are deposited as films on a transparent substrate, so a more accurate diagram would extend vertically several page heights. The band structure is shown in (b) . Not shown is the possibility that quasiparticles on the right can tunnel back to the left to break apart additional pairs.

Figure 7.27 shows such a junction arranged as a light-detecting diode: a positive bias voltage less than $2\Delta/q$ is applied to the right-hand superconductor, and a magnetic field is applied parallel to the junction. If the junction is very cold, all excited states are empty. In a normal Josephson junction it would be possible for the Cooper pairs to tunnel from left to right, but the magnetic field suppresses that current, so the diode does not conduct.

If the left-hand superconductor absorbs a single visible light photon of wavelength λ (energy hc/λ), it receives enough energy to break apart multiple Cooper pairs, promoting a maximum of $hc/\lambda\Delta$ electrons into excited states. These quasiparticles *can* tunnel across the insulator, and those that do produce a current pulse whose total charge is inversely proportional to the wavelength of the exciting photon.

Devices of this kind, called ***superconducting tunnel junctions*** **(*STJs*)**, operated with sufficient time resolution, can count individual incoming photons and determine the wavelength (from X-ray to near infrared) of each. The uncertainty of the wavelength determination depends on the wavelength. Although still very much in the development stages, a few experimental but practical multi-pixel STJ-based detectors have begun to appear at telescopes. See chapter 4 of Rieke (2003) and the references by Peacock *et al.* (1997) and Verhoeve *et al.* (2004).

Superconducting tunnel junctions promise to be the near-ideal astronomical detector: They can be fashioned into an array that produces an image yielding both spectroscopic information and high time resolution. Especially because they must operate at milli-kelvin temperatures, there are formidable engineering issues in developing them as practical and affordable astronomical tools, but there is no doubt about their superiority as detectors.

Summary

- Quantum mechanics accounts for a quantized pattern of permitted states for the energies, angular momenta, magnetic interactions, and spins of electrons bound to an isolated atom. Concepts

free state	*ground state*	*Pauli exclusion principle*
valence electron	*periodic table*	*quantum number*
fermion	*boson*	*spectroscopic notation*

- The outer (optical) electrons of an atom gain or lose energy by making transitions between permitted states. Concept:

excitation	*photo-emission*	*photo-ionization*
ground state	*photo-absorption*	*absorption edge*
thermal excitation	*Boltzmann distribution*	

- Permitted quantum states of isolated molecules are distinguished by the electronic states of their component atoms, but also by the quantized rotation and vibration states of the molecule. Concept:
  molecular absorption bands

- The energy states for electrons in solid-state crystals typically arrange themselves in continuous bands separated by forbidden band gaps. Concepts:

diamond lattice	*sp^3-hybrid orbitals*	*bonding state*
antibonding state	*valence band*	*conduction band*
holes	*semiconductor*	*intrinsic semiconductor*
electron degeneracy	*phonon*	*Fermi–Dirac statistics*
Fermi energy	*band-gap energy*	*cutoff wavelength*

- Adding small quantities of a selected impurity can produce desirable properties in a semiconductor. Concepts:

dopant	*extrinsic semiconductor*	*donor atom*
acceptor atom	*p-type*	*n-type*
impurity band		

- Photoconductors absorb a photon and create an electron–hole pair, thereby increasing the electrical conductivity of the material. Concepts:

absorption coefficient	*breakdown*	*absorption depth*

- The MOS capacitor is a block of extrinsic semiconductor separated from a metal electrode by a thin layer of insulation. With the proper voltage across the insulator, the device can store charges produced by photo-absorptions. Concepts:

SiO_2	*depletion region*	*potential well*
full-well capacity	*saturation*	

- The p–n junction produces a depletion region where photo-absorptions can generate charge carriers and an electric current. Concepts:

p–n junction	*recombination current*	*generation current*
diode	*forward bias*	*reverse (back) bias*
breakdown	*diode equation*	*p-n photodiode*
avalanche photodiode	*PIN photodiode*	*Schottky photodiode*

- Electrons can leave the surface of material in a vacuum if they have energies greater that the material's work function. Photons can supply the needed energy, and thus produce an electric current in a vacuum.

photocathode	*vacuum photodiode*	*anode*
electron affinity	*NEA photocathode*	

- A superconducting junction diode produces a number of conduction-band electrons that is proportional to the energy of the incoming photon. An SJD in pulse-counting mode can therefore measure both the intensity and the wavelength distribution of a source. Concepts:

Cooper pair	*Josephson junction*	*quasiparticles*

Exercises

1. Using the nl^x notation, write down the electron configuration for the ground state, first excited state, and third excited state of iron (atomic number 26) as suggested by Table 7.2.
2. There are several exceptions to the configuration-filling scheme presented in Table 7.2. The configuration of the ground state of copper is an example. Look up a table of electron configurations in atoms and find at least five other examples.
3. Suppose a certain diatomic molecule has an energy-level diagram similar to Figure 7.3 and consider only transitions within the $\Lambda = 0$ states. Suppose that relative to the ground state, state ($J = 1$, $v = 0$) has an energy of 1 eV. Suppose also that, no matter what the rotational state is, the relative energies of the lowest vibrational states are $v(v + 1)d$, where $d = 10^{-5}$ eV and v is the vibrational quantum number. (a) Compute the wavelengths of all permitted emission lines arising between levels $J = 0$ and $J = 1$, and involving vibrational states $v = 0, 1, 2, 3, 4$. The only permitted transitions are the ones in which $\Delta v = \pm 1$. (b) Sketch the emission spectrum for these lines.

4. Compute the relative probability of finding an electron at the bottom of the conduction band relative to the probability of finding an electron at the top of the valence band in a silicon crystal at a temperature of (a) 3 K and (b) 300 K. Use Fermi–Dirac statistics. Compare your answer with the one given by the Boltzmann equation.
5. Compute the fraction of incident photons absorbed by a 100-μm-thick layer of bare silicon if the photons have wavelength (a) 500 nm and (b) 800 nm. Assume the index of refraction of silicon is 4.4 at 500 nm and 3.8 at 800 nm.
6. How does an MOS capacitor made of an n-type semiconductor work? Why do you think p-type material is usually preferred for these devices?
7. Derive a relationship between the full-well capacity of an MOS capacitor and the maximum possible relative precision that the device can produce in a brightness measurement. What is the risk in planning to achieve this precision with a single measurement?
8. Assume you have a meter that measures electric current with an uncertainty (noise) of 100 picoamps. (One picoamp = 10^{-12} amp = 10^{-12} coulomb s^{-1}). You employ your meter with a photodiode in a circuit like the one in Figure 7.18a. You have a 2-meter telescope at your disposal, and use a filter to limit the light received to those wavelengths at which the detector is most sensitive. Compute the magnitude of the faintest star you can detect with this system. "Detect" in this case means the signal-to-noise ratio is greater than 3. Assume the photon flux from a zero-magnitude star in the bandpass you are observing is 10^{10} photons m^{-2} s^{-1}. Your photodiode detects 45% of the photons incident, and you may ignore any background signal.
9. In response to an incoming photon, a niobium-based STJ diode detects a pulse of 500 electrons. Assume tunneling operates with 100% efficiency, and the only source of noise is counting statistics. (a) Compute the energy of the incoming photon and its uncertainty. (b) What is the wavelength of the photon and its uncertainty? (c) Compute the spectroscopic resolution ($R = \delta\lambda/\lambda$) of this device as a function of wavelength. (d) Find the equivalent expression for a device in which the superconductor is hafnium instead of niobium.

Chapter 8
Detectors

> Honestly, I cannot congratulate you upon it. Detection is, or ought to be, an exact science, and should be treated in the same cold and unemotional manner. You have attempted to tinge it with romanticism, which produces much the same effect as if you worked a love-story or an elopement into the fifth proposition of Euclid.
>
> "But romance was there," I remonstrated.
>
> – Arthur Conan Doyle, *The Sign of the Four*, 1890

Astronomical detection, even more than the work of Sherlock Holmes, is an exact science. Watson, though, has an equally important point: no astronomer, not even the coldest and most unemotional, is immune to that pleasant, even romantic, thrill that comes when the detector *does* work, and the Universe *does* seem to be speaking.

An astronomical detector receives photons from a source and produces a corresponding ***signal***. The signal characterizes the incoming photons: it may measure their rate of arrival, their energy distribution, or perhaps their wave phase or polarization. Although detecting the signal may be an exact science, its characterization of the source is rarely exact. Photons never pass directly from source to detector without some mediation. They traverse both space and the Earth's atmosphere, and in both places emissions and absorptions may modify the photon stream. A telescope and other elements of the observing system, like correcting lenses, mirrors, filters, optical fibers, and spectrograph gratings, collect and direct the photons, but also alter them. Only in the end does the detector do its work. Figure 8.1 illustrates this two-stage process of signal generation: background, atmosphere, telescope, and instruments first modify light from the source; then a detector detects.

An astronomer must understand both mediation and detection if she is to extract meaning from measurement. This chapter describes only the second step in the measurement process, detection. We first outline the qualities an astronomer will generally find important in any detector. Then we examine a few important detectors in detail: the CCD, a few photo-emissive devices, the infrared array, and the bolometer.

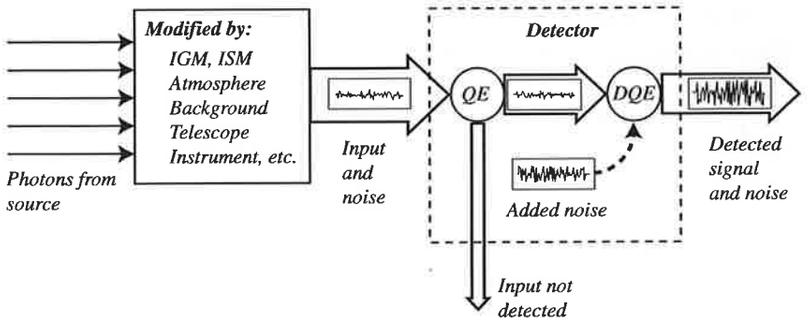


Fig. 8.1 Mediation and detection of a light signal (IGM = intergalactic medium, ISM = interstellar medium). The detection step may fail to record some of the mediated signal, and may introduce additional noise to the part of the signal that is recorded.

8.1 Detector characterization

Why does an astronomer choose one detector instead of another? Why did optical astronomers in the 1980s largely abandon photography, the then-dominant detector for imaging, in favor of solid-state arrays? Why are these same arrays useless for other purposes, such as measuring very rapid changes in brightness? Is there a *perfect* detector? We begin an answer with a list of several critical characteristics of any detector.

8.1.1 Detection modes

We can distinguish three distinct modes for detecting light.

Photon detectors produce a signal that depends on an individual photon altering the quantum-mechanical state of one or more detector electrons. For example, in the last chapter, we saw how a change in electron energy in a photoconductor or photodiode can produce a change in the macroscopic electrical properties like conductivity, voltage, or current. Other changes in quantum state might produce chemical reactions (as in photography) or a pulse of free electrons, as in vacuum photomultipliers. Photon detectors are particularly suited to shorter wavelengths (infrared and shorter), where the energies of individual photons are large compared to the thermal energies of the electrons in the detector.

Thermal detectors absorb the energy of the incoming photon stream and convert it into heat. In these devices the signal is the temperature change in the body of the detector. Although thermal detectors are in principle useful at all wavelengths, in practice, thermal detectors, especially a class called ***bolometers***, have been fundamentally important in the infrared and microwave regions, as well as very useful in the gamma and X-ray regions.

Wave detectors produce signals in response to the oscillating electric or magnetic field of the incoming electromagnetic waves, usually by measuring the interference effect the incoming fields have on a wave produced by a local oscillator. In principle, these detectors, unlike photon and thermal

detectors, can gauge the phase, intensity, and polarization of the detected wave. Wave detectors are especially useful in the radio and microwave parts of the spectrum.

8.1.2 Efficiency and yield

> Thou shalt not waste photons.
>
> – Anonymous, *c.* 1980

A good detector is efficient. We construct costly telescopes to gather as many photons as possible, and it seems perverse if a detector does not use a large fraction of these expensive photons to construct its signal.

Photography, for example, is relatively inefficient. The photographic detector, the emulsion, consists of a large number of tiny crystals, or ***grains***, of silver halide (usually AgBr) suspended in a transparent gelatin matrix. Photons can interact with a grain to eventually turn the entire grain into elemental silver. The more silver grains present in the emulsion after it has been processed, the stronger is the signal.

Why is the process inefficient? Some photons reflect from the surface of the emulsion and are not detected. Some pass right through the emulsion, while others are absorbed in its inactive parts without contributing to the signal. Nevertheless, silver halide grains absorb something like 40–90% of the incident photons. These absorbed photons produce photoelectrons that can induce a chemical change by reducing a silver ion to a neutral atom. The corresponding neutral bromine atom (the hole produced by photo-absorption) can vanish, either combining with the gelatin or with another bromine to form a molecule that escapes the crystal. Most holes do not vanish, however, and most photo-electrons recombine with holes before they can neutralize a silver ion. Some neutral silver atoms are created, but most are re-ionized by holes before the grain can be developed. Finally, it is only after three to six silver atoms drift and clump together at a spot on the grain that the crystal becomes developable. In the end, very few of the incident photons actually have an effect in photography. The process is inefficient.

The ***quantum efficiency***, ***QE***, is a common measure of detector efficiency. It is usually defined as the fraction of photons incident on the detector that actually contribute to the signal.

$$\mathrm{QE} = \frac{N_{\mathrm{detect}}}{N_{\mathrm{in}}} \tag{8.1}$$

In a perfect detector, every incident photon would be absorbed in a fashion that contributed equally to the signal, and the detector would have a QE of 100%.

Photographic emulsions have QE valuess in the range 0.5–5%.[1] Solid-state devices – like silicon photodiodes, superconducting tunnel junction (STJ) diodes, or metal-oxide-semiconductor (MOS) capacitors – have QE values in the 20%–95% range. Astronomers prefer these devices, in part, because of their high quantum efficiencies.

The quantum efficiency of a particular device is not always easy to measure, since (as in photography) the chain of events from incident photon to detection may be difficult to describe and quantify. ***Absorptive quantum efficiency*** is physically more straightforward, but somewhat less informative. It is defined as the photon flux absorbed in the detector divided by the total flux incident on its surface:

$$\eta = \frac{N_{abs}}{N_{in}}$$

Because absorbed photons are not necessarily detected, QE $\leq \eta$.

The ***quantum yield*** of a photon detector is the number of detection "events" per incident photon. For example, in silicon photoconductors, the detection event is the production of an electron–hole pair. If an incident photon has energy less than about 5 eV, it can produce at most one electron–hole pair, so the quantum yield is 1. For higher energy photons, a larger number of pairs are produced, around one e–h pair per 3.65 eV of photon energy. What happens in detail is that the first electron produced has so much kinetic energy that it can collide with the lattice to produce phonons that generate additional pairs. A 10-angstrom X-ray, therefore, will yield (on average) 34 photoelectrons. An STJ-based detector, you will recall, is particularly attractive because of its very large, wavelength-sensitive quantum yield.

8.1.3 Noise

> There are two kinds of light -- the glow that illuminates, and the glare that obscures.
>
> – James Thurber (1894–1961)

Although efficiency in a detector is important, what really matters in evaluating a measurement is its uncertainty. The uncertainty in the output signal produced by a detector is often called the ***noise***, and we are familiar with the use of the ***signal-to-noise ratio***, ***SNR***, as an indication of the quality of a measurement. It

[1] Quantum efficiency is a bit of a slippery concept in photography. For example, once a grain has formed a stable clump of three–six silver atoms, absorbed photons can make no further contribution to the signal, even though they create additional silver atoms. The entire grain is either developed or not developed depending only on the presence or absence of the minimum number of atoms. In photography, QE is thus a strong function of signal level – the highest efficiencies only apply if the density of developed grains is relatively low.

would seem that a perfect detector would produce a signal with zero noise. This is not the case.

You will recall that there is an uncertainly *inherent* in measuring the strength of any incident light ray. For a photon-counting device, this uncertainty arises from the Poisson statistics[2] of photon arrivals, and is just

$$\sigma = \sqrt{N}$$

where N is the number of photons actually counted. A perfect detector, with QE = 1, faithfully counts all incident photons and will therefore produce

$$(\text{SNR})_{\text{perfect}} = \frac{N_{\text{out}}}{\sigma_{\text{out}}} = \frac{N_{\text{in}}}{\sigma_{\text{in}}} = \sqrt{N_{\text{in}}}$$

Real detectors will differ from this perfect detector by either counting fewer photons (reducing the output noise, but also reducing both the output signal and the output SNR) or by exhibiting additional noise sources (also reducing the SNR). The ***detective quantum efficiency (DQE)*** describes this departure of a real detector from perfection. If a detector is given an input of N_{in} photons and has an output with signal-to-noise ratio $(\text{SNR})_{\text{out}}$, then the DQE is defined as a ratio:

$$\text{DQE} = \frac{(\text{SNR})^2_{\text{out}}}{(\text{SNR})^2_{\text{perfect}}} = \frac{N_{\text{out}}}{N_{\text{in}}} \tag{8.2}$$

Here N_{out} is a fictitious number of photons, the number that a perfect detector would have to count to produce a signal-to-noise ratio equal to $(\text{SNR})_{\text{out}}$. The DQE gives a much better indication of the quality of a detector than does the raw QE, since it measures how much a particular detector degrades the information content of the incoming stream of photons. For a perfect detector, DQE = QE = 1. For any detector, it should be clear from Equation (8.2) that DQE $\leq$ QE. If two detectors are identical in all other characteristics, then you should choose the detector with the higher DQE. If a parameter (wavelength of the observation, for example) affects both the input signal and the DQE, then you should choose a value that maximizes the value

$$(\text{Signal})_{\text{in}}\sqrt{(\text{DQE})} = (\text{SNR})_{\text{out}}$$

Returning to the example of the photographic emulsion, the noise in an image is experienced as ***granularity***: the microscopic structure of, say, a star image consists in an integral number of developed grains. Statistically, counting grains in an image is a Poisson process, and has an uncertainty and a SNR of $\sqrt{N_{\text{grains}}}$.

[2] Although we have been treating the photon-counting process as if it were perfectly described by Poisson statistics, both theory and experiment show this is not the case. Photon arrivals are not statistically independent – real photons tend to clump together slightly more than Poisson would predict. This makes little practical difference in the computation of uncertainties.

Since it takes something like 10–20 absorbed photons to produce one developed grain, the photographic process clearly degrades SNR. In addition, grains are not uniformly distributed in the emulsion, and some grains not activated by photons will nevertheless get developed to produce a background "fog." Both of these effects contribute noise, and thus reduce the DQE. A typical emulsion might have $\eta = 0.5$, QE = 0.04 and DQE = 0.02. Many solid-state detectors do not degrade the information in the input to anything like the degree that photography does, and their DQE values are close to their QE values – in the range 20–90%.

The DQE generally is a function of the input level. Suppose, for example, a certain QE = 1 detector produces a background level of 100 electrons per second. You observe two sources. The first is bright. You observe it for 1 second, long enough to collect 10,000 photoelectrons (so $SNR_{in} = 100$). For this first source, $SNR_{out} = 10,000/\sqrt{(10,100 + 10^2)} = 98$, and DQE = 0.96. The second source is 100 times fainter. You observe it for 100 seconds, and also collect 10,000 photoelectrons. For the second source, $SNR_{out} = 10,000/\sqrt{20,000 + 10,000} = 57.8$, and DQE = 0.33.

8.1.4 Spectral response and discrimination

The QE of a detector is generally a function of the wavelength of the input photons. Some wonderful detectors are useless or have low QE at some wavelengths. Silicon devices, for example, cannot respond to photons with $\lambda > 1.1$ μm since these photons have energies below the silicon band-gap energy. The precise relationship between efficiency and wavelength for a particular detector is an essential characteristic.

One can imagine an ideal detector that measures both the intensity and the wavelength distribution of the incoming beam. An STJ diode, operated in a pulse-counting mode, for example, ***discriminates*** among photons of different wavelength.

8.1.5 Linearity

In an ideal detector, the output signal is directly proportional to the input illumination. Departures from this strict linearity are common. Some of these are not very problematic if the functional relation between input and output is well known and well behaved. For example, in the range of useful exposures, the density of a developed photograph is directly proportional to the logarithm of the input flux. Figure 8.2 illustrates two very typical departures from linearity. At lower light levels, a detector may not respond at all – it behaves as if there were an input ***threshold*** below which it cannot provide meaningful information. At the other extreme, at very large inputs, a detector can ***saturate***, and an upper threshold limits its maximum possible response. Further increases in input will not move the output signal above the saturation level.

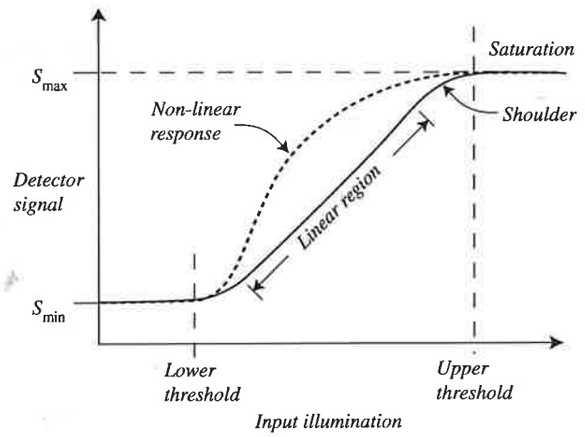


Fig. 8.2 Linear and non-linear regions in a typical detector response curve. The dashed response is completely non-linear.

8.1.6 Stability

The environment of a detector will change over time, perhaps because of variation in temperature, atmospheric conditions, or orientation with respect to gravity or to local magnetic fields. The detector itself may age because of chemical or mechanical deterioration, electrical damage, or radiation and particle exposure. Unrecognized changes can introduce systematic effects and increase uncertainties.

Two general approaches cope with detector instability. The first is to avoid or minimize anticipated changes: e.g. use thermostatic controls to maintain a constant temperature, keep the detector in a vacuum, shield it from radiation, use fiber-optic feeds so that the detector remains motionless. Basically, employ whatever strategies seem reasonable to isolate the detector from the environment. The second approach is to recognize that some changes are unavoidable and calibrate the detector to correct for the instability. For example, if the response of a detector deteriorates with age, make repeated observations of the same standard source so you can compute a correction that compensates for the deterioration.

Hysteresis is a form of detector instability in which the detector response depends on its illumination history. Human vision, for example, exhibits the phenomenum of positive and negative afterimage. Some solid-state detectors can continue to report ghost signals from bright objects long after the source has been removed.

8.1.7 Response time

How quickly can the detector make and report a measurement, then make and report the next measurement? The minimum time required is an important

parameter. Readout procedures for large CCDs, for example, can limit their response time to a hundred seconds or more, while STJs and photo-emissive devices have sub-millisecond response times.

8.1.8 Dynamic range

What is the maximum range in output signal that the detector will produce in response to input? From Figure 8.2, you might surmise (correctly) that the upper and lower detection thresholds limit the dynamic range. However, other details of the detection process can influence the dynamic range. For example, if the signal is recorded digitally as a 16-bit binary integer, then the smallest possible signal is 1, and the largest is 65, 535 ($= 2^{16} - 1$). Thus, even if the range set by saturation is larger, the dynamic range is limited by data recording to 1:65, 535.

8.1.9 Physical size and pixel number

The physical size of the detector can be very important. To measure the light from a single star in the telescope focal plane, for example, it will be advantageous to match the detector size with the image size produced by the telescope: if the detector is too small, it will not intercept all the light from the source; if it is too large, it will intercept unwanted background light and probably produce a higher level of detector noise. For some detectors, physical size is related to other properties like dynamic range and response time.

A ***single-channel*** detector measures one signal at a time, while a ***multi-channel*** detector measures several at once. An astronomer might use a simple two-channel detector, for example, to simultaneously measure the brightness of a source and the brightness of the nearby background sky. A ***linear array*** (a string of closely packed detectors arranged in a straight line) might be a good configuration for sensing the output of a spectrograph. A ***two-dimensional array*** of detectors can record all parts of an astronomical image simultaneously.

Clearly, the physical size of each detector of an array determines how closely spaced its elements, or ***pixels*** (for *picture element*) can be. Sometimes there must be some inactive area between the sensitive parts of the pixels, sometimes not. Large arrays are more easily manufactured for some types of detectors (e.g. MOS capacitors) than for others (e.g. bolometers and wave detectors). There is an obvious advantage in field of view for detectors with a large number of pixels.

Astronomers currently employ mosaics of solid-state arrays of up to one billion pixels, with the largest individual arrays (CCDs of up to 100 megapixels in size) finding application in the X-ray through optical regions. Somewhat smaller arrays (1–4 megapixel) are in use at near-infrared (NIR) and mid-infrared (MIR) wavelengths. Focal-plane arrays of hundreds of pixels are used on some far-infrared (FIR) and sub-millimeter telescopes. Radio detectors are almost always single-pixel or few-pixel devices. At the beginning of the CCD era, photographic plates had a clear advantage in pixel number: for a very moderate cost, a photographic plate had a very large area (tens of centimeters on a side), and thus, in effect, contained up to 10^9 pixels. Mosaics of CCD arrays, although quite expensive, now match the size of medium-sized photographic plates.

8.1.10 Image degradation

Astronomers go to extremes to improve the resolution of the image produced by a telescope – minimize aberrations, launch the telescope into space, and create active and adaptive optics systems. Two-dimensional detectors like arrays should preserve that resolution, but in practice can often degrade it. ***Sampling theory*** was originally developed to understand electronic communications in media such as radio broadcasting and music reproduction. The Nyquist theorem states that the sampling frequency of a waveform should be greater than two times the highest frequency present in the wave. Extending this theorem to the spatial domain means that to preserve maximum detail, pixel-to-pixel spacing should be less than the ***Nyquist spacing***. The Nyquist spacing is one-half the full width at half-maximum (FWHM) of the point-spread function of the telescope. If pixel spacing is larger than the Nyquist value, the resulting ***under-sampling*** of the image degrades resolution.

Other effects can degrade resolution. Signal can drift or bleed from its pixel of origin into a neighboring pixel, or photons can scatter within the array before they are detected.

8.2 The CCD

> One morning in October 1969, I was challenged to create a new kind of computer memory. That afternoon, I got together with George Smith and brainstormed for an hour or so. . . .When we had the shops at Bell Labs make up the device, it worked exactly as expected, much to the surprise of our colleagues.
>
> – Willard Boyle, Canada Science and Technology Museum, 2008

When Boyle and Smith (1971) invented the first ***charge-coupled devices*** at Bell Laboratories in 1969 they quickly recognized the CCD's potential as multi-pixel light detector instead of a computer memory. By 1976, astronomers had recorded the first CCD images of celestial objects.[3] Since that time, the CCD has become a standard component in applications that include scanners, copiers, mass-market still and video cameras, surveillance and medical imagers, industrial robotics, and military weapon systems. This large market has diluted the research and development costs for astronomy. The consequent rapid evolution

[3] The first CCD images reported from a professional telescope were of the planets Jupiter, Saturn, and Uranus, taken in 1976 by Bradford Smith and James Janesick with the LPL 61-inch telescope outside Tucson, Arizona.

of the scientific CCD has profoundly revolutionized the practice of optical observational astronomy. This section gives a basic introduction to the principles of operation of the CDD and its characteristics as a detector.

8.2.1 General operation

Recall how an MOS capacitor stores photoelectrons in a potential well. A CCD is an array of MOS capacitors (one capacitor per pixel) equipped with circuitry to read out the charge stored in each pixel after a timed exposure. This read-out scheme (called "charge-coupling") moves charges from one pixel to a neighboring pixel; pixel-by-pixel shifting is what makes the array a CCD, rather than something else.

The basic ideas behind the array operation are simple. Imagine a matrix of MOS capacitors placed behind a shutter in the focal plane of a telescope. To take a picture, we first make sure all the capacitor wells are empty, open the shutter for the exposure time, then close the shutter. While the shutter is open, each pixel accumulates photoelectrons at a rate proportional to the rate of photon arrival on the pixel. At the end of the exposure, the array stores an electronic record of the image.

Figure 8.3 describes how the CCD changes this stored pattern of electrons into a useful form – numbers in a computer. In Figure 8.3a we show the major components of the detector. There is the light-sensitive matrix of MOS capacitors: in this case an array three columns wide by three rows tall. A column of pixels in the light-sensitive array is called a ***parallel register***, so the entire light-sensitive array is known collectively as the parallel registers. There is one additional row, called the ***serial register***, located at the lower edge of the array and shielded from light. The serial register has one pixel for each column of parallel registers (in this case, three pixels). Both the serial and parallel register structures are fabricated onto a single chip of silicon crystal.

Reading the array requires two different charge-shifting operations. The first (Figure 8.3b) shifts pixel content down the columns of the parallel registers by one pixel. In this example, electrons originally stored in row 3 shift to the serial register, electrons in row 2 move to row 3, electrons in row 1 move to row 2. Just before this first shift is initiated, the serial register is cleared of any charges that may have accumulated before or during the exposure.

The second operation now reads the newly filled serial register by shifting its contents to the right by one pixel (Figure 8.3c1). The electrons in the rightmost pixel shift into a new structure – a series of ***output amplifiers*** – that ultimately converts the charge to a voltage. This voltage is in turn converted to a binary number by the next structure, the ***analog-to-digital converter (ADC)***, and the number is then stored in some form of computer memory. The CCD continues this shift-and-read of the serial register, one pixel at a time (Figures 8.3c2 and 8.3c3) until all serial register pixels have been read.

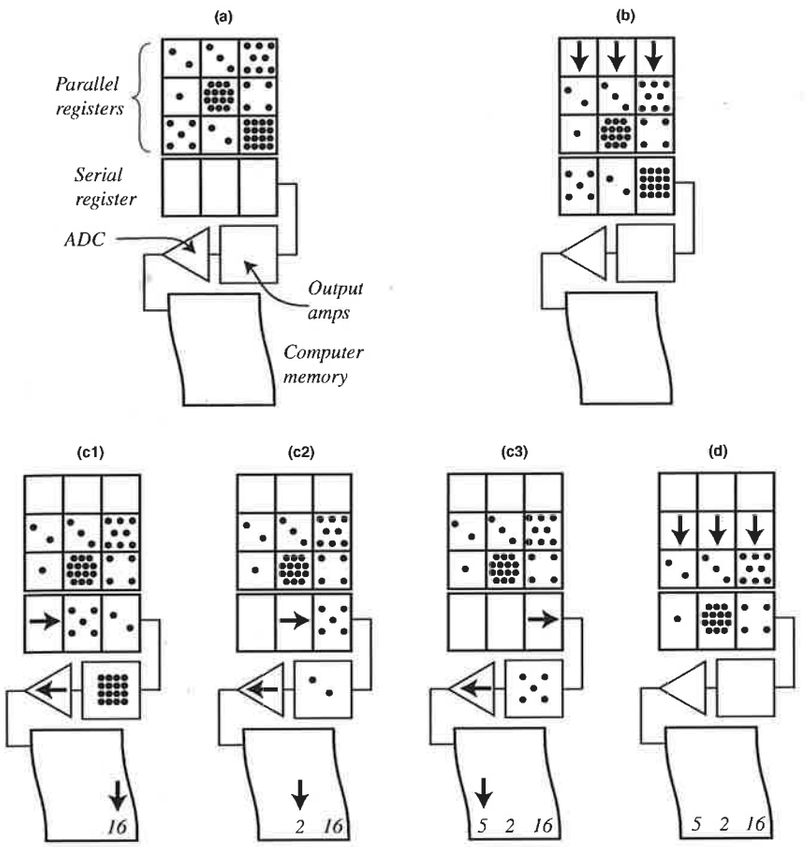


Fig. 8.3 CCD components and readout. (a) The accumulated photo-electrons in a 3×3 array of capacitors – the parallel register. (b) Shift of the bottom row into the serial register, all remaining rows shift down in the parallel register. (c) Read of the serial register one column at a time. (d) Next row shifts down into the empty parallel register.

Now the whole operation repeats for the next row: there is another shift of the parallel registers to refill the serial register with the next row (Figure 8.3d); the serial register is in turn read out to memory. The process continues (parallel shift, serial shifts, and reads) until the entire array has been read to memory. The first stage of the output amplifier is usually fabricated onto the same silicon chip as the registers. The subsequent amplifiers and the ADC are usually located in a separate electronics unit.

How does the CCD persuade the electrons stored in one capacitor to move to the neighboring capacitor? Many strategies are possible, all of which depend upon manipulating the depth and location of the potential well that stores the electrons. A parallel or serial register is like a bucket brigade. The bucket (potential well) is passed down the line of pixels, so that its contents (electrons) can be dumped out at the end. Figure 8.4 illustrates one strategy for moving the well. The depth of a potential well depends on the voltage applied to the metal, and is greatest at the Si–SiO_2 junction, closest to the metal layer. (See, however

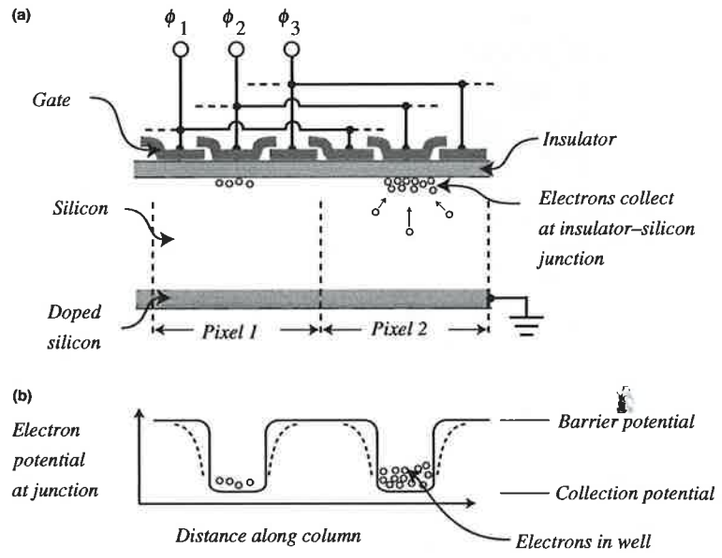


Fig. 8.4 Gate structure in a three-phase CCD. Two pixels are shown in cross-section. Collection and barrier potentials on the gates isolate the pixels from each other during an exposure. Overlapping gates produce a gradient in the barrier region (dashed curve in lower figure) that enhances collection.

the section below on the buried-channel CCD.) The figure shows two pixels in the same register (column) of a ***three-phase CCD***. In this device, the metal electrode is separated into three ***gates***, and these are interconnected so that gate 1 of every pixel connects to gate 1 of every other pixel, and likewise for gates 2 and 3. Thus, a single pixel can simultaneously have three separate voltages or ***phases*** applied to its front side, producing a corresponding variation in the depth of the potential well, as illustrated in the figure. The interconnection of gates insures that the pattern of well depth is identical in every pixel of the register.

Setting the correct voltages on three separate gates implements both charge-shifting and pixel isolation. For example, during an exposure, phase 2, the voltage on the central metal electrode, can be set to a large positive value (say 15 V), producing what is known as the ***collection potential*** in the semiconductor. The other two phases are set to a smaller positive voltage (say 5 V), which produces the ***barrier potential***. The barrier potential maintains the depletion region in the silicon, but prevents electrons from drifting across pixel boundaries. Photoelectrons generated in the barrier region of the silicon will diffuse into the nearest deep well under the collection phase and remain there. Each isolated pixel thus stores only charges generated within its boundaries.

To illustrate how the three gates might be used for charge shifting, assume again that the pixels are isolated during an exposure with collection under phase 2 ($\varphi_2 = +15$ V) and a barrier under the other phases ($\varphi_1 = \varphi_3 = +5$ V).

Figure 8.5 illustrates the three voltage changes that will shift charges by one pixel.

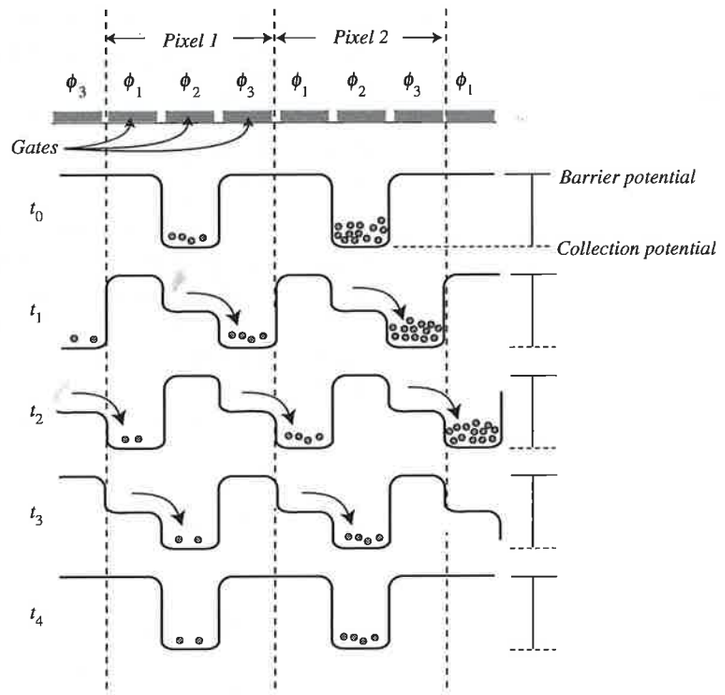


Fig. 8.5 Shifting potential wells in a three-phase CCD. See Figure 8.4 for the corresponding physical structure. Two pixels in the same register (either parallel or serial) are illustrated here. At the end of the shift, electrons stored in pixel 1 have shifted to pixel 2.

1. At time t_1, gate voltages change so that $\varphi_3 = 15$ V and $\varphi_2 = 10$ V. The electrons under φ_2 will diffuse to the right, and collect under φ_3.
2. At time t_2, after a delay that is long enough for all electrons to diffuse to the new location of the deep well, voltages change again, so that $\varphi_1 = 15$ V, $\varphi_3 = 10$ V and $\varphi_2 = 5$V. Stored electrons drain from phase 3 of the original pixel to phase 1 of the neighboring pixel.
3. A third cycling of gate voltages ($\varphi_1 = 10$ V, $\varphi_2 = 15$ V and $\varphi_3 = 5$ V) brings the electrons to the middle of the pixels at time t_3, and the one-pixel shift is complete.

The values of the barrier and collection potentials are somewhat arbitrary, but there are usually some fairly well-defined optimal values. These values, along with the properties of the insulator layer, determine required values of the ***clock voltages*** (the input values for φ_1, φ_2, and φ_3). An electronic system called the ***CCD controller*** or ***CCD sequencer*** sets the clock voltages and manages the very precise timing of their changes. The controller, usually built around a simple microprocessor, is generally housed in the same electronics box as the ADC and output amplifiers. Alternatively, the controller can be a program on a general-purpose computer. Besides manipulating the clock voltages, the controller also performs and coordinates several other functions, generally including:

- clearing the appropriate registers before an exposure and or a read;
- opening and closing the shutter;

- controlling the sequence of reads of the parallel and serial registers, including the patterns for special reads (see the discussions of on-chip binning and windowing below);
- controlling the parameters of the output amplifiers and the ADC (in particular, setting two constants called the ***bias level*** and the ***CCD gain*** discussed below);
- communicating with the computer that stores the data.

Two-phase and four-phase readout schemes are also sometimes used in CCDs. Most modern consumer digital cameras utilize arrays of (complementary-metal-oxide-semiconductor) ***CMOS*** capacitors, in which individual output amplifiers are fabricated onto the front side of each pixel. This design means that the pixels can be read out in parallel, rather than one at a time. The CMOS detectors are less expensive than CCDs of the same size, consume less power, and read out very rapidly (around 70 megapixels per second). They have not seen much use in astronomy, since they suffer from much higher read noise, dark current, pixel-to-pixel charge diffusion, and (usually) lower QE; however, they are gradually becoming more competitive with CCDs.

8.2.2 Channel stops, blooming, full well, and gain

The barrier potential prevents electrons from migrating from one pixel to another along a column in the parallel registers. What about migration along a row? In a classical CCD, shifts along a row are never needed, except in the serial register. The CCDs prevent charge migration along a row in the parallel registers by implanting (by heavily diffusing a dopant) a very highly conductive strip of silicon between columns. These ***channel stops*** held, say, at electrical ground, produce a permanent, extra-high barrier potential for stored electrons. Think of a pixel as a square bucket that holds water (or electrons). Two sides of the bucket, those that separate it from the adjacent columns, are maintained by the channel stop and are permanently tall and thin. The other two sides, the ones that separate it from its neighbors on the same column, are not as tall, and can be lowered or moved by "clocking" the gate voltages.

Consider what might happen if a pixel in an array fills with electrons during an exposure. As additional photoelectrons are generated in this saturated pixel, they will be able to spill over the barrier potential into the adjacent wells along their column, but cannot cross the channel stop. This spilling of charge along a column is called ***blooming*** (see Figure 8.6). Bloomed images are both unattractive and harmful: detection of photons in a pixel with a filled well becomes very non-linear; moreover, blooming from a bright source can ruin the images of other objects that happen to lie on the same CCD column. Nevertheless, in order to optimize the exposure of fainter sources of interest, astronomers will routinely tolerate saturated and bloomed images in the same field.

Fig. 8.6 Blooming on a CCD image: the saturated vertical columns are the bloom. The other linear spikes on the bright star image result from diffraction by the vanes supporting the telescope's secondary mirror.

There are designs for ***anti-blooming CCDs***. Recent designs utilize special clocking during the exposure in a buried-channel CCD (see below) to temporarily trap excess electrons at the oxide interface.

The maximum number of electrons that can be stored in a single pixel without their energies exceeding the barrier potential is called the CCD's ***full well***. The size of the full well depends on both the physical dimensions of the pixel, design of the gates, and the difference between the collecting and barrier potentials. Typical pixels in astronomical CCDs are 8–30 μm on a side and have full-well sizes in the range 25,000 to 500,000 electrons.

The final output from a scientific CCD is an array of numbers reported by the ADC to the storage computer. The number for a particular pixel is usually called its ***pixel content***, and is measured in ***ADU***s (analog-to-digital units). Pixel contents are proportional to the voltage the ADC receives from the output amplifier. The ***gain*** of the CCD is the number of electrons that need to be added to a pixel in order to increase the output contents for that pixel by one ADU.

For example, suppose a particular CCD has a full well of 200,000 electrons, and is equipped with a 16-bit ADC. The ADC is limited to digital outputs between 0 and 65,535 ($= 2^{16} - 1$). A reasonable value for the gain might be 200,000/65,535 = 3.05 electrons/ADU. A smaller gain would mean that the CCD is better able to report small differences in pixel content, but would reach ***digital saturation*** before reaching the electronic full well. One might do this intentionally to avoid the non-linear shoulder in Figure 8.2. At a larger gain, the CCD would reach full well before the output could reach the maximum possible digital signal, so dynamic range would be reduced.

8.2.3 Readout time, read noise, and bias

To maximize DQE, the amplifier and ADC of an astronomical CCD should introduce the smallest possible noise to the output signal. A technique called

correlated double sampling (***CDS***) is capable of very low noise operation – only a few electrons per pixel. The noise added by the CDS circuit depends crucially on how quickly it does its job – the faster, the noisier. Another consideration – the time needed for the analog-to-digital conversion – also limits the read time per pixel. Practical times correspond to a pixel sample frequency of 10 to 200 kHz, with higher frequencies producing higher noise. Except for low frequencies, noise added by the amplifier stage is proportional to the square root of the frequency.

The basis of charge-coupled readout is the one-pixel-at-a-time movement of the array contents through a single amplifier, and this is a bottleneck. A low-noise CDS stage in a scientific CCD must read out slowly, and the larger the array, the longer the read time. An important difference between scientific-grade CCDs and the commercial-grade CCDs and CMOS arrays in camcorders is the readout rate – to obtain real-motion video images, an array must read out about 30 times a second. The large read noise that results is usually not objectionable in a consumer camera because of the high input level. In contrast, the astronomical input signal is usually painfully low, and a low-noise, ***slow-scan*** CCD for astronomy may require many tens of seconds to read a single image.

There are some cases in astronomy where the large read noise of a rapid scan CCD is not objectionable, and in which time resolution is very important – observations of occultations of bright stars or rapid changes in solar features, for example. Also note that a rapid scan is not a problem if no data are being digitized. Thus, reading an array to clear it before an exposure can be done very quickly.

For the usual astronomical tasks, though, it is mainly lengthy readout time that puts a practical limit on the number of pixels in a CCD. (Time spent reading the detector is time wasted at the telescope!) Two strategies can speed read times. The first uses multiple amplifiers on a single array. Imagine, as in Figure 8.7a, an array with an amplifier at each corner. The CCD has two serial registers, at the top and bottom. The controller clocks the readout to split the parallel registers – they read out to both ends simultaneously – and does the same with each serial register. Each amplifier reads one quarter of the array, so the total read time is reduced by the same factor. The image can then be re-assembled in software. Multi-amplifier astronomical CCDs up to 9000 × 9000 pixels now (2010) exist.

A second strategy is to build a mosaic of several very closely spaced but electrically independent CCDs. Figure 8.7b shows an eight-element mosaic read by 16 amplifiers. An early device similar to this, the *Mosaic Imager*, was placed in service at the Kitt Peak National Observatory in 1998. It contained eight 2048 × 4096 CCDs arranged to form an 8196 × 8196 pixel (64 megapixel) detector that is 12 cm (5 inches) on a side. Gaps between the individual CCDs are about 0.6 mm (40 pixels). A relatively simple combination of shifted multiple exposures will fill in those parts of an image masked by the gaps on a single

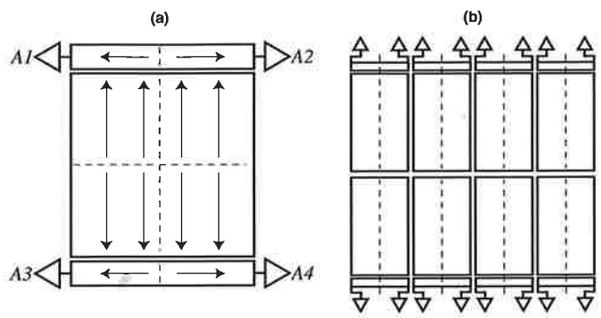


Fig. 8.7 Large-format CCD strategies. (a) A large monolithic detector with multiple serial registers and amplifiers (four, in this case). Read time is reduced by a factor equal to the number of amplifiers, and the total CTE is improved. (b) A mosaic of eight arrays butted to form a single large-area detector.

exposure. Mosaics have become so important that some modern CCDs are manufactured to be "almost-four-side-buttable" – so that the width of the gaps in a mosaic need be only to 30–100 pixels on all sides. At the present time (2009), there are several 50–120 megapixel mosaic arrays in service, and several observatories are about to introduce mosaics of over 100 devices and up to 3 gigapixels (see chapter 4 of Howell, 2006). These huge arrays expect to have not-very-objectionable readout times in the 20–60 second range. A major problem with these large-format arrays, in fact, may turn out to be simple data storage: an observer can expect to generate terabytes of image data in a few nights.

8.2.4 Dark current, cooling, and vacuum enclosures

At room temperature, a CCD is a problematic detector for astronomy. The energy of thermal agitation generates electron–hole pairs in the depletion zone and the resulting steady flow of electrons into the CCD potential wells is called ***dark current***. Dark current is bad for three reasons:

1. It adds some number of electrons, N_D, to whatever photoelectrons are produced in a pixel. You must make careful calibrations to subtract N_D from the total.
2. Dark current adds not only a background *level*, N_D, but also introduces an associated uncertainty or *noise* to any signal. Since the capture of dark-current electrons into the pixel wells is a random counting process, it is governed by Poisson statistics. The noise associated with N_D dark electrons should be $\sqrt{N_D}$. This noise is more insidious than the background level, since it can never be removed. Dark current *always* degrades SNR.
3. At room temperature, the dark current can saturate a scientific CCD in seconds, which makes it impossible to record faint objects. Not good.

Lower the temperature of the CCD, and you reduce dark current. The Fermi distribution provides an estimate for the rate at which dark charges accumulate in a semiconductor pixel:

$$\frac{dN_D}{dt} = A_0 T^{\frac{3}{2}} e^{-\frac{E_G}{2kT}}$$

Here T is the temperature in kelvins, A_0 is a constant that depends on pixel size and structure, and E_G is the band-gap energy. A large fraction of dark current in a pixel arises at the Si–SiO_2 interface of the capacitor, where discontinuities in the crystal structure produce many energy states that fall within the forbidden band. Electrons in these interface states have small effective band gaps, and hence produce a large dark current.

A common method for cooling a CCD is to connect the detector to a ***cryogen*** – a very cold material with a large thermal mass. A very popular cryogen is a bath of ***liquid nitrogen*** (***LN_2***), a chemically inert substance that boils at 77 K = –196 °C. Since it is generally a good idea to keep the CCD at a somewhat warmer temperature (around –100 °C), the thermal link between detector and bath is often equipped with a heater and thermostat.

A cold CCD produces difficulties. The CCD and the LN_2 reservoir must be sealed in a vacuum chamber for two reasons. First, a CCD at –100 °C in open air will immediately develop a coating of frost and other volatiles. Second, the vacuum thermally insulates the LN_2 reservoir from the environment, and prevents the supply of cryogen from boiling away too rapidly. Filling the CCD chamber with an inert gas like argon is a somewhat inferior alternative. Vacuum containers, called ***Dewars***, can be complicated devices (see Figure 8.8), but are quite common in observatories. At a minimum, the dewar must provide a transparent window for the input, a method for feeding electrical signals though the vacuum seal, a system for adding cryogen, and a method for periodically renewing the vacuum.

Another option for more modest cooling is dry ice (solid CO_2), which is less expensive than LN_2. Dry ice sublimates at –76 °C = 197 K.

Compact and relatively inexpensive thermoelectric (***Peltier junction***) coolers instead of cryogens require very small dewar sizes. These solid-state coolers can maintain a detector in the –30 to –50°C range, where the dark current of an ordinary CCD is still quite high, but where the dark current from an MPP CCD (see below) is acceptable for many astronomical applications. Such coolers are considerably more convenient to use than cryogens.

At the other extreme, superconducting junctions, many small band-gap detectors for the infrared, and most bolometers, require temperatures below what liquid nitrogen provides. ***Liquid helium***, which boils at 4.2 K, is an expensive cryogen that is difficult to handle. Liquid ^{3}He boils at 3.2 K, but is even more difficult and expensive. To avoid the expense of evaporating helium into the air, one option is a ***closed-cycle refrigerator*** that compresses and expands helium fluid in a cycle. If they employ two or three stages, these systems can cool detectors to the 10–60 K range. Special closed systems using helium-3 evaporation can bring small samples to temperatures in the 0.3–3.2 K range.

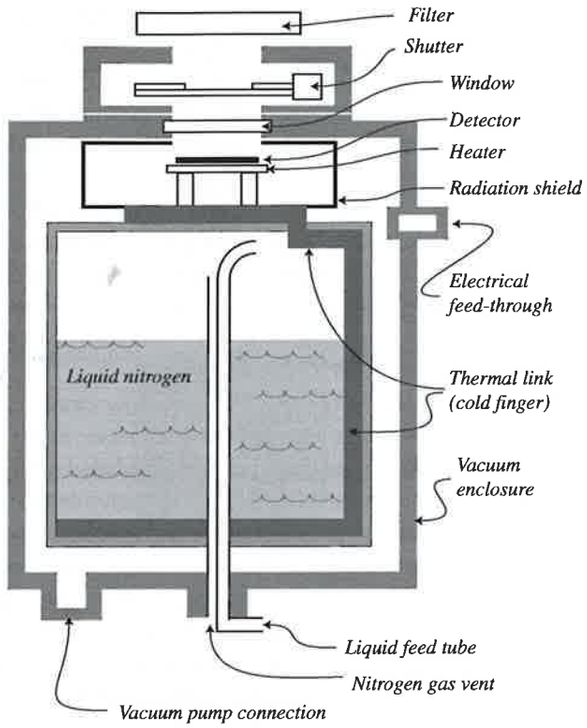


Fig. 8.8 A simple dewar for cooling a detector using liquid nitrogen. This design is common for devices that "look upward," and prevents cryogen from spilling out of the reservoir as the dewar is tilted at moderate angles.

8.2.5 Charge-transfer efficiency

The charge-coupled readout works perfectly only if all the electrons in a well shift from pixel to pixel. Disaster results if significant numbers of electrons are left behind by a shift. Images will appear streaked, and photometry becomes inaccurate. Signal loss because of charge-transfer inefficiency is greatest from the pixels furthest from the amplifier. The fraction of electrons in a pixel that are successfully moved during a one-pixel transfer is the ***charge-transfer efficiency***, or ***CTE***. Although one transfer will require three clock cycles and sub-pixel transfers in a three-phase device, CTE is always computed for a full pixel transfer. In a single-amplifier CCD, p is the actual number of full pixel transfers needed to read a particular charge packet. If the rows and columns of the parallel registers are numbered from the corner nearest the amplifier, then $p = R + C$, where R and C are the row and column numbers of the pixel in question. The fraction of the original charge packet that remains after p transfers (the total transfer efficiency, or TTE) is just

$$\mathrm{TTE} = (\mathrm{CTE})^p$$

The CTE needs to be very close to one. For example, suppose a 350 × 350 pixel array has a CTE of "three nines" (CTE = 0.999), which in this context is *not* very close to 1. Then $p = 350 + 350 = 700$, so TTE = $(0.999)700 = 0.49$; this device will lose over half the charge from the most distant pixel in the array before bringing it to the amplifier. Multi-megapixel arrays require CTE values approaching six nines.

What limits CTE? One issue is time – when CCD gate voltages change during a read, electrons need time to diffuse into the new location of the potential well. Usually, the required time is shorter than the time needed for the CDS and amplifiers to complete a low-noise read. However, at very low temperatures, electron velocities can be so small that CTE suffers because of slow diffusion, and so operation below about –100 °C is inadvisable.

Charge ***traps*** are a more serious limitation. A trap is any location that will not release electrons during the normal charge-transfer process. Some traps result from imperfections in the gates, channel stops, or the insulation of a pixel – flaws that deform the potentials during a read cycle to create unwanted barriers. Other traps are due to radiation damage, to unintended impurity atoms (usually metals like iron or gold), to structural defects in the silicon lattice, and to some effects not completely understood. The surface of the silicon layer in contact with the insulator will invariably have a large number of charge traps; these are such a serious problem that all modern CCDs are designed so that the potential well excludes the front surface (see the next section). Some traps affect only a few electrons each. If scattered throughout the entire body of a CCD, they produce a small decrease in the overall CTE. Other traps can render a pixel non-functional, so that it will not transfer charge in a meaningful way. This compromises the entire column upstream from the trap. Devices with a "bad column" or two are still very useful, but place additional demands on the observing technique.

Manufacturing defects can also cause a complete failure of charge transfer. The usual problems are short circuits and open circuits in the gate structure, or shorts between a gate and the semiconductor. Any of these can render a single pixel, a partial or complete column, or an entire device unreadable. The expense of a particular CCD is directly related to the ***manufacturing yield*** – if many devices in a production run need to be discarded, the cost of a single good device must rise. In the early days of CCD manufacture, yields of acceptable devices of a few percent were not uncommon.

8.2.6 The buried-channel CCD

The simple MOS/MIS (metal-insulator-semiconductor) capacitor we have been discussing up until now has its minimum electron potential (i.e. the bottom of the collection well) at the Si–SiO_2 interface. A CCD made of these capacitors is a ***surface-channel*** device, since charge transfer will require movement of electrons close to the interface. The high density of trapping

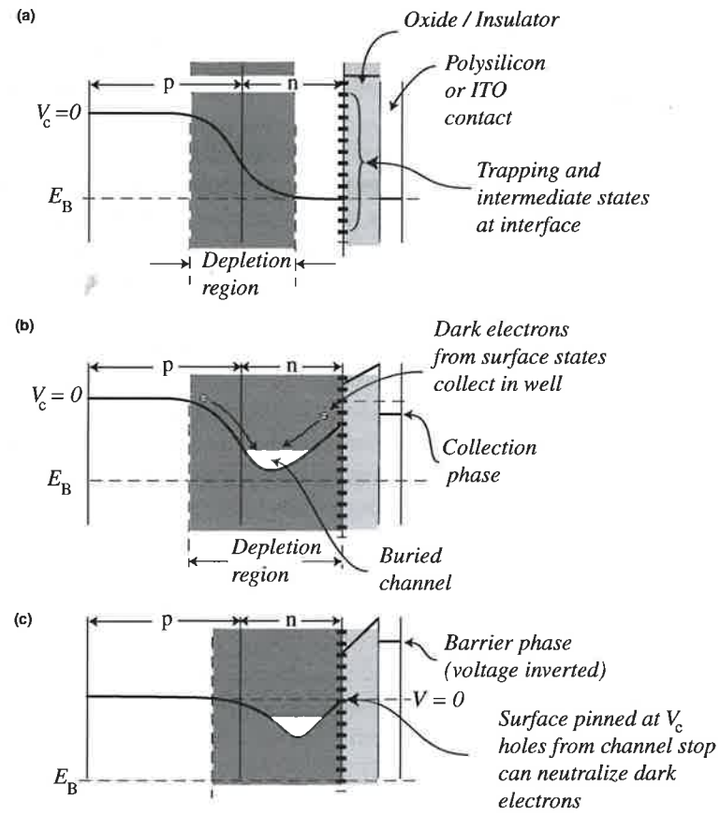


Fig. 8.9 A buried channel in a p–n junction capacitor. (a) There is no buried channel in the electron potential when the normal collection phase voltage is applied. If the gate voltage is reduced, as in (b), electrons collect away from the interface. (c) Inverting the voltage on the barrier-phase electrodes pins the surface potential to the channel-stop value and allows a current of holes to flow to neutralize dark-current electrons.

states at the interface makes it impossible to achieve acceptable charge-transfer efficiency in a surface-channel CCD. (Values of only 0.99 are typical.) All modern scientific CCDs are designed so that the transfer channel is located several hundred nanometers below the interface. In these ***buried-channel CCDs*** (***BCCDs***), all electrons collect in a region safely removed from the surface traps, and all charge transfers take place within the unperturbed interior of the semiconductor lattice.

Manufacturers can produce a buried channel by constructing a p–n junction near the semiconductor surface. Figure 8.9 illustrates the basic principle. Figure 8.9a shows the potential energy for electrons in an MOS or MIS device in which the semiconductor consists of a thin n-type region (perhaps 300–800 nm thick) layered on top of a much thicker p-type region. Within the semiconductor, the potential exhibits the basic pattern for a junction diode – there is a high-resistivity region depleted of majority charge carriers near the junction, and a potential difference, E_B, across the depletion zone. In Figure 8.9a we connect the p side to electrical ground, and set the gate voltage a relatively large positive voltage near

E_B. In this state, photoelectrons created in the depletion zone will be swept into the broad channel in the n region, where they can still interact with surface traps. Making the gate voltage even more positive will deepen the well and create a surface channel.

To create the buried channel, the voltage on the gate is made *more negative*. In Figure 8.9b, the gate voltage has been lowered so that electrons are repelled from the surface. This alters the shape of the potential and produces a minimum in the n-region, which is called the ***collection potential***. (The required voltage on the gate is the ***collection phase***.) Note two important features: First, electrons that collect in the well do not contact the surface. (This is good.) Second, the capacity of the well is reduced compared to a surface-channel device made from the same material. (This is not so good).

Figure 8.9c illustrates the electron potential under the barrier phase. Here the gate voltage is even more negative. The potential minimum in the semiconductor, although somewhat closer to the surface, is still buried. As a result, electrons generated under the barrier phase also avoid the surface traps as they move internally to the nearest collection potential.

8.2.7 Alternative CCD readout designs

You should be aware of several alternative methods for reading out the CCD that offer some specialized advantages. Consult Howell (2006) or the manufacturers and observatory websites (e.g. pan-STARRS, e2v, Kodak; see Appendix I) for further details.

The ***orthogonal-transfer CCD***, or ***OTCCD*** (see Tonry *et al.*, 1997) has a gate structure that permits charge-coupled shifting of pixel contents either along the row or along the column, on either the entire array or on subsections. Orthogonal-transfer CCDs can make small image shifts to compensate for tip–tilt seeing-disk motion during an exposure, and are being used for the 1 gigapixel mosaic of the pan-STARRS project.

Frame-transfer CCDs permit a very short time interval between successive frames. They recognize that it is the amplifier stage that limits the readout rate of a scientific CCD, so rapidly read an acquired frame into an inactive (shielded) set of parallel registers. The device then reads the shielded frame slowly through the amplifier while the next frame is being acquired.

Low-light-level CCDs or ***L3CCDs*** have additional extra-large, deep-well MOS capacitors in a "charge multiplication" extension of the serial register. The device clocks charges from the serial register into these capacitors at a very high voltage, so that the energy of a transferred electron can produce an additional electron–hole pair when it enters a multiplication capacitor. Several hundred multiplication transfers typically produce multiplication gains of 100–1000 before amplification, so read noise is insignificant, permitting rapid readout (1–10 MHz) and true photon-counting at low light levels.

8.2.8 The MPP CCD

Interface states at the Si–SiO_2 junction remain the major source of dark current in a simple BCCD. Thermal electrons can reach the conduction band by "hopping" from one interface state to another across the forbidden gap. You can eliminate this electron hopping by ***pinning*** a phase, as in Figure 8.9c. To pin the phase, you set the voltage on the gate to so negative a value that the potential at the interface ***inverts***, that is, it reaches the same potential as the back side of the p region, which is also the same, V_c, as the potential of the conductive channel stops. Any further reduction in the gate voltage has little effect on the interface potential, since the surface is now held at ground by holes that flood in from the channel stops. The abundance of holes means that thermal electrons are neutralized before they can hop through the interface states. Dark current in a pinned phase is reduced by several orders of magnitude.

A ***partially inverted*** three-phase CCD operates with one non-inverted phase (the collection phase, as in Figure 8.9b), and with the other two phases pinned and serving as the barrier phases, as in Figure 8.9c. Dark current in such a device is about one third of what it would be in a completely non-inverted mode. If all three phases are pinned, the CCD is a ***multi-pinned-phase (MPP)*** device, and dark current less than 1% the rate in non-inverted mode. The obvious difficulty with MPP operation is that there is no collection phase – the buried channel runs the entire length of a column. Multi-pinned-phase devices therefore require additional doping under one of the phases to make a permanent collection potential. This is possible because the value of E_B in Figure 8.9 depends on the density of dopants in the semiconductor. In an MPP device, for example, the surface under phase 2 might invert with the collection phase set at –5 V, while the other two (barrier) phases require –7 V for inversion.

With their remarkably low dark currents, MPP CCDs can operate at room temperature for several minutes without saturation. In recent designs, dark rates below 0.1 electron per second are routine at –40 °C, a temperature attainable with inexpensive thermoelectric coolers. An MPP CCD controlled by a standard personal computer is a formidable and inexpensive astronomical detector within the financial means of many small observatories, both professional and amateur. As a result, modern observers using telescope apertures below 1 meter are making quantitative astronomical measurements of a kind that would have been impossible at the very best observatories in the world in 1975.

The full-well capacity of an MPP device is a factor of two or three less than a partially inverted BCCD. Modern MPP devices nevertheless have respectable full wells. Appendix I gives the specifications for a few devices currently on the market. If the larger full well is more important than the reduced dark current, the proper selection of clock voltages makes it possible to run a device designed for MPP operation in a partially inverted mode.

8.2.9 Surface issues

We need to address the very practical question of getting light into the depletion region of the CCD pixels.

Frontside options. The most direct approach sends light through the metal gates. Since even very thin layers of most metals like copper or aluminum are poor transmitters, the "metal" layer of the CCD is usually made of highly doped ***polysilicon***: silicon in a glass-like, amorphous state – a random jumble of microscopic crystals. A thin (about 0.5 micron) layer of doped polysilicon is both relatively transparent as well as a good electrical conductor, but it does, however, absorb green, blue and (especially) ultraviolet light. Other conductive materials, like doped ***indium tin oxide*** (***ITO***) have better transparency properties than polysilicon; ITO electrodes are becoming common, but are somewhat harder to fabricate.

There are two general strategies for further improving the short wavelength QE of a front-illuminated CCD. The first is somehow to make the gate structure more transparent. The second is to change the wavelength of the incoming light to one at which the gates are more transparent.

Open-electrode architecture improves transparency with a gate structure that leaves part of the pixel uncovered. For example, the collection-phase electrode might be oversized and shaped like a hollow rectangle. It is even possible to fabricate pixel-sized ***microlenses*** over the frontside to redirect much of the incoming light to the uncovered area of each pixel.

A related approach is the ***virtual-phase CCD***, where a single gate covers half of the pixel, and a four-step potential profile is constructed by implanting dopants in the semiconductor. Changing the voltage on the single gate can produce pixel-to-pixel charge transfer similar to a four-phase CCD. Virtual-phase CCDs have even better blue QEs than open-electrode devices, especially if equipped with microlenses, but are more difficult to fabricate and generally have relatively poor CTE values.

A different strategy applies a thin coating of ***phosphor*** on top of the gates. The useful phosphors are organic molecules that absorb a short-wavelength photon to move to an excited state, then de-excite by emitting one or more longer-wavelength photons. Lumigen (or lumogen), for example, is a commercial compound that absorbs light shortward of 420 nm, and is otherwise transparent. Upon de-excitation, it emits photons at around 530 nm, which can easily penetrate polysilicon gates. Since phosphors emit in all directions, they will slightly degrade image resolution at short wavelengths. Another drawback is that some phosphors tend to evaporate in a vacuum, especially at high temperatures.

Backthinning. A completely different solution sends the light in through the back (from the bottom of Figure 8.10) of the device, avoiding the gates completely. This ***backside illumination*** has the advantage that green, blue, and ultraviolet, which would be absorbed by a polysilicon or ITO layer, will pass

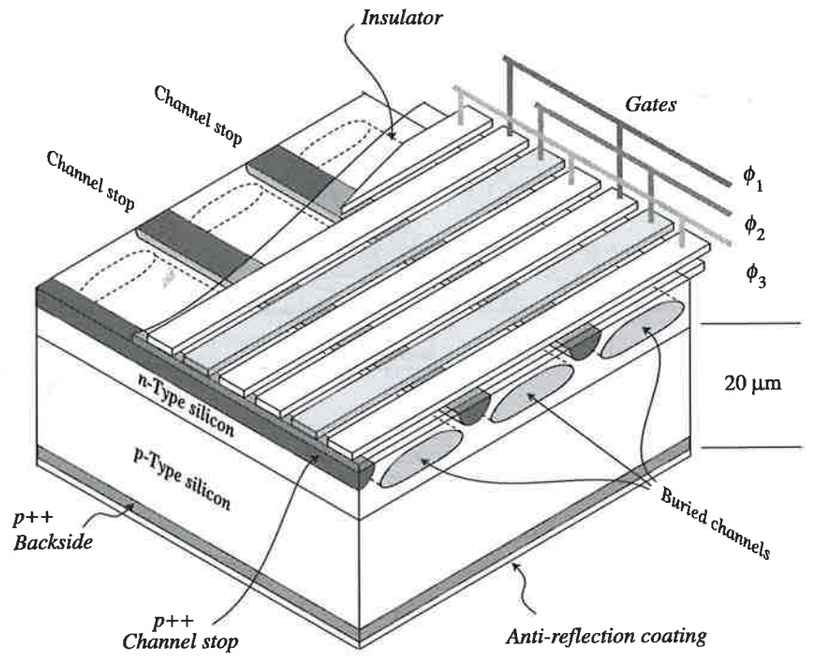


Fig. 8.10 Schematic of a thinned, three-phase CCD. In a conventional CCD, insulated gate electrodes usually overlap, while in an open architecture, gaps more closely follow the pixel pattern. This drawing is of a backthinned device. A front-illuminated device would have a much thicker silicon layer, with the AR coating above the gates.

directly into the silicon. Since these photons have a short absorption depth, they create photoelectrons mainly near the back face of the device. This is a serious problem. In order for the electrons to be able to diffuse from the back face into the depletion zone without recombining, the semiconductor layer needs to be very thin (10–20 μm). "Thinning" the silicon will in turn reduce its ability to absorb NIR photons, which have a large absorption depth. The final geometry needs to be something of a compromise. Nevertheless, astronomers have generally embraced backthinned CCDs, since they detect a considerably larger fraction of incident photons of all wavelengths than does any frontside-illuminated device (see Figure 8.11). Their main drawback is that they are difficult to manufacture and therefore expensive, if available at all.

If red and near-infrared QE is very important, the ***deep-depleted CCD*** offers some improvement over the normal backthinned device. Because the depth of the light-sensitive depletion zone is inversely proportional to the dopant concentration, use of a lightly doped (high resistivity) silicon layer means that the total layer thickness of the CCD can be increased to about 50 μm. The thicker detector has greater long-wavelength sensitivity, and is mechanically easier to fabricate. However, achieving the required resistivity can be difficult, and cosmetic quality to date has been inferior to thin devices.

Anti-reflection coatings. An anti-reflection (AR) coating is most effective for light of a particular wavelength, so a CCD designer must choose the coating

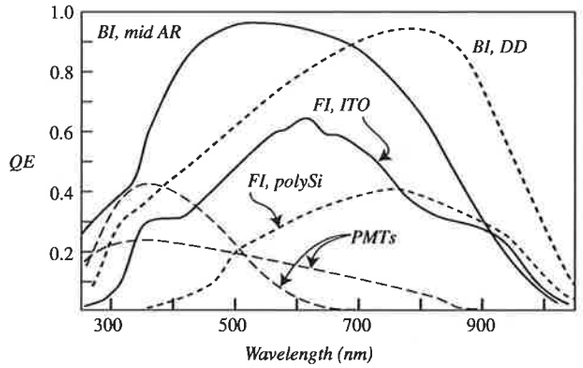


Fig. 8.11 Efficiencies of light detection for various illumination strategies in a CCD, and photocathode choices in a PMT. Curves are representative of the extremes. The abbreviations BI and FI indicate back- and front-illuminated CCDs. The figure shows the QE curves for a normal thinned device with a mid-band AR coating, a deep-depletion (DD) CCD with a near-infrared coating, ITO and polysilicon front-illuminated CCDs. The two photomultiplier tubes (PMTs) are very high efficiency bi-alkali photocathodes with different spectral sensitivities.

with the intended use of the detector in mind. Often CCD manufacturers offer a choice of coatings to enhance either the short-wavelength, mid-wavelength or NIR response. Figure 8.11 shows a selection of the QE characteristics of a few different CCD designs.

8.3 Photo-emissive devices

Researchers have developed the simple vacuum photodiode described in the last chapter from a detector of very limited capability (with poor QE in the red, very low signal levels, mechanical fragility, and single-channel operation) into devices that compete with or enhance CCDs in special circumstances. In this section we examine three astronomical detectors that depend upon the vacuum photoelectric effect.

8.3.1 The photomultiplier tube

One disadvantage of the simple vacuum photodiode described in the last chapter (Figure 7.24) is low signal level. The ***photomultiplier tube*** (***PMT***) is a vacuum device that increases this signal by several orders of magnitude. Figure 8.12 illustrates its operation. In the figure, a voltage supply holds a semi-transparent photocathode on the inside of the entrance window at large negative voltage, usually around one or two kilovolts. A photon hits the cathode and ejects a single electron. In the vacuum, this electron accelerates towards the more positive potential of a nearby electrode called a ***dynode***, which is coated with a material (e.g. Cs_3Sb, CsKSb, BeO, GaP) that can easily release electrons to the vacuum if hit by an energetic particle. Because the original photoelectron impacts the dynode with 100 eV or so of kinetic energy, it usually ejects several secondary electrons. The number of secondary electrons is a statistical quantity whose mean value, δ, usually lies between 2 and 10. The group of electrons

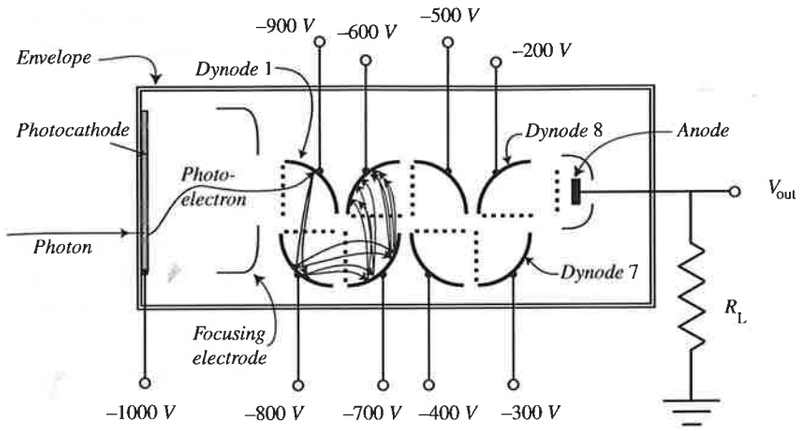


Fig. 8.12 A simple photomultiplier tube. The potential of the first dynode accelerates a single photoelectron emitted from the cathode. Its impact releases several secondary electrons, which accelerate and hit dynode 2, releasing another generation of secondaries. After (in this case) eight stages of amplification, a large pulse of electrons flows through the anode and load resistor to ground.

ejected from the first dynode then accelerates to the second dynode, where each first-dynode electron produces δ second-dynode electrons. The process continues through n dynodes, until the greatly multiplied pulse of electrons lands on the anode of the PMT. If each dynode is equivalent, the total number of electrons in a pulse generated by a single photoelectron is

$$N = A\delta^n$$

where the factor A accounts for inefficiencies in redirecting and collecting primary and secondary electrons.

In the figure, the signal is the average DC voltage measured across a load resistor. However, for weak sources, the large pulses of electrons that arrive at the anode are easily counted electronically, and the PMT can operate in a ***pulse-counting mode***: each pulse is generated by the arrival of a *single* photon at the cathode. In this mode, the QE of the PMT depends on the QE of the photocathode, which can be as high as 40–50% for some materials (see Figure 8.11).

The single-channel PMT was the detector of choice for precise astronomical brightness measurements from 1945 until the advent of CCDs in the early 1980s. The spectral responses of the available PMT photocathode materials defined, in part, some of the now-standard photometric band-passes (the U, B, and V bands in Table 1.2, for example). Since photomultipliers have few advantages over CCDs, they have become rare at observatories. One important advantage of the PMT, however, is response time. The temporal spread of a single pulse at the anode limits the shortest interval over which a PMT can sense a meaningful change in signal. Pulse widths are so narrow (5–10 nanosecond) for many PMTs that they can, in principle, detect signal changes as rapid as a few

milliseconds. The response time of a CCD, in contrast, is several tens of seconds for a standard slow-scan device, with quicker response possible with increased noise. Arrays of STJs, although still in the development phase, have the potential for response times similar to PMTs.

8.3.2 The microchannel plate

The upper part of Figure 8.13 shows an important variation on the PMT. Take a glass capillary with a diameter between 5 and 25 μm, and a length around 40 times its diameter. Coat the inside surface of this tube with a semiconductor that has good secondary electron-emitting properties, and connect the ends of the channel coating to the voltages as shown. You have created a ***microchannel***. Place this microchannel assembly in an evacuated chamber between a photocathode and an anode, and it can serve in place of the dynode chain of a PMT. A

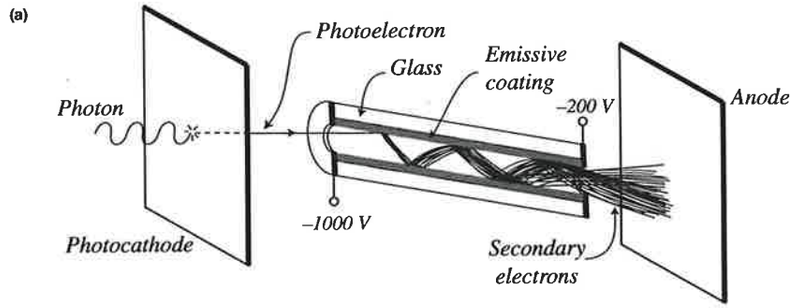


Fig. 8.13 (a) A single microchannel – a small glass tube whose interior is coated with dynode-type material. A large potential drop along the tube insures that an electron impact at one end will produce a burst of secondary electrons at the other. (b) A closely packed array of channels forming an MCP.

photoelectron from the cathode will accelerate towards the upper end of the channel, where it strikes the wall and generates a spray of secondary electrons. These secondary electrons will in turn strike the channel wall further down, multiplying their numbers. After several multiplications, a large pulse of electrons emerges from the end of the microchannel and accelerates to the anode.

A ***microchannel plate*** (MCP), as illustrated in Figure 8.13b, consists of an array of up to several million microchannels closely packed to form a plate or disk several millimeters in diameter and less than a millimeter thick. The electrical contact that coats the front surface can be made of a metal that has some secondary-electron emission capabilities, so that photoelectrons that do not strike the inside of a channel might still be detected via emission from the contact. You can make a high gain but very compact PMT by sandwiching several MCPs between a photocathode and anode in a vacuum enclosure, as in Figure 8.14. Such ***MCP PMTs***, operated as single-channel devices, have an advantage in size, power consumption, response time and stability in magnetic fields compared to dynode-based devices.

The MCP, however, is most valuable as a component in a two-dimensional detector. Various anode configurations or electron-detection devices can generate an output image that faithfully reproduces the input on the cathode of an MCP. The ***multi-anode microchannel array detector*** (***MAMA***) is an example. In the MAMA, the anode of the MCP PMT is replaced with two planes of parallel wires that form an x–y grid. A pulse of electrons emerging from a

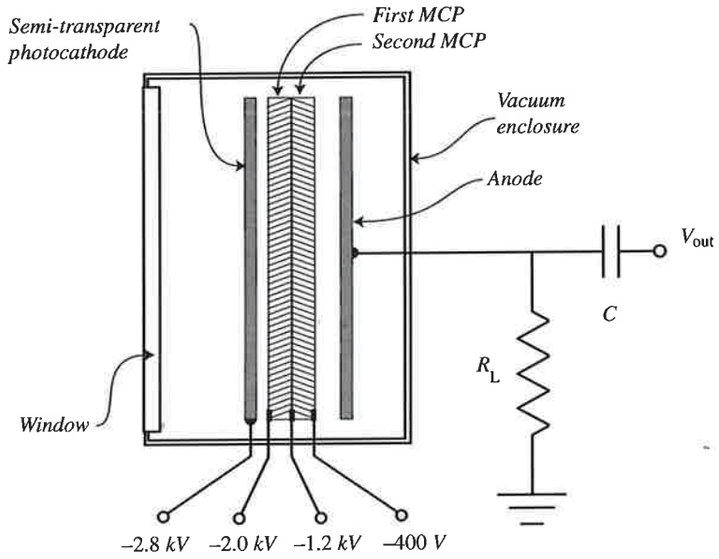


Fig. 8.14 A schematic MCP PMT. The dynodes of the photomultiplier are replaced by two microchannel plates. The capacitor at the output means the device is used in pulse-counting mode. Compare with Figure 8.11.

particular microchannel will impact with maximum intensity on one x-wire and one y-wire. Special circuitry then increments the signal count at the corresponding x–y address in the output image.

The MAMA detectors are especially useful at short wavelengths where the DQE of the device can be very high if it is equipped with a "solar-blind" photocathode insensitive to visual and infrared photons. Space astronomy has employed MAMA detectors to great advantage in for the detection of X-rays and far-ultraviolet light. Although silicon CCDs are also sensitive at these wavelengths, they suffer from high sky background levels from starlight and scattered sunlight that cannot be completely removed by filtering.

8.3.3 Image intensifiers and the ICCD

An ***image intensifier*** is not a detector, but a vacuum device that amplifies the brightness of an image. Because military interest in night vision drives the development of intensifiers, the military terminology (Generation I, II, III[4] etc.) for different designs has become standard. Figure 8.15 shows a Generation II intensifier coupled by optic fibers to a CCD. The intensifier resembles a MCP PMT, but it has a phosphor screen instead of an anode. A photoelectron leaving the cathode produces a pulse of high-energy electrons that excites multiple molecules in the phosphor. These then de-excite by photo-emission. The location of the phosphor emission maps the location of the original photo-absorption on the cathode. A single input photon can generate 10^4 to 10^7 phosphor photons.

As shown in Figure 8.15, an intensifier can be a useful first stage for an array detector like a CCD. It is important to understand, however, that although an intensifier will vastly increase signal strength and decrease exposure times, it will always decrease the *input* SNR for the CCD.

For example, consider a source that produces N_i photons at the photo-cathode of an image intensifier during an integration. If the input is dominated by photon noise (assume background is negligible) then the uncertainty in the input signal is just $\sqrt{N_i}$. The intensifier output at the phosphor is

$$N_{out} = gN_i$$

where g is the gain factor of the intensification. The variance of N_{out} is therefore

$$\sigma_{out}^2 = g^2\sigma_{in}^2 + \sigma_g^2 N_i^2 = \left(g^2 + \sigma_g^2 N_i\right)N_i$$

Here σ_g is the uncertainty in the gain. Thus, the SNR at the input and output are

[4] Generation I devices (now obsolete) used electric or magnetic fields to accelerate photoelectrons from a cathode and then re-focus them directly onto the phosphor. Generation II and III devices use an MCP to form the image as described in the text. Generation III devices have advanced photocathodes sensitive in the NIR.

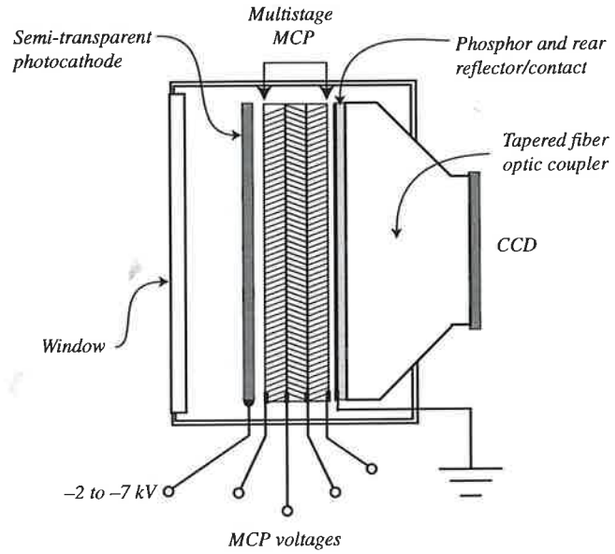


Fig. 8.15 The ICCD. The photocathode is in the image plane of the telescope. The image-intensifier stage produces an intensified image at the phosphor, which is transmitted by a fiber bundle or lens to the CCD. A very thin aluminum film behind the phosphor screen increases forward transmission and reduces the chance that light from the phosphor will feed back to the photocathode. Instead of the fiber coupler, a system of lenses can transfer the image from the phosphor to the CCD.

$$SNR_{in} = \sqrt{N_i}$$

$$SNR_{out} = gN_i\left[\left(g^2 + \sigma_g^2 N_i\right)N_i\right]^{-\frac{1}{2}} = \sqrt{N_i}\Big/\left[1 + \frac{\sigma_g^2 N}{g^2}\right]^{\frac{1}{2}} \leq SNR_{in}$$

So long as intensifier gain is uncertain, intensification will degrade the SNR.

Intensified CCDs (ICCDs) are thus useful in situations where the primary noise source is NOT photon noise in the signal, and/or where rapid signal changes need to be monitored. In such cases (e.g. read noise or dark-current noise dominant), using an intensifier can improve the DQE of the entire device by decreasing the required exposure times.

A related use for the image intensifier is as a ***signal conditioner*** for a CCD – the image on the phosphor is not only brighter than the one that arrives at its photo-cathode, it also emits photons of a different wavelength. You can select a cathode sensitive to ultraviolet light, for example, and a phosphor that emits near the CCD QE peak in the red. The ICCD thus detects sources that it would find absolutely invisible without the intensifier.

8.4 Infrared arrays

Although modern CCDs in some respects approach perfection as astronomical detectors, the large size of the forbidden band gap of silicon means that they are blind to all light with wavelength longer than 1.1 μm. The development of infrared-sensitive arrays has faced great technical difficulties, but advances have come quickly. These advances have had an even greater impact on

infrared astronomy than the CCD has had in the optical. This is because prior to the CCD, optical astronomers had an excellent, although inefficient, multi-pixel detector – the photographic plate. Prior to infrared arrays, infrared astronomers had only single-pixel devices. Different wavelength regions in the infrared place different demands on detector technology. We first make a brief survey of these differences, then examine the general method of infrared detector fabrication. Chapter 6 of Glass (1999) gives a general qualitative discussion of infrared technology, and chapter 11 of McLean (2008) gives a more technical treatment.

8.4.1 Detectors at different wavelengths

In the near infrared (1–5 μm; J, H, K, L, and M bands) practical arrays for astronomy first appeared at observatories in 1986. Although a number of materials were tried, the most successful eventually proved to be arrays of junction diodes made of indium antimodide (***InSb*** – often pronounced "ins-bee") or HgCdTe (mercury–cadmium telluride, or ***MCT***). The initial arrays contained only a few pixels but, by 2009, manufacturers were producing 2048 × 2048 pixel buttable arrays, and astronomers were assembling infrared-sensitive mosaics that were only an order of magnitude smaller (70 megapixels) in size than contemporary CCD-based devices. Modern NIR arrays have QE values less than 70%, read noise levels less than 10 electrons, and dark currents below 0.1 electrons per second.

In the MIR (5–28 μm) progress has been more modest, partly because high background levels limit ground-based observing at these wavelengths. At even the best sites in the MIR, the atmosphere is marginally transparent only in the N band (8–13 μm) and, to a lesser extent, the Q band (17–28 μm) so that any useful observations require a very large telescope or an orbiting observatory like Spitzer. At present, the most advanced MIR arrays are of blocked-impurity-band (BIB) photoconductors, usually fabricated from silicon doped with antimony (Si:Sb) or arsenic (Si:As). As of 2009, arrays in a 1024 × 1024 pixel format have begun to appear at the world's largest telescopes. Arrays of Schottky photodiodes (with PtSi as the metal layer) are easier to fabricate and have seen some use, but are limited by poor quantum efficiency.

In the far infrared (25–350 μm) the Earth's atmosphere is completely opaque (there is a weak and erratic window at 40 mm at high-altitude sites). Far-infrared detectors, therefore, must be flown in spacecraft or very high-altitude aircraft. In general, extrinsic detector arrays have been the most useful in this region. Doped silicon with the smallest band gap, Si:Sb, has a cutoff wavelength at around 30 μm, so for longer wavelengths, extrinsic germanium has been used. The Spitzer Space Telescope, for example, carries 32 × 32 pixel arrays of Ge:Ga (cutoff near

115 μm) and a 2 × 20 stressed[5] Ge:Ga array (cutoff near 190 μm). Mosaics of Spitzer-sized devices are under construction. For even longer wavelengths, observers have used small arrays of bolometers (discussed in a later section).

8.4.2 Infrared detector construction

Building an infrared array of photon detectors of any of the types discussed above is different from building a CCD in an important way. Charge-coupled devices are based on a mature technology. Buoyed by the ballooning market in computers and consumer electronics over the past forty years, manufacturers have refined their skill in the fabrication of electronic components in intrinsic silicon and devices based on p–n junctions in silicon. Expertise with more difficult materials like InSb, MCT, and extrinsic silicon and germanium is limited in comparison. That infrared arrays exist at all is due in large part to their applicability to battlefield imaging, surveillance, and remote sensing. Because building electronics is so much easier in silicon, almost all modern infrared arrays are built as two-layer ***hybrids***: one layer is composed of the infrared-sensitive material, the other, made of silicon, provides the electronics for reading the signal.

Figure 8.16 sketches some of the details in the construction of two pixels of a NIR hybrid array. The light-sensitive elements are junction photodiodes made of MCT – that is, the alloy $Hg_{(1-x)}Cd_xTe$. The MCT has an adjustable cutoff wavelength, and although it has appeared primarily in NIR detectors, it is a potentially useful material at longer wavelengths.[6] 256 × 256 Arrays of this design were installed in the NIR camera and multi-object spectrograph (NICMOS) of the Hubble Space Telescope. Successors to the NICMOS array (first the PICNIC and then the HAWAII arrays manufactured by Rockwell Scientific Corporation) have grown to 2048 × 2048 size.

The figure illustrates a cross-section of two pixels. The top layer is the silicon readout array, which contains several CMOS field-effect transistors (MOSFETs) at each pixel. The lower layer contains the infrared-sensitive material – in this case a p–n photodiode at each pixel. The total thickness of the MCT is

[5] Creation of a majority carrier in p-type material requires breaking an atomic bond and remaking it elsewhere (movement of an electron from the valence band to an acceptor state). It is easier to do the bond breaking (it takes less energy) if the crystal is already under mechanical stress. Thus, the cutoff wavelength of a stressed p-type crystal is longer than for an unstressed crystal. Maintaining the proper uniform stress without fracturing the material is a delicate operation.

[6] The pure form of HgTe behaves like a metal, while pure CdTe has a band gap of 1.6 eV. The band gap of the $Hg_{(1-x)}(CdTe)_x$ alloy depends on x, the cadmium telluride fraction, and is 0.31 eV (cutoff at 4 μm) at $x = 0.35$ and 0.1 eV (cutoff at 100 μm) at $x = 0.17$. So far, cutoff wavelengths longward of 20 μm have been difficult to achieve, but new techniques such as lattice fabrication by means of molecular beams may permit manufacturers to make low-x MCT of the required uniformity.

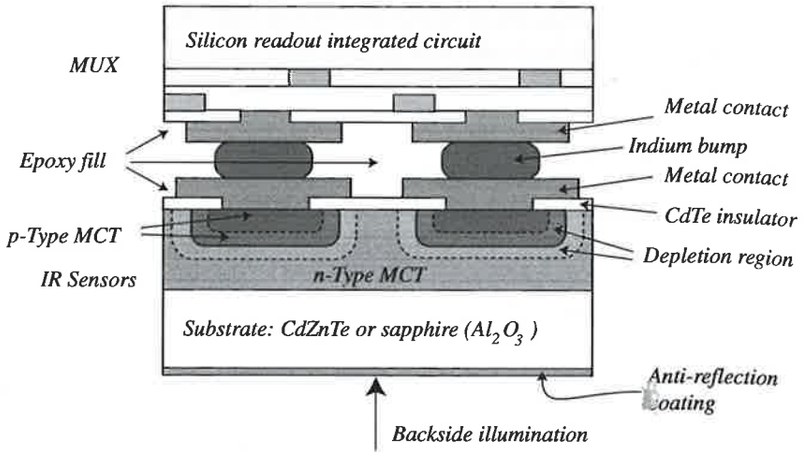


Fig. 8.16 Cross-section of two pixels of a hybrid array of MCT photodiodes. Each diode connects to the silicon readout circuits through an indium bump conductor.

quite small, and it is grown or deposited on a transparent substrate, like sapphire, to provide mechanical strength. Initially, the two layers are manufactured as separate arrays. A small bump of the soft metal indium is deposited on the output electrode of each photodiode. A matching bump of indium is deposited on the corresponding input electrode of each silicon readout circuit. The silicon and the infrared-sensitive arrays are then matched pixel-to-pixel and pressed together, so the indium bumps compress-weld against their mates to make good electrical contact. The spaces between bumps can then be filled with epoxy to secure the bond.

There are obvious and not-so-obvious pitfalls in making arrays using this "bump-bonding" approach, but the technique is becoming mature. Nevertheless, NIR arrays remain considerably more expensive than CCDs of the same pixel dimensions.

Reading an infrared array differs fundamentally from reading a CCD. There is no pixel-to-pixel charge transfer: each pixel sends output to its individual readout integrated circuit (***ROIC***) in the silicon layer. Since one of the tasks of the silicon layer is to organize the multiple signals from all pixels into a single stream of data from the amplifier, the layer is often called the ***multiplexer*** or ***MUX***. Many multiplexers, especially in large arrays, read to several (usually two or four, but sometimes many more) data lines simultaneously. Important differences from CCDs include:

- Since a pixel does not have to (nor is it able to) pass charge to and from its neighbors, a "dead" pixel (caused, for example, by a failure in the bump bond) will not kill the entire upstream column, as it might in a CCD. Although saturation occurs, the "blooming" penalty present in CCDs is not a feature of infrared arrays.
- Since readout is separate from sensing, reads can be non-destructive, and the same image read several times. Moreover, the array can be read out while the infrared layer is still responding to light.

- Very high background levels invariably hamper infrared observations from the ground. This forces very short (0.1–10 seconds) integration times to avoid saturation. To cope with the resulting data rate, controllers often co-add (average) many of the short exposure images and save only that result.
- Many infrared sensors are somewhat non-linear, so calibration for linearity is a much greater concern with an infrared array than it is with a CCD.
- Because of the smaller band gaps involved, dark currents in infrared arrays can be a severe problem, and these detectors must operate at low temperatures. Although some NIR arrays work well with liquid-nitrogen cooling, MIR arrays require refrigeration to much lower temperatures.
- Any detector sensitive to wavelengths longer than about 5 μm requires a cold enclosure to shield it from the infrared light flooding in from its warm (and therefore glowing) surroundings. These hot surroundings include the telescope structure and optics, so space telescopes that can be kept cold are superior infrared observing platforms. For similar reasons, the secondary mirrors of ground-based infrared telescopes are designed to be as small as possible. In the FIR, even the readout circuits are heat sources that need to be isolated from detectors.

8.5 Thermal detectors

Thermal detectors do not depend upon photons to move charge carriers directly from one band to another. They work, rather, as two-element devices: (1) a thermometer, which senses the temperature increase produced in (2) an absorber, when the latter is exposed to an incoming light beam. Figure 8.17 sketches a generalized thermal detector. In the figure, a heat sink at temperature T_0 encloses the absorber and thermometer in an evacuated cavity. A strip of material with conductance G connects the absorber and heat sink. To make an observation, the shutter is opened and the incoming light is allowed to deposit energy in the absorber at rate P_{in}. After a time, the temperature of the absorber will increase by amount ΔT, a quantity that therefore measures P_{in}.

For example, if the absorber were allowed to reach equilibrium with the shutter open, the following condition will apply:

$$\begin{aligned} P_{\text{in}} &= P_{\text{out}} = P_{\text{conduct}} + P_{\text{radiate}} \\ &\approx G\Delta T + \sigma A\left[(T_0 + \Delta T)^4 - T_0^4\right] \\ &= \Delta T\left\{G + \sigma A\left[4T_0^3 + 6T_0^2\Delta T + 4T_0(\Delta T)^2 + (\Delta T)^3\right]\right\} \\ &= \Delta T\{G + \sigma A R(T_0)\} = \Delta T G' \end{aligned} \tag{8.3}$$

In this equation, A is the total surface area of the absorber, and σ is Stefan's constant. We assume the solid angle subtended at the absorber by the shutter opening is small. To maximize the sensitivity, which is just $\Delta T/P_{\text{in}}$, the term in braces in the final expression must be kept as small as possible. To this end, it is

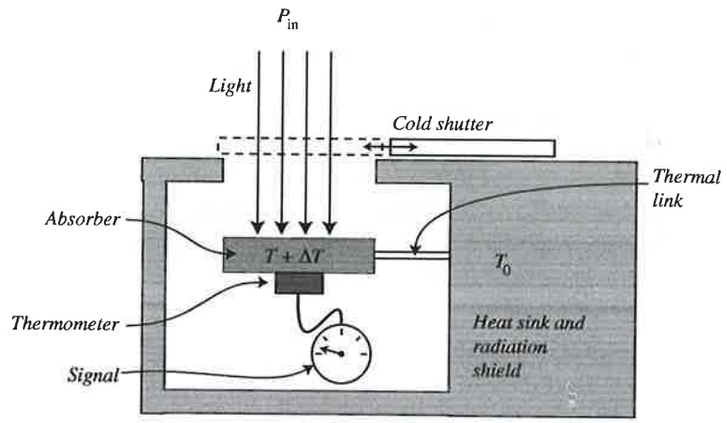


Fig. 8.17 General design for a thermal detector. A thermometer records the increase in the temperature of a light-absorber after it is exposed to a source. A strip of conducting material links the absorber to a large heat sink. Because it may require power to operate, the thermometer may contribute to P_{in}.

clear that both the conductance of the link and the area of the absorber should be small. Likewise, especially since it appears to the third power, the temperature of the sink, T_0, needs to be kept very low. An additional benefit of low temperature is that if the conductance term is large relative to the radiative term, the effective conductance, G', in W K^{-1}, will be nearly constant. Although it is not addressed in the equilibrium expression above, keeping the heat capacity of the absorber small will mean that the ***response time*** of the detector is short. The time dependence will be given by

$$\Delta T(t) = T_0 + \frac{P_{in}}{G'}[1 - \exp(-tG'/C)] \tag{8.4}$$

Here C is the heat capacity of the detector in J K^{-1} and the time constant is C/G'. In general, a thermal detector will employ an absorber that is black, with a small heat capacity. Physical size should be as small as possible, matching the size of the focal-plane image but not approaching the scale of the wavelength being observed. The absorber will be linked to a temperature sink that is kept as cold as possible (often, for example, with liquid helium).

In practice, the thermal detectors used in astronomy have almost always been ***bolometers***, defined as thermal detectors in which the temperature sensor is a ***thermistor*** (a contraction of the words *thermal* and *resistor*), a small mass of semiconductor or metal whose electrical conductivity is a strong function of temperature. The small signal levels usually characteristic of astronomical measurements have restricted instruments to just a few thermistor materials: extrinsic silicon has been used, but n-type extrinsic germanium (doped with gallium) is the most common choice. The gap between the donor states and the conduction band in Ge:Ga is around 0.01 eV, so a modest increase in temperature will excite donor electrons to the conduction band and decrease the resistance of the thermistor. As in any semiconductor with a band gap, the

resistance as a function of temperature will be given by an equation of the form

$$R(T) = R_0 T^{-\frac{3}{2}} e^{\frac{A}{kT}} \tag{8.5}$$

where A is a constant that depends on the band-gap energy and R_0 depends on both the band gap and the doping level.

The situation is somewhat different at temperatures below 5 K, where the mechanism of thermal excitation of electrons across the gap between donor and conduction states becomes unimportant. The practice there is to use very highly doped thermistors, in which conductivity is due to hopping within the impurity band. Equation (8.5) does not apply in this case, but other semi-empirical expressions relate a decrease in resistance to an increase in temperature.

Thermistor resistance is usually monitored in a simple circuit in which the voltage drop across the sensitive material is observed with the thermistor placed in series with a stable reference resistor, as illustrated in Figure 8.18a. The reference resistor must be large relative to the bolometer resistance; otherwise a positive heating feedback can run away in the bolometer.

One simple bolometer design, illustrated in Figure 8.18b, has an elegant feature – the absorber and the temperature sensor are one and the same element. The wires that connect this sensor to the conductivity-measuring circuit also provide the thermal link to the heat sink. Figure 8.18c shows a compound bolometer, in which the thermistor is fastened to a radiation-absorbing plate. Both these designs can be easily incorporated in a one-dimensional array. Modified geometries have led to two-dimensional arrays of a few hundred pixels.

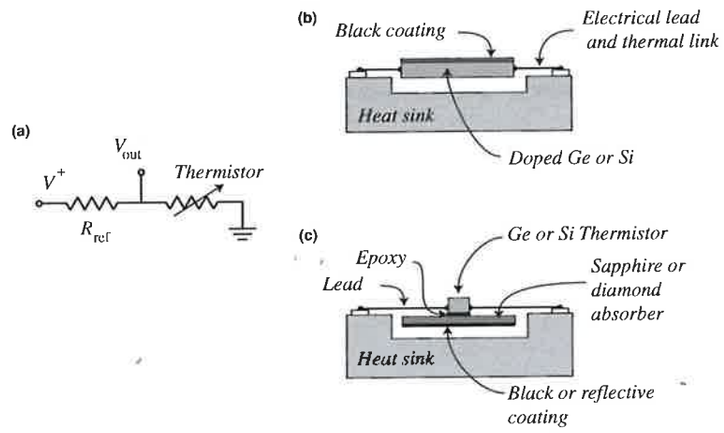


Fig. 8.18 Bolometers. (a) Electrical characteristics: a thermistor, whose electrical resistance depends on temperature, is connected in series with a reference resistor. Voltage across the reference changes with the temperature of the detector. (b) A bolometer in which the thermistor is also the absorber. Here the electrical lead is also the thermal link to the heat sink. Alternatively, in (c), the thermistor can be mechanically and thermally bonded to the absorber.

At very low temperatures, a bolometer can employ a type of superconducting thermistor called a ***transition edge sensor*** (***TES***) to attain very high sensitivity. See Rieke (2003), chapter 9, for a more detailed discussion of the TES and of bolometers in general.

Summary

- An important measure of detector quality is the detective quantum efficiency:

$$\mathrm{DQE} = \frac{(\mathrm{SNR})^2_{\mathrm{out}}}{(\mathrm{SNR})^2_{\mathrm{perfect}}}$$

- Characteristics of a detector include its mode (photon, wave, or thermal). Important detector concepts:

signal	*noise*	*quantum efficiency (QE)*
absorptive QE	*quantum yield*	*spectral response*
spectral resolution	*linearity*	*saturation*
hysteresis	*stability*	*response time*
dynamic range	*physical size*	*array dimensions*
Nyquist spacing	*image sampling*	

- The charge-coupled device, or CCD, is usually the preferred astronomical detector at visible wavelengths. Concepts:

parallel registers	serial register	*gates*
clock voltages	*barrier potential*	*ADU*
output amplifier	*collection potential*	*CCD gain*
blooming	*channel stop*	*CMOS capacitor*
full well	*digital saturation*	*read noise*
corrrelated double sampling	*multi-amplifier arrays*	*dark current*
orthogonal-transfer CCD	*cryogen*	*Dewar*
buried-channel CCD (BCCD)	*CTE*	*traps*
multi-pinned phase (MPP)	*inverted gate*	*L3CCD*
frontside illumination	*backthinned CCD*	*ITO*
deep-depleted CCD	*open electrode*	*microlens*
virtual-phase CCD		

- Several important astronomical detectors depend on the vacuum photoelectric effect. Concepts:

photomultiplier (PMT)	*dynode*	*pulse-counting*
microchannel plate	*MCP PMT*	*MAMA*
image intensifier	*signal conditioner*	*ICCD*

- Observational techniques and device performance with infrared arrays is highly dependent on the wavelength region observed. Concepts:

near-, mid- and far-infrared	InSb	*MCT*
BIB detectors	*Si:Sb*	*NICMOS*
HAWAII	*hybrid array*	*ROIC*
MUX	*indium bump bond*	

- Semiconductor thermistor resistance:

$$R(T) = R_0 T^{-\frac{1}{2}} e^{\frac{A}{kT}}$$

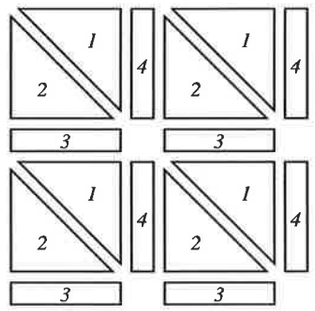
- Thermal detector concepts:

heat sink	*bolometer*
thermistor	*time constant*

Exercises

1. A photodiode has an overall quantum efficiency of 40% in the wavelength band 500–600 nm. The reflectivity (fraction of photons reflected) at the illuminated face of the detector in this band is measured to be 30%. If this face is treated with AR coatings, its reflectivity can be reduced to 5%. Compute the QE of the same device with the AR coating in place.
2. A certain detector measures the intensity of the light from a stable laboratory blackbody source. The signal in three identical trials is 113, 120 and 115 mV. From the blackbody temperature, the experimenter estimates that 10^4 photons were incident on the detector in each trial. Compute an estimate for the DQE of the detector.
3. A photon detector has a QE of q and a quantum yield of y. The uncertainty in y is $\sigma(y)$. Show that DQE = q if $\sigma(y) = 0$, but that DQE $< q$ otherwise.
4. A CCD has pixels whose read noise is 3 electrons and whose dark current is 1 electron per second. The QE of the detector is 0.9. Compute the DQE of a single pixel if 1000 photons are incident in a 1-second exposure. Compute the DQE for the same pixel if the same number of photons is incident in a 400-second exposure.
5. An MOS capacitor observes two sources in the band 400–600 nm. Source A has a spectrum such that the distribution of photons in the 400–600 nm band is given by $n_{\mathrm{A}}(\lambda) = A\lambda^3$. Source B has a distribution of photons given by $n_{\mathrm{B}}(\lambda) = B\lambda^{-2}$ in the same band. If the two sources generate photoelectrons at exactly the same rate, compute their (energy) brightness ratio. You may assume the detector's QE is not a function of wavelength.
6. Construction of a monolithic 8192 × 8192 pixel CCD array is technologically possible. How long would it take to read this array through a single amplifier at a pixel frequency of 25 kHz?

7. The gate structure for four pixels of a certain orthogonal transfer CCD is sketched at below. Propose a pattern for (a) assigning gate voltages during collection, (b) a method for clocking voltages for a one-pixel shift to the right, and (c) a method for clocking voltages for a one-pixel shift downwards. Gates with the same numbers are wired together.



8. At an operating temperature of 300 K, a certain CCD exhibits a dark current of 10^5 electrons per second. (a) Estimate the dark rate, in electrons per second, if this CCD is operated at –40 °C (233 K). (b) Compute the operating temperature at which the dark current will be 10 electrons per second.
9. A CCD has a CTE of "three nines" (i.e. CTE = 0.9990). What fraction of the charge stored in the pixel most distant from the amplifier actually reaches the amplifier if the array is (a) 128 pixels on a side or (b) 2048 on a side?
10. A rapid-scan CCD has a read noise of 200 electrons per pixel. You observe a source that produces 400 photoelectrons spread over 25 pixels. Dark current and background are negligible. (a) Compute the SNR for this measurement. (b) Suppose an image intensifier is available with a gain of 10^4 and a gain uncertainty of ±5%. Repeat the SNR computation for the intensified CCD. Should you use the bare or the intensified CCD for this measurement?
11. Consider the general situation in which a bare CCD would record N photoelectrons with a total read noise of R electrons in a given exposure time. An intensifier stage has a gain of g and a gain uncertainty of σ_g. If $g >> 1$, show that the intensifier will improve the overall DQE in the same exposure time if $R^2 g^2 > \sigma_g N$.
12. A single-element bolometer operates with a heat sink at 12 K. The thermal link has a conductance of $G' = 5 \times 10^{-7}\,\mathrm{W\,K^{-1}}$ and a heat capacity of $C = 3 \times 10^{-8}\,\mathrm{J\,K^{-1}}$. (a) Compute the time constant and temperature change after 2 seconds of exposure to a source that deposits 10^{-10} W in the bolometer. (b) If the bolometer is a doped germanium thermistor with a resistance of R_0 ohm at 12 K and effective energy gap of A = 0.02 eV, compute the fractional change in resistance due to the exposure in (a).

Chapter 9
Digital images from arrays

> All the pictures which science now draws of nature and which alone seem capable of according with observational fact are mathematical pictures.
>
> – Sir James Jeans, *The Mysterious Universe*, 1930

Astronomers normally present the output of a sensor array in the form of a digital image, a picture, but a mathematical picture. One appealing characteristic of a digital image is that the astronomer can readily subject it to mathematical manipulation, both for purposes of improving the image itself, as well as for purposes of extracting information.

Accordingly, the chapter will proceed by first presenting some general thoughts about array data, and some general algorithms for image manipulation. Because they are so useful in astronomy, we next examine some procedures for removing image flaws introduced by the observing system, as well as some operations that can combine multiple images into a single image. Finally, we look at one important method for extracting information: *digital photometry*, and derive the *CCD equation*, an expression that describes the quality you can expect from a digital photometric measurement.

9.1 Arrays

Astronomers usually use ***panoramic detectors*** to record two-dimensional images and, at optical wavelengths, they most often use a charge-coupled device (CCD). Unlike a photographic plate (until the 1980s, the panoramic detector of choice), a CCD is an ***array*** – a grid of spatially discrete but identical light-detecting elements. Although this chapter discusses the CCD specifically, most of its ideas are relevant to images from other kinds of arrays. These include broadband superconducting tunnel junctions (STJs), hybrid semiconductor arrays used in the infrared, ultraviolet-sensitive devices like microchannel plates, and bolometer arrays used in the far infrared and at other wavelengths.

9.1.1 Pixels and response

A telescope gathers light from some astronomical scene and forms an image in its focal plane. At each point (x', y') in the focal plane, the image has brightness, $B(x', y')$, measured in W m^{-2}. The function B is an imperfect representation of the original scene. Every telescope has limited resolving power and optical aberrations. Often a telescope will transmit some parts of the scene more efficiently than other parts, perhaps because of dust on a filter or some structural obstruction. In addition, some contributions to the brightness of the image do not originate from the remote source: the background glow from the atmosphere or (in the infrared) from the telescope, for example.

We introduce a panoramic detector (or ***focal-plane array***) to record this imperfect image. It is invariably a *rectangular* array, with elements arranged in N_X columns and N_Y rows. We denote the location of an individual detector element in this array as $[x, y]$, where it will be convenient to restrict x and y to integer values (running from 1 to N_X and 1 to N_Y, respectively). Instead of the phrase "individual detector element," we use the word ***pixel*** (from "picture element").

Figure 9.1 shows a few pixels of some array. The sensitive area of a single pixel is a rectangle of dimensions δ_{px} by δ_{py}, and the pixels are separated by distances d_{px} horizontally and d_{py} vertically. For most direct-imaging devices, pixels are square ($d_{px} = d_{py} = d_p$) and have sizes in the 5–50 μm range. Linear arrays ($N_x \gg N_y$), sometimes used in spectroscopy, are more likely to employ oblong pixel shapes.

If $d_p > \delta_p$ in either direction, each pixel has an insensitive region whose relative importance can be measured by the geometric ***fill factor***,

$$\frac{\delta_{px}\delta_{py}}{d_{px}d_{py}}$$

For many CCDs $\delta_p = d_p$, and the fill factor is unity.

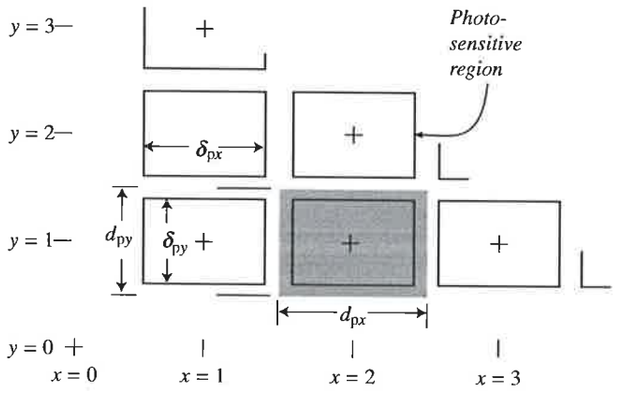


Fig. 9.1 Pixels near one corner of a detector array. The shaded region indicates the pixel at [2, 1], which consists of a photosensitive region surrounded by an insensitive border.

Our detector lies in the focal plane of the telescope, with the x (for the detector) and x' (for the function B) axes aligned. We are free to choose the origin of the primed coordinate system, so can make the center of a pixel with coordinates $[x, y]$ have primed coordinates:

$$x' = x \cdot d_p$$
$$y' = y \cdot d_p$$

The light falling on the pixel $[x, y]$ will have a total power, in watts, of

$$P[x,y] = \int_{\left(y-\frac{1}{2}\right)\delta_p}^{\left(y+\frac{1}{2}\right)\delta_p} \int_{\left(x-\frac{1}{2}\right)\delta_p}^{\left(x+\frac{1}{2}\right)\delta_p} B(x',y')\,dx'\,dy' \qquad (9.1)$$

In Equation (9.1), we use square brackets on the left-hand side as a reminder that the detector pixel takes a discrete sample of the continuous image $B(x', y')$ and that x and y can only take on integer values. This ***pixelization*** or ***sampling*** produces a loss of image detail if the pixel spacing, d_p, is less than half the resolution of the original image. Such ***undersampling*** is usually undesirable.

We expose the pixel to power $P[x, y]$ for a time interval, t. It responds by producing (in the case of the CCD) a number of photoelectrons. We call this the photo-response, $r_0[x, y]$. Note that for many photon detectors, including the CCD, $r_0[x, y]$ depends on the *number of incident photons*, not on the energy, $tP[x, y]$. To complicate matters, the photo-response signal usually mixes indistinguishably with that produced by other mechanisms (thermal excitation, light leaks, cosmic-ray impacts, radioactivity, etc. The pixel gives us, not $r_0[x, y]$, but $r[x, y]$, a total response to all elements of its environment, including $P[x, y]$; see Figure 9.2.

Although it is convenient to think of the CCD response on the microscopic level of individual electrons, this may not be the case for other devices. In some

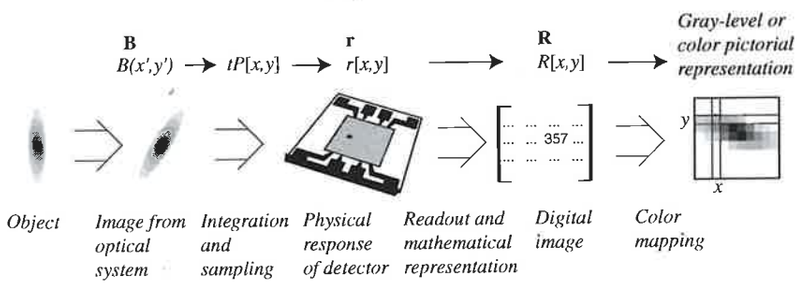


Fig. 9.2 Production of an astronomical digital image

arrays, it will be better to regard $r[x, y]$ as an analog macroscopic property like a change in temperature or conductivity.

9.1.2 Digital Images

Our instrument must communicate a quantification of $r[x, y]$ to the outside world. In the case of the CCD, the clock circuits transfer charge carriers through the parallel and serial registers, and one or more amplifiers convert each charge packet to a voltage (the ***video signal***). Another circuit, the analog-to-digital converter (ADC), converts the analog video signal to an electronic representation of an integer number, primarily because integers are much easier to store in a computer. We symbolize the integer output for pixel $[x, y]$ as $R[x, y]$, its ***pixel value***.

The entire collection of all $N_x \times N_y$ integers, arranged as a mathematical array to echo the column–row structure of the detector, is **R**, a ***digital image***. Sometimes we call a digital image a ***frame***, or an ***exposure***. We use boldface symbols for an entire array (or image), as in **R**, and the subscripts in square brackets to indicate one element of an array (a single pixel value), as in $R[x, y]$. The digital image, **R**, is the digital representation of the detector response, **r**. The relation between **R** and **r** may not be simple.

Digital images are simply collections of numbers interpreted as images, and they can be produced in a variety of ways – perhaps by scanning and digitizing an analog image, by an artistic effort with a computer "paint" program or by any method that assembles an ordered array of numbers. Often, to help interpret the array, we ***map*** the numbers onto a gray-scale or color-scale and form a pictorial representation.

For example, the "picture" of the nearby galaxy M51 in Figure 9.3 is a representation of a digital image in which a grid of squares is colored according to the corresponding pixel values. Squares colored with 50% gray, for example, correspond to pixel values between 2010 and 2205, while completely black squares correspond to pixel values above 4330. A mapping like Figure 9.3 usually cannot show all the digital information present, since pixel values are often 16-bit integers,[1] while human vision only distinguishes at most a few hundred gray levels (which code as 7- or 8-bit integers).

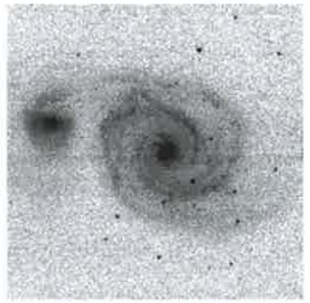


Fig. 9.3 A CCD image of the galaxy M51.

9.1.3 CCD Gain

We use some special terminology in the case where image **R** represents the response of an astronomical array. Quantifying detector response usually means measuring a voltage or current (i.e. an *analog* quantity) and subsequently expressing this as a *digital* quantity. Hence, each pixel of **R** is said give a count of how many ***analog-to-digital units*** (ADUs) were read from the detector. Each pixel value, $R[x, y]$, has "units" of ***ADU***. The terms ***data number*** (DN) and ***counts*** are sometimes used instead of ADU.

The differential change in $r[x, y]$that produces a change of one ADU in $R[x, y]$ is called the ***gain***

$$g[x, y] = \text{gain} = \frac{\mathrm{d}r[x, y]}{\mathrm{d}R[x, y]}$$

In the general case, gain will differ from pixel to pixel, and may even depend on the signal level itself. In the case of the CCD, gain is set primarily by the output amplifier and the ADC, and the astronomer might even set the gain with the controlling software. We expect approximately identical gain for all pixels. Moreover, CCD amplifiers are generally linear, so we usually assume $g[x, y]$ is independent of $r[x, y]$. The CCD gain has units of electrons per ADU:

$$g = \text{CCD gain} = \frac{\mathrm{d}r[x, y]}{\mathrm{d}R[x, y]} \text{ [electrons per ADU], independent of } r, x \text{ and } y$$

Gain may differ (by a small amount, one hopes) for each amplifier on a multi-amplifier CCD chip, or for the components in a mosaic.

9.1.4 Pictures lie

> The world today doesn't make sense, so why should I paint pictures that do?
>
> – Pablo Picasso (1881–1973)

Figure 9.3, the gray-scale map of a CCD image of the galaxy M51, imperfectly represents **R**, the underlying digital image. But even the underlying image is a lie. There are interstellar, atmospheric, and telescopic effects that mask, distort, and destroy information as light travels from M51 to the detector, as well as additions and transformations introduced by the detector itself – all information

[1] The number of bits (binary digits), n_B, in a computer memory location determines the value of the largest integer that can be stored there. (It is $2^{n_B} - 1$.) Thus, a 16-bit integer can have any value between 0 and 65,535, while an 8-bit integer can have values between 0 and 255.

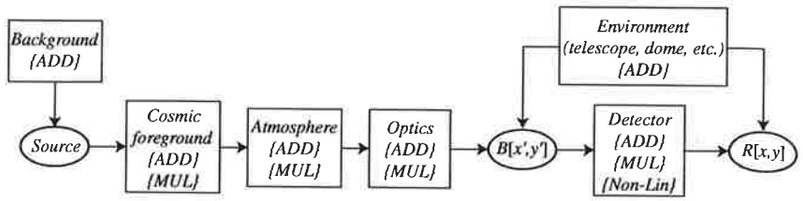


Fig. 9.4 Additive (ADD), multiplicative (MUL) and non-linear (Non-Lin) effects produce imperfections in detector output. Alterations by optics include intentional restrictions by elements like filters. The local environment may add signal by introducing photons (e.g. light leaks) or by other means (e.g. thermal dark current, electronic interference, cosmic rays).

gains and losses that we would rather not have. Figure 9.4 schematically represents the most obvious elements that might influence the raw digital image.

For the moment, imagine a "perfect" digital image, **R***. In **R***, the number of ADUs in a pixel is directly proportional either to the energy or number of photons arriving from the source located at the corresponding direction in the sky. The image **R*** is not influenced by any of the elements represented in Figure 9.4. Mathematically, three kinds of processes can cause **R**, the raw image, to differ from **R***, the perfect image:

Additive effects contribute or remove ADUs from a pixel in a way that is independent of the magnitude of $R^*[x, y]$. Examples include:

- background radiation emitted by the telescope, the Earth's atmosphere, foreground or background stars, or any other objects visible to the pixel;
- impacts of cosmic rays and other energetic particles;
- the ambient thermal energy of the pixel;
- a voltage intentionally added to the video signal to guarantee amplifier linearity.

Multiplicative imperfections change $R^*[x, y]$ to a value proportional to its magnitude. Examples include:

- spatial or temporal variations in quantum efficiency or in gain;
- absorption by the Earth's atmosphere;
- absorption, reflection, or interference effects by optical elements like filters, windows, mirrors and lenses, as well as dirt on any of these

Non-linear imperfections change $R^*[x, y]$ to a value that depends on a quadratic or higher power of its magnitude. An example would be a quantum efficiency or gain that depends on the magnitude of $R^*[x, y]$. ***Saturation***, a decrease in detector sensitivity at high signal levels, is a common non-linear imperfection.

All these imperfections are least troublesome if they are ***flat***, that is, if they have the same effect on every pixel. Subtracting a spatially uniform background is relatively easy. In contrast, if the imperfection has detail, removing it requires more work. Subtracting the foreground stars from an image of a galaxy, for example, is relatively difficult. Not every imperfection can be removed, and every removal scheme inevitably adds uncertainty. No image ever tells the complete truth.

9.2 Digital image manipulation

> If a man's wit be wandering, let him study the mathematics.
>
> – Francis Bacon (1561–1626)

One of the great benefits of observing with modern arrays is that data take the form of digital images – numbers. Astronomers can employ powerful and sophisticated computing tools to manipulate these numbers to answer questions about objects. We usually first find numerical answers, but eventually construct a narrative answer, some sort of story about the object. Our concern in the remainder of this chapter is to describe some of the computational and observational schemes that can remove the imperfections in astronomical images, and some schemes that can ***reduce*** those images to concise measurements of astronomically interesting properties. We begin with some simple rules.

9.2.1 Basic image arithmetic

First, some conventions. As before, boldface letters will symbolize complete digital images: **A**, **B**, **C**, and **D**, for example, are all digital images. Plain-faced letters, like h and k, represent single-valued constants or variables. As introduced earlier, indices in square brackets specify the location of a single pixel, and A [2,75] is the pixel value of the element in column 2, row 75, of image **A**.

If $\{op\}$ is some arithmetic operation, like addition or multiplication, then the notations

$$\mathbf{A} = \mathbf{B}\ \{op\}\ \mathbf{C}$$
$$\mathbf{A} = k\ \{op\}\ \mathbf{D}$$

mean that

$$A[x,y] = B[x,y]\ \{op\}\ C[x,y], \text{ and}$$
$$A[x,y] = k\ \{op\}\ D[x,y],$$
$$\text{for all indices, } 1 \le x \le N_X \text{ and } 1 \le y \le N_Y$$

That is, the indicated operation is carried out on a pixel-by pixel basis over the entire image. Clearly, all images in an equation must have the same size and shape for this to work. For example, suppose you take an image of Mars, but, by mistake, leave a lamp on inside the telescope dome. The image actually recorded, as depicted in Figure 9.5, will be

$$\mathbf{A} = \mathbf{M} + \mathbf{L}$$

where **M** is the image due to Mars alone and **L** is the image due to the lamp and everything else. You might be able to obtain a good approximation to **L** by

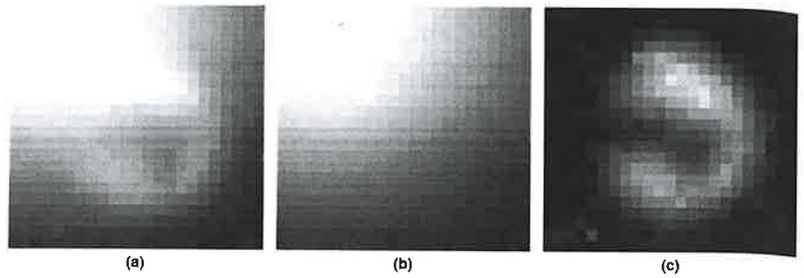


Fig. 9.5 Image subtraction. The image in (a) is of the planet Mars plus background and foreground illumination (**M** + **L**) . The image **L** in (b) is of the blank sky at the same altitude and azimuth, obtained after sidereal motion took Mars out of the field. The image in (c) is the difference between the two.

leaving the lamp on and taking an image of the blank sky. If so, as shown in the figure, you can computationally recover **M**:

$$\mathbf{M} = \mathbf{A} - \mathbf{L}$$

9.2.2 Image dimensions and color

We find it natural to think of digital images as two-dimensional objects – brightness arrayed in rows and columns. But a digital image is just a way to interpret a string of numbers, and there are many cases in which it makes sense to think of images with three or more dimensions. For example, you take a series of 250 images of the same star field to search for the period of a suspected variable. Each image has 512 rows and 512 columns. It makes sense to think of your data as a three-dimensional stack, with dimensions $512 \times 512 \times 250$. You will therefore encounter terms like ***data cube*** in the astronomical literature. Another common example would be the output of an array of STJ detectors, where spectral distribution would run along the third dimension. Higher dimensions also make sense. Suppose you take 250 images of the field in each of five filters – you then could have a four-dimensional "data-hypercube." We will not discuss any special operations for these higher-dimensional objects. How they are treated will depend on what they represent, and often will come down to a series of two-dimensional operations.

Color images are a special case. Digital color images pervade modern culture, and there are several methods for encoding them, most conforming to the device intended to display the image. For example, each pixel of a color computer monitor contains three light sources: red (R), green (G), and blue (B). The ***RGB color model*** represents an image as a three-dimensional stack, one two-dimensional digital image for each color. Each pixel value codes how bright the corresponding colored light source should be in that one pixel. The RGB is an additive color model: increasing pixel values increases image brightness.

Subtractive color models are more suited to printing images with ink on a white background. The most common, the ***CMYK model***, uses a stack of four two-dimensional images to represent amounts of cyan, magenta, yellow, and black ink in each pixel. In a subtractive model, larger pixel values imply a darker color.

Astronomers almost never detect color images directly, but will frequently construct ***false color*** images as a way of displaying data. For example, you might create an RGB image in which the R channel was set by the pixel values of a K-band (i.e. infrared) image, the G channel was set by the pixel values of a V-band (i.e. visual) image and the B channel was set by the pixel values of a far-ultraviolet image. The resulting image would give a sense of the "color" of the object, but at mostly invisible wavelengths.

Astronomers also use ***color mapping*** to represent the brightness in a simple digital image. In a color mapping, the computer uses the pixel value to reference a color look-up table, and then displays the corresponding color instead of some gray level. Since the eye is better at distinguishing colors than it is at distinguishing levels of gray, a color map can emphasize subtle effects in an image.

9.2.3 Image functions

We expand our notation to include functions of an image. In the following examples, each pixel in image **A** is computed from the pixels with the same location in the images in the right side of the equation:

$$\mathbf{A} = -2.5\log(\mathbf{C})$$
$$\mathbf{A} = h(\mathbf{B})^2 + k\sqrt{\mathbf{C}}$$
$$\mathbf{A} = \max(\mathbf{B}, \mathbf{C}, \mathbf{D})$$
$$\mathbf{A} = \text{median}(\mathbf{B}, \mathbf{C}, \mathbf{D})$$

The "max" function in the third example would select the largest value from the three pixels at each x, y location:

$$A[x,y] = \text{the largest of}\{B[x,y], C[x,y], D[x,y]\},$$

for each x and y in **A**.

Likewise, the fourth example would compute the median of the three indicated values at each pixel location. We can think of many more examples. We also introduce the idea of a function that operates on an entire image and returns a *single* value. For example, the functions maxP and medianP:

$$a = \text{maxP}(\mathbf{A})$$
$$b = \text{medianP}(\mathbf{A})$$

will treat the pixels of image **A** as a list of numbers, and pick out the largest value in the whole image and the median value of the whole image, respectively. Again, you can think of a number of other examples of functions of this sort.

9.2.4 Image convolution and filtering

The concept of ***digital filtration*** is a bit more complex. Image ***convolution*** is an elementary type of digital filtration. Consider a small image, **K**, which measures

$2V + 1$ rows by $2W + 1$ columns (i.e. the number of rows and columns are both odd integers). We define the convolution of **K** on **A** to be a new image, **C**,

$$\mathbf{C} = \text{conv}(\mathbf{K}, \mathbf{A}) = \mathbf{K} \otimes \mathbf{A}$$

Image **C** has the same dimensions as **A**, and its pixels have values

$$C[x,y] = \sum_{i=1}^{(2V+1)} \sum_{j=1}^{(2W+1)} K[i,j]A[(x - V - 1 + i), (y - W - 1 + j)] \qquad (9.2)$$

The array **K** is sometimes called the ***kernel*** of the convolution. For example, consider the kernel for the 3×3 boxcar filter:

$$\mathbf{B} = \begin{bmatrix} \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Figure 9.6 suggests the relationship between the kernel, the original image, and the result. (1) The center of the kernel is aligned over pixel $[x, y]$ in the original image. (2) The value in each pixel of kernel is multiplied by the value in the image pixel beneath it. (3) The sum of the nine products is stored in pixel $[x, y]$ of the filtered image. (4) Steps (1)–(3) are repeated for all valid values of x and y.

Figure 9.7 shows an image before and after convolution with a boxcar filter. What happens in the convolution is that every pixel in the original image gets replaced with the average value of the nine pixels in the 3×3 square centered on itself. You should verify for yourself that this is what Equation (9.2) specifies. The boxcar is a filter that blurs detail – that is, it reduces the high spatial frequency components of an image. Figure 9.7c shows that a larger-sized boxcar kernel, 7×7, has an even greater blurring effect.

Note that in convolution, there is a potential problem at the image edges, because Equation (9.2) refers to non-existent pixels in the original image **A**. The usual remedy is artificially to extend the edges of **A** to contain the required

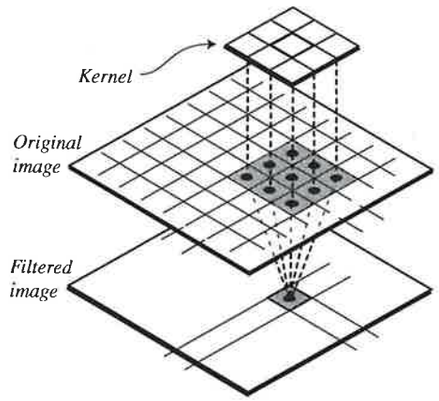


Fig. 9.6 Image convolution operation. The kernel is aligned over a set of pixels in the original centered on position x, y. The result is the sum of the products of each kernel pixel with the image pixel directly beneath it. The result is stored in pixel x, y of the filtered image.

pixels, typically setting the value of each fictitious pixel to that of the nearest actual pixel.

Convolutions that blur an image are called ***low-pass filters***, and different kernels will blur an image in different ways – a ***Gaussian kernel*** (whose values are set by a two-dimensional Gaussian function) can simulate some atmospheric seeing effects, for example. Other kernels are ***high-pass filters***, and emphasize image detail while suppressing large-scale (low spatial frequency) features. Representative of these is the ***Laplacian kernel***. The 3×3 Laplacian is

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

It essentially computes the average value of the second derivative of the intensity map – enhancing pixels that differ from the local trend. Figure 9.8 shows an example. Other filter kernels can produce image sharpening without loss of large-scale features, edge detection, gradient detection, and embossing effects.

A particularly useful filtering process is ***unsharp masking***. The filtered image is the original image minus an "unsharp-mask" image – this mask is a low-pass filtered version of the original. The unsharp mask enhances the high-frequency components and reduces the low-frequency components of the image, emphasizing detail at all brightness levels. Since convolution is distributive, unsharp masking can be accomplished by convolution with a single kernel. For example, convolution with the 5×5 identity kernel

$$\mathbf{I} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

leaves the image unchanged. Convolution with a 5×5 Gaussian ($\sigma = 1.25$ pixels)

$$\mathbf{G} = \frac{1}{3.58} \begin{bmatrix} 0.03 & 0.08 & 0.11 & 0.08 & 0.03 \\ 0.08 & 0.21 & 0.29 & 0.21 & 0.08 \\ 0.11 & 0.29 & 0.38 & 0.29 & 0.11 \\ 0.08 & 0.21 & 0.29 & 0.21 & 0.08 \\ 0.03 & 0.08 & 0.11 & 0.08 & 0.03 \end{bmatrix}$$

creates a blurred mask. An unsharp masking filtration operation would be

$$\mathbf{C} = 2(\mathbf{I} \otimes \mathbf{A}) - \mathbf{G} \otimes \mathbf{A}$$

or, using the distributive properties of the convolution operation,

$$\mathbf{C} = (2\mathbf{I} - \mathbf{G}) \otimes \mathbf{A} = \mathbf{F} \otimes \mathbf{A}$$

In this example,

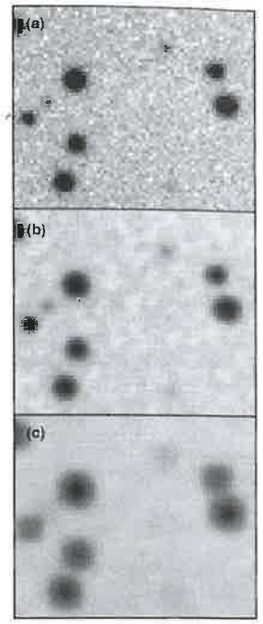


Fig. 9.7 Boxcar filter. (a) A CCD image of a few stars in the cluster M67, displayed as a negative gray-scale (stars are black). (b) The previous image convolved with a 3×3 boxcar. The smoothing effect is most obvious in the sky background. (c) The image in (a) convolved with a 7×7 boxcar.

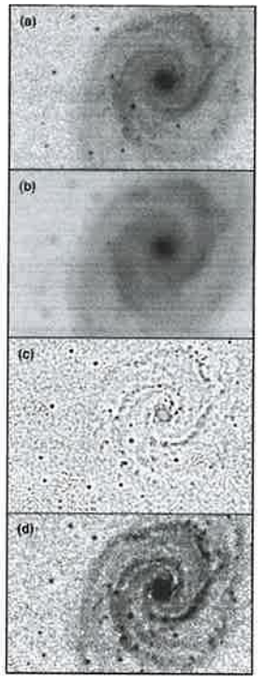


Fig. 9.8 Digital filtration. (a) Image of M51. Full width at half-maximum (FWHM) of the star images is about 2 pixels. (b) After application of a 7-pixel FWHM Gaussian filter to (a). (c) The original after application of a 5 × 5 Laplacian filter, which emphasizes features like stars and the higher-contrast spiral features. Note that sky noise is also enhanced. (d) After application of an unsharp mask based on the Gaussian in (b).

$$\mathbf{F} = 2\mathbf{I} - \mathbf{G} = \frac{1}{3.58}\begin{bmatrix} -0.03 & -0.08 & -0.11 & -0.08 & -0.03 \\ -0.08 & -0.21 & -0.29 & -0.21 & -0.08 \\ -0.11 & -0.29 & -6.78 & -0.29 & -0.11 \\ -0.08 & -0.21 & -0.29 & -0.21 & -0.08 \\ -0.03 & -0.08 & -0.11 & -0.08 & -0.03 \end{bmatrix}$$

Other forms of filtration are not convolutions as defined by Equation (9.2), but do utilize the idea illustrated in Figure 9.6 – the value of a pixel in the filtered image is determined by applying some algorithm to the neighboring pixels as described by a kernel. For example, a 3 × 3 ***local-median filter*** sets the filtered pixel value equal to the median of the unfiltered pixel and its eight neighbors. Other examples use conditional operators: you might map the location of suspicious pixels by computing the statistics, include the sample standard deviation, *s*, inside a 9 × 9 square surrounding a pixel, then apply the operation:

> if the original pixel differs from the local mean by more than 3s,
> then:
> set the value of the filtered pixel to zero,
> otherwise:
> set the value of the filtered pixel to one.

Any digital filtration destroys information, so use it with caution.

9.3 Preprocessing array data: bias, linearity, dark, flat, and fringe

When astronomers speak of ***data reduction***, they are thinking of discarding and combining data to reduce their volume as well as the amount of information they contain. A single CCD frame might be stored as a few million numbers – a lot of information. An astronomer usually discards most of this. For example, he may only care about the brightness or position of a single object in the frame – information represented by just one or two numbers. Ultimately, he might reduce several hundred of these brightness or position measurements to determine the period, amplitude, and phase of a variable star (just three numbers and their uncertainties) or the parameters of a planet's orbit (six numbers and six uncertainties).

Few astronomers enjoy reducing data, and most of us wish for some automaton that accepts what we produce at the telescope (raw images, for example) and gives back measurements of our objects (magnitudes, colors, positions, chemical compositions). In practice, a great deal of automation is possible, and one characteristic of productive astronomy is a quick, smooth path from telescope to final measurement. The smooth path is invariably paved with one or more computer programs working with little human intervention. Before using or writing such a program, the astronomer must get to know his data, understand their imperfections, and have a clear idea of what the data can or cannot reveal.

Eventually, data reduction permits ***data analysis*** and ***interpretation*** – for example, what kind of variable star is this, what does that tell us about how stars evolve? Properly, the boundaries between reduction, analysis, and interpretation are fuzzy, but each step towards the interpretation stage should become less automatic and more dependent on imagination and creativity.

The first and most automatic steps remove the most obvious imperfections. Data should characterize the astronomical source under investigation, not the detector, telescope, terrestrial atmosphere, scattered light, or any other perturbing element. This section examines the very first steps in reducing array data, and explains reductions that must be made to all CCD data (and most other array data), no matter what final measurements are needed. Other authors sometimes refer to these steps as the ***calibration*** of the image. I prefer to separate these steps into the ***preprocessing*** and then ***de-fringing*** of the image.

Consider, then, a raw image, **R**. Of the many imperfections in **R**, preprocessing attempts to correct for:

- **Bias**. If a detector is exposed to no light at all, and is given no time to respond to anything else in its environment, it may nonetheless produce positive values for a particular pixel, $R[x, y]$, when it is read out. In other words, even when $r[x, y]$, the response of the detector, is zero, $R[x, y]$, *is not.* This positive output from a zero-time exposure is called the bias level, or the ***zero level***, and will be present in every frame as a quantity added to the output.
- **Dark response**. If a detector is not exposed to a signal from the telescope, but simply sits in the dark for time *t*, it will in general respond to its dark environment so that $r[x, y]$ is not zero. In most detectors, this dark response is the result of thermal effects. In a CCD, electron–hole pairs are created from the energy present in lattice vibrations at a rate proportional to $T^{3/2} \exp(-a/kT)$, where *a* is a constant that depends on the size of the band gap. Like the bias, the dark response adds ADUs to the readout of every frame. Unlike the bias, dark response will depend on exposure time.
- **Linearity**. The response of a *linear* detector is directly proportional to incoming signal. All practical detectors are either completely non-linear or have a limited range of linearity. One of the appealing characteristics of CCDs is the large range of signal over which their response is linear. Even CCDs, however, ***saturate*** at large signal levels, and eventually cease to respond to incoming photons.
- **Flat field response**. Identical signals generally do not produce identical responses in every pixel of a detector array. Not all pixels in the array respond to light with equal efficiency. This defect can arise because of structural quantum-efficiency differences intrinsic to the array. It can also arise because of vignetting or other imperfections in the optical system like dust, fingerprints, and wildlife (insects turn up in unexpected locations) on filters or windows.

The observer wants to remove these instrument-dependent characteristics from her images in preprocessing. To do so, she must make some reference observations and appropriate image manipulations. We consider each of the four

preprocessing operations in turn. The books by Howell (2006) and by Martinez and Klotz (1998) treat CCD data reduction in greater detail.

9.3.1 Bias frames and overscans

If the observer simply reads her array with zero integration time (actually, the CCD first clears, then immediately reads out), never exposing it to light, she has obtained a ***bias frame***. The bias frame represents the electronic background present in every frame, no matter how short the integration time. The idea, of course, is that this is uninteresting information, and the astronomer needs to subtract the bias frame from every other frame she plans to use. In practice, one bias frame may well differ from another. For larger CCD arrays readout time may be long enough for much to happen, including cosmic-ray hits, local radioactivity, and electronic interference.

It is good practice to obtain many bias frames. For one thing, properly combining several frames will reduce uncertainty about the average level of the bias, as well as minimize the influence of cosmic-ray events. For another, the careful observer should monitor the bias level during an observing run, to guard against any drift in the average level, and to make sure any two-dimensional pattern in the bias is stationary.

Assume for the moment that the bias does not change with time, and that the astronomer takes N bias frames during the run. Call these $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N$. How should he combine these frames to compute $\mathbf{Z}$, the one representative bias image he will subtract from all the other frames? Here are some possibilities:

(1) **Mean**. Set $\mathbf{Z} = \text{mean}(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)$

This is a bad strategy if there are any cosmic-ray hits. Computationally easy, it will dilute the effects of cosmic rays, but not remove them.

(2) **Median**. Set $\mathbf{Z} = \text{median}(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)$

This works well, since the median is relatively insensitive to statistical outliers like the large pixel values generated by cosmic rays. It has the disadvantage that the median is a less robust and stable measure of central value than the mean, and is thus somewhat inferior for those pixel locations not struck by cosmic rays.

(3) **Indiscriminant rejection**. At each $[x, y]$, reject the largest pixel value, then use (1) or (2) on the remaining $(N-1)$ values.

This removes cosmic rays, but is possibly too drastic, since it skews the central values towards smaller numbers. An alternative is to reject both the largest and the smallest values at each location. This discards two entire images worth of data, and skews cosmic-ray pixels to slightly larger numbers.

(4) **Selective rejection**. At each $[x, y]$, reject only those pixels *significantly* larger than the mean, then apply (1) or (2) on the remaining values. To decide whether or not a pixel value is so large that it should be rejected, use a criterion like:

$$z_i[x,y] > \mu[x,y] + k\sigma[x,y]$$

where μ and σ are the mean and standard deviation of the pixel values, (a) at x, y, or (b) over a segment of the image near x, y, or (c) over the entire image. The value of the constant k determines how selective the rejection will be. For a normal distribution, $k = 3$ will reject 14 legitimate (non-cosmic-ray) pixels out of 10,000.

This is an excellent strategy, but is computationally intensive. Strategy 4b or 4c makes it possible to produce a "clean" $\mathbf{Z}$ from a single frame by replacing the rejected pixel value with the mean or median value of its neighbors.

You will undoubtedly think of other advantages or disadvantages to all these strategies, and also be able to compose alternatives. The exact strategy to use depends on circumstance, and we will use the notation

$$\mathbf{Z} = \text{combine}(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)$$

to indicate some appropriate combination algorithm.

What if the bias changes over time? The astronomer might compute different $\mathbf{Z}$ values for different segments of the run, but only if the changes are gradual. A common alternative strategy for CCDs is to use an ***overscan***. You produce overscan data by commanding the clocks on the CCD so that each time the serial register is read, the read continues for several pixels *after* the last physical column has been read out.[2] This produces extra columns in the final image, and these contain the responses of "empty," unexposed pixels, elements of the serial register that have not been filled with charge carriers from the parallel registers. These extra columns are the overscan region of the image and record the bias level during the read. The usual practice is to read only a few extra columns, and to use the median pixel values in those columns to correct the level of the full two-dimensional $\mathbf{Z}$ image. If ω_i is the overscan portion of image i, and Ω_Z is the overscan portion of the combined $\mathbf{Z}$, then the bias frame to apply to image i is

$$\mathbf{Z}_i = \mathbf{Z} + (\text{medianP}(\omega_i - \Omega_Z)) \tag{9.4}$$

Figure 9.9 shows a slightly more complicated application of an overscan. Here the zero level has changed during the read, and shows up in the image most clearly as a change in background in the vertical direction. To correctly remove

[2] It is also possible to continue to read beyond the last exposed *row*. This means the overscan of extra rows will include the dark charges generated during the full read time. For arrays operating with significant dark current, this may be significant. Some manufacturers intentionally add extra physical pixels to the serial register to provide overscan data.

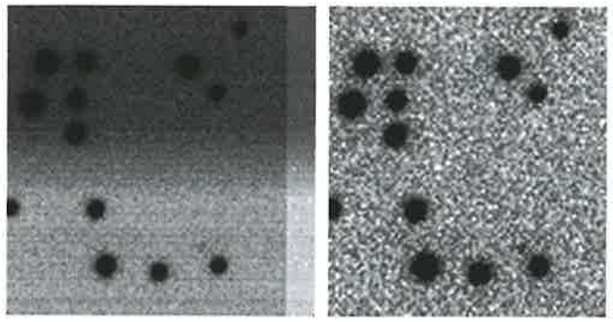


Fig. 9.9 Overscan correction. (a) This frame has a 10-column overscan region on its right edge. The frame in (b) results after the bias frame, corrected for the overscan, is subtracted and the overscan section trimmed from the image. Note that the frame in (b) is displayed with a different gray-scale mapping than in (a).

the bias, the astronomer fitted a one-dimensional function (in the y-direction) to the difference $(\omega_i - \Omega_Z)$, and added that function to **Z**.

9.3.2 Dark current

Even in the absence of illumination, a detector will generate a response during integration time t. This is called the ***dark response***. The rate at which the dark response accumulates is the ***dark current***. Although primarily a thermal effect, dark current will not be the same for every pixel because of inhomogeneities in fabrication. Some pixels, called "hot" pixels, differ from their neighbors not in temperature, but in efficiency at thermal production of charge carriers.

To calibrate for dark current, a long exposure is taken with the shutter closed – this is called a ***dark frame***, **d**. In view of the earlier discussion about cosmic-ray hits and uncertainties, it is best to combine several individual dark frames ($\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_M$) to produce one representative frame:

$$\mathbf{d} = \text{combine}(\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_M)$$

The dark frames should be obtained in circumstances (temperature, magnetic environment) as similar as possible to those prevailing for the data frames. If **d** has exposure time t, then you may compute the ***dark rate*** image as

$$\mathbf{D} = \frac{\mathbf{d} - \mathbf{Z}}{t}$$

or

$$\mathbf{D} = \frac{1}{t}\text{combine}([\mathbf{d}_1 - \mathbf{Z}_1], [\mathbf{d}_2 - \mathbf{Z}_2], \ldots, [\mathbf{d}_M - \mathbf{Z}_M])$$

The second form applies if you are using an overscan correction for each dark frame as in Equation (9.4). You may then correct for dark current and bias on every data frame by subtraction of the image $t\mathbf{D} + \mathbf{Z}$. The units for **D** in the above equations are ADUs per second. However, dark current for CCDs is almost always quoted in units of electrons per second as $\mathbf{D}_e = g\mathbf{D}$, where g is the detector gain in electrons per ADU.

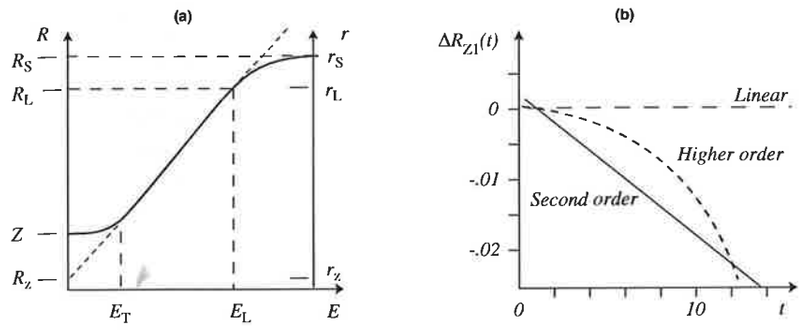


Fig. 9.10 Linearity: (a) A schematic of the output R, in ADU, and the response, r, in electrons, of a single pixel in a detector that is linear over a restricted input range. The sloped dashed line is Equation (9.5). (b) An experimental determination of a correction for non-linearity as explained in the text.

Observers routinely cool detectors to reduce dark current and its associated noise. In some cases (e.g. a CCD at –90 °C) the dark rate may be so low that you can omit the correction. A careful observer, though, will always take dark frames periodically, even if only to verify that they are not needed. It also appears that the dark current in some multi-pinned-phase (MPP) CCDs is somewhat non-linear, which means you must either model the non-linearity or take dark frames whose exposure times match those of the data frames.

9.3.3 Detector linearity

All practical devices depart from linearity. If a pixel receives photons at rate $P[x, y]$ for time t, it has a ***linear output*** if

$$R[x,y] = R_z[x,y] + t(D[x,y] + Q[x,y]P[x,y]) = R_z[x,y] + bE[x,y] \qquad (9.5)$$

where $R_z[x,y]$, $D[x,y]$ and $Q[x,y]$ are pixels in time-independent arrays, respectively a zero level, the dark rate, and the efficiency in ADUs per photon. A similar equation applies to the response in electrons, r. The CCD response curves resemble Figure 9.10a, where the horizontal variable, E, the exposure, is a quantity proportional to total input (photon count plus dark current). The detector in the figure is linear between a threshold exposure, E_T, and an upper limit, E_L. The output labeled Z is the ***bias*** of the pixel.

The pixel ***saturates*** at response r_S and output R_S. Recall that saturation in a CCD pixel results if its potential well fills and the MOS capacitor can store no additional charge carriers. Saturated pixels are insensitive to light and to dark current, but in many devices can ***bloom*** and spill charge carriers into neighboring pixels.

Charge-coupled devices have remarkably good linearity over several orders of magnitude. Equation (9.5) typically holds to better than 1% over a range $E_L \approx E_s \approx 10^5 E_T$. Moreover, the threshold effect for a CCD is very small, so that $R_z \approx Z$. Recall also that ADCs may be set so that pixels reach digital saturation at $r < r_L$.

Many other devices have significant non-linear behavior well short of their saturation levels, and the observer must remove this non-linearity. Figure 9.10b illustrates one method for empirically measuring a correction.

Suppose R_1 is the response of one pixel to a one-second exposure. If this pixel is a linear detector, satisfying Equation (9.5), then different integration times with a constant light source should give response $R_L = R_z + t(R_1 - R_z)$. However, the actual detector has the non-linear response, R. A series of exposures of different lengths can generate a plot of the quantity

$$\Delta R_{z1}(t) = \frac{1}{t}(R - R_L) = \frac{R - R_z}{t} - (R_1 - R_z) \tag{9.6}$$

as a function of exposure time t. If a linear fit to this has non-zero slope (e.g. the solid line in Figure 9.10b) that fit indicates a quadratic equation for the corrected linear response in the form

$$\mathbf{R}_L = \mathrm{lin}(\mathbf{R}) = a + b\mathbf{R} + c\mathbf{R}^2 \tag{9.7}$$

Strictly speaking, the constants a, b, and c could be different for different pixels in an array. In most cases, uncertainties in the pixel-to-pixel variation justify using average values for the entire array.

9.3.4 Flat field

Correcting for pixel-to-pixel variations in device sensitivity is both the most important and the most difficult preprocessing step. Conceptually, the correction procedure is very simple. The astronomer takes an image of a perfectly uniform (or "flat") target with the complete observing system: detector, telescope, and any elements like filters or obstructions that influence the focal-plane image. If the observing system is equally sensitive everywhere, every pixel in this ***flat-field image***, after correction for the bias and dark, should produce an identical output. Any departure from uniformity in the corrected flat-field image will map the sensitivity of the system, in the sense that pixels registering higher counts are more sensitive. Figure 9.11 shows a raw CCD image, a flat-field image, and the original image after the flat-field correction.

At least three practical difficulties hamper the kind of correction illustrated. First, it is difficult to produce a sufficiently (i.e. 0.5% or better) uniform target. Second, sensitivity variations are in general a function of wavelength. Therefore, the spectrum of the target should match that of the astronomical sources of interest. Spectrum matching becomes especially troublesome with multiple sources of very different colors (stars and background sky, for example) in the same frame. Finally, it is difficult to guarantee that the "complete observing system" remains unchanged between the acquisition of the flat field and acquisition of the data frames.

In common practice, observers employ three different objects as the flat-field target: (1) the bright twilight sky, (2) the dark night sky, and (3) a nearby

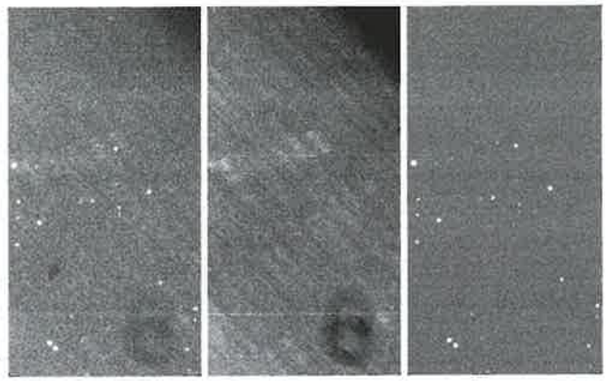


Fig. 9.11 (a) A section of an unprocessed CCD image of a star field. (b) A combined twilight flat for the same section. The two dark rings in the lower right are the shadows of two dust particles on the CCD window. The dark region in the upper right results from vignetting by the filter holder. The section is shown after preprocessing in (c).

object – usually an illuminated surface inside the observatory dome. Images of these sources are usually termed twilight, dark sky, and dome flats, respectively. Each has advantages and disadvantages.

Twilight flats

The clear twilight sky is not uniform: it is brighter all the way around the horizon than it is near the zenith, and, of course, brighter in the direction of the rising or recently set Sun. By pointing towards the zenith (the exact location of the "flat" spot – usually 5–10 degrees anti-solar from the zenith – is slightly unpredictable), the observer finds a target uniform to about 1% over a one-degree field. It is rare to do better than this. For narrow fields of view, this is acceptable. Clouds usually prohibit good flats.

The advantages of the twilight-sky target are that, for a brief interval, it is the right brightness, and relatively uniform. Moreover, observing in twilight means flat-field calibrations do not consume valuable nighttime hours. The disadvantages are:

- Large-scale uniformity is limited by the natural gradient in the twilight sky, and small-scale uniformity is limited by the gradual appearance of star images as twilight fades.
- The twilight sky has a spectrum that is quite different from that of most astronomical sources, as well as that of the night sky.
- Twilight brightness and spectrum both change rapidly. The duration of useable twilight is short, and with large arrays (long readout times), or with many filters, it becomes difficult to accumulate sufficient numbers of images.
- Scattered skylight near the zenith has a strong linear polarization, and the flat field of some systems may be polarization sensitive.

Dark-sky flats

The emission from the dark (moonless!) night sky is a tempting source for flat fields. Uniformity is perfect at the zenith and degrades to about two percent per

degree at a zenith angle near 70 degrees. Moreover, the spectrum of the night sky is identical to one source of interest: the background that will usually be subtracted from all data frames, an especially important advantage if measuring sources fainter than the background sky. High sky brightness is the rule in the ground-based infrared, where dark-sky flats are the rule.

Offsetting these attractive characteristics are some potent negatives for dark-sky targets. First, stars are everywhere. Any dark-sky flat will inevitably contain many star images, marring the target's uniformity. The observer can remove star images and construct a good flat with the ***shift-and-stare*** or ***dither*** method. The astronomer takes many deep exposures of the dark sky, taking care always to "dither" or shift the telescope pointing between exposures by at least many stellar image diameters. He then combines these in a way that rejects the stars. For example, take five dithered images of a dark field. If the density of stars is low, chances are that at any $[x, y]$ location, at most one frame will contain a star image; so computing the median image will produce a flat without stars. More sophisticated combination algorithms can produce an even better rejection of stellar images. The shift-and-stare method should also be employed for twilight flats, since (1) they will usually contain star images and (2) telescope pointing should be shifted back to the flat region near the zenith for each new exposure anyway.

Understand the limitations of shift-and-stare: the scattered-light halos of bright stars can be many tens of seconds of arc in radius and still be no fainter than one percent of the background. Removing such halos, or extended objects like galaxies, can require large shifts and a very large number of exposures.

A second difficulty is that the dark sky is – well – dark. In the visible bands, one typically requires 10^2 to 10^6 times as long to count the same number of photons on a dark-sky frame as on a twilight frame. Sometimes, particularly in broad bands with a fast focal-ratio telescope, this is not a serious drawback, but for most work, it is crucial. Each pixel should accumulate at least 10^4 electrons to guarantee one percent Poisson uncertainty; so dark-sky flats will typically require long exposure times. They are consequently very costly, since time spent looking at blank sky might otherwise be spent observing objects of greater interest.

A modification of shift-and-stare can sometimes help here. If the objects of interest occupy only a small fraction of the frame, then it should be possible to dither and collect many unaligned data frames. The median of these unaligned frames is the dark-sky flat, and no time has been "wasted" observing blank sky, since the flat frames also contain the science.

Dome flats

A source inside the dome is an attractive flat-field target, since the astronomer in principle controls both the spectrum and the intensity of the illumination, and observations can be taken during daylight. With very small apertures, it is possible to mount a diffusing light box at the top of the telescope tube, but most

telescopes are simply pointed at a white screen on the inside of the dome. In practice, in a crowded dome, it is often difficult to set up a projection system that guarantees uniform illumination, the shadow of a secondary may become important in the extrafocal image, and there is an increased possibility of introducing unwanted light sources from leaks or reflections. Nevertheless, dome flats are a very important flat-field calibration technique.

Computing simple flats

Assume you have collected N flat-field images, all taken through a single filter, using one of the targets discussed above. If $\mathbf{s}_i$ is one if these raw images, then the first step in creating the calibration frame is to remove its bias, dark, and non-linearities:

$$\mathbf{f}_i' = \text{lin}(\mathbf{s}_i) - \mathbf{Z}_i - t_i\mathbf{D}$$

As before, $\mathbf{D}$ is the dark rate, t_i is the exposure time, and $\mathbf{Z}_i$ is the overscan-corrected bias. Next, to simplify combining frames each, frame should be normalized so that the median pixel has a value of 1.0 ADU:

$$\mathbf{f}_i = \mathbf{f}_i'/\text{medianP}(\mathbf{f}_i')$$

Finally, all normalized frames should be combined to improve statistics, as well as to remove any stars or cosmic-ray events:

$$\mathbf{F}_C = \text{combine}(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_N)$$

A different calibration frame must be produced for each observing configuration. Thus, there must be a different flat for each filter used, and a different set of flats whenever the observing system changes (e.g. the detector window is cleaned, or the camera rotated).

Compound flats

Given the imperfections of all three flat-fielding techniques, the best strategy sometimes combines more than one technique, applying each where its strengths are greatest. Thus, one uses a dome flat or twilight flat to establish the response of the system on a small spatial scale (i.e. the relative sensitivity of a pixel compared with those of its immediate neighbors.) Then, one uses a smoothed version of a dark-sky flat to establish the large-scale calibration (e.g. the response of the lower half of the detector relative to the upper half). The idea is to take advantage of both the absence of small-scale non-uniformities (stars) in the dome or twilight target as well as the absence of large-scale non-uniformities (brightness gradients) in dark-sky targets. To create the compound flat-field calibration, assume that $\mathbf{F}_S$ and $\mathbf{F}_L$ are calibration frames computed as described in the previous section. Frame $\mathbf{F}_S$ is from a target with good small-scale uniformity, and $\mathbf{F}_L$ from one with good large-scale uniformity. Now compute the ratio image and smooth it:

$$\mathbf{c} = \frac{\mathbf{F_L}}{\mathbf{F_S}}$$
$$\mathbf{C} = \text{conv}\{\mathbf{b}, \mathbf{c}\}$$

The kernel in the convolution, **b**, should be chosen to remove all small-scale features from image **c**. Image **C** is sometimes called an ***illumination correction***. The corrected compound flat is just

$$\mathbf{F} = \mathbf{F_S} \cdot \mathbf{C}$$

9.3.5 Preprocessing data frames

Suppose a CCD has output $\mathbf{R_i}$ in response to some astronomical scene. Preprocessing corrects this image for non-linearity, bias, dark, and flat field:

$$\mathbf{R_{pi}} = \frac{\text{lin}(\mathbf{R_i}) - \mathbf{Z_i} - t_i\mathbf{D}}{\mathbf{F}} \tag{9.8}$$

Preprocessing non-CCD array data can differ slightly from the above procedures. For infrared arrays read with double-correlated sampling, the output is the difference between reads at the beginning and the end of an exposure, so bias values cancel and **Z** is numerically zero. Also in the infrared, emission from the variable background often dominates the images, so much so that raw images may not even show the location of sources before sky subtraction. A common observing practice then is to "chop" telescope pointing between the object investigated and the nearby (one hopes, blank) sky to track its variations. Many infrared-optimized telescopes employ ***chopping secondary mirrors*** that efficiently implement rapid on-source/off-source switching. ***Chopping*** is in this context different from ***nodding*** – manually moving the telescope in the shift-and-stare technique.

In the infrared, then, these high-signal sky frames are usually combined to form the flat-field image. A typical preprocessing plan might go like this: if $\mathbf{s_1}, \mathbf{s_2}, \ldots, \mathbf{s_n}$ are the sky exposures and **F** is the flat, then

$$\begin{aligned}
\mathbf{S} &= \text{combine}(\mathbf{s_1}, \mathbf{s_2}, \ldots \mathbf{s_n}) \\
\mathbf{f}' &= \mathbf{S} - \mathbf{d} \\
\mathbf{F} &= \mathbf{f}'/\text{medianP}(\mathbf{f}') \\
\mathbf{S_i} &= a_i\mathbf{f}' \\
\mathbf{R_{pi}} &= \frac{\mathbf{R_i} - \mathbf{S_i} - \mathbf{d}}{\mathbf{F}}
\end{aligned} \tag{9.9}$$

We assume that both $\mathbf{R_i}$ and **S** are first corrected to remove non-linearity. In the fourth equation, a_i is a scaling factor that matches the medianP or modeP of **S** to the sky level of the data frame, which might be computed from the adjacent (in time) sky frames.

9.3.6 Fringing

Monochromatic light can produce brightness patterns in a CCD image due to reflection and interference within the thin layers of the device. Fringing is usually due to narrow night-sky emission lines, and if present means that the image of the background sky (only) contains the superimposed fringe pattern. It tends to occur in very narrow band images, or in images in the far red where night-sky upper-atmospheric OH emission is bright. The fringe pattern is an instrumental artifact, and should be removed.

The fringe pattern depends on the wavelengths of the sky emission lines, but its amplitude varies with the ratio of line to continuum intensity in the sky spectrum, which can change, sometimes rapidly, during a night. Fringes will not appear on twilight or dome flats, but will show up on a dark-sky flat produced by the shift-and-stare method.

If fringing is present, you should *not* use the dark-sky image to create a flat (use twilight), but use the dark-sky image to create a fringe calibration. If **S** is the dark-sky image as in Equation (9.9) and **B** is the processed dark sky image, we have

$$\mathbf{B} = \frac{\mathbf{S} - \mathbf{d}}{\mathbf{F}} = \mathbf{B_c} + \mathbf{B_f} = \mathbf{B_c} + A\mathbf{b_f}$$

Here, $\mathbf{B_c}$ is the part of **B** due to the continuum, and $\mathbf{B_f}$ the part due to fringes. We treat $\mathbf{B_f}$ as the product of an amplitude, A, and normalized pattern, $\mathbf{b_f}$. If $\mathbf{B}'$ is a slightly smoothed version of **B**, then:

$$\begin{aligned}
B_c[x,y] &= \text{minP}(\mathbf{B}') \\
\mathbf{B_f} &= \mathbf{B} - \mathbf{B_c} \\
A &= \text{maxP}(\mathbf{B}') - \text{minP}(\mathbf{B}')
\end{aligned}$$

Removing the fringes from a processed science image $\mathbf{R_{pi}}$ is then simply a matter of measuring the fringe amplitude on the science image, A_i, and subtracting the calibration fringe pattern scaled to match:

$$\mathbf{R_{pfi}} = \mathbf{R_{pi}} - \frac{A_i}{A}\mathbf{b_f}$$

9.4 Combining images

After preprocessing, astronomers often combine the resulting images. You might, for example, have acquired a dozen images of an extremely fascinating galaxy, and reason (correctly) that adding all of them together digitally will produce a single image with superior signal to noise. The combined image should show features in the galaxy, especially faint features, more clearly than do any of the individual frames. In another example, you

may be trying to observe a nebula whose angular size is greater than the field of view of your CCD. You would like to assemble a complete image of the nebula by combining many of your small CCD frames into a large mosaic. Combining images is a tricky business, and this section provides only a brief introduction.

9.4.1 Where is it? The centroid

Suppose you want to combine images **A** and **B**. An obvious requirement is that the pixel location of a source in **A** must be the same as its location in **B**. But what exactly *is* the location of a source? We can compute the location by a two-step process, preferably through instructions to a computer:

1. Decide which pixels belong to the source.
2. Compute an appropriate centroid of those pixels.

It is not totally obvious how to complete step 1. In the case of point sources like stars, you can get a good idea of their approximate locations by applying a Laplacian filter (whose size matches the point-spread function – see the next section) to a digital frame and noting the maxima of the filtered image. To decide which pixels around these locations are part of a star image and which are not requires some thought. For example, if you ask which pixels in a typical CCD image receive light from a bright star in the center of the frame, the answer, for a typical ground-based point spread function, is: "all of them." A better question is: "which pixels near the suspected star image receive a signal that is (a) larger than (say) 3σ above the background noise and (b) contiguous with other pixels that pass the same test?"

Figure 9.12 illustrates this approach (there are others) – the bar heights indicate pixel values in a small section of a CCD frame. Although most of the pixels in the area probably registered at least one photon from the star in

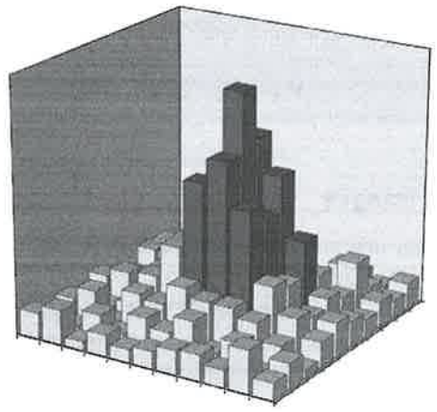


Fig. 9.12 Bar heights represent pixel values near a faint star image. Darker bars are high enough above the background to qualify as image pixels.

the center, only those colored dark gray stand out from the background according to the 3σ contiguous criterion.

With the "star pixels" identified, you can then compute their ***centroid***. Typically, you consider only that part of the dark gray volume in Figure 9.12 that is above the background level, and compute the (x, y) coordinates (fractional values permitted) of its center of mass. If $R[x, y]$ is a pixel value and if B is the local background level, then the centroid coordinates are:

$$x_{\text{cen}} = \frac{\sum_x \sum_y x(R[x,y] - B)}{\sum_x \sum_y (R[x,y] - B)}, \quad y_{\text{cen}} = \frac{\sum_x \sum_y y(R[x,y] - B)}{\sum_x \sum_y (R[x,y] - B)} \tag{9.10}$$

The sums include only star pixels. Depending upon the signal-to-noise ratio (SNR) in the sums in Equations (9.10), the centroid can locate the image to within a small fraction of a pixel.

9.4.2 Where is it, again? PSF fitting

Finding the centroid of an image is computationally simple, but works well only in cases where images are cleanly isolated. If images blend together the centroid finds the center of the blended object. Even if there is no confusion of images, one object may asymmetrically perturb the background level of another (a galaxy near a star, for example).

In situations like this, you can use knowledge of the ***point-spread function*** (*PSF*) to disentangle blended and biased images. The procedure is to fit each of the *stellar* (only) images on the frame with a two-dimensional PSF, adjusting fits to account for all the flux present. The actual algorithm may be quite complex, and special complications arise if there are non-stellar objects present or if the shape of the PSF varies from place to place due to optical aberrations or to anisoplanatism in adaptive systems. Despite the difficulties, PSF fitting is nevertheless essential for astrometry and photometry in crowded fields.

9.4.3 Aligning images: shift, canvas and trim

Figure 9.13 shows two CCD frames, **A** and **B**, of M33 at different telescope pointings. Each frame has dimensions $x_{\text{max}} = 256 \times y_{\text{max}} = 256$. We consider the problem of ***aligning*** the two images by applying a ***geometric transformation*** to each – a geometric transformation changes the pixel coordinates of image data elements. In this example, we make the transformation by first measuring the $[x, y]$ coordinates for three stars in the area common to both frames. Suppose that on average, we find for these objects that $x_{\text{B}} - x_{\text{A}} = \Delta x_{\text{B}} = -115$ and that $y_{\text{B}} - y_{\text{A}} = \Delta y_{\text{B}} = 160$. (Assume for now that coordinates are restricted to integers.) There are two possible goals in making the transformation.

First, we might wish to make a new image that contains data from *both* **A** *and* **B**, perhaps to improve the SNR. Do this by creating **A′** and **B′**, two small images that contain only the overlap area from each frame:

$$\begin{aligned} A'[\xi,\eta] &= A[\xi + \Delta x_B, \eta] \\ B'[\xi,\eta] &= B[\xi, \eta + \Delta y_B] \end{aligned} \tag{9.11}$$

The values stored in the pixels of **A′** and **B′** and are the same as the values in **A** and **B**, but they have different coordinates. The ***translation*** operation executed by Equation (9.11) simply slides **B** and **A** until coordinates match. An important step in making the new images discards or ***trims*** any pixels that fall outside the overlap region. Specifically, we trim all pixels except those with coordinates $1 \le x_{max} - |\Delta x_B|$ and $1 \le y_{max} - |\Delta y_B|$. Both trimmed images thus have the same size, which means we can combine them (add, average, etc.) using image arithmetic. For example:

$$\mathbf{C}_{AND} = \mathbf{A}' + \mathbf{B}'$$

More complicated combination algorithms might be appropriate, especially with large numbers of images.

A second mode of image combination arises when we note that the large galaxy in Figure 9.13 does not fit in a single frame, and we wish to combine the two frames to make a wide-angle view. We want an image that includes *every* valid pixel value from *either* **A** *or* **B**. The procedure is simple: we make two ***canvases***, $\mathbf{C_A}$ and $\mathbf{C_B}$, each with dimensions $x_{max} + |\Delta x_B|$ by $y_{max} + |\Delta y_B|$, large enough to include all pixels. Then, we "paste" each image in the appropriate section of its canvas, and then combine the large canvases into one large final image. In our example, the canvases have coordinates ξ', η', and the operations that paste the images onto their canvases are:

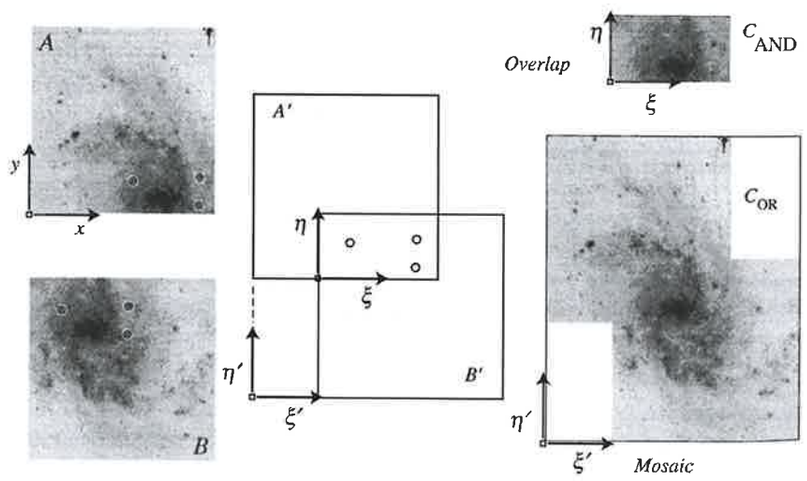


Fig. 9.13 Aligning and combining two images. Alignment and transformation are based on the coordinates of the three marked stars in the overlap region. See text for details.

$$C_A(\xi',\eta') = \begin{cases} A'(\xi',\eta' - \Delta y_B) = A(\xi',\eta' - \Delta y_B), & 1 \le \xi' \le x_{max}, 1 \le \eta' \le y_{max} + \Delta y_B \\ -10,000 & \text{otherwise} \end{cases}$$

$$C_B(\xi',\eta') = \begin{cases} B'(\xi',\eta' - \Delta y_B) = B(\xi' + \Delta x_B,\eta'), & 1 \le \xi' \le x_{max} + \Delta x_B, 1 \le \eta' \le y_{max} \\ -10,000 & \text{otherwise} \end{cases}$$

The arbitrary large negative value of –10,000 simply flags those pixels for which there are no data. Any value that cannot be confused with genuine data can serve as a flag. We can combine the two canvases

$$\mathbf{C}_{OR} = \text{combine}(\mathbf{C_A}, \mathbf{C_B})$$

with some appropriate algorithm. For example, the pseudo-code:

$$\text{IF } \{C_A[x,y] \ne -10,000 \quad \text{AND} \quad C_B[x,y] \ne -10,000\}$$

$$\text{THEN } C_{OR}[x,y] = \frac{1}{2}[C_A[x,y] + C_B[x,y]]$$

$$\text{ELSE } C_{OR}[x,y] = \max[C_A[x,y], C_B[x,y]]$$

will compute values for mosaic pixels for which there are some data, and put a flag (–10,000) in those where there is no data.

9.4.4 Aligning images: geometric transformations

Translations are only one of several kinds of geometric transformation. Suppose, for example, you wish to combine images from two different instruments. The instruments have different pixel scales[3] (in seconds of arc per pixel); so one set of images requires a scale change, or ***magnification***. The transformation is

$$\begin{aligned} x &= \xi / M_x \\ y &= \eta / M_y \end{aligned}$$

Again, $[\xi, \eta]$ are the coordinates in the new image, and the equations allow for stretching by different amounts in the x- and y-directions.

Small ***rotations*** of one image with respect to another might occur if a camera is taken off and remounted on the telescope, or if images from different telescopes need to be combined, or even as the normal result of the telescope mounting (e.g. imperfect polar alignment in an equatorial, an imperfect image rotator in an alt-azimuth, or certain pointing schemes for a space telescope). If

[3] Scale differences can have subtle causes: the same CCD–telescope combination can have slightly different scales because of focal-length changes caused by thermal effects on mirrors or chromatic effects in lenses.

A′ is the image produced when **A** is rotated about its origin counterclockwise through angle θ, then $A'[\xi, \eta]$ has the same pixel value as A[x, y] if

$$x = \xi \cos\theta + \eta \sin\theta$$
$$y = \eta \cos\theta - \xi \sin\theta$$

For wide fields, optical ***distortions*** can become significant (e.g. the Seidel pincushion or barrel distortion aberration, which increases with the cube of the field size). These require relatively complicated transformations.

In creating mosaics from images with different telescope pointings, ***projection effects*** due to the curvature of the celestial sphere also need to be considered. Such effects have long been an issue in photographic astrometry, and chapter 11 of Birney *et al.* (2006) outlines a simple treatment of the problem.

To derive any geometric transformation, the general approach is to rely on the locations of objects in the field. In the final transformed or combined image we require that a number of reference objects (1, 2, 3, . . . , N) have pixel coordinates $(\xi_1, \eta_1), (\xi_2, \eta_2), \ldots, (\xi_N, \eta_N)$. We can call these the ***standardized coordinates*** – they might be coordinates derived from the known right ascention (RA) and declination (Dec) of the reference objects, or might be taken from the actual pixel coordinates on a single image. Now, suppose one of the images you wish to transform, image **B**, contains some or all of the reference objects, and these have coordinates

$$(x_{B1}, y_{B1}), (x_{B2}, y_{B2}), \ldots, (x_{BM}, y_{BM}) \quad M \leq N$$

Your task is to find the transformations

$$x = f_B[\xi, \eta]$$
$$y = g_B[\xi, \eta]$$

that will tell you the pixel values in **B** that correspond to every pair of standardized coordinates. You specify the forms for the functions from your knowledge of how the images are related. You might, for example, expect that narrow-field images from the same instrument would require just a simple translation, while wide-field images from different instruments might need additional correction for magnification, rotation, distortion, or projection. For a given functional form, the usual approach is to use a least-squares technique to find the best values for the required constants $\Delta x_A, \Delta y_A, \theta, M_x$, etc. Note that some geometric transformations may not ***conserve flux*** (see the next section).

Reducing data from digital arrays very commonly involves a two-step ***align and combine*** procedure:

(a) apply geometric transforms on a group of images to produce a new set aligned in a common system of coordinates, correcting for flux changes if necessary, then
(b) combine the aligned images with an appropriate algorithm.

This procedure is often termed ***shift and add***. Basic observational issues make shift and add an indispensable technique, and we already discussed some of these in the context of the ***shift-and-stare*** observing technique for flat-field calibration images. (You do shift and *stare* at the telescope, shift and *add* the data-reduction computer.) To shift and stare, or ***dither***, the observer takes several exposures of the same scene, shifting the telescope pointing slightly between exposures. The aim is to produce a number of equivalent exposures, no two of which are perfectly aligned.

There are many reasons to take several short exposures rather than one long one. For one thing, all arrays saturate, so there may well be an exposure time limit set by the detector. Second, the only way to distinguish a pixel illuminated by a cosmic-ray strike from one illuminated by an astronomical object is to take multiple images of the scene. Astronomical objects are present in every image at the same standardized coordinate location; cosmic rays (and meteor trails and Earth satellites) are not. Similarly, bad pixels, bad columns, and the insensitive regions in array mosaics always have the same pre-transformation coordinate, but different standardized coordinates. When images are aligned, the bad values due to these features in one frame can be filled in with the good values from the others.

9.4.5 Interpolation

Geometric transforms set the values of the pixel at standardized coordinates $[\xi_j, \eta_j]$ in a new image to those at pixel at (x_j, y_j) in the original image; see Figure 9.14. Now, ξ_j and η_j must be integers, but x_j and y_j generally contain fractional parts. Therefore, we use round brackets (non-integers permitted) to write, symbolically

$$B'[\xi_j, \eta_j] = B(f_B[\xi_j, \eta_j], g_B[\xi_j, \eta_j]) = B(x_j, y_j)$$

Since we only know the pixel values for the image **B** at locations where x and y are integers, we must use the pixel values at nearby integer coordinates to *estimate* $B(x_j, y_j)$ – the value a pixel *would* have if it were centered precisely at the non-integer location (x_j, y_j).

We could, for example, ignore any image changes at the sub-pixel level, and simply round x_j and y_j up or down to the nearest integers, and set $B(x_j, y_j)$ equal to the value of the ***nearest pixel***. This is simple, and largely preserves detail, but will limit the astrometric accuracy of the new image.

Bilinear interpolation often gives a more accurate positional estimate. Figure 9.14 shows the point (x_j, y_j) relative to the centers of actual pixels in the original image: we compute x_0 and y_0, the values of x_j and y_j *rounded down* to the next lowest integers. Thus, the values of the four pixels nearest the fractional location (x_j, y_j) are

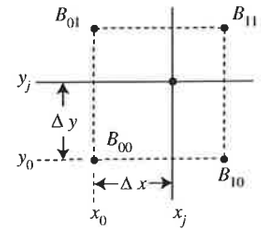


Fig. 9.14 Bilinear interpolation. The methods finds the value of the image intensity at point (x_j, y_j), given the nearest pixel values.

$$B_{00} = B[x_0, y_0], \qquad B_{10} = B[x_0 + 1, y_0]$$
$$B_{01} = B[x_0, y_0 + 1], \qquad B_{11} = B[x_0 + 1, y_0 + 1]$$

If we assume that **B** changes linearly along the axes, we can make two independent estimates for the value of $B(x_j, y_j)$. These average to the bilinear interpolated value:

$$B(x_j, y_j) \approx (1 - \Delta x)(1 - \Delta y)B_{00} + (\Delta x)(1 - \Delta y)B_{10} + (1 - \Delta x)(\Delta y)B_{01} + (\Delta x)(\Delta y)B_{11} \tag{9.12}$$

where

$$\Delta x = x_j - x_0, \quad \Delta y = y_j - y_0$$

Bilinear interpolation preserves astrometric precision and affects photometry in predictable ways. (Any geometric transformation in which the output grid does not sample the input grid uniformly will change the photometric content of the transformed image.) As you can see from Equation (9.12), the procedure essentially takes a weighted average of four pixels – as such, it *smoothes* the image. Bilinear interpolation chops off peaks and fills in valleys, so an interpolated image is never as sharp as the original; see Figure 9.15. Furthermore, the smoothing effect artificially reduces image noise.

If resolution is of great concern, it is possible to fit the pixels of the original image with a higher-order function that may preserve peaks and valleys. The danger here is that higher-order surfaces may also produce artifacts and photometric uncertainties, especially for noisy images. Nevertheless, it is not unusual for astronomers to use higher-order fitting techniques like bicubic interpolation or B-spline surfaces.

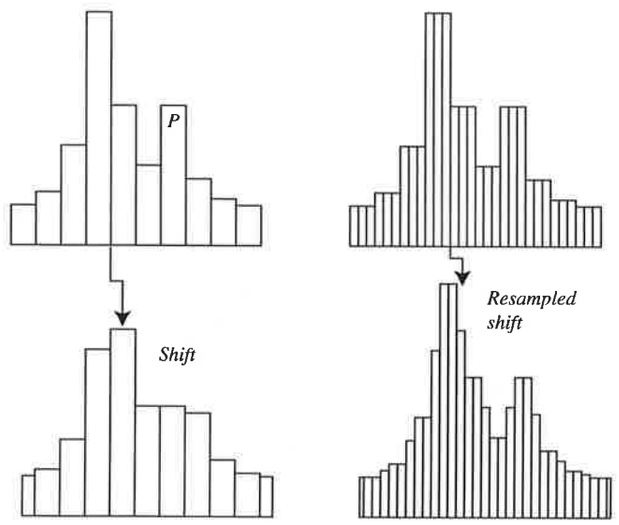


Fig. 9.15 The original pixel values in the upper left are shifted by 0.5 pixels to the right. In the lower left, the shift and linear interpolation smoothes the original, removing peak P, which may be due to noise or some real feature. At the lower right, 3 × resampling preserves more detail after shift and interpolation.

9.4.6 Resolution: resampling, interlace, and drizzle

Geometric transformations are essential for combining images with the shift-and-add image technique. Transformations, however, require either interpolation (which degrades resolution) or the "nearest-pixel" approximation, which degrades positional information. ***Resampling*** the original image at higher magnification circumvents some of the image degradation that accompanies interpolation, and in some cases can actually *improve* the resolution of the combined image over that of the originals.

The idea is to make the pixels of the output, or transformed, image smaller (in seconds of arc), and thus more closely spaced and numerous, than the pixels of the input image. In other words, the scale (in arcsec per pixel) of the standardized coordinates is larger than the scale of the original input coordinates. We discuss three resampling strategies.

The first is just a modification of the shift-and-add (and interpolate) algorithm. All that is done is to resample each input image by an integral number (e.g. each original pixel becomes nine pixels in the resampled version). After shifting or other transformations, resampling mitigates the smoothing effect produced by interpolation, since this smoothing effect is on the scale of the output pixels. Figure 9.15 shows a one-dimensional example. An image is to be shifted 0.5 pixels to the right from its position in the original. The left-hand column shows the result of the shift and linear interpolation without resampling, and the right column shows the same result if the output pixels are one-third of the size of the input. Linear interpolation in each case produces some smoothing, but the smoothing is less pronounced with the finer grid. Compared to using the original pixel sizes, aligning multiple images on the finer output grid will of course improve the resolution of their combined image.

The second method is usually called ***interlace***, and is in some ways analogous to the nearest-pixel approach described earlier. The interlace algorithm examines each input pixel (i.e. $B[x, y]$ at only integer coordinates), locates its transformed center in a particular output pixel in a finer grid (but again, only integer coordinates), and copies the input value to that single output pixel. There is no adjustment for fractional coordinates, nor for the fact that the input pixel may overlap several output pixels. Figure 9.16a gives an example of a shifted and rotated input grid placed on an output grid with smaller pixels. The center of each input pixel is marked with a black dot. Interlacing this single input places values in the output pixels (i.e. the dark-colored pixels), "hit" by the dots, and "no value" or "zero-weight" flags in the other pixels.

Interlace for a single image is a flawed approach. First, it creates a discontinuous image, since only some fraction of the output pixels will score a "hit", and the remainder will have zero weight. Second, we have introduced positional errors because we ignore any fractional coordinates.

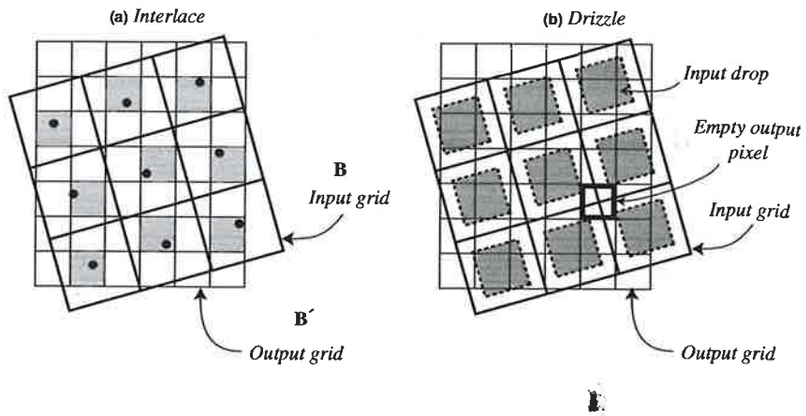


Fig. 9.16 Resampling an input grid. The interlace technique (a) regards values in the input grid as if concentrated at points. Grayed pixels on the output copy the values from the input points, white-colored output pixels have no value. The drizzle method (b) assumes values are spread over a square "drop" smaller than an input pixel. Most output pixels overlap one or more input drops, although some, as illustrated, may overlap none.

Both problems become less significant as more images of the same scene are added to the output. If each addition has a slightly different placement on the output grid, a few additions could well fill in most output pixels with at least one valid value. Moreover, positional information improves as the interlace fills and averaging reduces uncertainty in the brightness distribution.

The combined image is a weighted mean of all the shifted frames, with the weight, $w_i[\xi, \eta]$, of a particular pixel either one (if it is a hit) or zero (if no hit *or* if we decide the hit is by a cosmic ray or by a bad pixel). Thus, the combined image **C** is

$$C[\xi,\eta] = \frac{1}{\sum_{i=1}^{N} w_i[\xi,\eta]} \left(w_1[\xi,\eta]B'_1[\xi,\eta] + w_2[\xi,\eta]B'_2[\xi,\eta] + \ldots + w_N[\xi,\eta]B'_N[\xi,\eta]\right) \tag{9.13}$$

We cannot use Equation (9.13) for any pixel in **C** with a combined weight of zero. In this case, the pixel has no valid value. It is possible to interpolate such a missing value from the surrounding output pixels, but this will cause photometric errors unless the "no-value" status is due to masking cosmic rays or bad pixels.

Interlacing shifted images has the potential for actually improving image resolution in the case where the camera resolution is limited by the detector pixel size rather than by the telescopic image itself. Figure 9.17, shows the interlaced result for a one-dimensional example: a double source with a separation of 1.3 input pixels, with each source FWHM = 0.8 pixels. Three dithered input images are shown, none of which shows the double nature of the source, as well as the interlaced combination with 1/3-size output pixels. The combined image resolves the two components.

The interlace technique is powerful, but unfortunately difficult to execute observationally. Suppose, for example, a detector has 0.8 arcsec pixels. To effectively interlace images on an output grid of pixels half that size, the

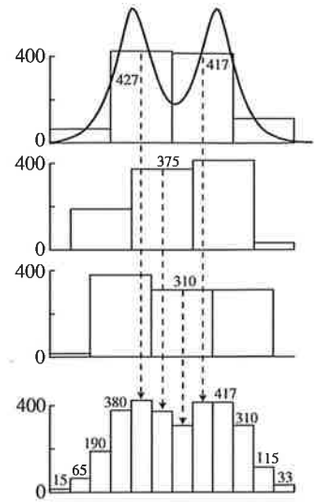


Fig. 9.17 The interlace method in a one-dimensional example. The actual brightness distribution of a double source is sampled with large pixels. In three samples (upper plots) displaced from one another by 1/3 of a pixel, no indication of the double nature of the source appears, yet the combined and interlaced image (bottom plot) does resolve the source.

astronomer must observe four frames, displaced orthogonally from one another by an odd multiple of 0.4 arcsec. Some observers can achieve the placement needed for an efficient interlace, but the limited precision of many actual telescope controls usually produces a set of exposures whose grids are dithered randomly at the sub-pixel level.

The ***variable-pixel linear reconstruction*** method, more commonly known as ***drizzle***, can be much more forgiving about input grid placement. Drizzle assumes that the flux in a square input pixel of size (length) d is not spread over the pixel, but is uniformly concentrated in a smaller concentric square, called a "drop," whose sides have length fd; see Figure 8.16b, where the drops are the shaded squares. The fractional size of the drops, i.e. the value of f, can be varied to accommodate a particular set of images. As $f \to 0$ the drizzle method approaches the interlace method, and as $f \to 1$ drizzle, drizzle approaches resampled shift and add.

We introduce a parameter, s, to measure the relative scale of the output pixels: for input pixels of length d, output pixels have length sd. The drizzle algorithm then runs as follows: Input pixel $B_i[x, y]$ in frame i will contribute to output pixel $B'_i[\xi, \eta]$ if any part of the input *drop* overlaps the output pixel. If the area of overlap is $a_i[x,y,\xi,\eta](fd)^2$, then the contribution will be

$$B_i[x,y]W_i[x,y]a_i[x,y,\xi,\eta]s^2$$

The factor s^2 conserves surface brightness in the final image, and the weighting factor $W_i[x, y]$ accounts for bad pixels and other effects (e.g. exposure time) in the input frame. Adding all contributions from the input image (up to four input drops can overlap a single output pixel), we assign the output value and weight as

$$B_i'[\xi,\eta] = s^2 \sum_{x,y} B_i[x,y] W_i[x,y] a_i[x,y,\xi,\eta]$$

$$w_i[\xi,\eta] = \sum_{x,y} W_i[x,y] a_i[x,y,\xi,\eta]$$

We make the final combination of images by computing the weighted mean of all the input frame contributions to each pixel as in Equation (9.13).

$$C[\xi,\eta] = \frac{\sum_{i=1}^{N} B_i'[\xi,\eta]}{\sum_{i=1}^{N} w_i[\xi,\eta]}$$

9.4.7 Cleaning images

Images inevitably have defects caused by bad detector pixels or by unwanted radiation events like cosmic-ray impacts or radioactive decays in or near the detector. Most methods for removing such defects require multiple dithered images of the same scene.

One familiar prescription works quite well. Start with $N>2$ dithered images $\{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_N\}$ whose intensities are scaled to the same exposure time. Align them (i.e. use a geometric transform to make all astronomical sources coincide):

$$\mathbf{R}_i' = \text{GXform}(\mathbf{R}_i),\ i = 1, \ldots, N$$

Now combine the transformed images to form the median image:

$$\mathbf{C} = \text{median}(\mathbf{R}_1', \mathbf{R}_2', \ldots, \mathbf{R}_N')$$

The median is relatively insensitive to pixel values (like many radiation events or bad pixels) that differ greatly from the central value, so it produces a "clean" version of the image. Although simple to execute, the median becomes less graceful with images of differing weights and does have some shortcomings:

1. At a location where all pixel values are good, the median is not as good an estimator of the central value as is the mean.
2. The median is not *completely* insensitive to deviant values: e.g. the median will be slightly biased towards higher values at the location of cosmic-ray hits.
3. The median will perform very poorly in special cases (e.g. if multiple values at the same location are bad).

A more sophisticated cleaning method is to ***flag*** the defects in the original images, sometimes by assigning the affected pixels a special value (a large negative number, for example) or by assigning them a weight of zero. In one technique of this sort, the astronomer generates a special companion image, the ***mask***, for each $\mathbf{R}_i$. The mask values (usually either one or zero) indicate whether the corresponding image pixel is to be included or excluded in any subsequent operations; see Figure 9.18.

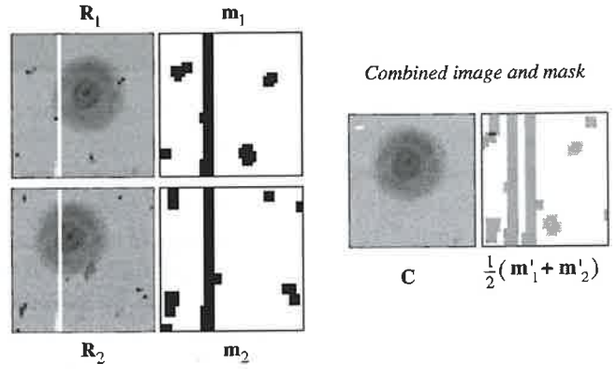


Fig. 9.18 Pixel masks. Two offset images of the planetary nebula NGC 2392 (The Eskimo) are marred by an insensitive column and many cosmic ray strikes. The mask next to each raw image on the left blocks (black pixels = 0, white = 1) every bad pixel and its immediately adjacent neighbor. The right-hand images show the combined image and mask after alignment. Since there are only two images, the combined image shows noticeably different noise levels in masked and unmasked regions. Two pixels in the upper left are masked in both images and have zero weight. They show as black in the right-hand image of the combined masks.

How can you generate a mask for a particular image? Usually, bad detector pixels or columns are well documented or are easily discovered on flat-field exposures. You can identify radiation events, which occur at random locations and can mimic images of astronomical objects, with the median-image method described at the start of this section. Once the complete mask is generated for an input image, a conservative approach might be to mask all pixels that are adjacent to bad pixels as well, since radiation events tend to spill over. At the end of this process, there will be a separate mask for each input image.

You then geometrically transform all input images, along with their masks, so that all are aligned. The final combination of these aligned images is a weighted mean in which all defective pixels are ignored. That is, if $\mathbf{m}_i$ is the mask for input image i, w_i is the image weight, and $\mathbf{m}_i'$ is the transformed mask:

$$C[\zeta,\eta] = \frac{\sum_{i=1}^{N} w_i m_i'[\zeta,\eta] R_i'[\zeta,\eta]}{\sum_{i=1}^{N} w_i m_i'[\zeta,\eta]}$$

Figure 9.18 illustrates a simple combination of two small images using masks.

9.5 Digital aperture photometry

We have discussed the preprocessing of individual images (the linearity, dark, bias, flat, and fringe corrections) and the combination of multiple frames to produce a deeper and possibly wider image. As a reminder, we summarize those steps here:

$$\mathbf{R}_{\text{pf}i} = \frac{\text{lin}(\mathbf{R}_i) - \mathbf{Z}_i - t_i\mathbf{D}}{\mathbf{F}} - \frac{A_i}{A}\mathbf{b}_f$$

$$\mathbf{R}'_i = \mathrm{GXform}(\mathbf{R}_{\mathrm{pf}i})$$
$$\mathbf{C} = \mathrm{combine}(\mathbf{R}'_1, \mathbf{R}'_2, \ldots, \mathbf{R}'_N)$$

Here we understand that the combination will be something like a median image or weighted mean, perhaps utilizing masks and a drizzle.

The next task in the reduction procedure is often measurement of the brightness of one or more objects. Measuring brightness is at heart a simple task – we did it in the exercises in Chapters 1 and 2. Start with the preprocessed image – an individual frame, $\mathbf{R}_{\mathrm{pf}i}$, or an aligned/combined accumulation of such frames, $\mathbf{C}$. Then just add up the emission from the object of interest, which usually is spread over many pixels. In doing so, remember to remove the background, which contains positive contributions made by sources both behind and in front of the object of interest. The latter include scattered light from other astronomical objects as well as the glow of the atmosphere and (especially in the thermal infrared) of the telescope. We will use the terms ***sky*** and ***background*** interchangeably for all this unwanted light, no matter where it originates. Once we have isolated the signal attributable to the source alone, we will need to quantify the uncertainty of the result.

Finally, the signal measured will only be meaningful if it is calibrated – expressed in units like magnitudes or watts per square meter. We consider the calibration process in the next chapter, and confine ourselves here to the tasks of separating signal from background and of estimating the uncertainty of the result.

9.5.1 Digital apertures and PSF fits

Consider a very common situation: from a digital image, you want to determine the brightness of a *point* source – a star, quasar, or small object in the Solar System. Define a circular area, the ***digital aperture***[4], that is centered on the centroid of the object (see Figure 9.19). The radius of the digital aperture should include a substantial fraction of the emission from the star. Now make three simple computations:

1. Add up all the pixel values inside the aperture. This sum represents the total emission from the aperture – the light from the star plus the light from the background. To deal with fractional pixels (see Figure 9.19) at the edges, multiply every value by $A[x, y]$, the fraction of the pixel's area that is inside the aperture.

$$\mathrm{Total} = \sum_{x,y} A[x,y] R_{\mathrm{p}}[x,y]$$
$$n_{\mathrm{pix}} = \sum_{x,y} A[x,y]$$

[4] Yes, aperture means "opening". The terminology recalls the days of photoelectric photometry, when it was necessary to place an opaque plate with one small clear aperture in the focal plane. This passed only the light from the star and very nearby sky through to the photocathode, and blocked all other sources.

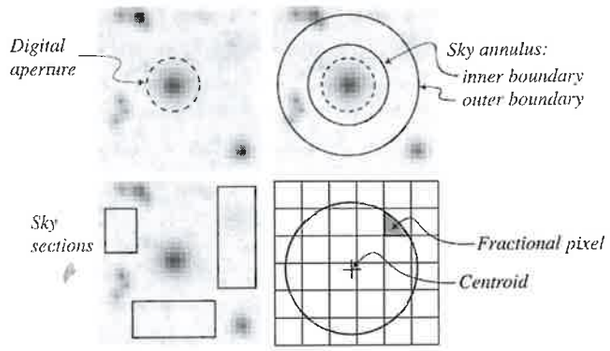


Fig. 9.19 Digital apertures. The upper left-hand image shows a circular aperture centered on a point source; lower left, rectangular apertures for sampling background emission. The image in the upper right shows an annular aperture for sampling sky emission near a point source. All curved apertures will require some strategy for dealing with pixels that contain some segment of the boundary, as in the image at the bottom right.

The sums are understood to extend over the entire x–y extent of the aperture. The number n_{pix} is just the area of the aperture in pixels.

2. Estimate $\bar{B}$, the value of the sky emission per pixel. Usually, you must estimate $\bar{B}$ from a source-free region near the object of interest (see the next section for details). Compute that part of the emission in the aperture that is due to the sky:

$$\mathrm{sky} = n_{\mathrm{pix}} \bar{B}$$

3. Subtract the sky emission from the total, and the remainder is the detector response attributable to the source alone; this is the signal in ADUs:

$$S_{\mathrm{ADU}} = \mathrm{total} - \mathrm{sky}$$
$$S_{\mathrm{ADU}} = \sum_{x,y} A[x,y] R_{\mathrm{p}}[x,y] - n_{\mathrm{pix}} \bar{B} \qquad (9.14)$$

In situations in which star images seriously overlap, digital aperture photometry fails, because it is impossible easily to estimate the polluting star's contribution to the background of the object of interest. We have already discussed (Section 9.4.2) the idea of fitting a ***PSF*** to each star image on a frame. Point-spread-function fitting is required in crowded-field photometry since (at the cost of considerable computational complexity) it can separate the contributions of individual overlapping images from one another and from the diffuse background. Once all overlapping images are accounted for, integration of the PSF fit of the image of interest gives the *signal* term in Equation (9.14).

The question of ***aperture size*** is important. For PSF fitting, the aperture size is often technically infinite, e.g. a typical PSF for a ground-based image is Gaussian-like – see Appendix J. You can determine the appropriate PSF *shape* by examining high-SNR star images and comparing their profiles (especially the central regions with good signal) to expected shapes or by constructing an empirical shape. For PSFs, well over 90% (usually much more) of emission is within a diameter of three times the FWHM.

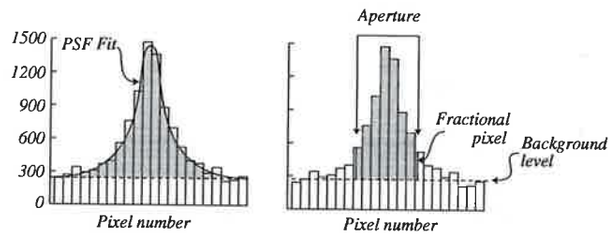


Fig. 9.20 PSF and aperture photometry. At left are values of pixels in the *x*-direction through the centroid of a star image. Data have been fit with a background sky level and a Gaussian. The shaded area gives the contribution to the total flux from this row of pixels. At right are the same data measured with a digital aperture, where the shaded area again gives the contribution to the total flux. The optimum aperture size will exclude some flux present in the wings of the profile.

If he does not use PSF fitting, the astronomer must choose the digital aperture size; see Figure 9.20. There are two conflicting considerations: he wants a *large* aperture because it includes as much light as possible from the star, yet he wants a *small* aperture because it excludes background light and, especially, its associated noise. An aperture that includes too much sky will decrease the SNR of the final measurement, as will an aperture that includes too little of the source. The optimum size varies with the relative brightness of the star. Since point-source photometry requires the same aperture size for all stars, this generally means the astronomer chooses the aperture size based on the faintest star observed. The choice is implemented in software, so it is easy to try a range of apertures (diameters somewhere between 0.75 and 4 times the FWHM of the image profile) and identify the aperture (usually, a diameter near 1.5–2 times the profile FWHM) that yields the best SNR.

Finally, note that a digital aperture need not be circular. Indeed, many objects have decidedly non-circular shapes, and invite equally non-circular apertures. Photometry via Equation (9.14) applies as well to such shapes.

9.5.2 Measuring the sky

Both PSF fitting and digital aperture photometry demand an accurate measure of the sky emission per pixel over of the source of interest. This, of course, is one area where it is impossible to measure the sky brightness, so we measure the sky *near* the source, and hope that sky brightness does not change with location. There are some cases where this hope is forlorn. A notorious example is the photometry of supernovae in other galaxies: the background due to the host galaxy changes drastically on the scale of a digital aperture size, so any "nearby" sky measurement is guaranteed to introduce uncertainty. (Fortunately, supernovae are temporary. An image of the galaxy obtained with the same instrument after the supernova has faded can provide the needed background measurement.)

A smooth background near the source should ameliorate many difficulties. In this case, the nearest possible sample should be the most accurate, and a sample symmetrically positioned around the source stands a chance of averaging out any trends. Figure 9.19 shows a digital aperture and a ***sky annulus***. The annulus

is a region between two circles centered on the source. The smaller circle (the inner boundary of the annulus) is as small as possible, but still large enough to exclude any appreciable emission from the source. The outer radius of the annulus is less strictly determined, but should be large enough to include a statistically significant number of pixels. If the outer radius is too large, it may sample sky that differs from the sky within the aperture.

The best estimate of the sky value in the annulus is clearly *not* the mean pixel value: the annulus is bound to contain images or halos of other stars. These bias the mean towards larger values. The median is less sensitive to the influence of this kind of pollution, and the mode is even better: the most common value in the annulus certainly sounds like the optimum measurement of the sky. Practical computation of the mode usually requires the construction of a smoothed histogram, with the sky value computed as the mean of the values in the most populous bin of the histogram.

Figure 9.19 illustrates a second approach to measuring the sky value. An astonomer selects one or more relatively star-free sections of the image, and computes the median or modal value. The advantage of this method is that it avoids the influence of nearby sources on background estimates, and if the field near the source of interest is crowded, this is the only alternative. The disadvantage is that the sky sections may be relatively far from the point of interest, and they may not sample uniformly enough to minimize the effects of large-scale trends in background brightness. As explained earlier, in the infrared, one generally obtains sky levels from separate (chopped) exposures.

9.5.3 Signal and noise in an aperture

Knowing the uncertainty of a digital photometric measurement is nearly as important as discovering its value. In this section, we develop an equation for the SNR in aperture photometry with a CCD. The general approach, if not the exact equation, will apply for photometry with all digital arrays.

For simplicity, we consider only the case of a single exposure, corrected for non-linearity, dark, bias, and flat. Recall the digital aperture photometry operation given in Equation (9.14):

$$S_{\mathrm{ADU}} = \left\{ \sum_{x,y} A[x,y] R_{\mathrm{p}}[x,y] \right\} - n_{\mathrm{pix}} \bar{B} \tag{9.15}$$

Here $A[x,y]$ is the fraction of the pixel inside the digital aperture, and $\bar{B}$ is the estimated average background emission per pixel. The pixel values are in ADUs (analog-to-digital units), values that we can convert to the number of electrons read out from the pixel by multiplying by g, the CCD gain factor. In terms of electrons, then, Equation (9.15) becomes

$$\mathrm{Signal} = N_* = g[S_{\mathrm{ADU}}] = \left\{ \sum_{x,y} A[x,y] r_{\mathrm{p}}[x,y] \right\} - n_{\mathrm{pix}} b_{\mathrm{e}}$$

Here the signal is N_*, the total number of electrons produced by the source in the aperture. The values $r_p[x,y]$ and b_e are the preprocessed pixel value and the estimated background value in electrons. The noise, or uncertainty in N_*, follows from an application of Equation (2.17) to (9.15). Although it is not always safe to do so, we assume uncertainties in pixel values are not correlated:

$$\sigma_N^2 = \left\{ \sum_{x,y} \{A[x,y]\}^2 \sigma_{r,p}^2[x,y] \right\} + n_{pix}^2 \sigma_{b,e}^2 \tag{9.16}$$

To evaluate $\sigma_{r,p}^2[x,y]$, the uncertainty in a preprocessed pixel value, we write out the preprocessing operation for a single pixel as described for a CCD in Equation (9.8):

$$r_p[x,y] = \frac{1}{f[x,y]} \{L(r[x,y])r[x,y] - d_e[x,y] - \zeta_e[x,y]\} \tag{9.17}$$

Here:

$f[x,y]$ = the normalized flat field response,

$d_e[x,y]$ = the estimated dark count in the pixel in electrons,

$\zeta_e[x,y]$ = the estimated bias level in the pixel, in electrons, and

$L(r[x,y])$ = the linearity correction for the pixel, expressed as multiplicative factor.

We will again assume that the uncertainties in each of the variables in Equation (9.17) are not correlated, so that we can apply Equation (2.17) to compute the variance of processed pixel value:

$$\sigma_{r,p}^2[x,y] = \sigma_r^2 \frac{L^2}{f^2} + \sigma_{d,e}^2 \frac{1}{f^2} + \sigma_{\zeta,e}^2 \frac{1}{f^2} + \sigma_L^2 \frac{r^2}{f^2} + \sigma_f^2 \left\{ \frac{Lr - d_e - \zeta_e}{f^2} \right\}^2$$

To simplify the notation, we have omitted the $[x, y]$ coordinate references for all the terms on the right-hand side. We can clean up this expression further by noting that $f \approx 1$ and $L \approx 1$:

$$\sigma_{r,p}^2[x,y] = \sigma_r^2 + \sigma_{\zeta,e}^2 + \sigma_{d,e}^2 + \sigma_L^2 r^2 + \sigma_f^2 (r - d_e - \zeta_e)^2 \tag{9.18}$$

We will examine each of the terms on the right-hand side in turn. The first term in Equation (9.18) is the square of the uncertainty in the raw pixel value itself. The unprocessed pixel value is just

$$r[x,y] = r'[x,y] + \zeta[x,y] = n[x,y] + b[x,y] + d[x,y] + \zeta[x,y]$$

where the four quantities on the extreme right are, respectively, the actual single-pixel values of the signal, background, dark, and bias expressed in electrons. The variance of the raw pixel value must be

$$\sigma_r^2[x,y] = \sigma_{r'}^2 + \sigma_\zeta^2 = r' + \rho^2$$

The variance of the pixel response, $r'[x,y]$, is equal to its mean because the response is Poisson-distributed. Noise present on every readout of the device,

independent of exposure time or illumination, produces the variance of the bias level, ρ^2. For a CCD, we distinguish two components:

$$\rho^2 = \sigma_{read}^2 + \sigma_{digit}^2$$

The first is the ***read noise*** – the uncertainty in the zero-level signal from the output amplifier. The second is the ***digitization noise*** – the uncertainty that results when the ADC circuit rounds the analog signal to an integer. If the analog values are uniformly distributed, the digitization noise is $g/\sqrt{12}$, where g is the CCD gain. Usually the CCD gain is adjusted so that the digitization component is smaller than σ_{read}^2, but we will include both in the parameter ρ, the "zero-level uncertainty" or "digital read noise".

The actual values for the background $b[x,y]$ and dark $d[x,y]$ levels in a particular pixel are unknown, and we will simply use estimated values: the dark is estimated from dark frames (or assumed to be zero if the detector is sufficiently cold) and the background is estimated from nearby "sky" pixels. Thus

$$\sigma_r^2[x,y] \approx n + b_e + d_e + \rho^2 \tag{9.19}$$

The second term in Equation (9.18) is the squared uncertainty in the "estimated" bias level (different from the read noise in a single pixel!). This estimate is usually computed by averaging a number of calibration frames. If the bias drifts, then $\sigma_{\zeta,e}$ might be large. If we obtain p_z bias frames, the minimum variance of the mean of the p_z values for the bias at pixel $[x,y]$ is given by

$$\sigma_{\zeta,e}^2 = \frac{\rho^2}{p_z} \tag{9.20}$$

where ρ is the digital read noise. If the bias is obtained from an overscan, and the base bias pattern is very well determined, then p_z is the number of columns in the overscan.

Likewise, the third term is the variance in the estimated dark count. If the estimate is an average of p_d dark frames, each of the same exposure time as the data frame, then the variance of the mean is

$$\sigma_{d,e}^2 = \frac{1}{p_d} \left\{ d_e + \left(1 + \frac{1}{p_z}\right) \rho^2 \right\} \tag{9.21}$$

The second term in the braces is there because dark frames are themselves processed by subtracting an estimated bias level. Note that we assume the same number of number bias frames, p_z, are used for the dark as for the data.

The fourth term in Equation (9.18) is $\sigma_L^2 r^2$, the variance in the linearity correction scaled by the square of the pixel value. Uncertainty in the linearity correction should not be of concern with most CCDs except near saturation

(usually no correction is made), but it can be an issue in infrared arrays where linearity properties may vary from pixel to pixel. The value of σ_L^2 can be measured by examining linearity calibration frames.

The fifth term in Equation (9.18), $\sigma_f^2(n+b)^2$ (where $n+b = r - d_e - \zeta_e$), arises from the uncertainty in the normalized flat field. In the ideal case, σ_f^2 should approach $(n_f[x,y])^{-1}$, where $n_f[x,y]$ is the total number of photoelectrons counted at the pixel location in all flat-field calibration exposures. One should thus be able to reduce this uncertainty to insignificance just by accumulating enough calibration frames. Real observational situations seldom approach the ideal, and one can investigate uncertainties in the flat by, say, comparing combined flats taken on two different nights, or by varying the color of the flat-field target.

Substituting Equations (9.19), (9.20), and (9.21) into (9.18) gives the variance in the value of a single preprocessed pixel.

$$\sigma_{r,p}^2[x,y] = n + b_c + a_d(d_c + a_z\rho^2) + \sigma_L^2 r^2 + \sigma_f^2(n + b_c)^2 \qquad (9.22)$$

where

$$a_d = 1 + \frac{1}{p_d}, \qquad a_z = 1 + \frac{1}{p_z}$$

Now return to Equation (9.16). We require a value for the uncertainty in the estimated background. We usually estimate the background by averaging $r_p[x,y]$ in a region of p_b pixels (e.g. the sky annulus) in which $n[x,y]$is zero. That is,

$$b_c = \frac{1}{p_b} \sum_{x,y}^{sky\ section} r_p[x,y]$$

The variance, then, is

$$\sigma_{b,c}^2 = \frac{1}{p_b^2}\sum\sigma_{r,p}^2[x,y] = \frac{1}{p_b}\left\{\bar{\sigma}_{r,p}^2\right\}$$

But we have just worked out $\sigma_{r,p}^2[x,y]$, the variance of a *single* preprocessed pixel. Substituting for the average variance $\bar{\sigma}_{r,p}^2$ from Equation (9.22) for the case $n[x,y] = 0$:

$$\sigma_{b,c}^2 = \frac{1}{p_b}\left\{b_c + a_d(d_c + a_z\rho^2) + \bar{\sigma}_L^2\bar{r}_b^2 + \bar{\sigma}_f^2(b_c)^2\right\} \qquad (9.23)$$

Now we can turn to Equation (9.16) one last time:

$$\sigma_N^2 = \left\{\sum_{x,y}^{aperture} \{A[x,y]\}^2\sigma_{r,p}^2[x,y]\right\} + n_{pix}^2\sigma_{b,c}^2$$

We know all terms on the right-hand side, so substituting

$$\sigma_N^2 = \left(\sum\{A[x,y]\}^2 n[x,y]\right) + \left(P^2 + \frac{n_{pix}^2}{p_b}\right)\left(b_e + a_d(d_c + a_z\rho^2)\right) + s_L^2 + s_f^2 \quad (9.24)$$

where

$$P^2 = \sum_{x,y}(A[x,y])^2 \le n_{pix}^2 = \left(\sum_{x,y} A[x,y]\right)^2$$

The parameter P is the properly weighted pixel count in the aperture for computing the uncertainty of a uniform signal. The two terms on the extreme right of Equation (9.24) are the contribution to the variance due to uncertainty in the linearity and the flat-field corrections. We can write formal expressions for these, but they can only yield values if one can examine the repeatability of flat and linearity calibrations; see Problems 9.4 and 9.5.

Equation (9.24) does not include some sources of uncertainty that could be important in a specific array, like uncertainties in corrections for charge-transfer inefficiency. If such effects can be well modeled, one could in principle represent them with additional terms.

As a tool for evaluating photometric uncertainty, the most serious problem with Equation (9.24) is its failure to account for systematic effects like those due to the non-uniformity of a flat-field target, color differences between sky, star, and flat, or variations in atmospheric transparency. You should *not* use this equation to evaluate the uncertainty in your digital photometry. ***As always, the primary information about the uncertainty of your photometry comes from the scatter in repeated observations and the disagreement of your results with those of others.***

But Equation (9.24) is far from useless. It gives you a way to compare the expected random error with the actual scatter in your data – if you get something unexpected, think hard to understand why. The equation is also a very important tool for *planning* observations, for answering questions like: "how many minutes at the telescope will I need if I want to measure the brightness of my $V = 22.5$ quasar with a precision of 1%?"

9.5.3 The CCD equation

We will use Equation (9.24) for the not-so-special case of aperture photometry on a star. For most reasonable apertures, the counts due to the star are very small at the edge of the aperture where the partial pixels are located. In that case, we will not be far off in making the approximation (which applies exactly if partial pixels are not employed):

$$\sum_{x,y} A^2[x,y]n[x,y] \approx \sum_{x,y} A[x,y]n[x,y] = N_*$$

The SNR then implied by Equation (9.24) is

$$\mathrm{SNR} = \frac{N_*}{\left\{N_* + (P + a_\mathrm{b})(b_\mathrm{c} + a_\mathrm{d}(d_\mathrm{c} + a_z\rho^2)) + s_\mathrm{L}^2 + s_\mathrm{f}^2\right\}^{\frac{1}{2}}} \tag{9.25}$$

This equation, in various approximations, is known as the ***CCD equation***. The usual approach is to simplify Equation (9.25) by assuming good preprocessing practices as well as good fortune, so that the system will remain stable and the observer will collect a very large number of bias and (if needed) dark frames, the flat-field and linearity corrections will not contribute significant errors, and that $P = n_\mathrm{pix}$. In other words, $a_\mathrm{d} = a_z = 1$, and $s_\mathrm{L} = s_\mathrm{f} = 0$, so that the CCD equation becomes

$$\mathrm{SNR} = \frac{N_*}{\left\{N_* + n_\mathrm{pix}\left(1 + \frac{n_\mathrm{pix}}{p_\mathrm{b}}\right)(b_\mathrm{c} + d_\mathrm{c} + \rho^2)\right\}^{\frac{1}{2}}} \tag{9.26}$$

Since the CCD equation is often used to estimate the required exposure time, t, we rewrite this as

$$\mathrm{SNR} = \frac{\dot{N}_* t}{\left\{\left[\dot{N}_* + n_\mathrm{pix}\left(1 + \frac{n_\mathrm{pix}}{p_\mathrm{b}}\right)(\dot{b}_\mathrm{c} + \dot{d}_\mathrm{c})\right]t + n_\mathrm{pix}\left(1 + \frac{n_\mathrm{pix}}{p_\mathrm{b}}\right)\rho^2\right]\right\}^{\frac{1}{2}}} \tag{9.27}$$

The dotted quantities give the electron rates for source photons, background photons, and dark current. The read-noise term is independent of the exposure time. Solving for the exposure time:

$$t = \frac{B + (B^2 + 4AC)^{\frac{1}{2}}}{2A} \tag{9.28}$$

where

$$P_\mathrm{b} = n_\mathrm{pix}\left(1 + {}^{n_\mathrm{pix}}/_{p_\mathrm{b}}\right)$$
$$A = \frac{\dot{N}_*^2}{(\mathrm{SNR})^2}$$
$$B = \dot{N}_* + P_\mathrm{b}(\dot{b}_\mathrm{c} + \dot{d}_\mathrm{c})$$
$$C = P_\mathrm{b}\rho^2$$

Figure 9.21 illustrates predictions based on the CCD equation in three different situations.

The ***bright-star*** or ***photon-noise-limited*** case, where the counting rate from the source, $\dot{N}_*$, exceeds all other terms in the denominator of (9.27). This case approaches the Poisson result:

$$\mathrm{SNR} = \sqrt{N_*} = \sqrt{\dot{N}_* t}$$

So in the bright-star case, the SNR improves as the square root of the exposure time, and the observer willing to devote sufficient time can produce

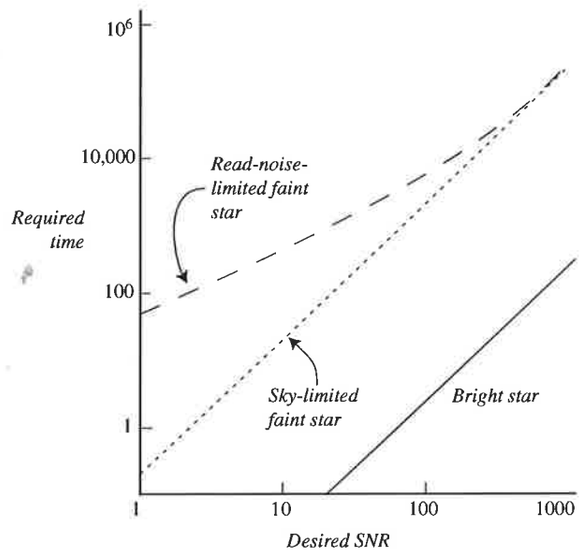


Fig. 9.21 The CCD equation. The plot shows the required time to reach a specified signal-to-noise ratio in the three limiting cases discussed in the text. Note the logarithmic scale.

measurements of arbitrarily high precision. However, actual precision attained may well be limited by the processes we ignored (e.g. the atmosphere) or the terms we eliminated in deriving Equation (9.27) – flat-field uncertainties, for example, scale as the first power of the exposure time, and can eventually dominate.

In the ***background-limited*** case, the background term

$$n_\mathrm{pix}\left(1 + \frac{n_\mathrm{pix}}{p_\mathrm{b}}\right)(\dot{b}_\mathrm{c} + \dot{d}_\mathrm{c})$$

is not insignificant compared to the electron rate from the source. This is usually the case that is most interesting to the observer, since it describes the limits of detection in a given situation. In the background-limited case, the SNR ratio still increases as the square root of the observing time, but now there is a penalty:

$$\mathrm{SNR} = \left\{1 + n_\mathrm{pix}\left(1 + \frac{n_\mathrm{pix}}{p_\mathrm{b}}\right)\frac{(\dot{b}_\mathrm{c} + \dot{d}_\mathrm{c})}{\dot{N}_*}\right\}^{\frac{1}{2}}\sqrt{\dot{N}_* t} = \{1 + B_*\}^{\frac{1}{2}}\sqrt{\dot{N}_* t} \tag{9.29}$$

The factor $\{1 + B_*\}$ becomes large under any occurrence of the following conditions: low source brightness, high sky brightness, high dark rate, large digital aperture, or small sky sample. A frequent situation is that of a faint source and bright sky, the ***sky-limited*** case. Remember also that Equation (9.29) ignores many sources of uncertainty due to both random and systematic effects. Detection limits derived from this expression should thus be taken to be optimistic.

The final case is the one in which the read noise is large. Here the SNR initially increases linearly with time, but eventually reaches the $\sqrt{t}$ dependence of either the bright-star or the sky-limited case.

Summary

- Digital images are ordered sets of numbers that can represent the output of an array of sensors or other data. Concepts:

pixel	*fill factor*	*undersampling*
pixel value	*gray-scale map*	*detector response*
ADU	*DN*	

- An important advantage of digital images is that they can be mathematically manipulated to remove defects and extract information. Concepts:

image arithmetic	*data cube*	*RGB color model*
CMYK	*false color*	*image functions*
digital filtration	*image convolution*	*kernel*
Gaussian kernel	*Laplacian kernel*	*boxcar*
unsharp mask	*median filter*	

- Digital images from a CCD can be processed to remove the effects of the detector and telescope. Concepts:

raw image	*bias frame*	*rejection algorithm*
overscan	*dark response*	*dark rate*
CCD gain	*linearity correction*	*chopping secondary*

- The flat-field correction very often limits photometric precision of a detector. Concepts:

flat-field image	*twilight flat*	*dark sky flat*
dome flat	*compound flats*	*shift and stare*
dither	*illumination correction*	

- Preprocessing images from an array requires subtraction of the dark signal and bias, then division by the normalized flat. Treatment of data from infrared arrays is slightly different because of the strong and variable sky background.
- Fringing is a variation in the sky background intensity due to interference effects in thin layers of a detector. Fringes can be removed if a flat without fringes is available.
- Combining images requires alignment, which requires both identification of feature coordinates and transformation of images. Concepts

centroid	*point-spread function*	*PSF fitting*
image alignment	*translation*	*rotation*
trim	*canvas*	*image mosaic*
magnification	*distortion*	*shift and add*
unsharp mask	*median filter*	

- Special methods for combining images can compensate for the loss of resolution due to interpolation, and can compensate for bad pixels. Concepts:

nearest pixel	*resampling*	*bilinear interpolation*
interlace	*pixel flag*	*drizzle*
image mask	*clean image*	

- Digital aperture photometry is a technique for measuring apparent brightness from a digital image. Concepts:

digital aperture	*sky annulus*	*PSF*

- The CCD equation gives the theoretical relation between the exposure time and expected SNR in digital aperture photometry, given source and sky brightness and detector and telescope characteristics. Concepts:

read noise	*digitization noise*	*background-limited*
photon-noise limited		

Exercises

1. Derive expressions for – and compute values of – the coefficients a, b, and c in Equation (9.7) for the detector whose calibration data fits the solid line labeled "second order" in Figure 9.10b.
2. The table at left below gives the coordinates and pixel values near a faint star on an array image. The small array at right is a sample of the nearby background. Find the x, y coordinates of the centroid of the star image using the criteria outlined in Section 9.4.1 of the text. Use a spreadsheet.

$y\backslash x$	1	2	3	4	5	6	7		
8	23	20	17	19	18	17	23		
7	18	25	20	18	26	18	19	16	19
6	20	27	33	30	27	23	18	14	16
5	19	31	40	34	28	22	25	13	11
4	26	29	53	51	28	28	21	21	18
3	22	26	40	32	33	18	24	16	17
2	23	30	26	24	26	23	14	20	18
1	16	19	20	18	18	17	16		

3. Suggest a strategy, similar to that in the latter part of Section 9.4.3, for combining N unaligned images to create a single mosaic image, $\mathbf{C}_{OR}$, that contains the combined data for every observed location in the collection
4. Show that in Equation (9.21), the formulae for the uncertainties in digital aperture photometry due to an uncertainty in the linearity correction will be given by

$$s_L^2 = \sum_{x,y} \left(\sigma_L(r) A[x,y] r[x,y]\right)^2 + \frac{n_{pix}^2}{p_b} \bar{\sigma}_L^2 \bar{r}_b^2 > P\sigma_L^2(\bar{r}_a)\bar{r}_b^2 + \frac{n_{pix}^2}{p_b} \bar{\sigma}_L^2 \bar{r}_b^2$$

where the quantities are those defined in Section 9.5.3. Explain why uncertainties in linearity are less troublesome if one is comparing stars of nearly equal brightness.

5. Show that the variance due to flat-field uncertainty in Equation (9.24) is

$$s_f^2 = \sum_{x,y} (A[x,y](n[x,y] + b_c))^2 \sigma_f^2 + \frac{n_{pix}^2}{p_b} \bar{\sigma}_f^2 b_c^2$$

6. On a 20-second exposure, a star with magnitude $B = 15$ produces an SNR = 100 signal with a small telescope/CCD combination. Assuming this is a photon-noise limited case, how long an exposure should be required to produce the same SNR for star with $B = 13.6$?
7. A star with $V = 21.0$ is known to produce a count rate of 10 electrons per second for a certain telescope/detector combination. The detector read noise is 4 electrons per pixel, and the dark rate is zero. Compute the exposure time needed to reach a SNR = 10 under the following conditions:
   (a) dark sky and good seeing: aperture radius = 3.5 pixels, sky brightness = 1.4 electrons per pixel per second;
   (b) moonlit sky and poor seeing: aperture radius = 5.0 pixels, sky brightness = 4 electrons per pixel per second.
8. A certain CCD has a gain of 2.4 electrons per ADU, a read noise of 7 electrons per pixel, and a dark current of 2.5 ADU per pixel persecond. In the V filter, the sky brightness averages 8 ADU per second. An astronomer wishes to observe a nebula whose average brightness is expected to be 7 ADU per pixel persec per second over a digital aperture area of 100 pixels. Compute the expected SNR for measurements of the nebula's brightness on exposures of (a) 1 second, (b) 10 seconds and (c) 100 seconds.

Chapter 10
Photometry

> The classification of the stars of the celestial sphere, according to different orders of magnitude, was made by ancient astronomers in an arbitrary manner, without any pretension to accuracy. From the nature of things, this vagueness has been continued in the modern catalogs.
>
> – François Arago, *Popular Astronomy*, Vol I, 1851

Astronomers have measured apparent brightness since ancient times, and, as is usual in science, technology has acutely influenced their success. Prior to the 1860s, observers estimated brightness using only their eyes, expressing the results in the uncannily persistent magnitude system that Ptolemy[1] introduced in the second century. As Arago notes, the results were not satisfactory.

In this chapter, after a brief summary of the history of photometry, we will examine in detail the surprisingly complex process for answering the question: how bright is that object? To do so, we will first introduce the notion of a defined bandpass and its quantitative description, as well as the use of such bandpasses in the creation of standard photometric systems. Photometry is most useful if it represents the unadulterated light from the object of interest, so we will take some pain to describe how various effects might alter that light: spectrum shifts, absorption by interstellar material, and the characteristics of the observing system. We will pay particular attention, however, to the heavy burden of the ground-based photometrist: the influence of the terrestrial atmosphere and the techniques that might remove it.

10.1 Introduction: a short history

The history of photometry is brief compared to that of astrometry, due to the symbiotic absences of scientific interest and appropriate instrumentation. John

[1] The magnitude system may very well predate Ptolemy. Ptolemy's catalog in the Almagest (*c.*137 CE) may be based substantially on the earlier catalog of Hipparchus (*c.*130 BC), which has not been preserved. It is unclear which astronomer – Ptolemy, Hipparchus, or another – actually introduced the scale. Moreover, Ptolemy is largely silent on the method actually used to establish the visual brightness estimates he recorded. Although Ptolemy tends to assign stars integral magnitudes, 156 stars (out of 1028) are noted as slightly (one third of a magnitude?) brighter or fainter than an integral value.

B. Hearnshaw (1996) provides a book-length history of astronomical photometry up to 1970. Harold Weaver (1946) gives a shorter and more technical account of developments up through World War II. A definitive history of the charge-coupled device (CCD) era remains unwritten.

To what degree will two stars assigned the same magnitude by a naked-eye observer actually have the same brightness? Modern measurements show pre-telescopic catalogs (e.g. Ptolemy and Tycho, both of whom were more interested in positions than in brightness) have an internal precision of about 0.5 magnitudes. Even the most skilled naked-eye observer can do little better: al Sufi in the ninth century devoted great attention to the problem and achieved a precision near 0.4 magnitudes. At the eyepiece of a telescope, several observers (e.g. the Herschels and, less successfully, the Bonner Durchmusterung observers Argelander and Schonfeld) produced better results (0.1 to 0.2 magnitudes) with a method of careful comparison to linked ***sequences*** of brightness standards.

After a suggestion by the French physicist François Arago (1786–1853), Karl Friedrich Zöllner (1834–1882) built the first optical/mechanical system for astronomical photometry in 1861. Many similar instruments soon followed. An observer using one of these ***visual photometers*** either adjusts the brightness of a comparison until it matches that of the unknown star, or dims the telescopic brightness of the unknown star until it disappears. Zöllner's instrument, for example, used crossed polarizers to adjust the image of an artificial star produced by a kerosene lamp.

Because the unknown need not be near a standard sequence in the sky, the visual photometer was efficient. Moreover, these devices were more *precise*, because brains are much better at judging equality (or complete extinction) than at making interpolations, especially interpolations based on memory of a sequence. Finally, the visual photometer was more *accurate* since making a mechanical adjustment gives a quantifiable measure fairly independent of a particular astronomer's eye and brain.

Astronomers got busy. Edward Pickering, at Harvard, for example, built a two-telescope "meridian photometer," which used crossed polarizers to equalize the images of two real stars. Between 1879 and 1902, Harvard visual photometrists measured the magnitudes of about 47,000 stars with a precision of about 0.08 magnitudes, and with an accuracy (based on modern measurements) of better than 0.25 magnitudes. Astronomers could now confidently examine the mathematical relationship between brightness and the ancient magnitude scale. Although several fits were proposed, by 1900 everyone had settled on the now familiar "Pogson normal scale":

$$\Delta m = -2.5 \log(b_1/b_2)$$

where b_1 and b_2 are the brightness of objects 1 and 2. The ancient scale turned out to be quite non-uniform in the logarithm: for example, the average brightness ratio between Ptolemy's magnitude 1.0 and 2.0 stars is 3.6, but between his

5.0 and 6.0 stars it is 1.3. The telescopic scales (e.g. Argelander) are closer to Pogson normal.

While the Harvard visual work progressed, photography matured. In 1850, William Cranch Bond and John Whipple, also at Harvard, photographed a few of the brightest stars. The invention of dry photographic plates (1871) increased convenience and sensitivity; eventually (around 1881) stars were recorded that were too faint to be seen by eye in any telescope. Many influential astronomers appreciated the vast potential of this new panoramic detector, and with virtually unprecedented international cooperation launched the *Carte du Ciel* project to photograph the entire sky and measure the brightness of every star below magnitude 11.0 (see Chapter 4). Astronomers soon learned to appreciate the difficulties in using photographs for quantitative photometric work, and it was not until the period 1900–1910 that several workers (notably Schwarzschild, Wirtz, Wilkins, and Kapteyn) established the first reliable ***photographic magnitude scales***. After the introduction (1910–1920) of physical photometers for objectively measuring images on plates, photography could yield magnitudes with uncertainties in the range 0.015–0.03 magnitudes. Such precision required very great care in the preparation, processing, and reduction of plate material, and could usually only be achieved in differential measurements among stars on the same plate.

In the first sustained photoelectric work, Joel Stebbins and his students at Illinois and Wisconsin performed extensive and precise photometry, first with selenium cells (1907), but soon with the vacuum photocell. Poor sensitivity at first limited the observations to very bright stars, but in 1932, when Albert Whitford and Stebbins added a vacuum-tube amplifier to the detector circuit, detection limits on their 0.5-meter telescope improved from 11th to 13th magnitude. The real revolution occurred in the 1940s, when the ***photomultiplier tube*** (***PMT***), developed for the military during World War II, became the astronomical instrument of choice for most precision work. It had very good sensitivity and produced uncertainties on the order 0.005 magnitudes in relative brightness.

The years from 1950 to 1980 were intensely productive for ground-based photoelectric work. Harold Johnson was an important pioneer in this era, first using the RCA 1P21 photomultiplier to define the UBV system, and later using red-sensitive photomultipliers to define an extended broadband system through the visual-near-infrared atmospheric windows.

Although astronomers still use photomultipliers for specialized work today, the CCD and other modern solid-state detectors have superceded them. In the optical, CCDs have superior efficiency, better stability, and a huge multiplex advantage (i.e. they can record many objects simultaneously, including standards). For ground-based differential work, CCD photometric precision on bright sources is generally set by photon-counting statistics (e.g. Equation (9.25)) or by uncertainties in calibration. For all-sky photometry and infrared work, the

atmosphere imposes more serious limitations –0.01 magnitude uncertainty is often regarded as routine. Photometry from spacecraft with solid-state devices, on the other hand, offers the potential of superb precision in both differential and all-sky work. For example, the Kepler space mission for detecting occultations by extrasolar planets, presently (2010) nearing launch, hopes to achieve uncertainties below 10 μmag over time scales of several weeks.

Observations from space are very, very costly, however, so ground-based photometry continues to be a central astronomical activity.

10.2 The response function

A photometric device is sensitive over a restricted range of wavelengths called its ***bandpass***. We distinguish three general cases of bandpass photometry to fit three different scientific questions.

10.2.1 Types of photometry

Single-band photometry. Suppose, for example, you suspect an extra-solar planet will move in front of a certain star, and you are interested in the occultation's duration and the fraction of the star's light blocked. You need only use a single band, since a geometric effect like the occultation of a uniform source will be identical at every wavelength. You would probably make a sequence of monitoring observations called a ***time series***, a tabulation of brightness as a function of time, and you would tend to choose a wide band to maximize signal and minimize the required exposure time and telescope size.

Broadband multi-color photometry. On the other hand you might want to know not just the brightness of a source, but also the general shape of its spectrum. Broadband multi-color photometry measures an ultra-low-resolution spectrum by sampling the brightness in several different bands. Although there is no strict definition, a "broad" band is generally taken to mean that the width of the band, $\Delta\lambda$, divided by its central wavelength, λ_c, is greater than 7%–10% , or, equivalently, the spectroscopic resolving power $R = \lambda_c/\Delta\lambda < 10$–15. Broadband systems choose bands that admit the maximum amount of light while still providing valuable astrophysical information. For example, the UBVRI system, the most common broadband system in the optical, uses bandwidths in the range 65–160 nm ($R = 4$–7). It provides information on surface temperature for a wide variety of stars, and more limited information on luminosity, metal content, and interstellar reddening.

The terminology recognizes each band as a "color", so "two-color photometry" measures magnitudes in two separate bands: B and V, for example. For both historical and practical reasons, one traditionally reports the results of n-color photometric measurements by giving one magnitude and $(n-1)$ color indices. The magnitude tells the apparent brightness, and the indices tell about other astrophysical variables like surface temperature. The term "color", as shorthand for "color index" has thus come to have a second meaning – ***color is the difference between two magnitudes***. So for example, the results of "*two*-color photometry" in B and V will be reported as a V magnitude and *one* $(B-V)$ color.

Narrow- and intermediate-band photometry. Although multi-color narrow-band photometry (roughly $R > 50$) can provide information about the shape of the spectrum, its intent is usually to isolate a specific line, molecular band, or other feature. The strategy here exchanges the large signal of the broadband system for a weaker signal with more detailed spectroscopic information. Common applications include the measurement of the strength of absorption features like Balmer-alpha or sodium D, or of the ratio of the intensities of emission lines in gaseous nebulae. Intermediate-band photometry ($15 < R < 50$) measures spectroscopic features that cannot be resolved with broader bands, but avoids the severe light loss of the very narrow bands. Examples of such features include discontinuities in spectra (for example, the "Balmer discontinuity" due to the onset of continuous absorption by hydrogen in stellar atmospheres at a wavelength of 364.6 nm), or very broad absorption features due to blended lines or molecular bands (for example, the band due to TiO in the spectra of M stars that extends from 705 to 730 nm).

10.2.2 Magnitudes

Recall that for some band (call it P), the ***apparent magnitude*** of the source as defined in Chapter 1 is just

$$m_P = -2.5\log(F_P) + C_P = -2.5\log\int_0^\infty R_P(\lambda) f_\lambda d\lambda + C_P \qquad (10.1)$$

Here m_P is the bandpass magnitude; F_P is the energy flux (the irradiance) within the band; f_λ is the monochromatic flux (also called the flux density or the monochromatic irradiance – it has units of watts per square meter of area per unit wavelength, or W m^{-3}). We choose the constant C_P to conform to some standard scale (e.g. the magnitude of Vega is zero in the visual system). The function $R_P(\lambda)$ describes the ***response*** of the entire observing system to the incident flux: it is the fraction of the energy of wavelength λ that will register on the photometer. We usually assume that f_λ is measured *outside* the Earth's atmosphere.

Photon detectors count photons, rather than measure energy directly. Recall that the ***monochromatic photon flux*** $\phi(\lambda)$ (number of photons per second per square meter of area per unit wavelength) is related to f_λ:

$$\phi(\lambda) = \frac{\lambda}{hc} f_\lambda$$

Photon detectors do *not* directly measure the quantity F_P in Equation (10.1) but report a signal proportional to the ***photon flux within the band***:

$$\Phi_P = \int_0^\infty R_{PP}(\lambda)\phi(\lambda)d\lambda = \frac{1}{hc}\int_0^\infty R_P(\lambda)f_\lambda\lambda d\lambda$$

Here $R_{PP}(\lambda)$ is the ***photon response***: the fraction of photons of wavelength λ detected by the system. This suggests that photon-counting detectors and energy-measuring detectors will measure on the same magnitude scale if

$$m_P = -2.5\log(\Phi_P) + C_{PP} = -2.5\log(F_P) + C_P$$

which requires

$$R_{PP}(\lambda) \propto \frac{R_P(\lambda)}{\lambda}$$

Although directly *measured* magnitudes are bandpass magnitudes, it makes perfect sense to talk about and compute a ***monochromatic magnitude***. This is defined from the monochromatic flux:

$$m_\lambda = -2.5\log(f_\lambda) + C'(\lambda) = -2.5\log\frac{hc\phi(\lambda)}{\lambda} + C'(\lambda) \qquad (10.2)$$

Here again, the value of the function $C'(\lambda)$is arbitrary, but is often chosen so that the monochromatic magnitude of Vega or some other (perhaps fictitious) standard is a constant at every wavelength. In this case, $C'(\lambda)$is a strong function of wavelength. Sometimes, however, the function $C'(\lambda)$ is taken to be a constant, and the monochromatic magnitude reflects the spectrum in energy units. You can think of the monochromatic magnitude as the magnitude measured with an infinitesimally narrow band. Conversely, you can think of intermediate or broadband photometry as yielding a value for m_λ at the effective wavelengths of the bands, so long as you recognize the energy distribution referenced is one of very low spectroscopic resolution.

10.2.3 Response function implementation

How is a band response implemented in practice? Both practical limits and intentional controls can determine the functional form of the responses $R_P(\lambda)$ or $R_{PP}(\lambda)$.

The ***sensitivity of the detector*** clearly limits the range of wavelengths accessible. In some cases, detector response alone sets the bandpass. Ptolemy, for example, based his magnitude system simply on the response of dark-adapted human vision, sensitive in the band 460–550 nm. In other cases, the detector response defines only one edge of the band. Early photographic magnitudes, for example, had a bandpass whose long-wavelength cutoff was set by the insensitivity of the photographic emulsion longward of 450 nm.

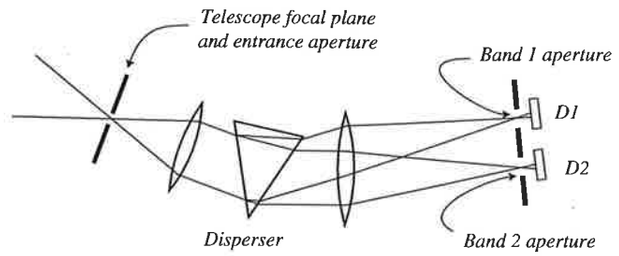


Fig. 10.1
A spectrophotometer. Each aperture defines the range of wavelengths that pass to its detector. It is possible to alter the wavelengths sampled by rotating the dispersing element or translating the apertures. In this case, the instrument is known as a spectrum scanner.

A ***filter*** – an element placed in the optical path to restrict transmission – is the usual method for intentionally delimiting a band. A ***bandpass filter*** defines both ends of the band by blocking all wavelengths except for those in a specific range. A filter can serve as a ***high-pass*** or ***low-pass*** element by defining only the lower or upper cutoff of a band. Filters that limit the transmission of all wavelengths equally are termed ***neutral-density filters***.

Another strategy for photometry is to use a dispersing element to create a spectrum. Sampling discrete segments of a spectrum with one or more photodetectors is equivalent to multi-band photometry. Such instruments are termed ***spectrophotometers***. A spectrophotometer (see Figure 10.1) generally defines bandpasses by using apertures, slots, or detectors of the proper size to select the desired segment of the spectrum. Multi-pixel solid-state detectors like CCDs blur the distinction between a spectrophotometer and a spectrograph: taking a CCD image of a spectrum is equivalent to letting each pixel act as an aperture that defines a band.

For ground-based observations, ***atmospheric transmission***, $S_{atm}(\lambda)$, limits the wavelengths that are accessible, and may completely or partially define a response function. Absorption in the Earth's atmosphere set the short wavelength cutoff of early photographic photometry at 320 nm, for example. In the infrared, absorption by water vapor is significant and variable. Figure 10.2 shows the approximate atmospheric transmission in the near infrared from 0.8 to 2.6 μm expected at a high elevation site. Also marked on the plot are the half-widths of the Johnson J and K bands as defined by filter transmission only. In these bands the atmosphere will set the long cutoff of J and the short cutoff of the K band, and variations in the atmosphere may change the shape of the overall photometric response function.

Normally, however, magnitudes are defined outside the Earth's atmosphere, and an astronomer must usually remove atmospheric effects during data reduction.

As an example of response definition, Figure 10.3 shows how four different factors interact to produce the response of the Johnson U band:

1. The transmissions of the filter – Corning glass number 9863 in Johnson's original definition.

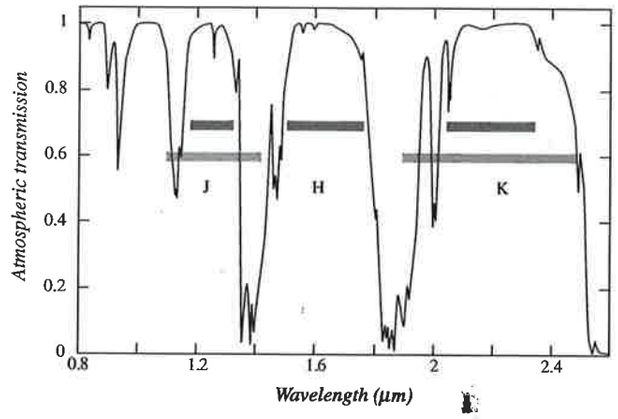


Fig. 10.2 Atmospheric transmission in the near infrared. Transmission curve is based on a model of the atmosphere at an elevation of 2.0 km, and will change with changes in water-vapor content. Light-gray lines locate the Johnson J and K photometric band-filter sensitivity (FWHM). Dark-gray lines show the sensitivity of the MKO filters for J, H, and K. The Johnson band definitions are much more susceptible to water-vapor variation than are the MKO definitions.

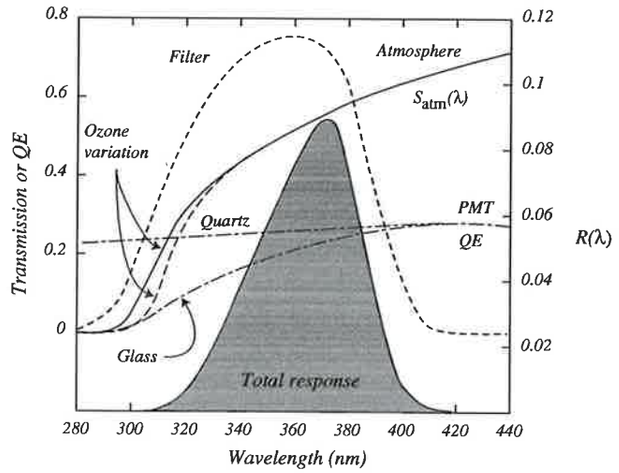


Fig. 10.3 Response function (shaded) for the Johnson U band. The function $R(\lambda)$ is the product of (1) the filter transmission, (2) the detector quantum efficiency with either a quartz or a glass window, and (3) the transmission of the atmosphere (two extremes, 4 mm and 2 mm of O_3, are indicated). The telescope and optics transmission usually do not affect the shape of $R(\lambda)$.

2. The quantum efficiency (QE) of the detector as a function of wavelength. In this case, the detector was a particular photomultiplier, the RCA 1P21 (now obsolete), which had an S-4 photocathode. The glass window of early tubes was later replaced with fused quartz, changing the short wavelength transmission.
3. The transmission of the atmosphere, $S_{atm}(\lambda)$. Photometry in this band assumes that the object is at the zenith, and that the ozone partial pressure is 3 mm. Changes in ozone concentration or zenith angle change the shape of $R_U(\lambda)$. For a PMT with a quartz window, the atmosphere sets the short wavelength cutoff. This feature of the U-band definition can be troublesome.

4. Transmission of the telescope optics. This is not plotted in the figure, since the reflectivity of freshly deposited aluminum is nearly constant in this region, with a value of around 0.92. Use of glass lenses, windows, or silver surfaces would change the shape of the response function.

10.2.4 Response function description

You will encounter various terms used to describe the response function. For example, for most response functions, there will be a single maximum value, R_{max}, which occurs at the ***peak wavelength*** λ_{peak}. Likewise, there are usually (only) two half-maximum points. These can be taken as specifications of the wavelengths at which transmission begins and ends, λ_{low} and λ_{high}:

$$R(\lambda_{peak}) = R_{max}$$

$$R(\lambda_{low}) = R(\lambda_{high}) = R_{max}/2$$

Given the half maxima, we can then define one measure for the width of the response by computing the ***full width at half-maximum***:

$$\text{FWHM} = \lambda_{high} - \lambda_{low}$$

The half-maximum points also determine the ***central wavelength*** of the band, which may be more representative of its mid point than λ_{peak}:

$$\lambda_{cen} = (\lambda_l + \lambda_{high})/2$$

A somewhat more sophisticated and possibly more useful measure of the width of a particular response function is the ***bandwidth***:

$$W_0 = \frac{1}{R_{max}} \int R(\lambda) d\lambda$$

Likewise, a somewhat more sophisticated measure of the center of a band is its ***mean wavelength***, which is just

$$\lambda_0 = \frac{\int \lambda \cdot R(\lambda) d\lambda}{\int R(\lambda) d\lambda}$$

Figure 10.4 illustrates these relations. For a symmetric function,

$$\lambda_{peak} = \lambda_{cen} = \lambda_0$$

Perhaps even more informative is the ***effective wavelength*** of the response to a particular source. The effective wavelength is a weighted mean wavelength (weighted by the source flux) and indicates which photons most influence a particular measurement:

$$\lambda_{eff} = \frac{\int \lambda \cdot f_\lambda \cdot R(\lambda) d\lambda}{\int f_\lambda \cdot R(\lambda) d\lambda}$$

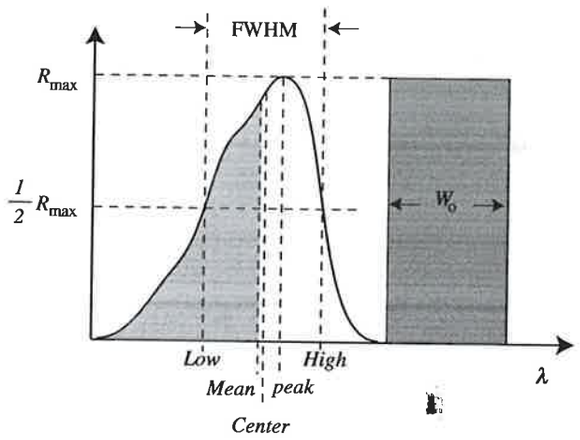


Fig. 10.4 Definitions of the middle and width of a band. The curve shows the function $R(\lambda)$. The mean wavelength divides the area under the curve into two equal parts (shaded and unshaded). The dark-gray rectangle has a width equal to the bandwidth and an area equal to the area under the curve.

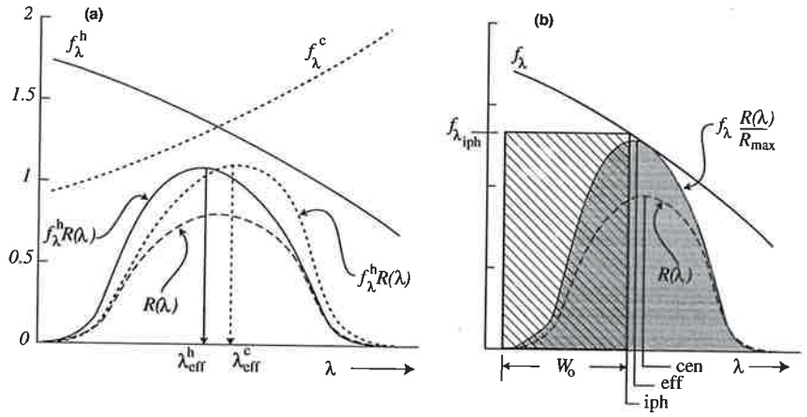


Fig. 10.5 (a) Effective wavelengths for two different sources in the same band. The solid curves apply to a hot source, and the dotted curves apply to a cool source with the same magnitude in the band. (b) Definition of the isophotal wavelength: the area of the hatched rectangle is the same as the shaded area under the curve. The dashed curve is the response function.

Figure 10.5 illustrates that different sources will in general have different effective wavelengths.

It is tempting to think of any bandpass measurement as equivalent to a measurement of the monochromatic flux at wavelength λ_{eff} multiplied by the bandwidth, W_0. This is nearly correct in practice, and for broadband photometry of stars (provided spectra are sufficiently smoothed) using this equivalence produces an error of a percent or less. To be strictly accurate with such an equivalence, we need to introduce yet another definition for the "middle" of the band. This one is called the ***isophotal wavelength***, λ_{iph}. The isophotal wavelength is the one for which we have

$$W_0 \cdot f_{\lambda_{iph}} = \frac{1}{R_{max}} \int f_\lambda \cdot R(\lambda) d\lambda$$

As with the effective wavelength, the exact value of the isophotal wavelength will depend on the spectrum of the source.

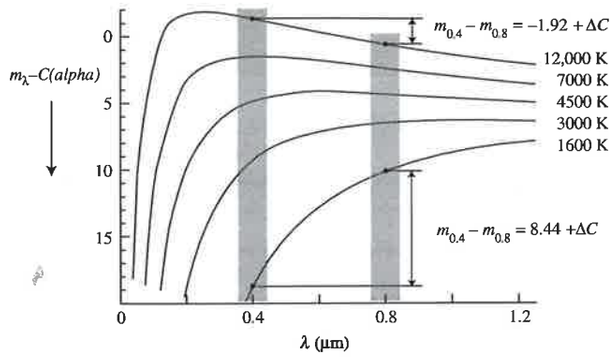


Fig. 10.6 Color indices for blackbodies. Curves are generated by taking the logarithm of the Planck function. Note that monochromatic magnitudes increase downwards. Spectra have been shifted vertically by arbitrary amounts for clarity. In this figure, $\Delta C = 0$.

10.2.5 Color indices

Multi-band photometry can measure the shape of an object's spectrum. It is convenient to think of the bands as sampling the monochromatic flux of a smoothed spectrum at their isophotal wavelengths. For example, Figure 10.6 shows the spectra of several blackbodies whose temperatures range from 1600 K to 12,000 K. The vertical scale of the figure shows the monochromatic magnitude in a system in which the constant in Equation (10.2) is independent of wavelength. Remember, this is *not* the usual case in astronomical photometry, where the spectrum of some standard object (e.g. Vega, which is similar to a blackbody with temperature of 9500 K), would be a horizontal line in a plot of m_λ as a function of λ. In the figure, we assume two bands, one with a mean wavelength at 0.4 μm, the other at 0.8 μm. It is clear that the arithmetical difference between these two magnitudes for a particular spectrum depends on the average slope of the spectrum, which in turn depends on the source's temperature. The convention is to speak of the difference between any two bandpass magnitudes used to sample the slope of the spectrum as a ***color index***.

For blackbodies, at least, the color index is not just useful, but definitive – its value uniquely measures the body's temperature. By convention, you compute the index in the sense:

$$\text{index} = m(\text{shorter}\lambda) - m(\text{longer}\lambda)$$

As mentioned earlier, astronomers usually symbolize the color index as the magnitude difference, sometimes enclosed in parenthesis. In the case of Figure 10.6, we might write the index as $(m_{0.4} - m_{0.8})$. In the case of the Johnson–Cousins red and infrared bands, the index would be written $(m_R - m_I)$, or more commonly $R - I$.

The behavior of the color index at the long and short wavelength extremes of the Planck function is interesting. In the Rayleigh–Jeans region (i.e. where $\lambda kT \gg hc$) you can show that

$$m_\lambda = \log T + C(\lambda) \tag{10.3}$$

so that the color index becomes

$$(m_{\lambda_1} - m_{\lambda_2}) = C(\lambda_1) - C(\lambda_2) = \Delta C$$

a constant independent of temperature. For example, in the Johnson broadband system, a blackbody of infinite temperature has color indices

$$(U - B) = -1.33, (B - V) = -0.46$$

At short wavelengths, the ***Wien approximation*** for the surface brightness of a blackbody holds:

$$B(\lambda, T) \approx \frac{2hc^2}{\lambda^5} \exp\left(-\frac{hc}{\lambda kT}\right)$$

So the color index is

$$(m_{\lambda_1} - m_{\lambda_2}) = \frac{a}{T}\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right) + C(\lambda_1) - C(\lambda_2) \tag{10.4}$$

Thus, at very small temperatures or wavelengths, the index is a linear function of $1/T$.

10.2.6 Line and feature indices

Real objects almost always have more complex spectra than do blackbodies, with features of astrophysical significance that may include absorption and emission lines, bands, and various discontinuities. Multi-band photometric indices can measure the strength of such features.

Two bands often suffice to measure the size of a discontinuity or the strength of a line, for example. In Figure 10.7a, bands C and D sample the continuum on the short and long wavelength sides of a sharp break in a spectrum. The index $(C-D)$ will be sensitive to the size of the break – but note two features of the index:

First, the actual relation between the size of the break and the numerical value of the $(C-D)$ index depends on the constants employed in the definition

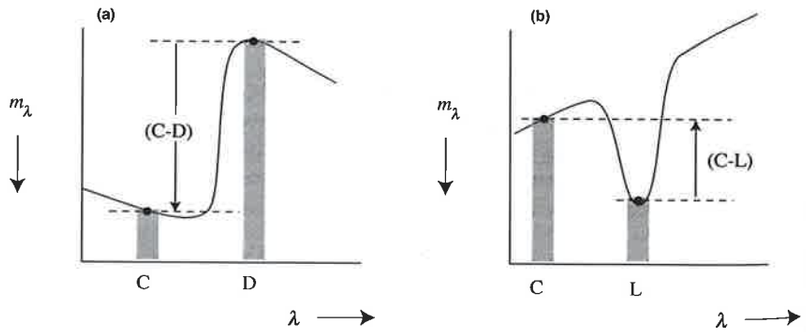


Fig. 10.7 Definition of indices to measure the strength of (a) a spectrum discontinuity, and (b) an absorption line. Monochromatic magnitudes are defined so that the constant in Equation (10.2) is independent of wavelength.

of the bandpass magnitudes in Equation (10.1). It might be convenient to have $(C-D) = 0$ when the break vanishes, but this may violate the convention that all indices should be zero for the spectrum of some standard object. (Examine Figure 1.5 – Vega has several non-zero spectrum discontinuities, yet all its indices are zero in some systems.)

Second, positioning the bands is important. The sensitivity of the index to the size of the break will diminish if either bandpass response includes light from the opposite side of the break. Likewise, if a band is located too far away from the break, unrelated features in the spectrum can affect the index. Obviously, it will be easier to position narrow bands than wide bands, but narrow bands give weaker signals.

A similar index can measure the intensity of an absorption or emission line (Figure 10.7b). Here one narrow band is centered on the feature, and the other on the nearby continuum. The magnitude difference measures the line strength. This strategy is common in detecting and mapping objects with strong emission lines in their spectra: for example, the astronomer takes two CCD exposures – one through a filter centered on the emission line in question, the second through one centered on the nearby continuum. Digital subtraction of the two registered and properly scaled images produces zero signal except in pixels where a source is emitting radiation in the line.

Figure 10.8 illustrates an alternative strategy for measuring a line index. Two bands – one broad, the other narrow – are both centered on the line. The narrow band is quite sensitive to the strength of the line, while the broad band is relatively insensitive, since most of the light it measures comes from the continuum. The index

$$\text{line index} = m_{\text{narrow}} - m_{\text{wide}}$$

tracks the strength of the absorption, in the sense that it becomes more positive with stronger absorption. One widely used line index of this sort is the β index, which measures the strength of the Balmer beta line of hydrogen, usually useful for luminosity or temperature classification of stars.

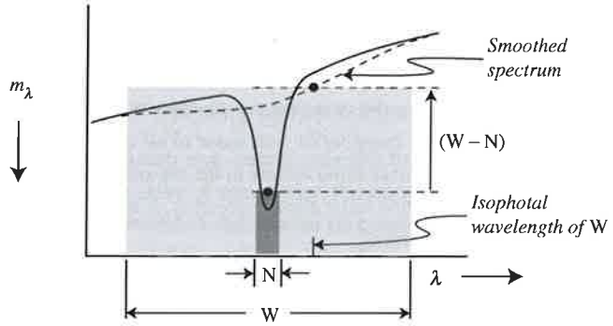


Fig. 10.8 A line index computed from wide and narrow bands centered on the same absorption line.

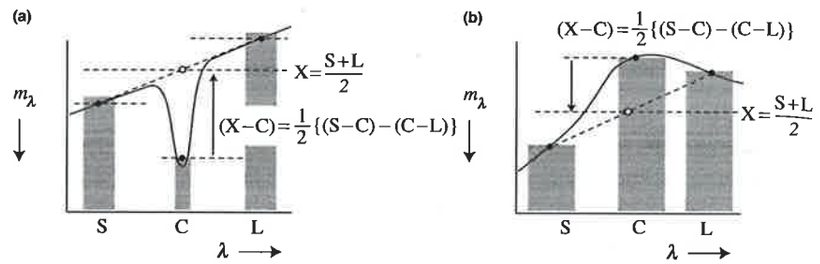


Fig. 10.9 Three bands can measure the curvature of the spectrum. In both (a) and (b), the index $2(X - C)$ tracks the monochromatic magnitude's departure from linearity.

Finally, consider a third kind of index. Three bands can measure the *curvature* (i.e. the second derivative, rather than the first) of a spectrum. Curvature can arise on a relatively small scale because of a sharp absorption or emission line, or on a large scale because of broad or diffuse features (molecular bands in gases, or absorption features in the reflection spectra of crystalline solids, for example). Figure 10.9 illustrates two situations with three (a) equally and (b) unequally spaced bands at a short, central, and long wavelength (S, C, and L). If we consider just the monochromatic magnitudes, and if the bands are equally spaced as in Figure 10.9a, the index

$$\text{curvature} = (m_S - m_C) - (m_C - m_L) = S + L - 2C$$

will be zero if the logarithmic spectrum is linear, and positive if the central band contains an absorption feature. *The curvature index depends on the difference between two color indices.* In practical systems the index will still track curvature even if bands are not equally spaced, and even if $C'(\lambda)$ in Equation (10.2) is *not* a constant.

10.3 The idea of a photometric system

The term ***photometric system*** implies at least two specifications:

1. The wavelength response for each band – that is, the shape of the function $R_P(\lambda)$ in Equation (10.1)
2. Some method for standardizing measurements made in those bands. This is important for two reasons:
   - Each observer needs to know the value for the constant C in Equation (10.1) that will assure agreement of his magnitudes with those of all other observers.
   - The differing hardware produces some variety in the response functions in practice, so a method for standardization must allow correction of the inevitable systematic effects due to imperfect matching.

The first specification, that of $R_P(\lambda)$, determines the ***instrumental*** or ***natural system***. The first and second together determine the ***standard system***.

Observations in the natural system alone can be quite useful (e.g. determining the period of a variable star), but only by placing magnitudes on the standard system can two astronomers confidently combine independent measurements.

Standardization might involve observations of laboratory sources, e.g. a blackbody of known temperature and therefore known absolute flux in W m^{-2}. Almost always, though, a single astronomical object or set of objects is a much more practical standardizing source. Almost all standard systems today rely upon some network of constant-brightness standard objects distributed around the sky. If everyone agrees on a list of stars and their corresponding magnitudes, anyone can calibrate measurements made in their instrumental system by observing the standards and the unknowns with the same procedures. Because systematic differences will most likely arise if the spectrum of the star observed is different from the spectrum of the standard star, most systems strive to define a set of standards that includes a wide variety of spectral types.

As we have seen with the Johnson U, K, and J bands, the definition of the bandpass sometimes involves the atmosphere. Atmospheric absorption, however, is variable, and removing this variation should be part of the standardization procedure.

Because standardization is so essential to a photometric system, some astronomers have devised ***closed photometric systems***, in which a relatively small group of observers carefully controls the instruments and data reduction, maximizing internal consistency. Many space-based observations (e.g. HIPPARCOS), and many ground-based surveys (e.g. the Sloane Digital Sky Survey) constitute closed systems. An ***open photometric system***, in contrast, is one in which all astronomers are encouraged to duplicate the defined natural system as best they can, and, through reference to a published list of standard stars, add to the pool of observations in the system.

10.4 Common photometric systems

Astronomers have introduced several hundred photometric systems. Bessel (2005) gives an extensive review of the most common systems. Here we examine only a few of the most widely used as an introduction to the operation of most.

10.4.1 Visual and photographic systems

The dark-adapted human eye determines the band of the ***visual photometric system***. In the earliest days of astronomy, the standardization procedure required that magnitudes measured in the system be consistent with the ancient catalogs (e.g. Ptolemy, al Sufi, and Bayer). The introduction of optical/mechanical visual photometers led to the establishment of ***standard sequences*** of stars, including (initially) the ***north polar sequence*** and (later) many secondary sequences (the

Table 10.1. *Bandpasses of historical importance*

Band	Symbol	Band definition	λ_{peak}, nm	FWHM
Visual	m_{vis}	Mesotopic* human eye	515–550	82–106
International photographic	m_{pg}, IPg	Untreated photographic emulsion + atmosphere	400	170
International photovisual	m_{pv}, IPv	Orthochromatic emulsion + yellow filter	550	100

* Visual photometry of stars uses a mixture of photopic (color, or cone) and scotopic (rod) vision, with the shift from cones to rods occurring with decreasing levels of illumination. The effective wavelength of the eye thus shifts to the blue as light levels decrease (the Purkinje effect); see Appendix B3.

48 Harvard standard regions and the 115 Kapteyn selected areas were perhaps the best studied).

In the early twentieth century, astronomers defined two bands based on the properties of the photographic emulsion (Table 10.1). The poor properties of the photographic emulsion as a photometric detector, and lack of very specific definitions, limited the success of this system. The ***international photographic band*** is sensitive in the near ultraviolet–blue region. The response of the ***international photovisual band***, somewhat fortuitously, roughly corresponds to that of the visual band (i.e. the human eye, sensitive to green–yellow). The IAU in 1922 set the zero point of both magnitudes so that 6th magnitude A0 V stars[2] in the north polar sequence would have (roughly) the same values as on the old Harvard visual system. This meant that the color index,

$$\text{color index} = m_{pg} - m_{pv}$$

should be zero for A0 stars, negative for hotter stars, and positive for cooler stars.

Many other photographic systems exist. The photovisual magnitude originally depended on "orthochromatic" plates, which were made by treating the emulsion with a dye to extend its sensitivity to about 610 nm. Other dyes eventually became available to extend photographic sensitivity to various cutoffs ranging through the visible and into the near infrared. Twentieth-century astronomers devised many filter–emulsion combinations and set up standard

[2] A0 V is the spectral type of Vega, which is *not* in the north polar sequence. Because of the early decision to keep visual magnitudes roughly consistent with the ancient catalogs, the photographic and photovisual magnitudes of Vega turn out to be close to zero. The importance of Vega stems in part from its brightness, which makes it a good candidate for absolute (i.e. watts per square meter per meter of wavelength) measurement of specific irradiance.

star sequences in a variety of photography-based systems. All these are mainly of historic interest.

10.4.2 The UBVRI system

By far the most widely used ground-based photometric system prior to the present has been the Johnson–Cousins UBVRI system (Table 10.2 and Figure 10.10). Johnson and Harris (1954) defined the UBV portion first, based on the response of the RCA 1P21 photomultiplier, a set of colored glass filters, and a list of magnitudes for a relatively small number of standard stars scattered around the celestial sphere. The V band closely corresponds to the international photovisual band and its zero point was set so that $V = m_{pv}$ for the standards in the north polar sequence. The U and B bands correspond to short- and long-wavelength segments of the photographic band, and to be consistent with the international system, their zero points are set so that the colors $U-B$ and $B-V$ are zero for A0 V stars.

After some pioneering work at longer wavelengths by Stebbins, Kron and Whitford, Harold Johnson and his collaborators in the period 1960–1965 extended the UBV system to include bands in the red (R_J) and near infrared (I_J), as well as the longer infrared bands (JHKLMNQ) discussed in the next section. Modern work with CCDs, however, has tended to replace the R_J and I_J with the R_C and I_C bands specified by Cousins and his collaborators (see Table 10.2 for the differences). In current practice, the lists of Arlo Landolt (1983, 1992) and Menzies *et al.* (1989, 1991) define the standard stars for the UBV$(RI)_C$ system.

Modern CCD observers sometimes have difficulty replicating the original photomultiplier-based instrumental system. A complicating factor is the great variation in CCD spectral response due to differing surface treatments, gate material, gate structure, backside illumination, etc. The U band causes the

Table 10.2. *The Johnson–Cousins UBVRI system.* The R_J and I_J data are from Colina *et al.* (1996). All other widths are from Bessel (1992). Effective wavelengths and monochromatic fluxes for a zero-magnitude, zero-color star are from the absolute calibration of Vega and Sirius by Bessell *et al.* (1998). Vega has $V = 0.03$ on this system

	U	B	V	R_C	R_J	I_C	I_J
λ_{eff}, nm	366	436	545	641	685	798	864
FWHM	66	94	88	138	174	149	197
f_λ at λ_{eff} in units of 10^{-12} W m^{-2} nm^{-1} for $V = 0$	41.7	63.2	37.4	22.6	19.2	11.4	9.39

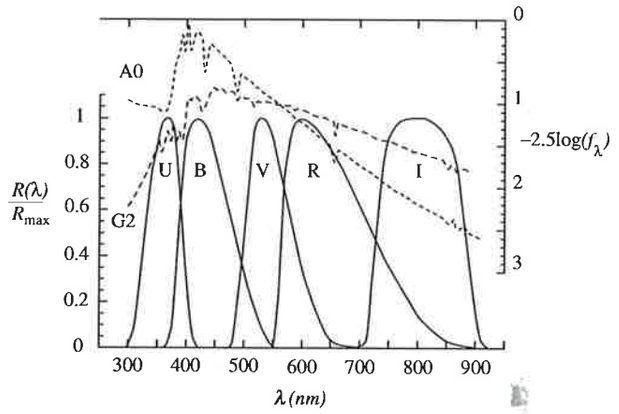


Fig. 10.10 Normalized response functions for the UBVRI system. Also shown are the monochromatic magnitudes for a representative A0 and G2 dwarf. Note the importance of the Balmer discontinuity near 370 nm in the A0 spectrum, and the break due to metal absorption near 400 nm in the G2 spectrum.

most trouble, partly because it is defined by the atmosphere at a particular altitude (see Section 10.2.3), and partly because of generally poor CCD response in the ultraviolet. Close matches are possible with a good knowledge of the individual CCD response and a careful choice of filters. For details, see Bessel (1990).

This multi-band system was designed with the rough spectral classification of stars in mind. Figure 10.10 shows the responses of the normalized UBV(RI)$_C$ bandpasses superimposed on spectra of an A0 and a G2 dwarf (i.e. matching, respectively, Vega and the Sun). The $U-B$ index is clearly sensitive to the Balmer discontinuity (present very obviously in the A star at 370 nm, and much reduced in the G star). The discontinuity – and hence the $U-B$ index – depends upon luminosity, at least for hot stars. The other indices are primarily sensitive to temperature (and therefore spectral type). The $B-V$ color is more sensitive to metal abundance than are $V-R$ or $R-I$, and fails as a useful index for M stars because of molecular band absorption. (In astrophysics, a "metal" is any element other than hydrogen or helium.) Because of its long baseline and relative insensitivity to chemical abundances, the $V-I$ index is the most purely temperature-sensitive index in this system ($V-K$ is even better, for the same reason). Appendix A10.1 tabulates the colors of various spectral types. The system is useful for measuring the photometric properties of objects besides normal stars: Solar System bodies, supernovae, galaxies, and quasars have all been extensively observed.

10.4.3 The broadband infrared system: JHKLMNQ

The broadband infrared system (Table 10.3) might be regarded as an extension of the UBVRI system, and shares a common zero point (so the colors of an unreddened A0 V star are zero). Detectors in this region cannot be silicon CCDs, but must be infrared arrays or single-channel infrared-sensitive devices.

Table 10.3. *The broadband infrared system.* JHKL from Bessell *et al.* (1998), M band from Rieke and Lebofsky (1985), and N and Q from Rieke *et al.* (1985)

	J	H	K	L	M	N	Q
λ_{eff}, μm for A0 stars	1.22	1.63	2.19	3.45	4.8	10.6	21
FWHM	0.213	0.307	0.39	0.472	0.46	3–6	6–10
f_λ at λ_{eff} in units of 10^{-11} W m^{-2} μm^{-1} for $V=0$	315	114	39.6	7.1	2.2	0.96	0.0064

Table 10.4. *Mauna Kea (MKO) filter characteristics.* Central wavelengths of L′ and M′ are significantly different from L and M, hence the renaming. Note that these are filter characteristics: actual bandpass responses will depend on detector, atmosphere, telescope optics, etc

	J	H	K	L′	M′
λ_{cen}, μm	1.24	1.65	2.20	3.77	4.67
FWHM	0.16	0.29	0.34	0.70	0.22

Subtraction of background can be a very serious problem in the infrared, as discussed in the previous chapter.

A more important complication is the fact that, for the ground-based infrared, bandpass definitions can depend very critically on atmospheric conditions (mainly the amount of water vapor encountered along the line of sight). Different observatories with identical hardware can experience different infrared window sizes and shapes if they are at different altitudes (extending to space observatories). The same observatory can experience similar bandpass variations due to changing humidity.

Different observatories have thus defined infrared bands differently, and the values in Table 10.3 merely represent the typical choices prior to the twenty-first century. The IAU in 2000 recommended a preferred natural system for JHK – the Mauna Kea Observatory near-infrared system (see Table 10.4). The MKO system attempts to minimize sensitivity to water vapor while optimizing the signal-to-noise ratio (SNR), usually by narrowing the FWHM.

A second important characteristic of infrared photometric systems stems partly from their relative immaturity: the standard star magnitudes for these bands are not as well defined as in the CCD region. The situation is best in the JHK bands, where at least three different and largely non-overlapping (and

Table 10.5. *The four-color and β system*

Name	u Ultraviolet	v Violet	b Blue	y Yellow	H_β wide	H_β Narrow
λ_{eff}, nm for A0 stars	349	411	467	547	489	486
FWHM, nm	30	19	18	23	15	3.0

still evolving) lists of standard stars have been in common use. These are beginning to converge on a common system consistent with measurements derived from the MKO-near-infrared bandpasses.

10.4.4 The intermediate-band Strömgren system: uvbyβ

Bengt Strömgren designed this intermediate-band system in the late 1950s, and David Crawford and many others developed it observationally in the 1960s and 1970s. They published several lists of standard stars during these years. The system avoids many of the shortcomings of the UBV system, and aims to classify stars according to three characteristics: temperature, luminosity, and metal abundance. Classification works well for stars of spectral types B, A, F, and G, provided the photometry is sufficiently accurate. Photometrists frequently supplement the four intermediate-band colors, uvby, with a narrow band index, β, which tracks the strength of absorption in the Balmer beta line. The β index greatly improves the luminosity classification for hotter stars, and is a good temperature indicator for cooler stars.

Emission in all of the four intermediate bands depends on temperature, but in addition, emission in the u and v bands is depressed by the presence of metals in a star's atmosphere. Also, the u band is depressed by the Balmer discontinuity, a temperature-dependent feature which is strongest for A0 stars, but which also depends on luminosity. To indicate astrophysical information, then, Strömgren photometry is generally presented as a y magnitude, a $(b-y)$ color, and two curvature indices. The $(b-y)$ color closely tracks temperature in the same way as the Johnson $B-V$ (in fact, $b-y \approx 0.68(B-V)$ over a large range of stellar types), but $(b-y)$ is somewhat less sensitive to abundance effects and is more useful at lower effective temperatures than is $B-V$. The two curvature indices are

$$c_1 = (u-v)-(v-b)$$

$$m_1 = (v-b)-(b-y)$$

The c_1 index measures the strength of the Balmer discontinuity, and in combination with temperature from $(b-y)$ yields information about luminosity. It is an improvement over the Johnson $(U-B)$, partly because the U filter straddles the Balmer discontinuity. The m_1 index measures metal abundance. The precise relationships between the indices and the astrophysical parameters are more complex than suggested here, but they have been well calibrated for spectral types hotter than K0.

10.4.5 Other systems

Many other photometric systems find less widespread use than those just described, and it is helpful to describe a few examples.

Photometry from space need not contend with any of the atmospheric and many of the background issues that complicate photometry from the ground. Within the parameters of a given detector, space observatories permit much greater freedom to base bandpass design on purely astrophysical considerations. The NICMOS2 camera on the Hubble Space Telescope (HST), for example, carried about 30 filters, many centered at bands completely inaccessible from the ground.

It is nevertheless very important to be able to tie space observations to ground-based measurements. The HIPPARCOS space mission, for example, used a two-filter broadband system closely related to B and V, while some of the NICMOS filters correspond to the JKLMN bands. The primary CCD camera for the HST (the WFPC/WFPC2), had slots for 48 filters, but those most commonly used closely matched the UBVRI system. The HST standard magnitudes, incidentally, are defined so that a source with constant f_ν has zero colors.

We can expect the introduction of novel ground-based systems to continue. New CCD-based systems might even replace well-established photomultiplier-based systems. For example, the Sloan Digital Sky Survey (SDSS), the automated ground-based program that is expected to produce photometry for over 10^8 stellar and non-stellar objects, uses a five-color system (see Table 10.6) designed to make optimal use of silicon CCD sensitivity. The SDSS database will be larger than all the PMT-based UBVRI observations accumulated since the 1950s. Since the SDSS colors give as good or better astrophysical information, the SDSS may (or may not) eventually displace UBVRI as the dominant broadband system in the visual.

Table 10.6. *The bands for the SDSS five-color system*

	u′	g′	r′	i′	z′
λ_{ccn}, nm	354	477	623	762	915
FWHM	57	139	137	153	95

10.5 From source to telescope

> . . .Slowly the Emperor returned –
> Behind him Moscow! Its onion domes still burned. . .
> Yesterday the Grand Army, today its dregs!
> . . .They went to sleep ten thousand, woke up four.
>
> – *Victor Hugo*, Russia, 1812, Trans. Robert Lowell

A grand army of photons leaves a source, but many are lost on their march to our telescope. This section follows one regiment of that army to consider its fortunes in detail. The goal of photometric reduction will be to reconstruct the original regiment from its dregs – to account for all losses and transformations during its long journey in the cold.

At least four different effects can alter the photons on their way to the telescope:

- wavelength shifts
- extragalactic absorption
- Galactic and Solar System absorption
- atmospheric absorption.

10.5.1 Wavelength changes

The regiment that leaves the source is $\phi_E(\lambda_E)d\lambda_E$, that is, all those photons with wavelength between λ_E and $\lambda_E + d\lambda_E$ emitted in one second in the direction that would place them in a unit area of our telescope aperture. The subscript E just means "emitted". The dregs of the regiment are the members of that original group that actually survive at the top of our atmosphere, which we will call $\phi(\lambda)d\lambda$. In the general case, we allow for the possibility that both numbers and their wavelengths can change.

We first consider the consequences of wavelength change by itself. Because of the Doppler effect, or because of the expansion of the Universe, or because of various relativistic effects, the wavelength of each photon on arrival may differ from its original value, λ_E. The new value is given by

$$\lambda_o = (1 + z)\lambda_E$$

where z is the ***redshift parameter*** ($z = (\lambda_o - \lambda_E)/\lambda_E$) of the source. Because of this wavelength change, the photons emitted into wavelength interval $d\lambda_E$ will arrive spread out over wavelength interval $d\lambda_o = (1 + z)d\lambda_E$. Since we consider only the effect of wavelength change, and ignore the absorption effects listed above, we can say that the *number* of photons is conserved, that is:

$$\phi(\lambda)d\lambda = \phi_E(\lambda_E)d\lambda_E$$

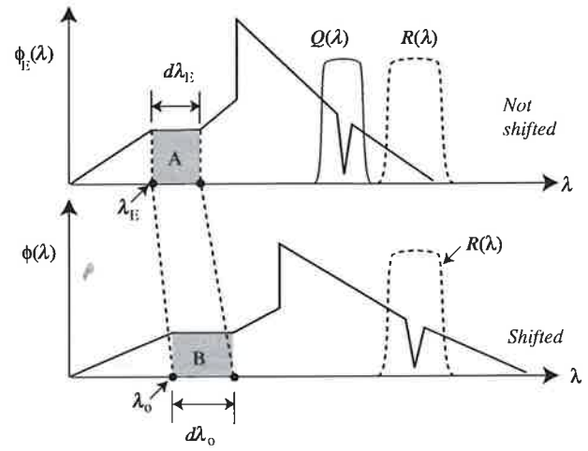


Fig. 10.11 Wavelength shifts and photometry. The upper panel shows an unshifted spectrum, and the lower panel shows the same spectrum shifted in wavelength by a redshift parameter $z = 0.2$. Photons originating in area A arrive in area B. Photons measured in band R originate in bandpass Q.

We have dropped the subscript for the observed wavelength. Thus, the observed and emitted monochromatic photon flux are related by (see Figure 10.11)

$$\phi(\lambda) = \frac{1}{1+z}\phi_E\left(\frac{\lambda}{1+z}\right) = f_\lambda\frac{\lambda}{hc} \qquad (10.5)$$

The function f_λ is the observed monochromatic flux density of Chapter 1. Although photon number is conserved, the monochromatic flux density is not:

$$f_E(\lambda_E) = f_E\left(\frac{\lambda}{1+z}\right) = \frac{hc\phi_E(\lambda_E)}{\lambda_E} = (1+z)^2 f_\lambda \qquad (10.6)$$

In practice, we must deal with an observer who makes magnitude measurements in a band. An observer who uses a bandpass with photon response $R(\lambda)$ will measure the magnitude

$$m_R = -2.5\log\int R(\lambda)\frac{\phi(\lambda)}{\lambda}d\lambda + C_R$$
$$C_R = -2.5\log\int R(\lambda)\frac{g_\lambda}{hc}d\lambda$$

where g_λ is the spectrum of a photometric standard of magnitude zero. We would like to understand how m_R relates to a magnitude measured for these same photons before their wavelength shift. These photons began their journey in a wavelength band different from R; call this band Q (again, see Figure 10.11). Measuring a magnitude in band Q on the unshifted spectrum gives

$$m_Q = -2.5\log\int Q(\lambda)\frac{\phi_E(\lambda)}{\lambda}d\lambda + C_Q$$
$$= -2.5\log\left[(1+z)\int Q(\lambda)\frac{\phi(\lambda(1+z))}{\lambda}d\lambda\right] + C_Q$$
$$C_Q = -2.5\log\left[\int Q(\lambda)\frac{g_\lambda}{hc}d\lambda\right]$$

If $m_R = m_Q$ then the observation in band R directly gives the magnitude the source would have in band Q if there were no shift. This equality is generally possible if R has the shifted shape of Q (i.e. $R((1+z)\lambda) = Q(\lambda)$). However, since R and Q must often be standard bands in some system, it is hard to meet this condition, so we must consider the difference:

$$m_R - m_Q = 2.5\log(1+z) + C_R - C_Q + 2.5\log\left[\frac{\int Q(\lambda)\frac{\phi(\lambda(1+z))d\lambda}{\lambda}}{\int R(\lambda)\frac{\phi(\lambda)}{\lambda}d\lambda}\right] \quad (10.7)$$

For objects in our own galaxy, z is quite small, and one almost always uses $R(\lambda) = Q(\lambda)$, $C_R = C_Q$. In that case, the first three terms in Equation (10.7) add to zero. The last term describes the effect of photons shifting into, out of, and within the band. In the case of narrow bands near sharp spectral features, even small Doppler shifts can produce large differences between $\phi((1+z)\lambda)$ and $\phi(\lambda)$. In such a case, one could compute m_Q from an observed m_R using Equation (10.7), provided enough is known about $\phi(\lambda)$.

For distant objects, z becomes large because of the expansion of the Universe. Equation (10.7) again suggests that it should be possible, given knowledge of z and $\phi(\lambda)$, to use an observed bandpass magnitude to compute the magnitude *that would be observed* (in the same or in a different band) *if the source had redshift* z = 0. Hubble called this kind of magnitude correction the ***K correction***. Although different authors define the K correction in slightly different ways, Hogg *et al.* (2002) give a good general introduction.

Wavelength shifts will affect the colors of galaxies with large z. Application of Equation (10.7) for two different bands gives an expression for the color change as a function of z. Having observed a color, you can solve that expression for z, so long as you can approximate the spectrum of a galaxy. These ***photometric redshifts*** from observed colors are valuable estimators of galaxy distance because they do not require observationally difficult spectroscopy of very faint objects.

10.5.2 Absorption outside the atmosphere

Space is not empty. Interstellar gas and dust in our own galaxy absorb and scatter light. Absorption (in which the photon ceases to exist) and scattering (in which the photon changes direction) are physically distinct processes, but they have the same effect on the regiment of photons headed towards our telescope – they remove photons from the beam. It is common to refer to both processes simply as "absorption." Absorption not only reduces the overall number of photons that arrive at the telescope – an effect sometimes called ***extinction*** – but it also alters the shape of the spectrum.

Diffuse gas absorbs photons to produce ***interstellar absorption lines and bands***. In the optical, the sodium D doublet is usually the strongest interstellar line, and in the ultraviolet, the Lyman-alpha line at 121.6 nm is usually strongest. Especially at short wavelengths, gas will also produce continuous absorption and absorption edges due to ionization. A strong feature at 91.2 nm due to ionization of hydrogen is very prominent, for example. Absorption by dust will generally alter the overall shape of the spectrum, and depending on its composition, add a few very broad features. In the region 0.22–5.0 μm, dust scatters short-wavelength photons more strongly than long-wavelength photons, so the resulting change in the shape of the spectrum is termed ***interstellar reddening***. In our notation, assume that

$S_{ism}(\lambda)$ = the fraction of photons of wavelength λ that are transmitted by the interstellar medium within our own galaxy.

$S_{cxg}(\lambda)$ = the fraction of photons arriving at observed wavelength λ that are transmitted by the interstellar medium outside our own galaxy

Note that because of the cosmological redshift, absorptions described by $S_{exg}(\lambda)$ involve photons that had wavelength $\lambda/(1+z')$ when they were absorbed by material with redshift parameter z'. (This produces the phenomenon of the ***Lyman-alpha forest*** in the spectra of distant objects: multiple absorption lines due to Ly α at multiple redshifts.) The photon flux that reaches the top of the Earth's atmosphere, then, is just

$$\phi(\lambda) = S_{ism}(\lambda)S_{cxg}(\lambda)\phi_0(\lambda) = S_{ism}(\lambda)S_{cxg}(\lambda)\frac{1}{1+z}\phi_E((1+z)\lambda) = f_\lambda\frac{\lambda}{hc}$$

We will call $\phi(\lambda)$ the *photon flux outside the atmosphere* and $\phi_0(\lambda)$ the *photon flux outside the atmosphere corrected for interstellar absorption*. In extragalactic astronomy, it is frequently difficult to estimate $S_{exg}(\lambda)$, so it may be important to distinguish between $\phi_0(\lambda)$ and the flux

$$\phi_G(\lambda) = S_{cxg}(\lambda)\phi_0(\lambda) \quad ,$$

where $\phi_G(\lambda)$ is the flux corrected for Galactic absorption but not for extragalactic absorption.

10.5.3 Absorption by the atmosphere

The Earth's atmosphere removes photons from the stream directed at our telescope, both through scattering and through true absorption. As before, we refer to both processes as "absorption", and note that atmospheric absorption will

both reduce the apparent brightness of the source spectrum as well as alter its shape. We therefore refer to ***atmospheric extinction*** and ***atmospheric reddening***. The atmosphere also introduces some sharper features in the spectrum, the ***telluric lines and bands***.

Extinction is a strong function of wavelength. At sea level, three opaque regions define two transmitting windows. Rayleigh scattering and absorption by atoms and molecules cause a complete loss of transparency at all wavelengths shorter than about 300 nm. This sets the short end of the ***optical–infrared window***. The second opaque region, from absorption in molecular bands (primarily due to H_2O and CO_2), begins at around 0.94 μm, has a few breaks in the near infrared and mid infrared, then extends from 30 mm to the start of the ***microwave–radio window*** at around 0.6 cm. The radio window ends at around 20 m because of ionospheric absorption and reflection.

Atmospheric extinction has a profound influence on life. The atmospheric infrared opacity prevents the Earth's surface from radiating directly into space and cooling efficiently. This so-called greenhouse effect is responsible for maintaining the average surface temperature at about 30 K higher than it would be without the atmosphere. Short-wavelength electromagnetic radiation is quite detrimental to biological systems, and none of the forms of life presently on Earth could survive if exposed to the solar gamma-ray, X-ray, and shortwave-ultraviolet radiation that is presently blocked by the atmosphere. Had life here originated and evolved to cope with an environment of either low temperatures or hard radiation, we would all be very different creatures indeed.

The wavelength dependence of extinction has an equally profound effect on astronomical life. Astronomy began by peering out at the Universe through the narrow visual window and evolved over many centuries to do a better and better job in that restricted region of the spectrum. Astronomy only discovered the radio window in the middle of the twentieth century. Yet later in that century, spacecraft (and aircraft) finally provided access to the entire spectrum. Only with the introduction of decent infrared arrays in the 1980s could astronomers take advantage of the gaps in the near-infrared atmospheric absorption available at dry high-altitude sites. Atmospheric absorption has made optical astronomy old, radio astronomy middle-aged, and gamma-ray, X-ray, and infrared astronomy young.

Quantitatively, we can postulate an atmospheric transmission function

$$S_{\rm atm}(\lambda, t, e, a) = \text{the fraction of photons of wavelength } \lambda \text{ that are transmitted by the Earth's atmosphere at time } t, \text{ elevation angle } e, \text{ and azimuth } a$$

The photon flux that actually reaches the telescope is then

$$\phi_{\rm A}(\lambda) = S_{\rm atm}(\lambda, t, e, a)\, S_{\rm ism}(\lambda)\, S_{\rm cxg}(\lambda) \frac{1}{1+z} \phi_{\rm E}((1+z)\lambda) = f_\lambda^{\rm A} \frac{\lambda}{hc}$$

The rate at which energy ultimately gets detected in an infinitesimal band will be

$$dE_{\rm sig} = aT'_{\rm P}(\lambda) f_\lambda^{\rm A} d\lambda = aT'_{\rm P}(\lambda) \frac{\phi_{\rm A}(\lambda)}{\lambda} d\lambda$$

Here a is the effective collecting area of the telescope, and $T'_{\rm P}(\lambda)$ is a function that quantifies the overall wavelength-dependent efficiency of the instrument. It includes such things as wavelength-sensitive reflectivity and transmission of optical elements, filter transmission, and detector quantum efficiency. Integrating the previous equation, we can express the raw ***instrumental magnitude*** measured inside the atmosphere:

$$\begin{aligned} m_{\rm P}^{\rm A} &= -2.5 \log \int T'_{\rm P} S_{\rm atm}(\lambda) \frac{\phi(\lambda)}{\lambda} d\lambda + C'_{\rm P} \\ &= m_{\rm P}^{\rm O} + A_{\rm atm} + A_{\rm ism} + A_{\rm cxg} + C_{\rm P}^{z} \\ &= m_{\rm P} + A_{\rm atm} \end{aligned}$$

Here the A parameters represent the atmospheric, Galactic, and extragalactic absorption, in magnitudes; $C_{\rm P}^{z}$ is the correction for wavelength shift; and $C'_{\rm P}$ is the constant that sets the zero point of the instrumental magnitude scale. The quantity $m_{\rm P}$, the ***instrumental magnitude outside the atmosphere***, depends on the telescope and photometer but is independent of the atmosphere. The quantity $m_{\rm P}^{\rm O}$ is the instrumental magnitude in the emitted frame corrected for all absorption effects.

We can write $m_{\rm P}$ as

$$m_{\rm p} = -2.5 \log \int T_{\rm P}(\lambda) \frac{\phi(\lambda)}{\lambda} d\lambda + C_{\rm P}$$

Here, $T_{\rm P}(\lambda)$ and $C_{\rm P}$ characterize the instrumental system located outside the atmosphere.

10.5.4 Photometric data: reduction strategy

The idea now is to remove all the effects outlined above and reconstruct the original stream of photons. Assume we have array data from a ground-based observatory. The data reduction steps will be:

1. Preprocess images to remove instrumental effects: non-linearities, sideband effects, dark, bias, and flat field. Correct for charge-transfer efficiency (CTE), fringing, and geometric effects like scale variation as needed. If appropriate, process (combine, shift and add, drizzle) to improve the SNR, and to remove flaws like bad pixels, mosaic gaps, and cosmic rays.
2. Perform digital aperture or area photometry on the processed images. This will measure the instrumental magnitude inside the atmosphere, $m_{\rm P}^{\rm A}$.

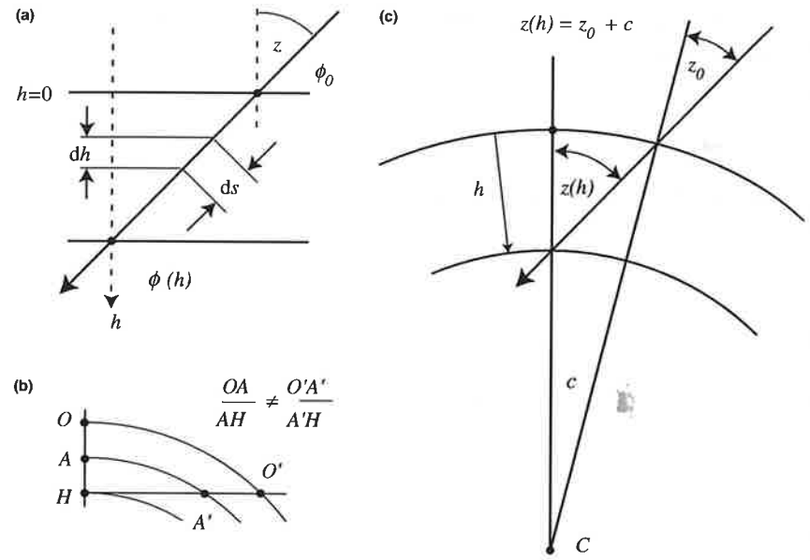


Fig. 10.12 Absorption geometries. (a) A plane-parallel slab. We assume the top of the atmosphere is at $h = 0$. Note that $ds = \sec(z)dh$, where z is the local zenith angle Figure (b) illustrates the fact that lower layers are more important in a spherical atmosphere. Figure (c) shows that the angle z increases with depth in a spherical shell. Refraction effects have been ignored.

3. Remove the effects of atmospheric absorption: compute the instrumental magnitudes outside the atmosphere, m_P.
4. Transform instrumental magnitudes and indices to a standard system, m_P^{STD}, if needed.
5. Derive astrophysical and astronomical corrections and parameters, as needed:
   (a) corrections for absorption: A_{ism}, A_{exg};
   (b) the correction for wavelength shift, C_P^z;
   (c) astrophysical conditions of the source: temperature, metallicity, stellar population age, distance, diameter, etc.

We described the first two steps in Chapter 9. Step 5 is beyond the scope of an introductory text, so we will concern ourselves now with steps 3 and 4.

10.6 The atmosphere

After the basic preprocessing and aperture photometry, the ground-based photometrist may wish to remove, or at least minimize, the effects of absorption by the terrestrial atmosphere. Under some circumstances, this is an impossible task. In other cases, astronomers can be confident that their inside-the-atmosphere measurements will yield excellent estimates of the outside-the-atmosphere values.

10.6.1 Absorption by a plane-parallel slab

Figure 10.12a shows a stream of photons traversing a horizontal slab of absorbing material (air, for example). The photons travel at angle z with respect to the vertical. We assume that the density and absorbing properties of the material change with h, the depth in the material, but are independent of the other coordinates. We assume

that if a flux of $\phi'(\lambda, h)$ travels over a path of length ds, the material will absorb a certain fraction of the photons. We write this absorbed fraction as

$$\frac{d\phi'(\lambda,h)}{\phi'(\lambda,h)} = -\alpha(\lambda,h)ds = -\sec(z)\alpha(\lambda,h)dh \tag{10.8}$$

where we introduce the function $\alpha(\lambda, h)$ to describe the absorption per unit distance. We can apply this result to the Earth's atmosphere by identifying z as the zenith angle of the source. However, the geometric and optical properties of the real, spherical atmosphere mean that the value of the ratio $d\phi'/\phi'$ is more complicated than the extreme right-hand side of Equation (10.8) suggests. There are two effects involved. First, as you can see from Figures 10.12b and c, because the atmosphere has spherical rather than plane symmetry, the angle z is not a constant, but is an increasing function of h. Second, the actual angle at any height will be even greater than that given by the spherical model because of atmospheric refraction. Taking both effects into account and assuming we have an observatory at depth H in the atmosphere, the solution to Equation (10.7) is

$$\phi_\Lambda(\lambda) = \phi(\lambda)e^{-\int_o^H \sec(z(h))\alpha(h)dh} = \phi(\lambda)e^{-\tau(\lambda,H)X} \tag{10.9}$$

Here $\phi_A(\lambda)$ and $\phi(\lambda)$ are the monochromatic photon fluxes inside and outside the atmosphere, respectively. We introduce two new functions on the right-hand side of Equation (10.9). First, the ***optical depth at the zenith***:

$$\tau(\lambda,H) = \int_0^H \alpha(h)dh$$

Physically, this definition implies that the monochromatic brightness at the zenith changes by the factor $S_{atm} = \exp(-\tau)$ due to absorption. Second, we introduce the concept of the ***air mass***:

$$X(\lambda,z) = \frac{1}{\tau(\lambda)}\int_0^H \sec(z(h))\alpha(h)dh \approx X(z) \approx \sec(z)$$

The air mass along a particular line of sight is a dimensionless quantity. It tells how much more absorbing material lies along that line than lies towards the zenith. The approximation $X = \sec(z(H))$ is good for small zenith angles. (The error is less than 1% for $z < 70°$, corresponding to an air mass of less than 3.) For larger zenith distances, the formula

$$X(z') = \sec(z')\left[1 - 0.0012\left(\sec^2 z' - 1\right)\right]$$

is a much better approximation. Here z' is the "true" zenith angle – the angle, $z(h = 0)$, between the observer's vertical and the optical path outside the atmosphere – which can be computed from the object coordinates and the sidereal time.

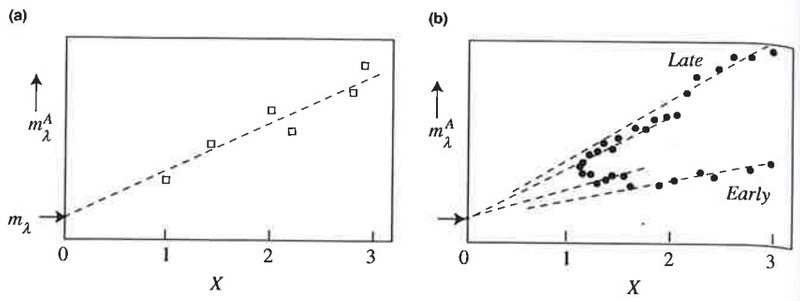


Fig. 10.13 Bouguer's law. (a) A linear fit to extinction star observations gives a measure of the extinction coefficient (slope) and the outside-the-atmosphere instrumental magnitude (intercept). (b) Multiple observations of an extinction star during a time in which the extinction coefficient is increasing.

10.6.2 Bouguer's law

From Equation (10.9), we can represent the monochromatic magnitude on the instrumental scale as

$$m_\lambda^\Lambda = -2.5\log\left[\frac{hc}{\lambda}\phi_\Lambda(\lambda)\right] = -2.5\log\left[\frac{hc}{\lambda}\phi(\lambda)\right] + 2.5\tau(\lambda)X(z)\log(e)$$
$$m_\lambda^A = m_\lambda + 1.086\tau(\lambda)X$$

We have omitted the constant for magnitude zero-point definition; m_λ^A is the magnitude as observed inside the atmosphere, and m_λ is the magnitude in the same system outside the atmosphere. Finally, we define the ***monochromatic extinction coefficient***, $k(\lambda) = 1.086\tau(\lambda)$, and rewrite the previous equation as

$$m_\lambda^A(X) = m_\lambda + k(\lambda)X \qquad (10.10)$$

This expression, which states that the apparent magnitude is a linear function of air mass, is known as **Bouguer's[3] law** (or sometimes, Lambert's law). Bouguer's law suggests the method for determining the value of the extinction coefficient, and thus a method for converting apparent magnitudes inside the atmosphere to magnitudes outside the atmosphere. The astronomer simply measures the brightness of some steady source (the ***extinction source***) at at least two different air masses – then, in a plot of magnitude as a function of air mass, Equation (10.10) tells us that the slope of the straight-line fit is $k(\lambda)$ and the y-intercept is m_λ; see Figure 10.13. Once she knows $k(\lambda)$, the astronomer can compute outside-the-atmosphere magnitudes for any other stars by making a

[3] Pierre Bouguer (1698–1758), a French Academician, was celebrated in his day for leading an expedition to Peru in 1735 to measure the length of a degree of latitude. The expedition conclusively demonstrated Newton's hypothesis that the Earth was oblate. Bouguer derived his law for atmospheric absorption by investigation of the general problem of light transmission through a medium. He also holds the distinction of being the first quantitative photometrist in astronomy – in 1725 he measured the relative brightnesses of the Sun and Moon by comparison to a candle flame.

single observation and applying Bouguer's law. With one powerful and elegant stroke, the astronomer has removed the absorbing effects of the atmosphere.

The power and elegance of Bouguer's law depends on the persistence of two conditions during the time over which observations are made:

(1) that $k(\lambda)$ is stationary – does not change over time,
(2) that $k(\lambda)$ is isotropic – does not change with location in the sky.

If both these conditions hold, observers will say that the atmosphere is ***photometric*** and feel powerful and capable of elegance. If the conditions are violated (visible clouds are one good indication), observers will recognize that certain kinds of photometry are simply impossible. There are intermediate situations – Figure 10.13b shows observations in which condition (1) is violated – extinction here changes gradually over several hours. As long as the changes are carefully monitored, the astronomer can still hope to recover the outside-the-atmosphere data via Bouguer's law.

Condition (2) is always violated because of the spherical nature of the atmosphere: absorption by lower layers becomes relatively more important at large zenith angles (Figure 10.12b), and total extinction as well the extinction versus wavelength function will change. This effect is not significant at smaller (< 3) air masses, so usually can be (and is) ignored. In general, it is a good idea to avoid *any* observations at very large air masses – the likelihood of encountering non-uniformities is greatly increased, as are all other atmospheric effects like seeing distortions, differential refraction, and background brightness levels.

10.6.3 Sources of extinction

Figure 10.14 plots $k(\lambda)$ for a typical clear (cloud-free) sky in the 0.3–1.4 μm region. As illustrated, the value is the sum of contributions from four different

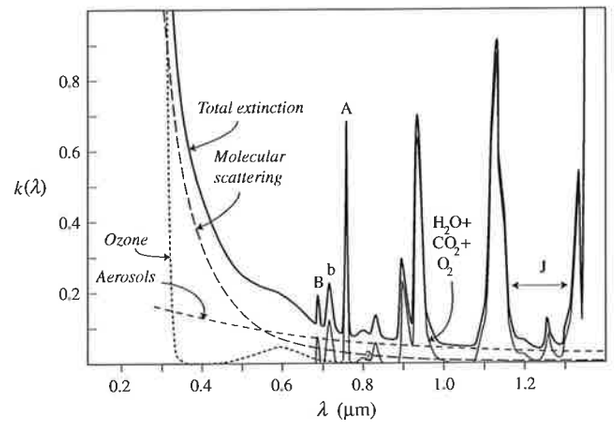


Fig. 10.14 A model for the contributions to the extinction coefficient. Aerosol and water-vapor absorption are highly variable. This is a low-resolution plot, so band structure is smoothed. Letters mark strong telluric Fraunhoffer features and the photometric J-band window.

processes, each of which has a characteristic spectral dependence. The processes are:

Rayleigh scattering by molecules. In this process, a photon encounters a molecule of air and is redirected. The probability of scattering is much greater for short-wavelength photons (for pure Rayleigh scattering, extinction is proportional to λ^{-4}). Molecular scattering explains why the sky is blue, since multiply scattered photons from the Sun will tend to be those of shortest wavelength. Molecular scattering is a component of the extinction that is stable over time, and its magnitude scales directly with the atmospheric pressure – higher altitudes will have more transparent skies.

Absorption by ozone. Continuous absorption by the O_3 molecule in the ultraviolet essentially cuts off transmission shortward of 320 nm. Ozone also absorbs less strongly in the Chappuis bands in the visible near 600 nm. Atmospheric ozone is concentrated near the stratopause at around 48 km above sea level, so the benefit of high-altitude observatories is limited with respect to ozone extinction. Ozone abundance is subject to seasonal and global variations, but does not appear to vary on short time scales.

Scattering by aerosols. Aerosols are suspensions of small solid or liquid particles (particulates) in air. Particulates range in diameter from perhaps 50 μm down to molecular size. Aerosol particulates differ from water cloud drops by their much longer natural residence time in the atmosphere. In fact, the way most aerosols are removed is by reaction with water droplets in clouds[4] and subsequent precipitation. Several different processes inject particulates into the atmosphere. Sea spray and bursting bubbles introduce salt. Winds over deserts introduce dust. Volcanoes inject ash and sulfur dioxide (a gas that interacts with water vapor to form drops of sulfuric acid). Burning fossil fuel and biomass introduce ash, soot, smoke, and more sulfur dioxide. The wavelength dependence of aerosol scattering depends largely on the size of the particle, and the typical wide range of sizes present (salt particles tend to be large; smoke particles, small) usually produces a relatively "gray" extinction (a λ^{-1} dependence is typical). A pale-blue (rather than deep-blue) sky indicates high aerosol extinction. Sometimes aerosols can produce striking color effects, including the lurid twilight colors from stratospheric volcanic ash and the "green sky" phenomenon due to Gobi Desert dust. Aerosol scattering can be quite variable, even on a short time scale, and different components reside at different atmospheric levels. Although salt, dust, and industrial pollution mainly stay in the lower layers (a scale height of 1.5 kilometers is representative), some volcanic eruptions and intense forest fires can inject aerosols into the stratosphere, where they may persist for weeks or even years, since there is no rain in the stratosphere.

Molecular-band absorption. The main molecular *absorbers* are water vapor and carbon dioxide, although oxygen has a few relatively narrow features, and we have already discussed the ozone bands. Water vapor and CO_2 bands demarcate the relatively transparent windows in the near and middle infrared. Carbon dioxide is well mixed with altitude in the atmosphere, but water vapor is concentrated near the surface and varies with temperature, time, and location. At sea level, the amount of vapor in one air mass corresponds to about 10 mm of liquid, on average. On Mauna Kea, one of the best conventional sites, the average is about 1 mm of precipitable water. At the south pole, which benefits from both high elevation and low temperature, values approach 0.15 mm. Stratospheric observatories carried by balloons or aircraft enjoy even lower values.

[4] Aerosol particles are crucial to the *formation* of water clouds – water vapor condenses into droplets much more readily if aerosols provide the "seed" surfaces on which condensation can proceed. Without such seeds, very clean air can supersaturate and reach a relative humidity well over 100%.

10.6.4 Measuring monochromatic extinction

Ground-based astronomers use a variety of methods for removing the effects of atmospheric extinction. We look at a few cases here, and then examine the complications introduced by heterochromatic (broadband) systems. To keep things simple at the start, we assume that we are observing a monochromatic magnitude and that our instrument is sensitive to exactly the same wavelength as some standard system.

Case 1: assume a mean extinction

At remote high-altitude sites, the extinction at shorter wavelengths is usually due almost entirely to Rayleigh scattering, and is therefore stable. Under these conditions, it may be safe simply to use the average (or better still, the median) extinction coefficient determined by other observers for the same site over the past years. This is a particularly reasonable approach if one is doing differential photometry, or if standard and program stars are observed at nearly the same air mass at nearly the same time.

Case 2: use known outside-the-atmosphere magnitudes

If you wish to determine the extinction yourself from the Bouguer law, this is extremely simple method applies – *if* you happen to know m_λ, the magnitude outside the atmosphere of some constant star. In this case, just point your telescope at the star, note the air mass, X, and take an image. You measure $m_\lambda^A(X)$, you know m_λ, so just solve Equation (10.10) for $k(\lambda)$. The method becomes more robust if you know values of m_λ for more than one star.

Case 3: draw the Bouguer line from observations

If you don't know m_λ, then you need to do a little more work and generate the Bouguer line. You take two or more exposures of the same field of stars over a wide range of air masses. This, of course, requires waiting for the zenith distance of your field to change. Many observers record the extinction field every 90 minutes or so. You plot the resulting instrumental magnitudes as a function of air mass, and if the night is ***photometric,*** you obtain a plot like Figure 10.13a.

Choose your extinction field carefully. With a good choice and a wide-field CCD, you may well have several stars (let's say 30) that yield good SNRs on a short exposure. Your data then will have greater statistical significance than a plot for just a single star in determining $k(\lambda)$. To do the computation of the Bouguer line, you might combine all 30 fluxes at each air mass and fit *one* straight line. You will often find it economical to make the extinction field identical to the program field or to a field containing standard stars.

Case 4: variable extinction and multi-night data

What if the extinction changes? If you observe over many consecutive nights, change is likely. If you are fortunate, the extinction will change slowly over time and uniformly over the entire sky. In this case, some modification of the previous method will yield $k(\lambda, t)$. For example, if you are sure of the outside-the-atmosphere magnitudes of some of your constant stars, then simply monitoring the constant stars will give instantaneous values of $k(\lambda, t)$.

If you do not have the instrumental outside-the-atmosphere magnitudes, there is still hope: if you have the *standard* magnitudes of two constant stars, observing them at different air masses measures $k(\lambda, t)$. The usual practice is to take one frame containing a standard star near the meridian (the "D" frame), and then immediately take a frame containing the second standard at large air mass in the east (the "M" frame). Provided that the extinction coefficient is the same in all directions, Bouguer's law gives the difference in the instrumental magnitudes of M and D, measured inside the atmosphere, as

$$\Delta m^{\Lambda}_{\mathrm{MD}} = m^{\Lambda}_{M\lambda}(X_M) - m^{\Lambda}_{D\lambda}(X_{\mathrm{D}}) = m_{M\lambda} - m_{D\lambda} + k(\lambda, t)(X_{\mathrm{M}} - X_{\mathrm{D}})$$
$$\Delta m^{\Lambda}_{\mathrm{MD}} = \Delta m_{\mathrm{MD}} + k(\lambda, t)\Delta X_{\mathrm{MD}}$$

For monochromatic magnitudes, the magnitude difference in the instrumental system should be equal to the difference in the standard system, i.e. $\Delta m^{\mathrm{STD}}_{\mathrm{MD}} = m_{M\lambda} - m_{D\lambda}$, so that we can write

$$k(\lambda, t) = \frac{\Delta m^{\Lambda}_{\mathrm{MD}} - \Delta m^{\mathrm{STD}}_{\mathrm{MD}}}{\Delta X_{\mathrm{MD}}}$$

All the quantities on the right-hand side are either known or measured. Because of the possibility of extinction change, many photometrists adopt the strategy of observing "MD" pairs every few hours through the night.

Case 5: use all the data

The most general methods make use of all available information, and include data for all nights in which sources are observed in common. Every frame taken during the run of several nights is affected by extinction and therefore contains information *about* extinction. One approach might work as follows: derive values for $k(\lambda)$ from the best nights – those for which it is possible to make good linear fits to the extinction data – and compute the outside-the-atmosphere magnitudes for *every* constant star (not just those used for the fits). You then should have a large set of extra-atmospheric magnitudes that you can use to find the extinction as a function of time for the more marginal nights. Cloudy nights in which the extinction changes rapidly will be suitable only for differential work. The extinction problem is well suited to a least-squares solution with constraints imposed by standard stars. See the discussion in chapter 10 of Sterken and Manfroid (1992) for a good introduction.

10.6.5 Heterochromatic extinction

The previous discussion strictly applies only for monochromatic magnitudes. For bandpass magnitudes, we must rewrite Equation (10.10) as

$$k_{\mathrm{P}}X = m^{\Lambda}_{\mathrm{P}} - m_{\mathrm{P}} = -2.5\log\left\{\frac{\int T_{\mathrm{P}} f_\lambda S_{\mathrm{atm}}\,\mathrm{d}\lambda}{\int T_{\mathrm{P}} f_\lambda\,\mathrm{d}\lambda}\right\}$$

or

$$k_{\mathrm{P}}X = -2.5\log\left\{\frac{\int T_{\mathrm{P}} f_\lambda \exp\left[\frac{-1}{1.086}k(\lambda)X\right]\mathrm{d}\lambda}{\int T_{\mathrm{P}} f_\lambda\,\mathrm{d}\lambda}\right\} \qquad (10.11)$$

where k_{P} is the extinction coefficient for band P. In general, the function on the right-hand side of Equation (10.11) is *not* linear in X, and will depend strongly on the shape of the function f_λ. We should expect, therefore, that it must be the case that $k_{\mathrm{P}} = k_{\mathrm{P}}(X, SpT)$, where the variable SpT indicates the spectral shape of the object observed. For most narrow-band and some intermediate photometry, the variation of k_{P} with X and SpT is so small that it can safely be ignored. For wider bands, however, this is not the case.

Think of the variation in k_{P} as due to two different but related effects. First, as we had seen earlier, the effective and isophotal wavelengths of a bandpass depend on the spectrum of the source. This effect is present even in outside-the-atmosphere photometry. We should expect that the extinction measured, say, for a red star will differ from the extinction measured for a blue star, since the center of the bandpass is different for the two. Because of this effect, the Bouguer plot of apparent magnitude versus air mass will give straight lines of different slopes for stars of different spectral shapes. The second, more

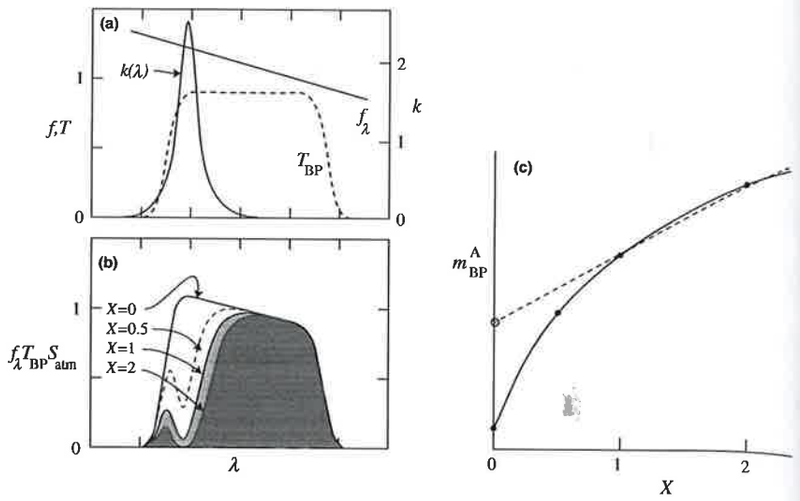


Fig. 10.15 The Forbes effect. (a) A model source spectrum, bandpass transmission T_{BP}), and extinction coefficient. The extinction is due entirely to a strong feature near the blue edge of the band. (b) The flux actually detected at the telescope is shown as a function of wavelength for four different air masses. Note the relatively small change between $X = 1$ and $X = 2$ (shaded regions) compared to the change between $X = 1$ and $X = 0$. (c) The Bouguer diagram for the data in (b), illustrating the non-linear relationship and the difference between the actual extra-atmospheric magnitude (filled circle) and the intercept of a linear fit to observable data (open circle).

invidious problem arises because atmospheric extinction itself changes the shape of the spectrum that reaches the telescope. This effect – called the ***Forbes effect*** – means you actually observe different spectra for the same star at different air masses. (Alternatively, you can think of the effect as changing the shape of the bandpass response as air mass changes.) A Bouguer plot of apparent magnitude versus air mass will therefore give a *curved* line. The Forbes effect, as illustrated in Figure 10.15, is particularly problematic if strong atmospheric absorption affects some parts of the photometric band more than others. This is the case for Johnson U and many of the wider infrared bands (review Figures 10.1 and 10.3). In such cases, the magnitude change in going from $X = 2$ to $X = 1$ can be considerably less than in going from $X = 1$ to $X = 0$. In some cases, both the width of the band and its effective wavelength can change dramatically at the smaller air masses. Use of outside-the-atmosphere magnitudes when the Forbes effect is present is a little tricky, and sometimes depends on having a good model of the response function, the atmosphere, and the unknown source. For precise work, therefore, it is best to use bands that exclude strong atmospheric absorptions (e.g. the MKO near-infrared system).

10.6.6 Second-order extinction coefficients

A solution for Equation (10.11) would be possible if we could make a good approximation of both the monochromatic extinction function $k(\lambda)$ and the shape of the spectrum f_λ, either numerically, or perhaps with a Taylor series expansion of each function. The required functions or their derivatives are rarely known, so what is usually done is to assume that the photometric color, which

gives some sense of the shape of the source spectrum, will account for most variations in the broadband extinction, and write

$$k_P = k'_P + k''_P \cdot (ci) \tag{10.12}$$

where (ci) represents some "appropriate" photometric index like $(B-V)$, and k''_P is called the ***second-order extinction coefficient***. For example, the coefficient for the Johnson V band might be written as

$$k_V = k'_V + k''_{V,BV} \cdot (B - V)$$

The color in this case is $(B-V)$, but might also have been chosen to be (V – R) or $(V-I)$. It is also quite common to use the instrumental colors instead of the standard colors, since instrumental colors will be available for many more extinction stars. Instrumental magnitudes are usually written in lower case for the Johnson–Cousins system, so the instrumental color in the above equation, for example, would be $(b-v)$. For broadband extinction, therefore, one has for the ith observation of star j

$$m^A_{p,i,j} = m_{p,j} + \left(k'_p + k''_p \cdot (ci)\right) X_{i,j} \tag{10.13}$$

Various approaches can then lead from the data (the $(X_{i,j}, m^A_{p,i,j})$ pairs) to values for the outside-the-atmosphere instrumental magnitudes, $m_{P,j}$, and for the two extinction coefficients, k'_p and k''_p. The second-order coefficient is difficult to determine from a few nights' observations. Fortunately, this coefficient should not vary much if the primary sources of extinction do not include absorption bands due to water vapor. It is common practice, therefore, to use mean values – established over time for a particular instrument and site – for the second-order coefficients.

All the cases discussed in the previous section on monochromatic data can be applied to correcting heterochromatic data for extinction, with the understanding that a second-order extinction coefficient may need to be determined. Also, understand that if you use an outside-the-atmosphere color index in Equation (10.13), you have not accounted for the Forbes effect at all. If the Forbes effect is severe, the form of Equation (10.13) is probably inadequate. You may then need to fit the extinction data with a function of X whose form has been derived from a model (similar to Figure 10.15) of how the flux in the bandpass changes with air mass. This requires detailed knowledge of both the bandpass shape and absorption behavior, as well as the spectrum of the source.

10.6.7 Indices or magnitudes?

The traditional method for reporting n-color photometric data is to give one magnitude and $n-1$ indices. It has also been traditional to make the extinction computations not for n magnitudes, but for one magnitude and $n-1$ indices.

Suppose, for example, we observe in the Johnson B and V bands. If we have a B frame and a V frame taken at about the same air mass, we can write Equation (10.13) once for each band, and then subtract the equations, yielding

$$(b-v)^{\Lambda} = (b-v) + \left\{\left(k'_B - k'\right)_V + \left(k''_{\mathrm{B,BV}} - k''_{V,\mathrm{BV}}\right)\cdot(B-V)\right\}X$$

Combining the coefficients, we have

$$(b-v)^{\Lambda} = (b-v) + \left\{k'_{\mathrm{B-V}} + k''_{\mathrm{B-V}}\cdot(B-V)\right\}X \qquad (10.14)$$

Here the new first- and second-order coefficients describe the effects of extinction *on the index*. There is an objective reason for analyzing extinction data via Equation (10.14) rather than via Equation (10.13): if either instrument sensitivity or atmospheric aerosol extinction drifts during an observing run, the effect on the observed color indices will be minor compared to the effect on the individual magnitudes.

This reasoning is less compelling in modern observing. For one thing, CCDs and infrared arrays are much less prone to sensitivity drift than photomultipliers. Moreover, the requirement that all instrumental bands be observed at the same air mass is somewhat restrictive, and may prevent the use of all the available extinction data.

10.7 Transformation to a standard system

Assume we have observed an outside-the-atmosphere instrumental magnitude for some source, either by correcting a ground-based observation for extinction or by direct observation with a space telescope. If we wish to compare our result with those of other observers, we must all use the same photometric system. Almost always, this means executing the next step in the photometric reduction procedure, the transformation from the instrumental to the standard system. This ***transformation*** will depend on (1) differences between the response of our instrumental system and that of the standard system and (2) the shape of the spectrum of the source.

10.7.1 The monochromatic case

Suppose we observe with indefinitely narrow-band filters whose central wavelengths, λ_1 and λ_2, may be slightly different from the standard wavelengths, λ_{S1} and λ_{S2}. The difference between the first standard magnitude and our instrumental result is

$$m^{\mathrm{STD}}_{\lambda_{S1}} - m_{\lambda_1} = -2.5[\log f_{\lambda_{S1}} - \log f_{\lambda_1}] + C_1^{\mathrm{STD}} - C_1 \qquad (10.15)$$

The last two terms in Equation (10.15) quantify the difference between the efficiency of our telescope/photometer and the standard instrument. These will be the same for every source observed. The first term on the right, in brackets, you will recognize as something like a color index. This prompts us to rewrite (10.15) as:

$$m^{\mathrm{STD}}_{\lambda_{S1}} \approx m_{\lambda_1} + \alpha_{12}(m_{\lambda_1} - m_{\lambda_2}) + \alpha_1 = m_{\lambda_1} + \alpha_{12}(ci)_{12} + \alpha_1 \qquad (10.16)$$

In this equation, α_{12} is called the ***color coefficient*** for the transformation and α_1 is called the ***zero-point constant***. Since indices are more closely related to physical variables than are individual magnitudes, in an n-band system, astronomers traditionally work with one magnitude transformation and $n-1$ index transforms, of the form

$$(m_1 - m_2)^{\mathrm{STD}} \approx m_{\lambda_1} - m_{\lambda_2} + (\alpha_{12} - \alpha_{21})(m_{\lambda_1} - m_{\lambda_2}) + \alpha_1 - \alpha_2$$

Redefining constants and simplifying the notation slightly:

$$(ci)^{\mathrm{STD}}_{12} \approx \gamma_{12}(ci)_{12} + \beta_{12} \qquad (10.17)$$

We can justify Equations (10.16) and (10.17) if the first term of Equation (10.15) is a linear function of some color index of the instrumental system (or of the standard system). From Figure 10.16, you can see that Equation (10.16) is valid only if $(ci)_{12}$ measures the slope of the logarithmic spectrum *between* λ_1 *and the standard wavelength*, λ_{S1}. Unfortunately, at the resolution of narrow-band photometry, many spectra have almost discontinuous features like lines. Many narrow-band systems may have indices that give a good value for the local derivative at each of the standard wavelengths. Thus, the relation in Equation (10.16) is rarely exact.

In practice, astronomers almost always assume a transformation equation like (10.16) or (10.17) but will use empirical methods to establish the color coefficient and zero point, or even the functional form for the transformation itself. The function might include a higher-order color term, or use some other photometric criterion to improve the accuracy of the transformation.

The general approach is simple: you measure m_{λ_1} and $(ci)_{12}$ for a range of standards whose spectra you expect to resemble those of the unknown sources under investigation. Your plot of $(m^{\mathrm{STD}}_{\lambda_1} - m_{\lambda_1})$ as a function of $(ci)_{12}$, as in Figure 10.17, then establishes a relation like Equation (10.16). In the example in the figure, the data invite a quadratic fit, so you would add a term with a ***second-order color coefficient*** to the transformation equation. Another reasonable empirical fit might use two different linear relations, one for lower and another for higher values of the color index. If a grid of standards establishes the transformation, then the transformation can introduce errors if the spectrum of an unknown source is sufficiently different from any element of the grid.

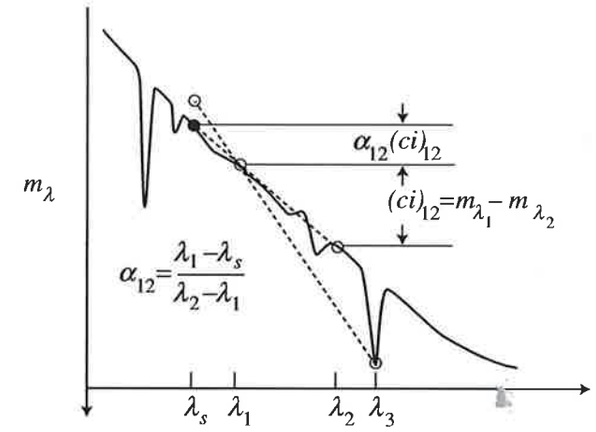


Fig. 10.16 Color transformation. Narrow-band measurements at wavelengths 1 and 2 can predict the monochromatic magnitude at the standard wavelength (filled circle at wavelength *S*) if the color index measures the slope between wavelengths *S* and 1. If the (*ci*) does not measure the slope – as in wavelengths 1 and 3, an error results.

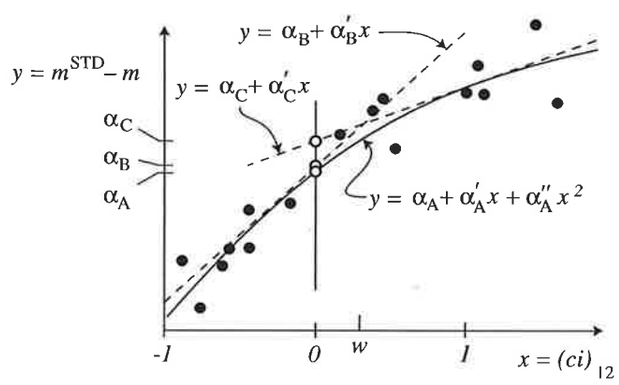


Fig. 10.17 Empirical determination of the transform coefficients from a plot of standard minus instrumental magnitudes as a function of color index. This data could be fitted with a quadratic function (A) or with two different linear functions: (B) for $ci < w$ and (C) for $ci > w$.

10.7.2 The heterochromatic case

The problem of transformation to a broadband standard system is similar to the monochromatic case: we just replace the fluxes in Equation (10.15) with the appropriate integrals:

$$m_P^{STD} - m_P = -2.5\left[\log \int T_{P,STD} f_\lambda d\lambda - \log \int T_P f_\lambda d\lambda\right] + C_P^{STD} - C_P$$

Except for the zero-point terms, the transformation to the standard system will reflect differences between the instrument response of your system, T_P, and the instrument response of the standard system, $T_{P,STD}$. In the broadband case, differences can be due to any differences in the detailed shape of the response function, especially those resulting in different central wavelengths and bandwidths. Because a broadband system examines a smoothed version of the spectrum, and because the bands tend to be close together relative to the scale of the smoothing, the approximation of Equation (10.16) will tend to be better in broadband than in most narrow-band systems.

Procedures are similar to those discussed for the monochromatic case: you plot, as a function of color, the difference between instrumental and standard magnitudes for a set of standards that spans the range of spectra you expect in the unknowns. If the functions T_P and $T_{P,STD}$ are identical, you will find a horizontal line at the level of the zero-point correction. If the functions differ (the usual case!), you fit the data to find the color term or terms in the transformation. The fit may be a straight line, or may require higher-order terms, and will apply safely only to those kinds of spectra sampled by the standard stars you have observed.

Summary

- The history of photometry has imposed the magnitude scale and the definition of several important broadband photometric systems.
- Photometric bandpass response functions are generally categorized as broad-, intermediate-, or narrow-band. A response can be implemented by filters, detector sensitivity, atmospheric transmission, or some combination of these. Concepts:

resolving power	*response function*	*photon response function*
high-pass filter	*peak wavelength*	*central wavelength*
mean wavelength	*effective wavelength*	*isophotal wavelength*
FWHM	*bandwidth*	*photon flux*
zero point		*bandpass magnitude*

- Photometric indices, which are linear combinations of bandpass magnitudes, quantify characteristics of an object's spectrum. Concepts:

color index	*blackbody spectrum*	*monochromatic magnitude*
line index	*curvature index*	*feature index*

- A standard photometric system specifies both the response functions of its bands as well as some method for standardizing measurements. Concepts:

open system	*closed system*	*instrumental system*
visual magnitude	*standard sequence*	*north polar sequence*
ci	*international system*	*photovisual magnitude*
UBVRI	*Cousins system*	*JHKLMNQ*
MKO filters	*Strömgren system*	*uvbyβ*
c_1 and m_1		*SDSS system*

- A shift in an object's spectrum caused by the Doppler effect or cosmological expansion will produce a photometric change that can be corrected if an astronomer has sufficient information (the *K* correction).

(*continued*)

Summary (*cont.*)

- Conversely, the color change observed in a very distant object can lead to an estimate of its redshift.
- Absorption by material outside the atmosphere can produce both reddening and absorption lines and bands in a spectrum.
- Absorption by material inside the atmosphere can produce both reddening and telluric absorption lines and bands. Concepts:

  optical–infrared window *microwave–radio window*
  instrumental magnitude *magnitude outside the atmosphere*

- Photometric data reduction proceeds in steps: (1) preprocessing, (2) digital photometry, (3) atmospheric extinction correction, (4) transformation to a standard system, and (5) further corrections and analysis.
- Bouguer's law is the basis for the correction for atmospheric extinction:

$$m_\lambda^\Lambda(X) = m_\lambda + k(\lambda)X$$

  Concepts:

  optical depth *air mass* *extinction coefficient*
  ozone bands *molecular bands* *Rayleigh scattering*
  aerosols *Bouguer line* *mean extinction*
  second-order extinction *monochromatic extinction* *heterochromatic extinction*
  Forbes effect

- Transformation to the standard system requires observation of standard objects using instruments identical to those used for the unknowns. Concepts:

  zero-point constant *standard star/extinction star*
  color coefficient *second-order color coefficient*

Exercises

1. Show that for a response function with a boxcar or triangular profile, the bandwidth = FWHM, but that for a Gaussian, the bandwidth < FWHM.
2. The table below gives the response function for a photometric bandpass, as well as the flux distributions for two sources. Characterize this system by computing (use a spreadsheet) all of the following:
   wavelength at peak transmission
   the FWHM
   bandwidth
   mean wavelength
   effective wavelength for each source
   isophotal wavelength for each source

λ (nm)	$R_{BP}(\lambda)$	f_λ^Λ	f_λ^B
500	0	1.70	0.37
505	0.04	1.56	0.47
510	0.24	1.43	0.57
515	0.41	1.31	0.67
520	0.5	1.20	0.78
525	0.55	1.10	0.89
530	0.64	1.00	1.00
535	0.77	0.92	1.12
540	0.88	0.84	1.24
545	0.96	0.77	1.37
550	0.99	0.70	1.50
555	1	0.64	1.64
560	0.81	0.57	1.78
565	0.5	0.52	1.92
570	0	0.46	2.07

3. A photometer on a spacecraft employs a grating-and-slot arrangement (see Figure 10.1) such that all radiation with wavelength between 2.0 and 4.0 μm is detected.
   (a) Assume the detector is a perfect bolometer, so that 50% of the energy between 1.0 and 3.0 μm is detected, independent of wavelength. In other words, $R_{BP}(\lambda)$ is a "boxcar" with mean wavelength 3.0 μm and bandwidth 2.0 μm. Compute the effective wavelength of this band for a hot star with $f_\lambda = A\lambda^{-4}$.
   (b) Now assume the detector is replaced with an infrared photon detector with uniform quantum efficiency such that 50% of the incident photons at each wavelength in the band are detected. Again, compute the effective wavelength of this band for a hot star with $f_\lambda = A\lambda^{-4}$. Note that you will need to devise an expression for the energy response function of the system.
4. Show that Equations (10.3) and (10.4) follow from the Rayleigh–Jeans and Wien approximations to the Planck law.
5. Gabriel very carefully constructs a filter for his CCD photometer so that the response function matches the standard bandpass of the Johnson V color very precisely. He observes two very well-established standard stars whose catalog data are given below. Gabriel discovers that with his CCD, no matter how carefully he observes, he always finds one star is brighter than the other: its image always contains more total analog-to-digital units (ADUs) on the CCD. Liz suggests to him that this is

because the CCD is a photon-counting device. (a) Explain her reasoning. (b) If Liz is correct, which star should be the brighter on the CCD and why?

	V	$B-V$
Star 1	9.874	0.058
Star 2	9.874	0.861

6. An astronomy student obtains two images of a galaxy, one in the B band, the other in the V band. Outline the image arithmetic operations the student would execute in order to produce a map of the $(B-V)$ color index for the galaxy. Failure to subtract the constant background sky for each image would cause problems in the map. For which parts of the map would these problems be most serious? On the other hand, would subtracting the background sky introduce any problems in the map? If so, which parts, and why?
7. Investigate the website for the Sloan Digital Sky Survey. In what ways is the SDSS five-color system superior to the UBVRI system?
8. An MOS capacitor observes two sources in the band 400–600 nm. Source A has a spectrum such that the distribution of photons in the 400–600 nm band is given by $n_A(\lambda) = A\lambda^3$. Source B has a distribution of photons given by $n_B(\lambda) = B\lambda^{-2}$ in the same band. If the two sources generate photoelectrons at exactly the same rate, compute their brightness ratio. You may assume the detector's quantum efficiency is not a function of wavelength.
9. Speculate, in terms of the Forbes effect, why it might be useful to define the standard magnitude as one measured at 1 air mass, rather than at zero air mass. What difficulties might be inherent in this choice?
10. An observer uses the B and V filters to obtain four exposures of the same field at different air masses: two B exposures at air masses 1.05 and 2.13, and two V exposures at air masses 1.10 and 2.48. Four stars in this field are photometric standards. Their standard magnitudes are given in the table below, as are the instrumental magnitudes in each frame.

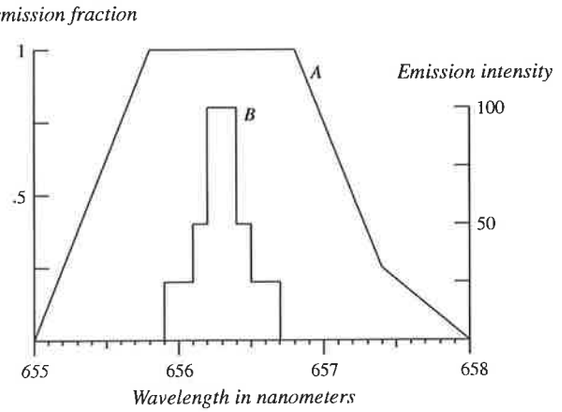
	$(B-V)$	V	$b(1)$	$b(2)$	$v(1)$	$v(2)$
Air mass			1.05	2.13	1.10	2.48
Star A	–0.07	12.01	9.853	10.687	8.778	9.427
Star B	0.36	12.44	10.693	11.479	9.160	9.739
Star C	0.69	12.19	10.759	11.462	8.873	9.425
Star D	1.15	12.89	11.898	12.547	9.522	10.002

(a) Compute the extinction coefficients for the instrumental system: k'_b, k''_b, k'_v, and k''_v. Hint: at each air mass, Equation (10.10) holds. Write an equation for the difference between the magnitudes at the two air masses (e.g. an equation for $b(2) - b(1)$). Examination of this equation will suggest a method for computing the coefficients. You may find it helpful to enter the data from the table into a spreadsheet in performing the computations.

(b) Compute the instrumental magnitudes of each star at zero air mass

(c) Compute the transformation coefficients, $\alpha_V, \alpha_{B-V}, \beta_{B-V}$, and γ_{B-V}, using the method outlined in Section 10.7.

11. A photometric bandpass whose response function is shown in curve A below measures the strength of the emission feature shown in curve B. In the figure, the source has zero velocity. Compute the change in the brightness measurement, in magnitudes, that would result if the source were given a radial velocity of 300 km s^{-1}.



To Measure the Sky
An Introduction to Observational Astronomy

Frederick R. Chromey
Vassar College

CAMBRIDGE
UNIVERSITY PRESS