4 Radiometry and Photometry: What you Get and What you See

What the eye or photographic film or instruments such as photomultiplier tubes *get* is incident radiant energy over a time interval (radiant power). But film and instruments cannot be said to *see* anything. Only the eye can see, and seeing is mostly the processing of radiant energy by the organs of seeing – retina, optic nerve, and, most important, the brain – not the simple formation of images on the retina. It is sometimes said that the eye is just like a camera. It would be more accurate to say that the eye is *not* just like a camera except in the most trivial sense, namely, they both contain arrangements of lenses to form images. Yet even the most complicated camera, a mere chunk of metal and glass, cannot begin to perform the feats routinely done with ease by the human and animal eye–brain combination. Our brains create a visual world beginning with light as the raw material.

Wavelength is not a synonym for color nor is radiant power a synonym for brightness. But we begin with these strictly physical properties of light (radiometry), then proceed to the sensations it produces in the human observer (photometry).

4.1 The General Radiation Field

Light is a superposition of electromagnetic waves, intertwined electric fields $\mathbf{E}$ and magnetic fields $\mathbf{H}$. Because these fields are vectors, so are electromagnetic waves. They satisfy vector wave equations similar to the scalar wave equation derived in Section 3.3 for the vibrating string. We usually are most interested in the rate at which radiant energy is transported by electromagnetic waves. The electric and magnetic fields determine this transport rate by way of the *Poynting vector*

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}, \tag{4.1}$$

where $\mathbf{E}$ and $\mathbf{H}$ are the fields at a point. If this point is on a surface (real or mathematical), the rate at which radiant energy is transported across unit area of that surface is

$$\mathbf{S} \cdot \mathbf{n}, \tag{4.2}$$

where $\mathbf{n}$ is a unit vector normal to the surface. The derivation of Eq. (4.1), boiled down to its essence, is fundamentally no different from, although more complicated than, the derivation in Section 3.3 of the energy flux vector for a vibrating string. We determined the time rate of change of kinetic and potential energy of a finite length of string, then noted that this was

Fundamentals of Atmospheric Radiation: An Introduction with 400 Problems. Craig F. Bohren and Eugene E. Clothiaux

ISBN: 3-527-40503-8

equal to the difference in energy fluxes at its end points. Similarly, Eq. (4.1) is obtained by determining the time rate of change of electric and magnetic energy within a bounded volume and noting that this is equal to the integral of the Poynting vector over the bounding surface. The energy flux vector for the string, Eq. (3.27), is the product of two functions. Similarly, the energy flux vector for the electromagnetic field, Eq. (4.1), is the (vector) product of two fields.

The scalar quantity Eq. (4.2), with the dimensions of power per unit area, is what we are after, but to get it we need two vector fields ($\mathbf{E}$ and $\mathbf{H}$) in order to determine a third ($\mathbf{S}$). This is possible in principle, but, except in very restricted circumstances, is essentially impossible in practice. We don't live in a world of simple electromagnetic waves. Just look around the room in which you are reading these words. You receive light from all directions, differing in amount and in its spectrum for each direction, a complex mosaic that changes when you move. How on earth could anyone ever sum the electric and magnetic fields originating from so many sources: walls, floor, ceiling, furniture, this book itself? Fortunately, we often don't have to determine these vector fields but can go straight to the desired scalar quantity, radiant energy transport. We can circumvent the electromagnetic field because most sources of visible and near-visible radiation are incoherent: there is no fixed phase relation between radiation from the walls and from the floor, from one part of a wall, and from another. Thus we can add the radiant energy from each source and ignore phase differences because they wash out when integrated over space or time. Another reason we can circumvent electromagnetic fields is that the wavelengths of the radiation of interest usually are much smaller than the objects with which the radiation interacts. If our radiation environment consisted of waves with wavelengths of order meters or more we might be in trouble. As we show in Section 3.4.2, the lateral coherence length of a source increases with its wavelength.

Radiometry is based on approximating electromagnetic radiation as a gas of photons. Like gas molecules, photons may be distributed in energy and in direction, and the properties of a photon gas may vary from point to point and from moment to moment. But photons do *not* interfere (see Sec. 3.4): they don't have phases. Or perhaps it would be more correct to say that phase differences wash out when we take averages over space or time, and so it is *as if* they had no phases. Ignoring phases results in an enormous simplification. Indeed, it makes radiometry (the measurement of radiation) possible. But all simplification comes with a price. We have to be alert to possible discrepancies between theory and observations. A theory in which phase is tossed out the window cannot possibly account for phenomena in which interference plays an important role.

The fundamental radiometric quantity is *radiance*, which to understand requires a thorough grasp of solid angle, which we turn to next. Solid angle is almost entirely absent from electromagnetic theory but plays a central role in radiometry.

4.1.1 Solid Angle

Any direction in a plane can be specified by a vector. Suppose that a vector $\mathbf{r}$ in the plane is rotated about a point O into another direction. The tip of the rotated vector traces out an arc of length s. The angle between these two directions, denoted by ϑ, is *defined* as the ratio s/r, and lies between 0 and 2π *radians* (the circumference of a circle of radius r is $2\pi r$). But another interpretation of angle is that it is the *measure* (size) of the set of *all* directions from O to points on the arc. Length is the measure of the set of all points lying on a line between

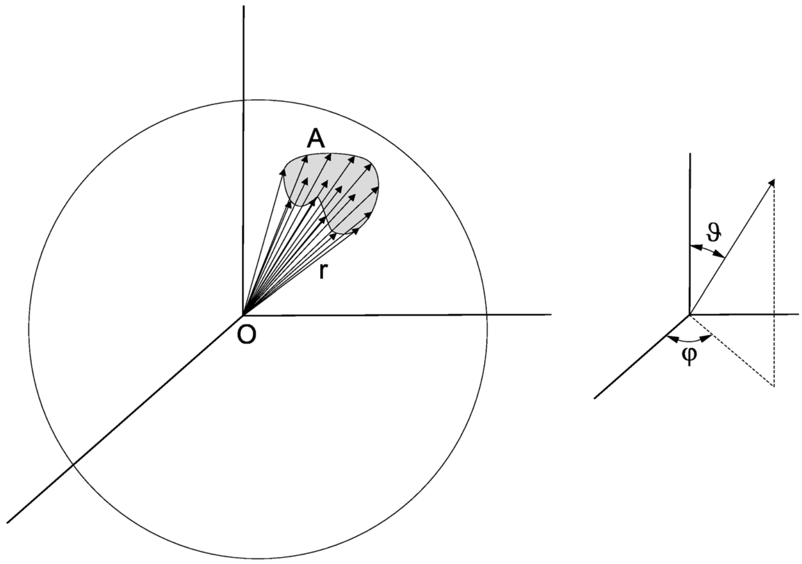


Figure 4.1: The solid angle of the infinite set of directions originating at O and ending on a surface with area A on a sphere of radius r is A/r^2.

two points. Area is the measure of the set of all points within a closed curve on a surface. And volume is the measure of the set of all points in space within a closed surface. These measures provide a way of comparing the size of one set of points with that of another. Far from being abstract, they are the means by which prices are assigned to rope, parcels of land, and gasoline (the gallon and liter are volumetric measures). Similarly, angle provides a way of comparing sets of directions in the plane. But directions are not confined to two-dimensional space. What is the measure of directions in three-dimensional space?

Consider a spherical surface of radius r on which a closed curve is inscribed (Fig. 4.1); the area of that part of the surface within this curve is A. Every vector from the origin O to a point on A specifies a direction in space. The measure of the set of all these directions is its *solid angle*

$$\Omega = \frac{A}{r^2}, \tag{4.3}$$

which lies between 0 and 4π *steradians* (abbreviated as sr). In principle, Ω can be determined by evaluating a surface integral:

$$\Omega = \frac{1}{r^2}\iint r^2 \sin\vartheta \; d\vartheta \; d\varphi = \iint \sin\vartheta \; d\vartheta \; d\varphi, \tag{4.4}$$

where ϑ and φ are spherical polar coordinates (co-latitude and azimuth, respectively) and the limits of integration are determined by A. Let's use Eq. (4.4) to determine the solid angle subtended by the sun at the surface of Earth, by which we mean the solid angle of the set of all directions from a point on Earth to the sun. We need this quantity later. The sun is azimuthally symmetric and its angular width, ϑ_{s}, is about $0.5°$ ($\pi/360$ rad). The solid angle subtended by the sun is therefore

$$\Omega_{\mathrm{s}} = \int_0^{2\pi} \int_0^{\vartheta_{\mathrm{s}}/2} \sin\vartheta \, d\vartheta \, d\varphi = 2\pi \left(1 - \cos\frac{\vartheta_{\mathrm{s}}}{2}\right) \approx 6 \times 10^{-5} \text{ sr.} \tag{4.5}$$

Because $\vartheta_{\mathrm{s}} \ll 1$, we can expand the cosine in Eq. (4.5) and truncate after the first two terms to obtain

$$\Omega_{\mathrm{s}} \approx \pi\vartheta_{\mathrm{s}}^2/4 \text{ sr.} \tag{4.6}$$

We could have obtained Eq. (4.6) by dividing the area of a small (planar) disc of radius $r\vartheta_{\mathrm{s}}/2$ by r^2.

Any direction in space can be specified by a unit vector $\mathbf{\Omega}$, and so we sometimes write Eq. (4.4) symbolically as

$$\Omega = \int_{\Omega} d\mathbf{\Omega}. \tag{4.7}$$

This does *not*, however, denote integration over the variable $\mathbf{\Omega}$ just as a volume integral written as

$$\int_{\mathrm{V}} dV \tag{4.8}$$

does not denote integration over the variable V. Equation (4.7) is simply a more compact way of writing Eq. (4.4).

4.1.2 Radiance

Before tackling radiance we need one more result. Consider a monodirectional, monochromatic, uniform beam of light. To obtain a measure of the amount of radiant energy transported by the beam we imagine a surface A to be placed in the beam with the normal to the surface parallel to it (Fig. 4.2). We can in principle determine how many photons in unit time N_{o} cross A ; N_{o} multiplied by the photon energy is the amount of radiant power (energy per unit time) crossing A. Divide this quantity by A to obtain the radiant power crossing unit area. This quantity is not solely a property of the beam (the radiation field): it also depends on the orientation of A. Tilt A so that its normal makes an angle ϑ with the direction of the beam (chosen so that $\cos\vartheta$ is positive). Now we measure N photons crossing A per unit time. This is related to the previous number by

$$N_{\mathrm{o}} \cos\vartheta = N. \tag{4.9}$$

It follows from this equation that

$$\frac{N_{\mathrm{o}}}{A} = \frac{N}{A\cos\vartheta}, \tag{4.10}$$

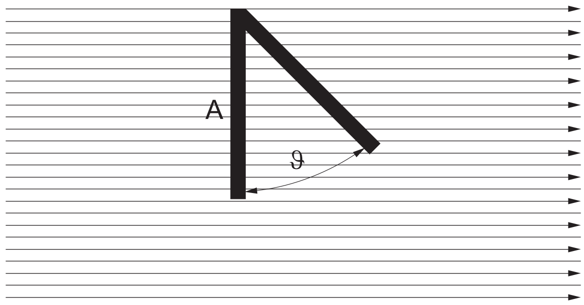


Figure 4.2: The rate at which radiant energy is transported across an area A in a monodirectional radiation field depends on the orientation of A.

and hence the quantity

$$\frac{N}{A\cos\vartheta}, \tag{4.11}$$

where N is the rate at which photons cross A for any ϑ, is a property solely of the radiation field. Geometrically, $A\cos\vartheta$ is the area of the surface projected onto a plane perpendicular to the direction of the beam.

We use the term radiation *field* to describe the beam. For our purposes a field is any physical quantity that varies in space and time, usually continuously except possibly across surfaces. Field quantities often satisfy differential equations.

The (scalar) radiation field (not to be confused with the underlying vector electromagnetic field) is specified by the *radiance* L, a non-negative distribution function much like the distribution functions discussed in Section 1.2. As we show in Section 6.1.2, L satisfies an integro-differential equation, another reason for saying that L specifies a radiation field. Like all distribution functions, L is defined by its integral properties, and in general depends on position, direction, frequency, and time, so we sometimes write it as $L(\mathbf{x}, \mathbf{\Omega}, \omega, t)$ to explicitly indicate these dependencies. At any point in space consider a planar surface of area A, a set of directions with solid angle Ω, a set of frequencies between ω_1 and ω_2, and a time interval between t_1 and t_2. The total amount of radiant energy confined to this set of frequencies and directions, and crossing this surface in the specified time interval is given by

$$\int_{t_1}^{t_2}\int_{\omega_1}^{\omega_2}\int_{\mathrm{A}}\int_{\Omega} L(\mathbf{x}, \mathbf{\Omega}, \omega, t)\cos\Theta \; d\mathbf{\Omega}\; dA\; d\omega\; dt, \tag{4.12}$$

where Θ is the angle between the normal to the surface and the direction $\mathbf{\Omega}$. The cosine factor is introduced so that L is a property solely of the radiation field, not of the orientation of A [see Eq. (4.11)]. The dimensions of L are power per unit area, per unit solid angle, per unit frequency. The radiance defined by Eq. (4.12) is sometimes called the spectral or monochromatic radiance, and its dependence on frequency or, equivalently, wavelength sometimes indicated by a subscript: $L_\omega,\ L_\nu,\ L_\lambda$. The total or integrated radiance is the integral of L over a range of frequencies. Unless specified otherwise, by radiance we mean spectral radiance.

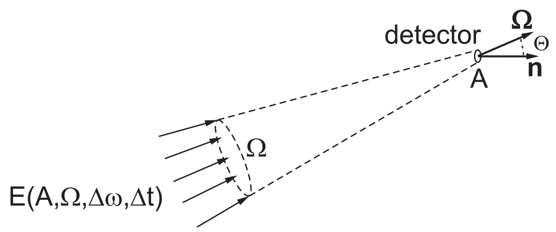


Figure 4.3: Radiant energy $E(A, \Omega, \Delta\omega, \Delta t)$, confined to a solid angle Ω around the direction $\mathbf{\Omega}$, is incident on a detector with area A. The unit vector $\mathbf{n}$ is normal to A.

Before writing Eq. (4.12) in a different way, we review the mean-value theorem of integral calculus, which we invoked in Section 1.2. According to this theorem (for a one-dimensional integral), between any two values x_1 and x_2, there is some intermediate value, call it $\overline{x}$, such that

$$\int_{x_1}^{x_2} f(x)\, dx = f(\overline{x})(x_2 - x_1), \tag{4.13}$$

where f is any continuous and bounded function. The mean-value theorem doesn't tell us how to find $\overline{x}$, only that it exists. The geometrical interpretation of this theorem is straightforward. The integral in Eq. (4.13) is the area of the region bounded by the continuous curve $y = f(x)$, the x-axis between x_1 and x_2, and lines perpendicular to the x-axis of length $y_1 = f(x_1)$ and $y_2 = f(x_2)$. According to the mean-value theorem there is some value $\overline{y}$ on this curve such that the area is $\overline{y}(x_2 - x_1)$.

The mean-value theorem also holds for multiple integrals such as Eq. (4.12), which therefore is equal to

$$\overline{L}\; \overline{\cos\Theta}\; A\Omega\; \Delta\omega\; \Delta t, \tag{4.14}$$

where $\overline{L}$ and $\overline{\cos\Theta}$ indicate some value of L and $\cos\Theta$ over the domain of integration, $\Delta\omega = \omega_2 - \omega_1$, and $\Delta t = t_2 - t_1$. Equation (4.14) provides a means by which we can (in principle) measure the radiance at a point and in a particular direction. Place a detector with area A at the point where L is to be measured (Fig. 4.3). The detector is collimated in that it receives radiation only over a set of directions with solid angle Ω, around some direction $\mathbf{\Omega}$, and is equipped with a filter that passes only radiation in some frequency interval $\Delta\omega$. Measure the total radiant energy $E(A, \Omega, \Delta\omega, \Delta t)$ received by this detector over some time interval Δt. Divide this energy by $A\cos\Theta\Omega\Delta\omega\Delta t$ to obtain an estimate for L. (A can be oriented relative to $\mathbf{\Omega}$ so that $\cos\Theta$ has the limiting value 1 as Ω shrinks.) Form the quotient

$$\frac{E(A, \Omega, \Delta\omega, \Delta t)}{A\cos\Theta\Omega\Delta\omega\Delta t} \tag{4.15}$$

for ever-decreasing values of A, Ω, $\Delta\omega$, and Δt until it no longer changes, then stop. At this point a fractional change in either A, Ω, $\Delta\omega$, or Δt leads to the same fractional change in $E(A, \Omega, \Delta\omega, \Delta t)$. The radiation field is now uniform over the geometric quantities A and Ω

and the intervals Δt and $\Delta\omega$. The quotient so obtained is L at space point $\mathbf{x}$, in the direction $\mathbf{\Omega}$, for frequency ω, at time t.

We cannot let A become indefinitely small because at some point ($A < \lambda^2$) the concept of a continuous radiance becomes invalid. But this is hardly cause for concern because we run into this kind of limitation all the time. For example, the density ρ at a point in a fluid often is defined as

$$\rho = \lim_{V \to 0} \frac{M}{V}, \tag{4.16}$$

where V is a volume containing the point and M the mass of fluid within this volume. But the limit of V here is not literally 0. When V shrinks to molecular dimensions (or smaller), the quotient in Eq. (4.16) undergoes wild fluctuations depending on whether V contains molecules or not. So the limit in Eq. (4.16) is interpreted to mean that we shrink V until the quotient no longer changes, then stop and call that quotient the density at a point. But a better definition of density, in our view, is that it is a distribution function. Its integral over any arbitrary (within limits) volume is the mass enclosed by that volume.

The only way to truly grasp radiance, or indeed any physical concept, is to become familiar with its properties, to observe how it behaves in as many contexts as possible. Defining radiance is only a first small step toward understanding it. One essential property of radiance is that it is *additive*: if several incoherent sources contribute to the radiance at a particular point and in a particular direction, the total radiance is the sum of the radiances from each source as if it were acting alone. Another property of radiance is its *invariance*, which we turn to next.

4.1.3 Invariance of Radiance

If absorption and scattering by the medium in which radiation propagates is negligible, radiance is invariant along a particular direction. By this is meant the following. Go to any point in the radiation field and determine the radiance there in the direction $\mathbf{\Omega}$. Now proceed along this direction. At any point on this path the radiance in the direction $\mathbf{\Omega}$ is the same. The proof is as follows.

At any point insert a planar surface with area A, sufficiently small that L in the direction $\mathbf{\Omega}$ is the same over every point of A, oriented so that its normal is parallel to $\mathbf{\Omega}$. The quantity LA is therefore the amount of radiation crossing A per unit solid angle around the direction $\mathbf{\Omega}$. A surface A'' at a distance r from A receives an amount of radiant energy

$$LA\Omega' = LA\frac{A''}{r^2}, \tag{4.17}$$

where Ω' is the solid angle subtended by A'' at A (both of which are $\ll r^2$). This is shown schematically in Fig. 4.4. If A'' is sufficiently small, the power intercepted by it (per unit area) is uniform and equal to

$$\frac{LA}{r^2}. \tag{4.18}$$

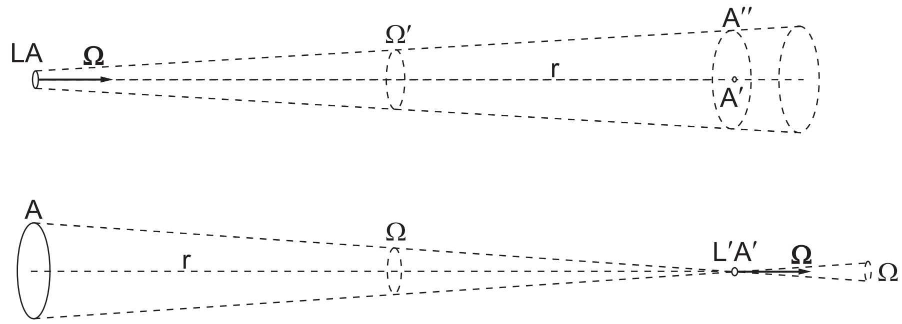


Figure 4.4: Invariance of radiance. The top figure shows the amount of radiation intercepted by A'' a distance r from A where the radiance is L. A smaller area A' within A'' is shown in the bottom figure. The scale is changed between top and bottom. The radiant energy intercepted by A' is LAA'/r^2, which equals $L'A'A/r^2$ in the absence of attenuation.

The power intercepted by a smaller area A' within A'' is therefore

$$P' = \frac{LA}{r^2}A'. \tag{4.19}$$

A is represented as much larger than A' in the bottom half of Fig. 4.4 because of the difficulty of representing all distances and areas to the same scale. At A' the radiance in the direction $\mathbf{\Omega}$ is L'. In the absence of attenuation all of P' must have originated from A and only from A, and hence

$$L' = \frac{P'}{A'\Omega} = \frac{LA}{r^2\Omega} = L, \tag{4.20}$$

where $\Omega = A/r^2$ is the solid angle subtended by A at A'. Equation (4.20) is the invariance principle for radiance, which also follows from the equation of radiative transfer (see Sec. 6.1.2) when absorption and scattering are negligible. This principle may come as a surprise in light of the oft-repeated mantra that radiation decreases as the inverse square of distance. Some radiometric quantities under some circumstances obey this law, but radiance does not. And radiance is usually what we observe, as we show in the following section.

4.1.4 Imaging Devices and Radiance

An imaging device, such as the human eye, maps radiances onto a surface. To show this, consider a uniformly luminous disc with radius h at a distance $s(\gg h)$ from a simple lens with area A_{lens} (Fig. 4.5). This object disc, with area $A = \pi h^2$, is centered on the optic axis of the lens, and the normal to the disc is parallel to the optic axis. The lens forms an inverted image of the disc at a distance s' from the lens. The radius h' of the image disc is given by

$$\frac{h'}{s'} = \frac{h}{s}, \tag{4.21}$$

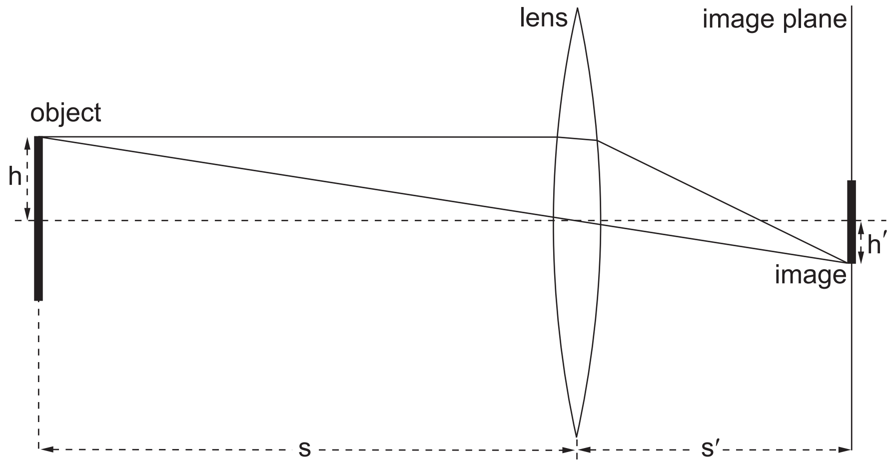


Figure 4.5: The size of an image in the image plane of a simple lens depends on the distance s between object and lens and its focal length s'.

and hence the area A' of the image disc relative to that of the object disc is

$$\frac{A'}{A} = \frac{s'^2}{s^2}. \tag{4.22}$$

The solid angle subtended by the lens at the object is approximately

$$\Omega_{\text{lens}} = \frac{A_{\text{lens}}}{s^2} \tag{4.23}$$

if $s \gg \sqrt{A_{\text{lens}}}$. Thus the radiant power from the disc intercepted by the lens is

$$LA\Omega_{\text{lens}} = LA\frac{A_{\text{lens}}}{s^2}, \tag{4.24}$$

where L is the radiance from the object disc (in the direction of the optic axis). Implicit in Eq. (4.24) is the assumption that $\cos\Theta \approx 1$ for the angle Θ between the normal to the object disc and any ray from the object disc to the lens. If reflection and absorption by the lens are small relative to transmission, almost all the radiant energy [Eq. (4.24)] illuminates the image disc. Thus the amount of radiant power per unit area of the image disc is

$$L\frac{A_{\text{lens}}}{s^2}\frac{A}{A'}, \tag{4.25}$$

which combined with Eq. (4.22) yields

$$L\frac{A_{\text{lens}}}{s'^2} = L\Omega'_{\text{lens}} \tag{4.26}$$

for the radiant power per unit area of the image disc. At each point of the image plane, the radiant power per unit area depends only on the radiance at the object point and a fixed quantity, the solid angle Ω'_{lens} subtended by the lens at the image plane. This comes about because of two opposing factors. As the object is moved farther from the lens, the radiant power it intercepts decreases as the inverse square of distance s [Eq. (4.24)]. But the area of the image also decreases in the same way [Eq. (4.22)], and hence the power per unit area of the image does not change.

Suppose that the lens is the human eye and the image plane is the retina. We can imagine any scene to be made up of small discs with different radiances, high radiances corresponding to what we call bright parts of the scene, and low radiances corresponding to dark parts. The radiances here are those in the direction toward the eye of an observer. Each disc in the scene is mapped onto its image on the retina. The size of the image disc decreases with distance but not the amount of radiant power per unit area of each image disc, which is proportional to the radiance of the object disc.

You can verify this for yourself by observing a white disc at ever-increasing distances on a clear day. Although the (angular) size of the disc decreases with distance, you perceive it to be just as bright.

But now we have to add some psychological and physical caveats. Although the size of retinal images decreases with distance, the human observer perceives the size of familiar objects to be more or less the same regardless of distance (within limits). This is called *size constancy*, a mechanism built into our brains and of which we are not directly aware. If we were to interpret our visual surroundings according to the changing sizes of retinal images, we might go mad. Familiar objects would appear to shrink as they recede from us. But this is not what we perceive, which you can verify by walking away from a familiar object such as a dog and noting that its perceived size does not markedly change. You do not gasp in amazement as you move away from a Doberman Pinscher and it shrinks to a Miniature Pinscher, then back to a Doberman as you return.

Another caveat is that the perceived brightness of an object does depend on its angular size. Again, you can verify this by cutting out circular discs of different sizes from the same sheet of white paper and pasting them onto the same large, uniform backdrop (say a large sheet of gray paper) illuminated uniformly by bright sunshine. Observe the discs at different distances and note which seems brighter. Objectively the radiance from all discs is the same if they are composed of the same material, against the same backdrop, and illuminated by the same light, but subjectively they are different. The objective pattern of radiant energy on the human retina is just the beginning of visual perception. The retina is connected by way of the optic nerve to the brain, which processes retinal images to create a visual world. Radiance variations are the bricks and mortar out of which the brain builds this world.

Now the physical caveat. As the distance s increases the solid angle subtended by the disc (source) at the lens decreases to the extent that the lateral coherence length (see Sec. 3.4.2) is greater than the diameter of the lens, which then produces an interference pattern. And as soon as interference comes into play, the assumptions underlying Eq. (4.26) break down. For example, Eq. (4.21) for the size of retinal images is no longer valid. From Eq. (3.50) it follows that we begin to run into the limitations of theory when

$$\frac{s}{h} > \frac{d_{\mathrm{lens}}}{\lambda}, \tag{4.27}$$

where $d_{\rm lens}$ is the diameter of the lens. This is why Eq. (4.26) is not valid for stars, which are essentially point sources in that their details cannot be resolved with telescopes. The image size of a star is not determined by Eq. (4.21), which predicts (incorrectly) that this size steadily decreases with increasing distance s (for fixed h). When s/h is such that Eq. (4.27) is satisfied, the size of the image becomes greater than that predicted by Eq. (4.26), less distinct, and its size constant. The amount of radiant power received by the lens, however, is inversely proportional to distance squared, and because the image size is constant, the amount of radiant power per unit area of the image decreases similarly.

A distinction also must be made between an imaging device and a flat-plate detector. Such a detector receives light from scenes in which radiance may vary markedly with direction. The power recorded by the detector is proportional to an average radiance because each point on the plate receives light from many directions. If we want the distribution of radiance, we would have to collimate the detector to exclude light except that from a small set of directions. The purpose of lenses in imaging devices (eye, camera, slide projector) is to ensure that each point of the detector (retina, film, screen) receives light from only one point of the scene to be imaged. As we note in Section 1.4.7, the essence of imaging is one-to-one mapping.

4.1.5 A Simple Lens Cannot Increase Radiance

Many of us as children have used a magnifying glass to burn a piece of paper or to incinerate some unfortunate ant or even to start a fire. Indeed, the focal point of a lens is called a *caustic*, and something that is caustic burns: caustic soda burns skin, a caustic remark burns your ears. This common experience with lenses has unfortunately engendered the misconception that lenses can increase radiance.

Consider a piece of white paper illuminated by sunlight on a clear day. If we neglect attenuation by the atmosphere, the solar radiance at the paper is the same as that at the surface of the sun. From the definition of radiance, and given that the solid angle subtended by the sun at the surface of Earth is small, the radiant power incident on unit area of the paper is $L_{\rm s}\Omega_{\rm s}$, where $L_{\rm s}$ is the radiance of the sun. For simplicity we take the direction of the sun to be normal to the paper. Now suppose we interpose a lens between the paper and the sun at a distance f, the focal length of the lens, above the paper. Thus the sun is imaged onto the paper. If $A_{\rm s}$ is the (projected) area of the sun, the area of its image $A_{\rm i}$ follows from Eq. (4.22):

$$A_{\rm i} = \frac{f^2}{r^2}A_{\rm s} = f^2\Omega_{\rm s}, \tag{4.28}$$

where r is the Earth–sun distance.

The radiant power intercepted by the lens is $L_{\rm s}\Omega_{\rm s}A_{\rm lens}$. We neglect reflection by the lens, which reduces this power by a few percent. The radiant power per unit area of the image of the sun on the paper is therefore

$$\frac{L_{\rm s}\Omega_{\rm s}A_{\rm lens}}{A_{\rm i}} = \frac{L_{\rm s}\Omega_{\rm s}A_{\rm lens}}{\Omega_{\rm s}f^2} = L_{\rm s}\Omega_{\rm lens}, \tag{4.29}$$

where $\Omega_{\rm lens}$ is the solid angle subtended by the lens at its focal point. With no lens in place, each unit area of the paper receives $L_{\rm s}\Omega_{\rm s}$. With a lens in place, a unit area of the image of the sun (*not* the region surrounding this image) on the paper receives $L_{\mathrm{s}}\Omega_{\mathrm{lens}}$. In the limited image region, therefore, the radiant power (per unit area) increases if $\Omega_{\mathrm{lens}} > \Omega_{\mathrm{s}}$. But the radiance in this region is that of the sun, which follows from Eq. (4.29), the radiant power per unit area of the sun's image. To convert this to radiance, divide by the solid angle of the lens, which yields L_{s}.

To make this clearer, imagine that we can shrink ourselves to the size of an ant and position ourselves on the image of the sun on the paper. What would we experience? First, we would be uncomfortably hot because the radiant power per unit area increases. But we would see the sun just as bright as without the lens. That is, the radiance of the sun would not change, although there would be an observable difference in the appearance of the sun: its angular width would be greater seen through the lens. The function of magnifying lenses is to increase *angular* sizes. Thus an ant at the focal point would see a sun of larger angular size. How much larger depends on the magnification of the lens $(2\times,\ 5\times,\ 10\times\ldots)$.

The origins of the misconception about lenses increasing radiance are not difficult to discover. A uniform surface illuminated by sunlight is uniformly bright. Take a lens and focus an image of the sun onto such a surface. The result is a bright spot (the image of the sun) surrounded by a much larger, darker area. All a lens can do is redistribute solar radiation, and hence if the radiant power per unit area of the image increases, the radiant power per unit area surrounding the image must decrease accordingly so that total radiant power is conserved. You can demonstrate this for yourself with a magnifying lens. Image the sun onto a uniform surface. If you are indoors, an incandescent lamp will do. What you'll observe is a bright spot surrounded by a darker area. It is this increase in *contrast* that is mistaken for an increase in the incident radiance. Humans respond to relative, not absolute differences. The radiance of gray paper illuminated by sunlight is likely to be greater than the radiance of white paper illuminated indoors by an incandescent lamp, and yet we perceive the white paper to be brighter than the gray because the white paper is brighter than its surroundings.

4.1.6 Radiance Changes Upon (Specular) Reflection and Refraction

Denote by L_{i} the radiance in a particular direction $\boldsymbol{\Omega}_{\mathrm{i}}$ incident on an optically smooth interface between two optically homogeneous media. This direction is specified by the angle ϑ_{i} between $\boldsymbol{\Omega}_{\mathrm{i}}$ and the normal to the surface. Let A be an area on this surface, sufficiently small that the radiance over A is approximately constant. The radiant power incident on A is

$$L_{\mathrm{i}} A \cos\vartheta_{\mathrm{i}}\Omega_{\mathrm{i}}, \tag{4.30}$$

where Ω_{i} is the solid angle of a set of directions centered around the direction of incidence and sufficiently small that the radiance is approximately constant. Let L_{r} be the radiance in the direction of reflection, which makes an angle $\vartheta_{\mathrm{r}} = \vartheta_{\mathrm{i}}$ with the normal. If the reflectivity is 100%, the reflected power is

$$L_{\mathrm{r}} A \cos\vartheta_{\mathrm{r}}\Omega_{\mathrm{r}}, \tag{4.31}$$

where Ω_{r} is the solid angle of that set of reflected directions corresponding to all the incident directions in Ω_{i}. These two solid angles are equal, which follows from the law of specular

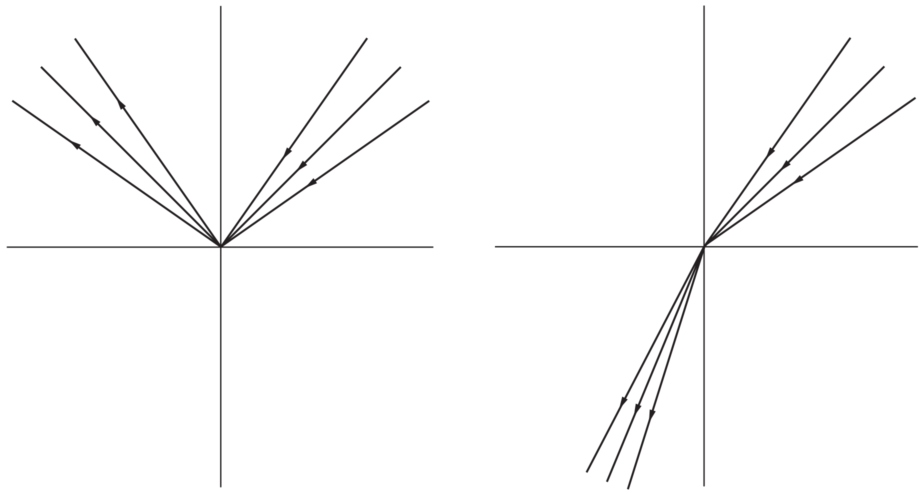


Figure 4.6: In specular reflection (left), incident rays in a set of directions with a given solid angle are mapped by reflection into a set of directions with the same solid angle. But in specular refraction (right), the solid angle of the set of refracted rays is less than that of the incident rays if the refractive index of the illuminated medium is greater than that of the medium containing the source.

reflection. That is, a set of incident directions is mapped into an equal set of reflected directions (Fig. 4.6). Conservation of radiant energy requires that Eqs. (4.30) and (4.31) be equal:

$$L_{\mathrm{r}} = L_{\mathrm{i}}. \tag{4.32}$$

This equation can be verified by looking at an object (a piece of white paper, say) and its reflection (image) in a mirror simultaneously. The image will be seen to be almost as bright as the object. We say "almost" because ordinary mirrors have reflectivities of not much more than 90% at visible wavelengths. If the reflectivity is denoted by $R(\vartheta_{\mathrm{i}})$, Eq. (4.32) becomes

$$L_{\mathrm{r}} = R(\vartheta_{\mathrm{i}})L_{\mathrm{i}}. \tag{4.33}$$

Equation (4.33) can be verified, at least qualitatively, as follows. The reflectivity of *any* surface approaches 100% as ϑ_{i} approaches 90° (glancing incidence). Take a black container – we used a black frying pan – and fill it with water to its brim. Hold a white object – we used a pill bottle – just above the surface of the water and observe the reflection of the object from different directions. If you crouch down so that your line of sight almost grazes the surface, you'll see a reflected image almost as bright as the object. Because the reflectivity of water at near-normal incidence at visible wavelengths is about 0.02 (see Fig. 7.6), the image is much less bright if you look down at the reflection.

The 50-fold reduction of radiance at near-normal incidence upon specular reflection by water is appreciable but not nearly so great as the reduction upon *diffuse* reflection (e.g., by snow and clouds), which we discuss in Section 4.2.

At visible wavelengths, the reflectivity of ice is almost the same as that of liquid water. You often can see from the windows of an airplane the consequences of specular reflection of sunlight by ice crystals. Although you might think of a mirror as being a more or less continuous surface, this is not necessarily so. A mirror can be distributed in space (smash a mirror and sprinkle the shards on the floor and you'll see what we mean). An example of this in nature is a collection of small ice crystals (e.g., plates) falling in air with their tip angles (the angle between the normal to a plate and the vertical) distributed around some small average angle. Perfect alignment would require all crystals to have a tip angle of zero. These crystals reflect an image of the sun, called a *subsun* (because it lies below the sun in angle), different from the sun in two respects. The radiance of the image is about 50 times less than that of the sun; the image is not perfectly circular because of the distribution of tip angles. Subsuns are one of the most frequently seen UFOs, although they are neither "unidentified" nor "flying" nor "objects". They are distorted images of the sun, perfect candidates for UFOs because they are much brighter than clouds, elliptical, and no matter how fast an airplane flies, always keep pace with it (as long as the supply of ice crystals lasts).

Now consider transmitted (refracted) radiance L_{t}. For simplicity we take the incident light to be propagating in air. The refractive index of the (negligibly absorbing) medium on which the light is incident is n. If we assume that the transmissivity of the medium is 1 (no reflection), energy conservation again requires that

$$L_{\mathrm{i}} A \cos\vartheta_{\mathrm{i}} \Omega_{\mathrm{i}} = L_{\mathrm{t}} A \cos\vartheta_{\mathrm{t}} \Omega_{\mathrm{t}}, \tag{4.34}$$

where ϑ_{t} is the angle of refraction and Ω_{t} is the solid angle of the set of transmitted directions corresponding to the set of incident directions. The two solid angles follow from Eq. (4.4):

$$\Omega_{\mathrm{i}} = \sin\vartheta_{\mathrm{i}} \Delta\vartheta_{\mathrm{i}} \Delta\varphi_{\mathrm{i}}, \;\; \Omega_{\mathrm{t}} = \sin\vartheta_{\mathrm{t}} \Delta\vartheta_{\mathrm{t}} \Delta\varphi_{\mathrm{t}}. \tag{4.35}$$

From the law of refraction we have

$$\sin\vartheta_{\mathrm{i}} = n \sin\vartheta_{\mathrm{t}}. \tag{4.36}$$

Square both sides of Eq. (4.36) and take the derivative of the result with respect to ϑ_{t}:

$$2\sin\vartheta_{\mathrm{i}} \cos\vartheta_{\mathrm{i}} = 2n^2 \sin\vartheta_{\mathrm{t}} \cos\vartheta_{\mathrm{t}} \frac{d\vartheta_{\mathrm{t}}}{d\vartheta_{\mathrm{i}}}. \tag{4.37}$$

Now approximate the derivative in this equation as the ratio of differences to obtain

$$\sin\vartheta_{\mathrm{i}} \cos\vartheta_{\mathrm{i}} \Delta\vartheta_{\mathrm{i}} \approx n^2 \sin\vartheta_{\mathrm{t}} \cos\vartheta_{\mathrm{t}} \Delta\vartheta_{\mathrm{t}}. \tag{4.38}$$

This result, Eqs. (4.34) and (4.35), and azimuthal symmetry ($\Delta\varphi_{\mathrm{i}} = \Delta\varphi_{\mathrm{t}}$) yield

$$n^2 L_{\mathrm{i}} = L_{\mathrm{t}}. \tag{4.39}$$

Thus for $n > 1$ the transmitted radiance is *greater* than the incident radiance (assuming no reflection). The reason for this is that upon refraction a set of incident directions is mapped into a smaller set of refracted directions (Fig. 4.6). Radiance is power per unit solid angle (and per

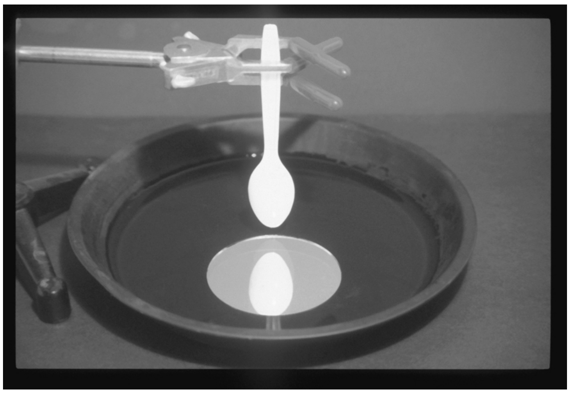


Figure 4.7: A white spoon suspended above a black pan filled with water and with a mirror lying on the bottom of the pan. The image of the spoon reflected by the mirror, and twice refracted, is almost as bright as the spoon. If you look carefully you can see the faint image of the spoon specularly reflected by the water.

unit area), and so a decrease in solid angle, all else being equal, results in an increase in radiance. If the rays in Fig. 4.6 are reversed, what is meant by incident and transmitted is reversed and the transmitted radiance is *less* than the incident radiance. Consider a 100%-reflecting mirror under water, which, as shown in Section 2.1, is negligibly absorbing at visible wavelengths over distances of tens of centimeters. What is the radiance seen by someone above the water looking at the mirror under water? If transmission from air to water increases radiance, reflection by the mirror doesn't change it, and transmission from water to air decreases it, the net effect should be zero. The radiance transmitted into the water is

$$L_{\mathrm{t1}} = n^2 L_{\mathrm{i}}. \tag{4.40}$$

Reflection does not change the radiance

$$L_{\mathrm{r}} = L_{\mathrm{t1}} = n^2 L_{\mathrm{i}}. \tag{4.41}$$

L_{i} is the radiance incident at the air–water interface, and hence L_{r} is the radiance at the water–air interface. Thus the radiance transmitted from water to air is

$$L_{\mathrm{t2}} = \frac{L_{\mathrm{r}}}{n^2} = L_{\mathrm{i}}, \tag{4.42}$$

as expected.

A simple demonstration of the reversibility implied by Eq. (4.42) can be obtained with a mirror under water in a black container. Suspend a white object (as a light source) above the surface of the water, and observe this object together with its transmitted-reflected-transmitted image (Fig. 4.7).

The assumptions of no reflection and $n \neq 1$ are inconsistent because the latter implies a nonzero reflectivity. But this often is not a serious error provided that we confine ourselves to near-normal angles of incidence. For example, the transmissivity of water (from air) at visible and near-visible wavelengths is about 0.98 (Fig. 7.6) for angles within about $40°$ of normal incidence, whereas n^2 is about 1.77.

Because Eq. (4.39) can be written

$$L_{\mathrm{i}} = \frac{L_{\mathrm{t}}}{n^2}, \tag{4.43}$$

one sometimes finds the assertion that radiance is not invariant in a particular direction (in the absence of attenuation) but rather radiance divided by refractive index squared. This needs to be qualified by the caveat that departures of the transmissivity from unity are neglected.

4.1.7 Luminance and Brightness

Up to this point we implicitly assumed that human observers respond to radiance, which is not incorrect but is incomplete. We respond to *integrated* radiance. Moreover, light of each wavelength over the visible spectrum does not evoke the same sensation of brightness even for equal radiance. The peak *luminous efficiency* of the human eye occurs in the middle of the visible spectrum (550 nm) and drops to nearly zero at 400 nm on the short-wavelength side of the peak and at 700 nm on the long-wavelength side. We show the luminous efficiency in Section 1.2 in our discussion of distribution functions, but for convenience reproduce it in Fig. 4.8. Strictly, this figure shows the luminous efficiency for *photopic* vision, when the eye is adapted to what loosely may be called normal light levels (e.g., daylight) in contrast with *scotopic* vision, when the eye is adapted to low light levels (e.g., moonlight). In photopic vision brightness matches are determined mostly by the cones in the retina, whereas in scotopic vision they are determined mostly by the rods; *mesopic* lies between these two extremes. But it is difficult to specify the exact conditions under which each type of vision predominates. The scotopic luminous efficiency curve has about the same shape as the photopic luminous efficiency but its peak is shifted about 50 nm toward the blue, giving rise to the *Purkinje effect*. This is manifested by changes in brightness of red and blue objects under different levels of illumination. If they appear equally bright in daylight, the blue object may be perceived to be brighter after the sun has set. Figure 4.8 is not that for a particular human observer but rather an average for many observers.

The eye integrates the spectral radiance over the visible spectrum weighted by the luminous efficiency. This integrated quantity, called the *luminance*, is

$$K \int LV \, d\lambda, \tag{4.44}$$

where V is the (dimensionless) luminous efficiency of the human eye (also called luminosity, relative luminance, etc.). The constant K is 683 lumens per watt ($\mathrm{lm\,W^{-1}}$), and hence luminance has the units of lumens per square meter per steradian, called the *nit*, an appropriate term (although not likely chosen to be humorous) given the level of nitpicking in photometry. The *candela* is the lumen per steradian, and hence another unit for luminance is candela per square meter ($\mathrm{cd\,m^{-2}}$). The integral in Eq. (4.44) is over wavelength but could just as well be over frequency. Luminance, unlike radiance, takes into account the spectral response of

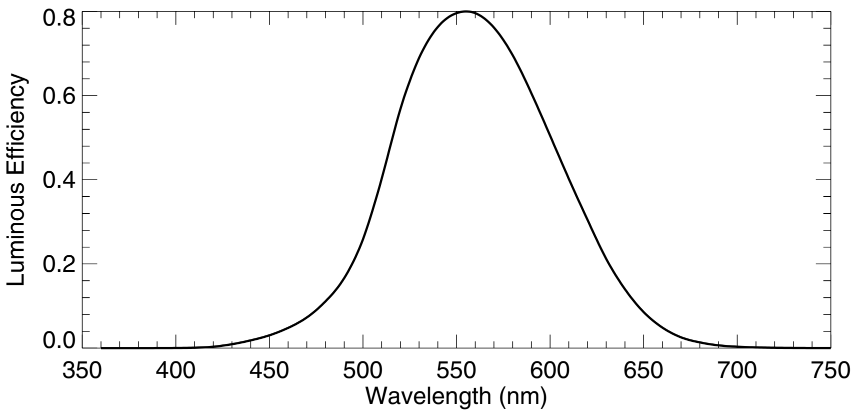


Figure 4.8: Luminous efficiency for the (average) human observer under photopic conditions.

a particular detector, the human eye (by which we mean the retina connected to the brain). Luminance is a *photometric* quantity; its counterpart, radiance, is a *radiometric* quantity. The qualifier photometric signals that the response of the human eye is taken into account. Luminance is a *psychophysical* quantity. The physical part comes from L in Eq. (4.44); the psychological part comes from V, which is obtained, in essence, by asking human observers about what they see.

The observable consequences of Fig. 4.8 and Eq. (4.44) were driven home to us when a student came to class wearing a white jacket with red and green fluorescent stripes on its sleeves. At a glance it appeared that these stripes might be brighter (higher luminance) than the rest of the jacket. We had a photometer near to hand, and so were able to make the necessary measurements. Indeed, the stripes were brighter. These measurements fly in the face of most of our everyday experiences with colored objects. Dye a white T-shirt, for example. The dye preferentially absorbs light of different visible wavelengths, which reduces the reflected radiance relative to that of the white shirt under the same illumination. Thus the luminance of the dyed shirt decreases, and is perceived to be darker than its white cousin. This is what usually happens. But there is a catch. Suppose that some light on the short-wavelength side of the peak of the luminous efficiency curve is shifted to longer wavelengths instead of being absorbed. The amount of reflected radiant energy is not increased but its spectrum is shifted to a region of higher luminous efficiency, and hence a higher luminance. This can happen with *fluorescent* objects, especially ones with high fluorescent yield (one incident photon gives rise to about one photon but of a different frequency).

This experience caused us to wonder about the function of the fluorescent orange vests worn by highway workers and hunters. The spectrum of such a vest (in sunlight) is shown in Fig. 4.9 together with that of an adjacent piece of white paper and of an orange (fruit), the paragon of orangeness. The vest is perceived to be orange, but its luminance is about 20% greater than that of the orange. The kinds of orange objects found in nature (in the woods during deer hunting season, for example) are appreciably darker than fluorescent vests, which

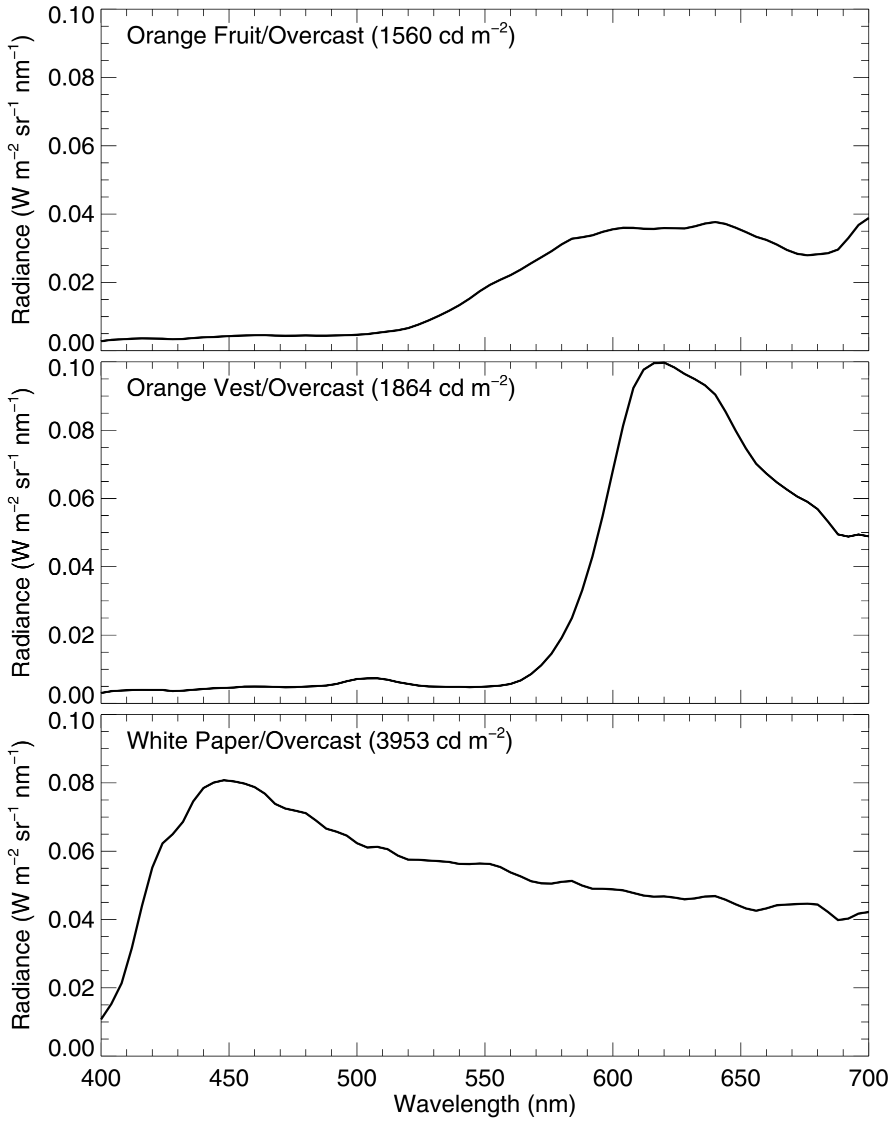


Figure 4.9: Measured spectra and luminances of an orange, an orange fluorescent vest, and white paper, all illuminated by an overcast sky.

by their unusually high luminances are seen to be unnatural, or perhaps extra-natural would be a better term. A deer hunter in the woods does not want to blend in with the surroundings but rather to stand out. A white object satisfies this criterion, but, alas, so does the rear end of a whitetail deer. An orange fluorescent vest or cap is bright and not likely to be confused with less bright orange vegetation.

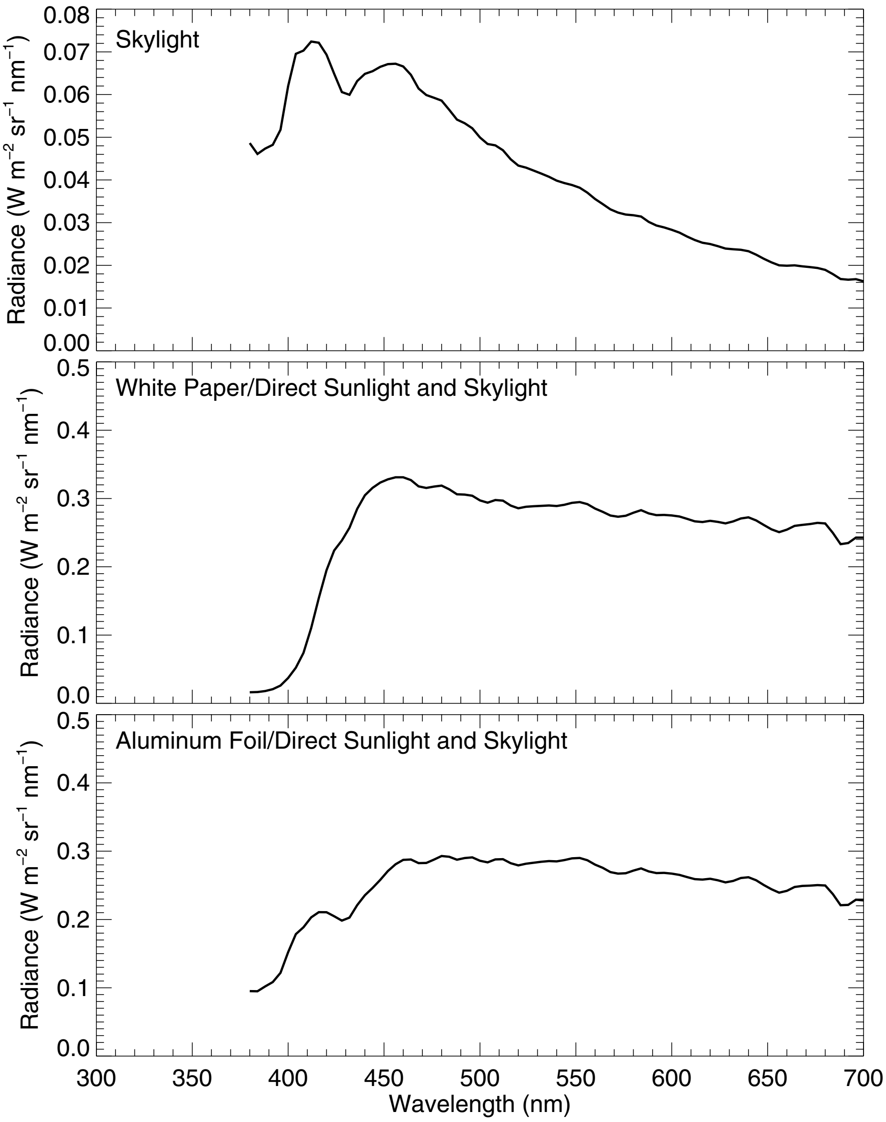


Figure 4.10: Measured spectra of skylight, white paper illuminated by daylight (direct sunlight and skylight), and aluminum foil illuminated by daylight. Note that the dip in the skylight spectrum at around 430 nm is absent in the spectrum of the white paper but present in the spectrum of the aluminum foil. The paper fluoresces whereas the foil does not.

Not only are fluorescent orange vests unnaturally orange, some white paper is unnaturally white, which we discovered by accident. To reduce direct sunlight, we measured reflection by white paper illuminated by sunlight and skylight (daylight); skylight spectra could be obtained without the white paper intermediary. The overhead skylight spectrum in a clean environment is approximately the solar spectrum modulated by a smooth function, the inverse-fourth-power law (Fig. 8.1). Thus peaks and valleys in the sunlight and skylight spectra must originate in the solar spectrum. To our surprise, however, features around 400–450 nm in skylight spectra disappeared in light reflected by white paper (Fig. 4.10). To discover why we measured daylight reflected by crumpled aluminum foil, which restored the skylight spectral features. Clearly, white paper can distort spectra. To dig deeper we used an ultraviolet source with a narrow ($\approx$30 nm) peak around 375 nm. Light reflected by crumpled aluminum foil faithfully follows this source spectrum, whereas that reflected by white paper bears almost no resemblance to it because of fluorescence (Fig. 4.11). Incident light at short wavelengths is transformed into light in wavelength regions where the luminous efficiency of the human eye is greater. Fluorescent whiteners are added to laundry detergents to make clothes "whiter than white" and to white paper to make it brighter. To obtain a reflector that doesn't distort spectra requires inorganic (negligibly absorbing) powders of various oxides and carbonates.

The eye is neither a radiance detector nor a luminance detector but rather a *brightness* detector. Radiance has units (e.g., $\mathrm{W\,m^{-2}\,sr^{-1}}$) and can be measured in princip. Luminance also has units and can be obtained from radiance by integration. But brightness, which has no units, is the "term most commonly used for the attribute of sensation by which an observer is aware of differences of luminance." What this means is that the eye can tell (within limits) when the luminances of two objects are equal but cannot tell their absolute luminances. The luminance of an object we would call gray is more than a hundred times greater when illuminated by direct sunlight than when illuminated by indoor lighting, and yet we still call it gray because we compare it to objects in its surroundings, not with some absolute standard tucked away in our heads.

4.1.8 A Few Words about Terminology and Units

By conscious design, not accident, we gradually introduced radiometric and photometric terms, as few as possible and only as needed. The tragedy of radiometry (and photometry) is that an inherently interesting subject – the best scientific instrument we carry with us everywhere is our eyes, and we are constantly immersed in a world of ever-changing brightness and color – has been made dreadfully boring by wallowing in a mire of terminology and units. This harsh view is shared by others, expressed with biting wit by R. C. Hilborn, who defined radiometry as an acronym for revulsive, archaic, diabolical, invidious, odious, mystifying, exotic terminology regenerating yawns. Terms are multiplied without end. Units in photometry border on the fantastic (foot-candle, talbot, nit, troland, candela, lux, lumen, stilb, foot-Lambert, nox, skot, and so on *ad nauseam*). Radiometry comes across as the science of terminology, its seeming objective being to multiply distinctions endlessly and thereby coin as many terms as possible. To make matters worse, the symbols are ghastly, quantities that are not derivatives yet written as derivatives, and strange-looking ones at that. Mass density rarely, if ever, is written as the derivative of mass with respect to volume [see Eq. (4.16)], but this sort of thing is done all the time in radiometry even though radiometric quantities are distribution functions fundamentally no different from mass density.

Other than in this paragraph we do not use the term *intensity* in this book. Intensity is an Alice-in-Wonderland term that means whatever you want it to mean. It is not uncommon

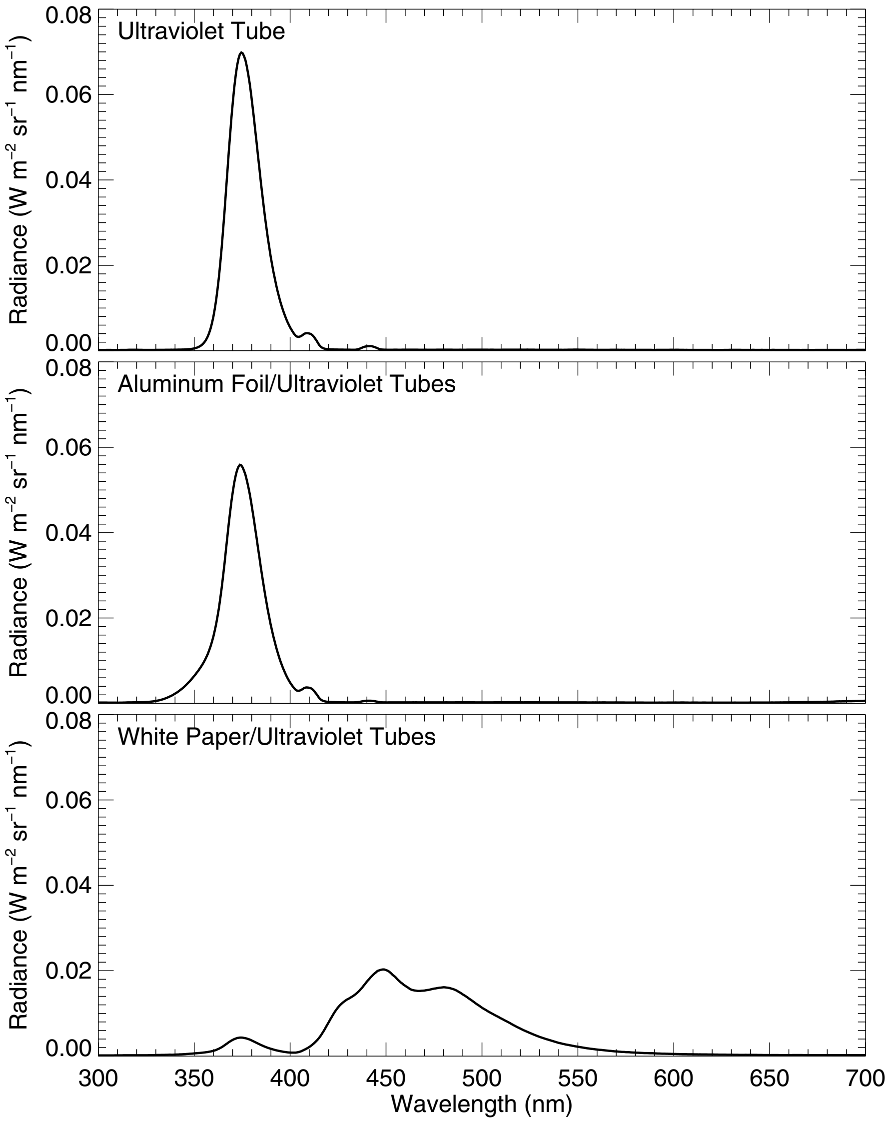


Figure 4.11: Measured spectra of an ultraviolet tube, tube light reflected by aluminum foil, and tube light reflected by white paper. The spectrum reflected by the foil faithfully follows that of the source, whereas because of fluorescence by the paper, its reflected spectrum bears little relation to the source.

to find it used for two (or more) different radiometric quantities in the same paper or book. A humorous yet serious attack on the misuse of intensity was published by J. C. Palmer in a journal devoted to measurement. The first sentence of his abstract hints at what follows: “The current misuse of the term *intensity* in physics is deplored." Palmer notes that he has seen intensity used in five different ways in optics and, like us, has encountered authors who change their meaning as the mood strikes them. After a tongue-in-cheek search for even more terms, Palmer halts abruptly and shouts (in boldface type) that "intensity is an SI base quantity." That is, (luminous) intensity "takes its place alongside the other six SI base quantities: length (metre), mass (kilogram), time (second), electric current (amperes), thermodynamic temperature (kelvin), and amount of substance (mole)." Intensity used in its proper sense is radiant (or luminous) power per steradian. Palmer's parting shot is, "The message is clear; those who use *intensity* in a context **other** than W/sr are either uninformed or just plain careless and sloppy. I can't comprehend a 'special reason' to redefine an SI base quantity, can you?" Using intensity to mean radiance is akin to using length to mean what to the rest of the world is area.

We do not especially like the term radiance, if for no other reason than that one has to be sober in order to pronounce it clearly enough for it to be distinguished from irradiance (Sec. 4.2). But we accept this term because it has been recommended by the Optical Society of America (OSA) for more than 50 years and endorsed by other international organizations devoted to standardization. OSA is the largest optical sciences society in the world, and the bulk of papers on optical science (including atmospheric optics) has been and continues to be published in its journals, the editors of which are supposed to frown on manuscripts in which intensity is used for radiance (or for whatever else strikes the mood of their authors).

The recommendation by OSA's Committee on Colorimetry that brightness be used to mean "that attribute of sensation by which an observer is aware of differences of luminance" goes back more than 50 years, although brightness seems to be used by astronomers for what optical scientists would call radiance. But brightness also has been used as a synonym for luminance. To distinguish between the two, the term photometric brightness has been suggested for luminance.

We have found that for our purposes two radiometric quantities (radiance and irradiance) and two photometric quantities (luminance and brightness), possibly a third, illuminance (the photometric counterpart of irradiance), are all that we need. We do not use intensity because we rarely encounter a genuine need for it and because it has ceased to have any meaning given that it has so many. Irradiance is often called flux, but this ought to be flux density.

Rather than scrap all the units in photometry they could be recycled as characters in a kind of Lord of the Rings fantasy. Once upon a time, in the land of the Nits, dwelt a king named Troland with his beautiful daughter Candela. And so on.

4.2 Irradiance

Although irradiances made brief appearances in previous chapters, the radiation fields were restricted to monodirectional beams. Now we remove that restriction. Unlike radiance, irradiance is not solely a property of the radiation field but depends, in general, on the orientation of the detector. Consider a plane surface in a radiation field. We call one side of this plane the upper side, the other the lower side. Radiation in all directions crosses this plane, so we restrict ourselves to a hemisphere of directions from lower to upper. Construct a coordinate system centered on a small area A of the plane; ϑ and φ are the spherical polar coordinates in this system. The radiance in each direction $\mathbf{\Omega}_\mathrm{i}$ contributes (approximately) an amount

$$L(\mathbf{\Omega}_\mathrm{i}) A \cos\vartheta_\mathrm{i} \Omega_\mathrm{i} \tag{4.45}$$

to the total radiant power incident on A, where ϑ_i is the angle between $\mathbf{\Omega}_\mathrm{i}$ and the normal to A (directed in the sense lower to upper) and Ω_i is the solid angle of a small set of directions around $\mathbf{\Omega}_\mathrm{i}$. Thus the total radiant power crossing unit area is (approximately)

$$\sum_i L(\mathbf{\Omega}_\mathrm{i}) \cos\vartheta_\mathrm{i} \Omega_\mathrm{i}, \tag{4.46}$$

where the sum is over all upward directions. In the limit of smaller and smaller solid angles, this sum becomes an integral

$$F_\uparrow = \int_{2\pi} L(\mathbf{\Omega}) \cos\vartheta \; d\Omega, \tag{4.47}$$

where 2π indicates integration over a hemisphere of directions. As shown in Section 4.1.1, integration over solid angle is the same as integration over a unit sphere:

$$F_\uparrow = \int_0^{2\pi} \int_0^{\pi/2} L(\vartheta, \varphi) \cos\vartheta \sin\vartheta \; d\vartheta \; d\varphi. \tag{4.48}$$

We may call $F_\uparrow$ the upward irradiance, but keep in mind that up and down are arbitrary designations for two opposite directions. For our purposes radiation does not know that gravity exists.

We have to be careful in specifying the downward irradiance as a positive quantity. In Eq. (4.48) the limits of integration of ϑ are such that $\cos\vartheta$ is positive. For ϑ from $\pi/2$ to π (lower hemisphere) $\cos\vartheta$ is negative. This is fixed by interchanging the limits of integration so that the integral giving the downward irradiance is positive:

$$F_\downarrow = \int_0^{2\pi} \int_\pi^{\pi/2} L(\vartheta, \varphi) \cos\vartheta \sin\vartheta \; d\vartheta \; d\varphi. \tag{4.49}$$

Two properties of irradiance must be kept in mind. First, the upward and downward irradiances depend, in general, on the orientation of the reference plane dividing the set of all directions into upward and downward hemispheres. Second, although radiance determines irradiance, the converse, in general, is not true: if we know the irradiance, we cannot uniquely determine the corresponding radiance. The one exception is an *isotropic* radiation field. A completely isotropic radiation field is one for which radiance does not depend on direction. A radiation field also can be isotropic over the upward or downward hemispheres. Suppose that the upward radiance is independent of direction. From Eq. (4.48) it therefore follows that

$$F_\uparrow = L \int_0^{2\pi} \int_0^{\pi/2} \cos\vartheta \sin\vartheta \; d\vartheta \; d\varphi = \pi L. \tag{4.50}$$

We can write this another way:

$$L = \frac{F_\uparrow}{\pi}. \tag{4.51}$$

At first glance Eq. (4.51) may seem to be missing a factor 2. After all, a hemisphere subtends 2π sr. The apparently missing factor is accounted for by the cosine of the angle between the normal to the surface and $\mathbf{\Omega}$. Another way to look at this is to recognize that the area of a hemisphere of unit radius projected onto a plane is π. Even in an isotropic radiation field the rate at which radiation crosses unit area per unit solid angle still depends on direction because of the cosine factor [Eq. (4.45)].

In Section 1.2 the Planck function $P_{\rm e}$ [Eq. (1.11) or (1.20)] was taken as that for irradiance to avoid a digression on radiance, which also would have required a preceding digression on solid angle. Within an opaque container at constant temperature the (equilibrium) radiance $P_{\rm e}/\pi$, sometimes denoted as B, is isotropic.

4.2.1 Diffuse Reflection

Suppose that a diffuse reflector such as snow is illuminated solely by direct sunlight (we neglect illumination by skylight). The direction of the sunlight makes an angle $\Theta_{\rm s}$ with the normal to the surface of the reflector. The downward irradiance is

$$F_{\downarrow} = \int L_{\rm s} \cos\vartheta \, d\Omega, \tag{4.52}$$

where $L_{\rm s}$ is the radiance of the sun. Because the solid angle the sun subtends is small, this integral is approximately

$$F_{\downarrow} \approx L_{\rm s} \cos\Theta_{\rm s}\Omega_{\rm s}. \tag{4.53}$$

The upward irradiance is a consequence of reflection of the downward irradiance (emission by snow is negligible at visible and near-visible frequencies):

$$F_{\uparrow} = R_{\rm d}F_{\downarrow} = R_{\rm d}L_{\rm s} \cos\Theta_{\rm s}\Omega_{\rm s}, \tag{4.54}$$

where the reflectivity (also called albedo) $R_{\rm d}$ of the diffuse reflector is that for reflection into all directions (in the upward hemisphere), although the contributions to $R_{\rm d}$ are, in general, different in different directions (per unit solid angle). Also, $R_{\rm d}$ depends, in general, on $\Theta_{\rm s}$. Assume that the upward radiance is isotropic. This is never true – there are no perfectly diffuse reflectors – but sometimes is a good approximation for a large range of directions (but not an entire hemisphere). With this assumption of isotropy the radiance of the diffuse reflector is

$$L_{\rm d} = L_{\rm s}R_{\rm d} \cos\Theta_{\rm s} \frac{\Omega_{\rm s}}{\pi}. \tag{4.55}$$

Even with the sun directly overhead illuminating clean, fine-grained, highly reflecting snow (see Fig. 5.16), its radiance is almost 10^5 times smaller than the sun's radiance, which is high only in a small set of directions. The radiance of snow is much less, but over an entire hemisphere of directions.

Now consider a flat-plate detector. Such a detector responds to the net amount of radiation it receives. The radiation illuminating the detector is sunlight (again, we neglect the contribution from the sky, which we could exclude with a suitable collimator). As in the previous example, the downward irradiance is given by Eq. (4.53) and the upward irradiance by

Eq. (4.54), where R_d is the reflectivity of the detector. The detector responds to the difference between these two irradiances, the rate at which radiant energy is absorbed by the detector:

$$F_\downarrow - F_\uparrow = L_\mathrm{s} \cos\Theta_\mathrm{s}\Omega_\mathrm{s}(1 - R_\mathrm{d}). \tag{4.56}$$

The response of the detector is proportional to this quantity. If we neglect the dependence of L_s on solar zenith angle Θ_s because of atmospheric attenuation, the response of the detector follows a cosine law if R_d is independent of Θ_s. Such an idealized detector is sometimes called a perfectly diffuse or *Lambertian* detector. Real detectors are never Lambertian, and so a correction has to be made for their departure from the ideal.

Equipped with an understanding of the difference between radiance and irradiance, we can return to the problem of the sun and a lens. We now recognize Eq. (4.29) as specifying the incident (downward) irradiance on the paper at the image of the sun. Assume that this paper is 100% reflecting (it is not) and is oriented so that sunlight normally illuminates it. Assume further that the paper is Lambertian (also not strictly true). With these assumptions, the radiance of the image of the sun on the piece of paper is

$$L_\mathrm{s}\frac{\Omega_\mathrm{lens}}{\pi}. \tag{4.57}$$

From this equation we might conclude that the radiance of the sun's image could be made greater than that of the sun if the solid angle of the lens subtended at its focal point were greater than π. This is a difficult order to fill. It requires a lens with a large collecting area and a short focal length. Unfortunately, the focal length of a lens is proportional to its radius of curvature, so small lenses (with small radii of curvature) have short focal lengths and large lenses (with large radii of curvature) have long focal lengths. Thus it would be difficult to make a lens that subtends an angle greater than π at its focal length.

We also now can revisit the problem of the radiance of an object under water, this time an isotropic diffuse reflector (diffuser) rather than a specular reflector (mirror). The diffuser is parallel to the air–water interface. First take the diffuser to be in air, illuminated by a beam with radiance L_i and solid angle Ω_i incident at an angle ϑ_i. The radiance of the diffuser in air is

$$L_\mathrm{a} = \frac{L_\mathrm{i}\Omega_\mathrm{i}\cos\vartheta_\mathrm{i}}{\pi}, \tag{4.58}$$

where for simplicity we take the reflectivity to be 1. According to Eq. (4.39), the diffuser under water is illuminated by a beam with radiance $L_\mathrm{i}n^2$. The radiance of the diffuser in the water is

$$L_\mathrm{r} = \frac{L_\mathrm{i}n^2\Omega_\mathrm{t}\cos\vartheta_\mathrm{t}}{\pi}. \tag{4.59}$$

An implicit assumption underlying this expression is that the reflectivity of the diffuser is the same in air as in water. The radiance L_r [Eq. (4.59)] is transmitted to an observer (in air), who detects radiance $L_\mathrm{w} = L_\mathrm{r}/n^2$. All these results can be combined to yield

$$L_\mathrm{w} = \frac{L_\mathrm{r}}{n^2} = \frac{L_\mathrm{i}\Omega_\mathrm{t}\cos\vartheta_\mathrm{t}}{\pi} = \frac{L_\mathrm{i}\Omega_\mathrm{i}\cos\vartheta_\mathrm{i}}{\pi n^2} = \frac{L_\mathrm{a}}{n^2}, \tag{4.60}$$

where we also used

$$\Omega_{\mathrm{i}} \cos\vartheta_{\mathrm{i}} = n^2 \Omega_{\mathrm{t}} \cos\vartheta_{\mathrm{t}}, \tag{4.61}$$

which follows from Eqs. (4.35) and (4.38).

Subject to all the assumptions we made, explicit and implicit, the visible radiance of a diffuser under water, unlike that of a mirror, is about 0.56 times its radiance above water under the same illumination. You can observe this by immersing a diffuser (white plastic spoons serve well) partly in water. There will be a distinct brightness change at the interface between the part of the diffuser in air and the part in water. But the precise geometry of the diffuser and even the nature of the water container will affect what is observed. For best results (greatest contrast), the container should be black so that light reflected by it does not illuminate the diffuser. And if the object is not flat, it should be oriented so that light reflected by one part of it does not illuminate other parts. Figure 4.12 shows spoons in different orientations partly immersed in water in black and white pans. Equation (4.60) should be looked upon as specifying the *maximum* reduction in radiance. Moreover, although the reflectivity of an ordinary silvered mirror does not change appreciably from air to water, that of glass does. And, as we show in Section 5.3.1, the reflectivity of a porous material (such as sand) decreases upon wetting. But the radiance decrease described by Eq. (4.60) is solely a consequence of the geometrical properties of radiance and irradiance.

You can see the same kind of brightness change when you partly immerse your hand in a basin filled with water. The part of your hand below the water is darker than that above, and the difference is more striking if the basin is stainless steel or dark porcelain rather than white porcelain.

Scientific progress undoubtedly has been retarded by the switch from taking baths to taking showers. Nothing of great moment has been discovered in the shower, whereas the bathtub is a natural laboratory. Archimedes did not burst from his shower shouting Eureka. If you take leisurely baths instead of showers (the fast food of bathing) you can't help but notice a darkening of your outstretched thighs at the interface between air and water.

To underscore the difference between radiance and irradiance consider the downward irradiance above and below the surface of water in a black container (so that no radiation is reflected by it). Assume that the downward radiance above the water is isotropic. With this assumption the downward irradiance is

$$\int_0^{2\pi} \int_0^{\pi/2} L_{\mathrm{i}} \cos\vartheta_{\mathrm{i}} \sin\vartheta_{\mathrm{i}} \; d\vartheta_{\mathrm{i}} \; d\varphi_{\mathrm{i}} = \pi L_{\mathrm{i}}. \tag{4.62}$$

If we neglect the departure of the transmissivity from 1, the downward radiance in the water is $L_{\mathrm{t}} = L_{\mathrm{i}} n^2$. This downward radiance is constant over a range of directions but not over a complete hemisphere. As we saw in Section 4.1.6, the hemisphere of downward incident directions is not mapped into a complete hemisphere of downward transmitted directions. That is, there are no rays in the water making an angle greater than ϑ_{c} with the normal to the surface, where this *critical angle* is given by

$$\sin\vartheta_{\mathrm{c}} = \frac{1}{n}. \tag{4.63}$$

This follows from the law of refraction for an angle of incidence $\pi/2$.

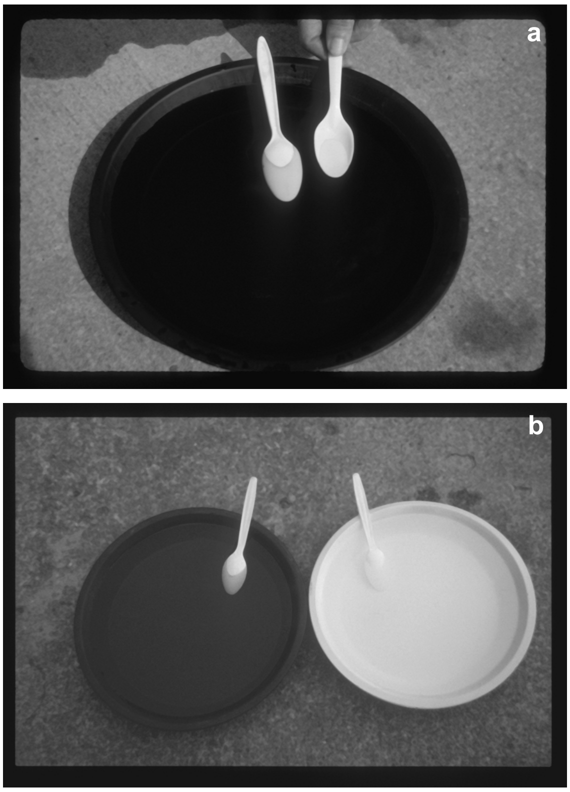


Figure 4.12: a) White plastic spoons oriented differently and partly immersed in water in the same black pan. Note the marked change of brightness of the spoons across the air–water boundary, especially for the spoon with its bowl oriented convex. b) White plastic spoons partly immersed in water in white and black pans. Both spoons are oriented with their bowls convex, and yet the brightness change across the air–water boundary is much greater for the spoon in the black pan.

The downward irradiance in the water is therefore

$$\int_0^{2\pi}\int_0^{\vartheta_\mathrm{c}} L_\mathrm{t}\cos\vartheta_\mathrm{t}\sin\vartheta_\mathrm{t}\,d\vartheta_\mathrm{t}\,d\varphi = 2\pi L_\mathrm{i} n^2 \int_0^{\vartheta_\mathrm{c}} \frac{1}{2}\frac{d}{d\vartheta_\mathrm{t}}(\sin^2\vartheta_\mathrm{t})\,d\vartheta_\mathrm{t} = \pi L_\mathrm{i}, \quad (4.64)$$

which is equal to the downward irradiance above the water. Thus irradiance is conserved in going from air to water whereas radiance is not. Again, we emphasize that this result is based on the assumption that reflection because of the air–water interface is negligible. This is not a bad assumption given that at visible wavelengths the ratio of reflected to (isotropic) incident irradiance is about 6%.

4.2.2 Flux Divergence

Consider a closed (imaginary) surface S in a radiation field. The outward unit normal at each point is $\mathbf{n}$. Radiation crosses this surface from outside to inside the region enclosed by S and vice versa. At any point the net rate at which radiation crosses unit area of the surface is

$$-\int_{4\pi} \mathbf{n} \cdot \boldsymbol{\Omega} L \, d\Omega. \tag{4.65}$$

The minus sign comes about because radiation transported into the region enclosed by S is counted as positive and radiation transported out is counted as negative. We may define the vector irradiance (or vector flux) at a point as

$$\mathbf{F} = \int_{4\pi} \boldsymbol{\Omega} L \, d\Omega. \tag{4.66}$$

With this definition, the total net radiation transported into the region enclosed by S is

$$-\int_{S} \mathbf{F} \cdot \mathbf{n} \, dS. \tag{4.67}$$

From the divergence theorem we have

$$\int_{S} \mathbf{F} \cdot \mathbf{n} \, dS = \int_{V} \nabla \cdot \mathbf{F} \, dV, \tag{4.68}$$

where V is the volume enclosed by S. Thus the rate of radiant energy deposition per unit volume is the negative of the vector flux divergence

$$-\nabla \cdot \mathbf{F}. \tag{4.69}$$

This is a generalization of the result we obtained previously for a monodirectional beam propagating in a purely absorbing medium. Equation (4.69) is not restricted to such a beam or to such a medium.

From Eq. (4.66) we have

$$\nabla \cdot \mathbf{F} = \int_{4\pi} \nabla \cdot (\boldsymbol{\Omega} L) \, d\Omega = \int_{4\pi} \boldsymbol{\Omega} \cdot \nabla L \, d\Omega. \tag{4.70}$$

A necessary condition for deposition (transformation) of radiant energy is that the gradient of the radiance cannot vanish for all directions. But this is not sufficient. The absorption coefficient must also be nonzero, which makes physical sense and is proved in Section 6.1.2.

4.3 Color

Color used as a synonym for wavelength (or frequency) is a common crutch of popular science writers, who seem to think that wavelength cannot be swallowed unless sugar-coated with color, thereby impeding understanding of both color and wavelength as well as depriving readers of something marvelous. Colors are produced in our brains. Sometimes these

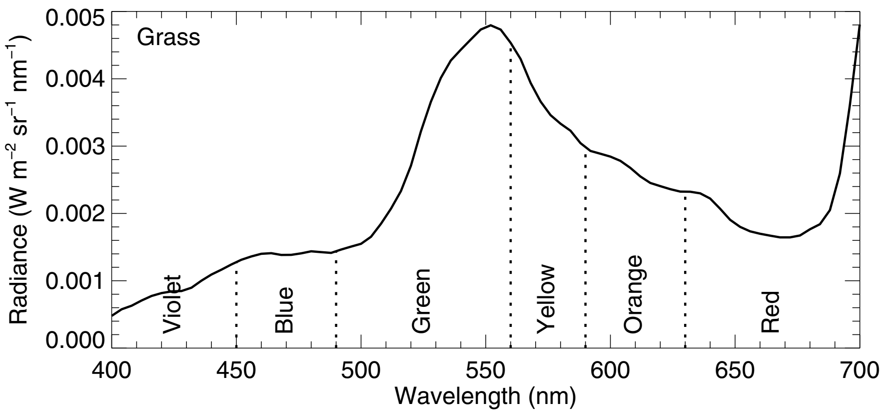


Figure 4.13: Measured spectrum of grass illuminated by daylight. The visible spectrum is divided into six regions to which color names are given.

colors originate from external stimulation, sometimes not; sometimes the stimulation is light, sometimes not. We may dream in color despite the absence of light on our retinas. We may see colors by taking psychotropic drugs or by being hit in the head. Even when light is the stimulus that produces color, there is not a one to one correspondence between a perceived color and the spectrum of the light that produces it. The histologist Santiago Ramón y Cajal, arguably Spain's most famous scientist, succinctly and beautifully captured the essence of human vision. "By virtue of a marvelous alchemy, begun in the retina and completed in the central nervous system, that which in the surrounding ether is simple undulatory motion, is converted in the brain into something completely new and purely subjective: sensations, perceptions, visual memories, associations of images, ideas, and wills."

Physicists are thoroughly steeped in spectra, especially line spectra because of their central role in the evolution of quantum mechanics. But much of what passes for knowledge about color among physicists is demonstrably wrong. Take, for example, the following assertion by a doctor of physics who is also a doctor of medicine, although the two doctorates seem to reside in different parts of his brain: "A leaf looks green because it is absorbing all the colors of the spectrum except green, which it is reflecting." This is not a carefully chosen exotic blooper but rather common currency. Consider this assertion about a green leaf in light of a measured visible spectrum of grass (Fig. 4.13). Although we label six bands in this spectrum with color names, they are somewhat arbitrary in that there are no sharp boundaries between different perceived colors. Note that some yellows and oranges are more intense than some greens. Moreover, light of *all* wavelengths is in light from green grass. And this is true of most colored objects in our surroundings. An apple is not red because, as we have heard many times, it reflects only red light. Red is merely the perceptually dominant color. We are incapable of looking at a source of visible light and assessing how much red light, green light, and so on it is composed of. We are not spectrometers for electromagnetic radiation although we are to a degree for acoustic radiation. We can distinguish the separate instruments in an

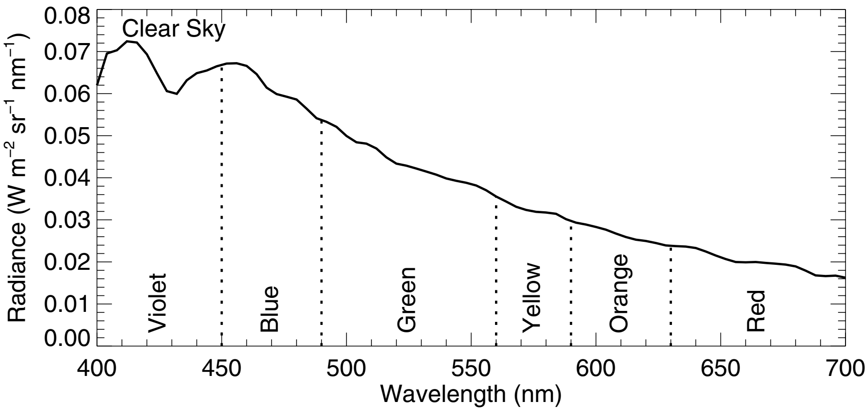


Figure 4.14: Measured skylight spectrum (near zenith). The visible spectrum is divided into six regions to which color names are given. Although the spectrum peaks in the violet, skylight is perceived to be blue.

orchestra, from the high-frequency piccolos to the low-frequency bassoons, and sense their relative contributions to the composite sound of all instruments playing simultaneously.

But, you might protest, the spectrum in Fig. 4.13 *peaks* in the green. So it does, which is largely irrelevant, as evidenced by Fig. 4.14, a measured spectrum of a clear, blue zenith sky. Although we give the color name blue to the sky, its spectrum peaks in the violet and, as with the spectrum of green grass, contains light of all wavelengths. Indeed, there is more light that is *not* blue in skylight than light in the band labeled blue. An even more striking example is provided by the display of a computer adjusted to show yellow (Fig. 4.15). Light in the band labeled yellow is overshadowed by that in the bands labeled green and orange. The sensation we call yellow can indeed be produced by monochromatic sources with wavelengths in the band 560 to 590 nm but also by mixing red and green beams containing *no* light of these wavelengths. What about that paragon of yellowness, the banana? Surely it must be a source of only yellow light. Alas, measurements show otherwise. The spectrum of a banana (Fig. 4.16) contains as much green and orange and red as yellow, if by these colors is meant light confined to certain bands of wavelengths.

None of these measured spectra are contrived even though they handily refute assertions that have been made many times about color, much of the confusion about which stems from a failure to distinguish between wavelength, color, and color names.

Countless students have been introduced to that mythical character Roy G. Biv as a way of learning the names of the seven colors supposedly identified by Newton. What did Newton really say? In his *Opticks*, Prop. II, Theor. II, he notes that "the Spectrum…did…appear tinged with this Series of Colors, violet, indigo, blue, green, yellow, orange, red, *together with all their intermediate Degrees in a continual Succession perpetually varying. So that there appeared as many Degrees of Colours, as there were sorts of Rays differing in Refrangibility*" [emphasis added]. The widespread notion that there are seven and only seven colors in nature

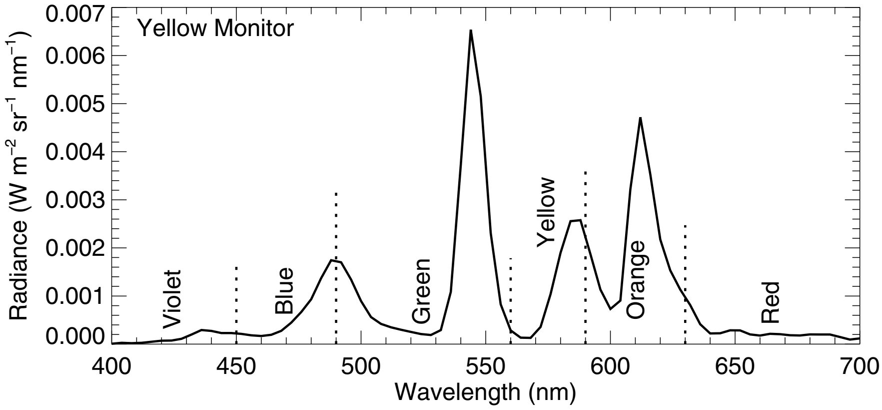


Figure 4.15: Measured spectrum of computer display yellow pixels. The visible spectrum is divided into six regions to which color names are given. Although the amount of radiation in the band labeled yellow is overshadowed by radiation in the bands labeled green and orange, the display is still perceived to be yellow.

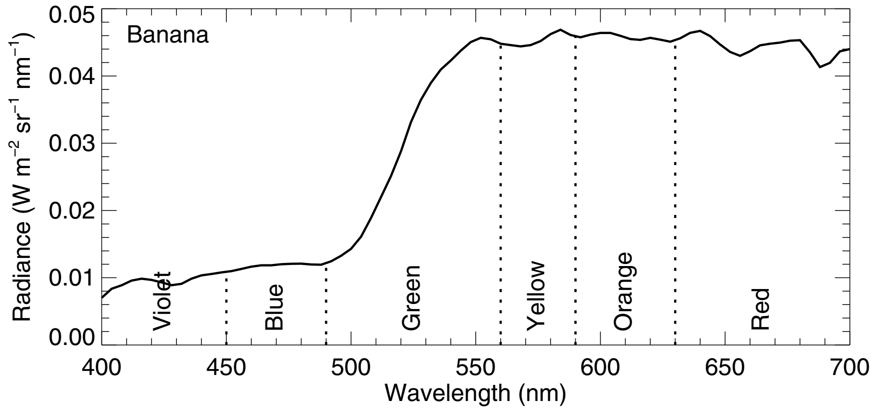


Figure 4.16: Measured spectrum of a banana illuminated by daylight. The visible spectrum is divided into six regions to which color names are given. A yellow banana is far from being a reflector of only yellow light.

is a consequence of the failure to read what Newton wrote and to come to grips with the nature of color. Newton did not believe that there are only seven colors, he merely gave seven color names. Why did he stop at seven? This is the magic number par excellence: seven pillars of wisdom, days of the week, deadly sins, wonders of the world, intervals of the diatonic scale, and so on. In a famous essay about Newton, John Maynard Keynes noted that "Newton was not the first of the age of reason. He was the last of the magicians..." Even today, three hundred years after Newton, many people still believe that numbers have magical powers, some being lucky, others unlucky. Yet we supposedly live in a scientific age. Newton lived in an age when magic, including number magic, still held sway over people's minds.

Every language has a finite number of color words, ranging from three to eleven, by which we mean ones used in everyday speech. We do not count the exotic words confected by paint manufacturers or fashion designers. Depending on our native language and our experience we use a small number of color names, most of them learned as children. And we can determine when two objects are perceived to have the same color, the basis of colorimetry, to which we turn next.

4.3.1 Colorimetry: The CIE Chromaticity Diagram

Colorimetry, the measurement of color, is based on color matching by human observers. All we can do is match two colors, and different observers are not likely to agree when colors exactly match. Suppose that we look at a large black screen with two identical circular holes in it. On the other side is a white screen that can be illuminated by two different sources. A partition is arranged so that we see only light from one source through one hole, light from the other source through the other hole. We wait long enough for our eyes to be fully adapted to photopic illumination, then adjust the sources until we judge the two discs of light, side by side, to be indistinguishable.

One source, the *sample*, has a narrow spectrum centered on some visible wavelength. We can adjust this source so that its radiant power is the same for every wavelength. The other source is a composite obtained by superposing (adding) light from three lamps (*primaries*), which we call red, green, and blue because these are the color sensations each separately evokes. We can independently adjust the power of each primary. For samples corresponding to each visible wavelength, we adjust the three primaries until we get a color and brightness match, then record the luminances of the primaries. For each visible wavelength we obtain three luminances, the *tristimulus* values, denoted as $\overline{r}$, $\overline{g}$, and $\overline{b}$. Each set of tristimulus values yields light indistinguishable from the sample. We do this for all visible wavelengths, thereby obtaining three *tristimulus curves*. But there is a catch. We cannot beg, borrow, or steal three real lamps that enable us to match all visible wavelengths. Despite the blather about three primary colors, they do not exist. The only way to match some sample wavelengths is to take one of the three lamps and add its light to that of the sample, the combination of which then can match the remaining two lamps. This results in some tristimulus values that are negative because one of the three lamps has been removed from the trio. This is deemed unacceptable, or at least inconvenient. But this inconvenience is removed once we recognize that the only essential aspect of color matching is that it requires *three* numbers, the absolute values of which are irrelevant. Thus any linear transformation

$$\overline{x} = a_{11}\overline{r} + a_{12}\overline{g} + a_{13}\overline{b}, \tag{4.71}$$

$$\overline{y} = a_{21}\overline{r} + a_{22}\overline{g} + a_{23}\overline{b}, \tag{4.72}$$

$$\overline{z} = a_{31}\overline{r} + a_{32}\overline{g} + a_{33}\overline{b}, \tag{4.73}$$

yields an acceptable set of tristimulus values, and we can choose the coefficients a_{ij} such that $\overline{x}$, $\overline{y}$, and $\overline{z}$ are never negative. The $\overline{g}$ curve obtained with human observers is remarkably sim-

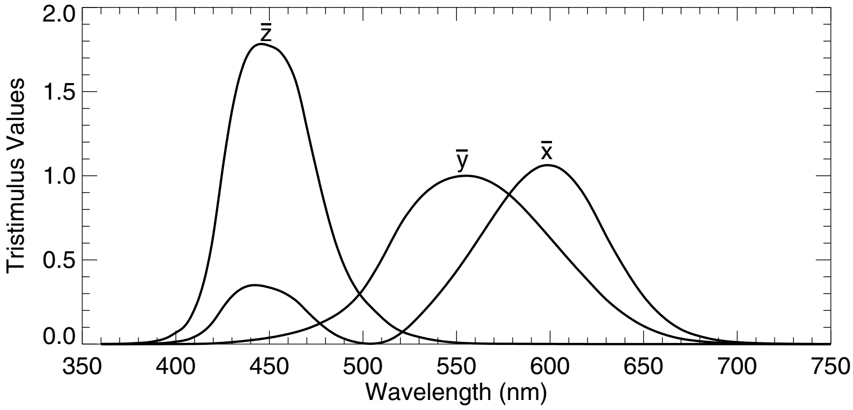


Figure 4.17: CIE 1931 tristimulus functions. The area under each curve is the same; $\overline{y}$ is the luminous efficiency (Fig. 4.8). These functions are obtained by transformations of data obtained from color matching experiments with human observers.

ilar to the curve of luminous efficiency (Fig. 4.8), so we choose the coefficients in Eq. (4.72) such that $\overline{y}$ is *exactly* the luminous efficiency. Then we choose the coefficients in Eqs. (4.71) and (4.73) such that the area under all three tristimulus curves is the same. The result, shown in Fig. 4.17, is the set of three curves for the CIE (2°) 1931 *standard observer*. CIE is the abbreviation for Commission Internationale de l'Eclairage (International Commission on Illumination), 2° is the field of view for which the color matching measurements were made, and 1931 is the year in which these values were adopted. The standard observer is not kept in a cage at the National Institute of Standards and Technology in Boulder, Colorado, to be trotted out on ceremonial occasions, but is rather a composite of many observers with normal color vision.

From the tristimulus values $\overline{x}$, $\overline{y}$, $\overline{z}$ for each visible wavelength follow the tristimulus values X, Y, Z for *any* source of visible light with (spectral) radiance L:

$$X = k \int L\overline{x}\, d\lambda, \tag{4.74}$$

$$Y = k \int L\overline{y}\, d\lambda, \tag{4.75}$$

$$Z = k \int L\overline{z}\, d\lambda, \tag{4.76}$$

where k is a constant we choose to suit our fancy. Note that because $\overline{y}$ is the luminous efficiency, Y is proportional to the luminance of the source [see Eq. (4.44)], which could be a *primary* source (e.g., lamp, direct sunlight) or a *secondary* source (light excited by a primary source).

What exactly do X, Y, Z signify? Two light sources with the same tristimulus values for standard conditions (field of view, absolute luminance, surroundings, adaptation, etc.) will be indistinguishable to the standard observer (who does not exist). In general, any set X, Y, Z can be obtained in an *infinite* number of ways, that is, by an infinite set of spectra. A spectrum determines a set of tristimulus values, but not the converse. This undeniable, experimentally verifiable fact, called *metamerism*, shatters fatuous notions about color being simply a synonym for wavelength.

Even though the standard observer is fictitious, real observers with normal color vision will agree that two light sources with the same tristimulus values are nearly indistinguishable. But the tristimulus values do not provide a complete description of the visual appearance of colored objects. The tristimulus values in Eqs. (4.74)–(4.76) are based on radiances, which depend on direction, whereas they could just as well be based on irradiances. Two colored objects may have the same (irradiance) tristimulus values and yet still appear different because of their different textures (and hence different radiances in different directions). For example, a glossy and a matte surface with identical tristimulus values usually would be perceived to be different. And then there is *simultaneous color contrast*: two colored objects identical in all respects may be perceived to have different colors if their surroundings are different. Simultaneous color contrast is often subtle but can be demonstrated dramatically. In a room illuminated only by red light, we asked several people to describe the color of a sheet of orange construction paper on a black backdrop. The unanimous response was red. But when we placed a sheet of red paper beside the orange sheet, everyone described it as orange. The spectrum of the light from the sheet had not changed, only its surroundings.

Because one of the tristimulus values of a source is proportional to its luminance, the other two must specify its chromatic characteristics. This in turn implies that color (as opposed to brightness) has *two* qualities, which could be dominant wavelength and purity. According to the Optical Society of America's definition in *The Science of Color* (p. 42) "*. . . dominant wavelength. . .* is the wavelength that appears to be dominant in the light. . . *purity* may be said to be the degree to which the dominant wavelength appears to predominate in the light." Although easy to grasp, dominant wavelength and purity are not used much by color scientists. Two associated terms are *hue* and *saturation*. Again we turn to *The Science of Color* (p .101) for definitions: "Hue is the attribute most commonly associated with the wavelength or dominant wavelength of the stimulus and designated by such terms as red, yellow, green, blue. Saturation is the degree to which a chromatic color sensation differs from an achromatic color sensation of the same brightness. For instance, a typical pink is a red of low saturation and high brightness, whereas the light from sodium vapor is a yellow of high saturation".

We can make the qualitative concepts of dominant wavelength and purity quantitative by means of the tristimulus values. Because Y is the luminance, we could factor it out by dividing X and Z by Y. Although this is not what is done, dividing X and Y by the sum of all three tristimulus values to obtain the *chromaticity coordinates*

$$x = \frac{X}{X+Y+Z}, \tag{4.77}$$

$$y = \frac{Y}{X+Y+Z}, \tag{4.78}$$

is equivalent. The set of all possible points x and y constitutes a two-dimensional, bounded color space, the *chromaticity diagram*, a geometrical representation of the gamut of colors perceived by someone with normal color vision. The chromaticity diagram is that region

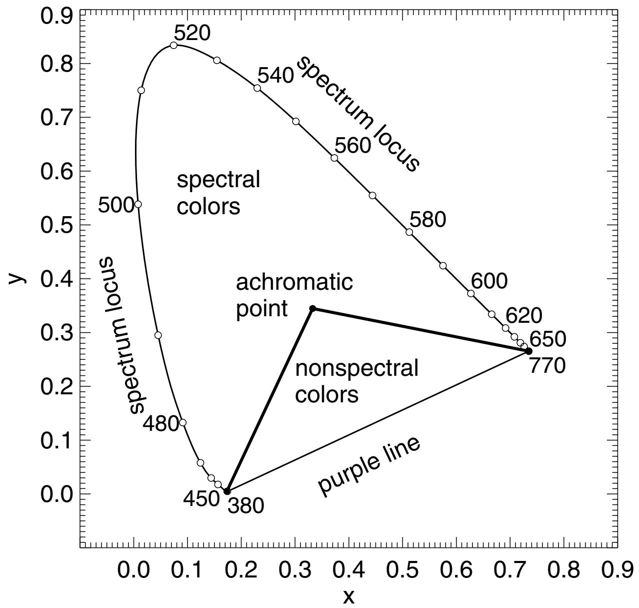


Figure 4.18: The CIE 1931 chromaticity diagram is a color space showing the gamut of colors perceived by the standard human observer. Regions of this color space are labeled by their salient characteristics: spectrum locus (set of all pure colors), spectral colors (those to which a dominant wavelength can be assigned), nonspectral colors (those to which a dominant wavelength cannot be assigned), the purple line (mixtures of violet and red), and the achromatic (or white) point.

in Fig. 4.18 bounded by the curved *spectrum locus*, the set of points corresponding to light sources of a single wavelength (a few wavelengths are labeled for reference), and the *purple line*. Purple is not a spectral color in that it cannot be obtained with a monochromatic source, whereas red, for example, can. Purple is inherently a mixture of violet and red light. The *achromatic* or *white point* is specified by the coordinates of a source we agree to call white. As we show in the following section, there is no such thing as absolute white, and hence the position of the achromatic point is not absolute.

Color space can be divided into two regions, one representing *spectral colors*, the other *nonspectral colors*. The spectral region is that part of the chromaticity diagram defined by all possible straight lines through the achromatic point and connecting two points on the spectrum locus. These two points correspond to *complementary colors*, in the sense that adding them in the proper proportion yields the sensation white. For example, a pure green with wavelength around 492 nm and a pure orange with wavelength around 605 nm are complementary colors: add two monochromatic sources of these wavelengths in the proper proportion and the result is white light (Fig. 4.19). The nonspectral region is that part of the chromaticity diagram defined

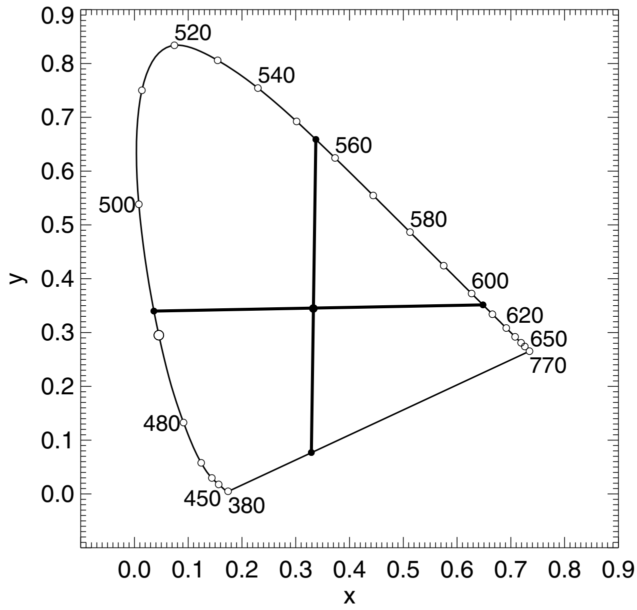


Figure 4.19: The CIE 1931 chromaticity diagram showing complementary wavelengths 492 nm and 605 nm. A suitable mixture of two sources of these wavelengths yields the sensation white. Although points on the purple line do not correspond to specific wavelengths, the complementary wavelength of such a point is defined by that of the wavelength (555 nm in the example shown) that must be subtracted from white light to yield the color on the purple line.

by all possible straight lines through the achromatic point and connecting a point on the purple line with a point on the spectrum locus. Because monochromatic sources of purple light do not exist, we cannot add a pure purple to its complementary pure color (the intersection point on the spectrum locus) to obtain white. But we can *subtract* from white light the complementary color, a green with wavelength 555 nm for the purple shown in Fig. 4.19.

Suppose that a source is represented by a point on the chromaticity diagram. From the achromatic point draw a straight line that intersects the source point and extend this line to the spectrum locus. The wavelength at the intersection point is the dominant wavelength. If a monochromatic source of this wavelength were to be added to white light in the proper proportion, the result would be indistinguishable in color from the source. The term *purity* is used somewhat carelessly (we confess to having committed this sin) without qualification, but there are two purities, one with a simple geometrical definition, the other with a simple photometric definition. The *excitation purity* is the length of the line from the achromatic point to the source point relative to the distance from the achromatic point to the spectrum locus point (dominant wavelength), usually expressed as a percentage. An achromatic source has an excitation purity of 0%, a monochromatic source an excitation purity of 100%. This

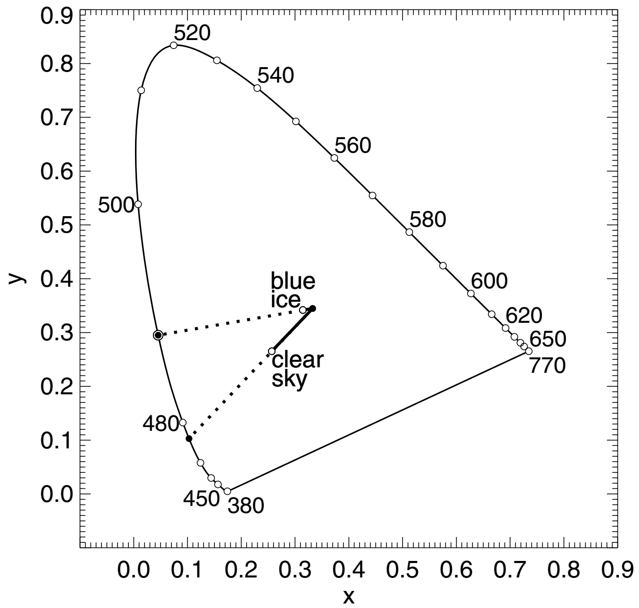


Figure 4.20: CIE 1931 chromaticity diagram with measured colorimetric points for a clear sky and blue ice. The dominant wavelength for the blue ice is 490 nm; that for the blue sky is 477 nm.

purity is a measure of how close the source point is to the spectrum locus, and its value is evident from a glance at the CIE diagram. But it lacks a simple photometric definition. The *chromatic purity* is not evident from the CIE diagram but is defined in what might be called the natural way: if Y_0 is the luminance of an achromatic source and Y_1 that of a monochromatic source such that the sum of these two sources matches the luminance Y and color of the source of interest, its chromatic purity is Y_1/Y. The two purities coincide at the end points (0 and 100%) but elsewhere are different.

Figure 4.20 shows chromaticity points obtained from spectral measurements of blue ice (frozen waterfall) and a clear sky near zenith (see Fig. 5.17 for the spectrum of the blue ice). The excitation purity of the sky is about 33%, its chromatic purity about 13%, and its dominant wavelength about 477 nm, solidly in the blue. The excitation purity of the blue ice is considerably lower, about 6%, its chromatic purity about 5%, and its dominant wavelength about 490 nm, what we might call blue–green. And indeed these calculations are in accord with what can be observed in winter on a clear day. Bubbly ice (a frozen waterfall in a road cut, for example) is a much less vivid blue than the zenith sky, and the ice has a noticeably more greenish cast than the sky.

The CIE diagram helps to support the assertion in Section 1.4.4 about the restricted validity of color temperature. Figure 4.21 shows the curve of chromaticity points for blackbodies with a large range of absolute temperatures. Despite this range, blackbodies can't come close

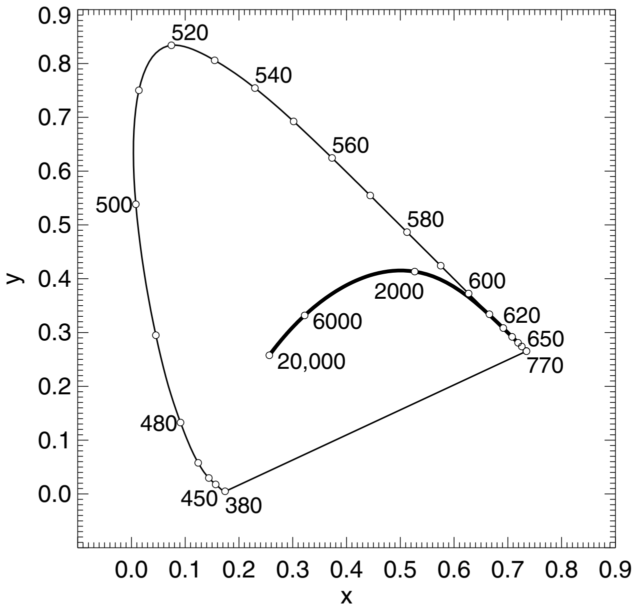


Figure 4.21: CIE 1931 chromaticity diagram showing the locus of colorimetric points for blackbodies with temperatures in the range 1000–20,000 K.

to matching the entire gamut of colors. At temperatures less than about 1000 K, the light from a blackbody is a red of high purity, whence the term red hot. The melting point of carbon is about 3800 K, that of tungsten about 100 K lower, which gives upper limits on the color temperature of heated carbon or tungsten filaments. The boiling point of tungsten is about 6300 K, approximately the color temperature of sunlight. As the absolute temperature increases without limit (don't try this at home) the color of a blackbody eventually becomes blue although never of high purity. A step beyond color temperature is the *correlated color temperature* of a source, the temperature of a blackbody with chromaticity point closest to that of the source. But even correlated color temperature becomes nearly meaningless for points far from the blackbody curve.

Figure 4.22 shows chromaticity points for more than a dozen sources: white light such as daylight reflected by magnesium oxide powder, snow, cloud, paper, fruit, a green dish, dirt and a fluorescent vest; zenith and horizon skylight; black light; the display of a computer set for various colors. The chromaticity points for all the white sources are different but clustered in the same neighborhood. And note that a red apple is far from being a pure red. Even the colors on the display, set to the highest purity, are not on the spectrum locus. The purest color is that for the orange fluorescent vest. The pattern that emerges here is clear: the colors in our everyday lives, even those purposely designed to be vivid, are usually far from being pure. Our visual world is mostly pastel.

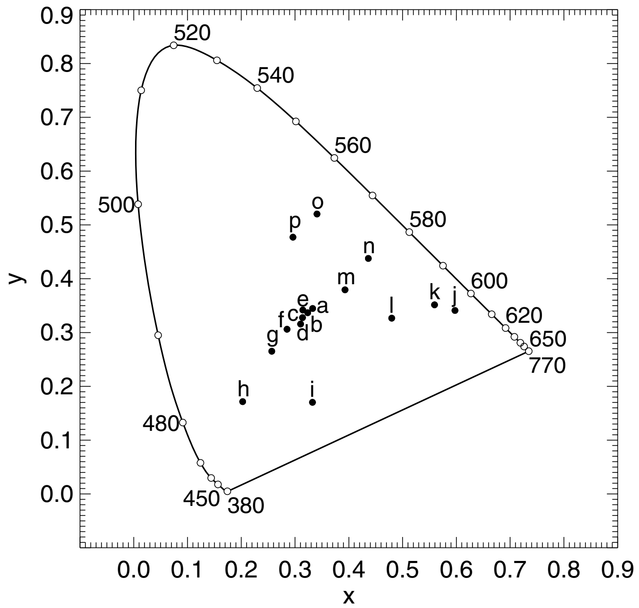


Figure 4.22: CIE 1931 chromaticity diagram with colorimetric points for several different sources: (a) MgO; (b) snow; (c) cloud; (d) thesis paper; (e) blue ice; (f) clear horizon sky; (g) clear zenith sky; (h) blue computer display pixels; (i) black light; (j) orange vest; (k) red computer display pixels; (l) apple; (m) dirt; (n) banana; (o) green dish; (p) green computer display pixels. Note that none of these points are on or even close to the spectrum locus.

Although the CIE 1931 diagram has long been the workhorse of colorimetry, it is being replaced gradually by the CIE 1976 UCS diagram, where UCS stands for *uniform chromaticity scale*. This diagram represents an attempt to make equal displacements correspond to equal perceptual differences. The coordinates (u', v') in this system are obtained from the tristimulus values:

$$u' = \frac{4X}{X + 15Y + 3Z}, \tag{4.79}$$

$$v' = \frac{9Y}{X + 15Y + 3Z}. \tag{4.80}$$

Because of the one-to-one relation between (x, y) and (u', v'), the CIE 1976 UCS diagram contains no new information but just displays old information in a new way. Figure 4.23 shows the CIE 1976 UCS diagram with a curve of blackbody chromaticity points.

Because chromaticity diagrams are infinite sets of points, we might be tempted to conclude that the number of colors is infinite, which is what Newton implied in his statement about colors "perpetually varying." Although wavelength may be said to vary perpetually, color cannot because no human observer can detect differences between sources represented by arbitrarily

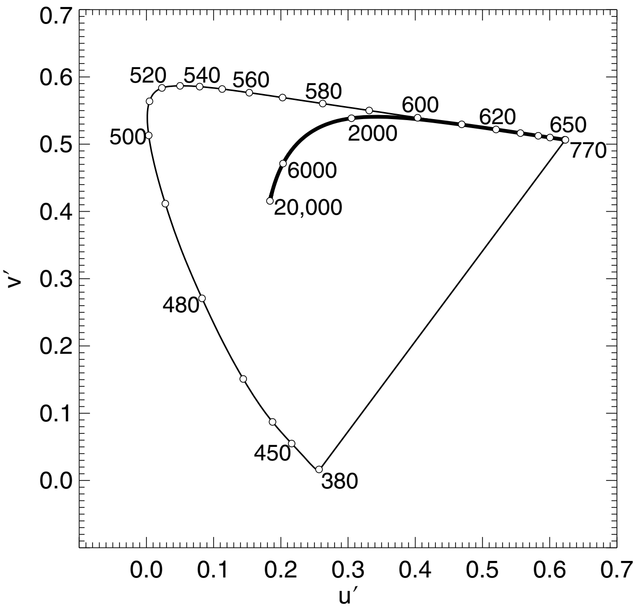


Figure 4.23: CIE 1976 UCS chromaticity diagram with blackbody points as in Fig. 4.21. The 1976 diagram contains no more information than the 1931 diagram but displays it differently.

close points on a chromaticity diagram. Around any point on a chromaticity diagram is a neighborhood of points representing sources indistinguishable to almost all observers. These neighborhoods are the *MacAdam ellipses*, elliptical subsets of color space specified by a fixed precision in color matching. Because of the finite area of these ellipses, the number of perceptually different colors is finite. Ellipses in different regions of color space do not have the same area or shape, and UCS diagrams are attempts to distort color space such that all ellipses are transformed into equal-area circles, an elusive goal not yet or ever likely to be reached. If to the two colorimetric coordinates is added luminance, we obtain ellipsoids in a three-dimensional space that specify the inherent fuzziness of the human observer's ability to distinguish between sources of different color and brightness.

There are an infinite number of visible wavelengths, a large but indeterminate number of perceptually different colors corresponding to fuzzy regions in color space, and a small range of colors words in everyday use in all languages.

4.3.2 The Nonexistence of Absolute White

To end, we return to the seemingly heretical statement in the previous section about the nonexistence of absolute white. White light often is defined as a source with all visible wavelengths in equal proportions, that is, a flat spectrum. We call this physicist's white. It doesn't exist in nature nor can you readily buy a source of such light, although with effort it might be created with carefully chosen filters. More important, this definition is largely irrelevant to the perception of white. This is most easily demonstrated by comparing different spectra of indisputably white light sources. Figure 4.24 shows four such spectra: snow illuminated by daylight, and white paper illuminated by an incandescent bulb, a fluorescent bulb, and a fluorescent tube. Despite the great differences in these spectra, to the human observer they all are perceived as white. Note in particular how much the spectrum of the fluorescent bulb departs from white light defined as a flat spectrum.

If you look at a sheet of white paper illuminated by an incandescent lamp, then take the paper into a room illuminated by a fluorescent lamp, then take the paper outdoors, you still see it as white. It doesn't change color although its spectrum changes. A spectrophotometer does not have a brain, whereas you do, and without your conscious effort it continually maintains color constancy so that under any illumination white paper is perceived to be white. Color constancy was strikingly demonstrated to us on a visit to the National Institute of Standards and Technology in Gaithersburg, Maryland. Scientists there had made a video of a scene illuminated by a source the color temperature of which could be varied continuously over a large range. As we watched this scene with the color temperature of the ambient light rapidly changing by thousands of degrees we barely noticed any differences. But when we compared (by splitting the screen) the scene illuminated by a source at 2000 K with the same scene illuminated by a source at 20,000 K, the difference was striking. The scene at 2000 K was noticeably redder than that at 20,000 K.

Only when you compare two sources of white light side by side can you detect any differences. The spectrum of an incandescent bulb is considerably richer in red light than is the spectrum of, say, snow (Fig. 4.24). To observe this look at the reflection (to reduce the luminance) of an incandescent bulb in a window that looks out onto snow. Here the bulb is noticeably yellowish-orange.

A striking demonstration can be obtained in a room illuminated solely by red light. Despite the red illumination of everything in the room, familiar objects in it known to be white are perceived to be more or less white because they are the brightest objects.

Although the human observer does not perceive any differences between white light sources with markedly different spectra, a camera does because it is not equipped with a brain. Thus if you use outdoor film for indoor photography (or vice versa), your camera cannot adapt, and the result may be photos of your friends and family with toothy, greenish grins. To avoid unnatural colors you must use daylight film for outdoor photography and tungsten film for indoor photography or use daylight film indoors with a blue filter on the camera. Digital cameras face the same problem of reproducing colors as faithfully as possible regardless of the ambient illumination, and are equipped with circuits and algorithms to achieve this end. And photos taken with digital cameras can be processed afterwards to obtain more faithful reproductions of what the human observer sees. The digital camera represents an attempt to more closely duplicate with silicon chips what the eye–brain system does.

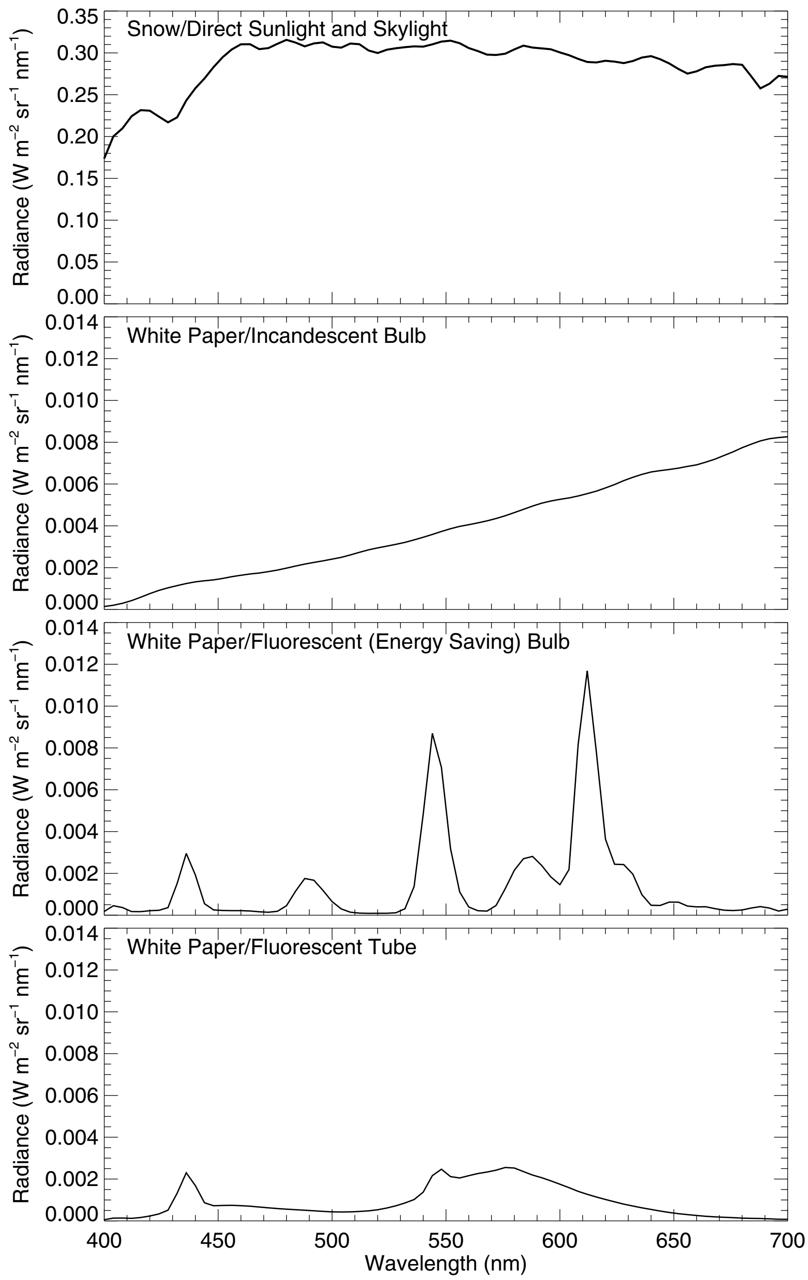


Figure 4.24: Measured spectra of snow illuminated by daylight and white paper illuminated by an incandescent bulb and by two different fluorescent bulbs. All these spectra are objectively quite different but subjectively the same: white.

References and Suggestions for Further Reading

In denoting $\mathbf{E}$ as the electric field (also called electric vector, electric field intensity, electric field strength) and $\mathbf{H}$ as the magnetic field (also called magnetic vector, magnetic field intensity, magnetic field strength), we follow the current convention despite its historical and physical incorrectness. The fundamental fields, the movers and shakers of charges, are $\mathbf{E}$ and $\mathbf{B}$, as evidenced by the Lorentz equation for the force on a charge q with velocity $\mathbf{v}$: $\mathbf{F} = q(\mathbf{E} + \mathbf{v}\times\mathbf{B})$. $\mathbf{B}$ is now often called the magnetic induction, whereas this term should be used for $\mathbf{H}$. In free space $\mathbf{B}$ is proportional to $\mathbf{H}$ by way of a universal (scalar) constant. In material media, however, there is no such general relation, but rather many (assumed) particular relations. For more on this see W. G. Dixon, 1978: *Special Relativity: The Foundation of Macroscopic Physics*, Cambridge University Press, p. 200, also Werner S. Weiglhofer and Akhlesh Lakhtakia, Eds., 2003: *Introduction to Complex Mediums for Optics and Electromagnetics*, SPIE Press, pp. 31 & 348. One author gets around this notational confusion by referring simply to the "$\mathbf{B}$-field and $\mathbf{H}$-field" (Günther Scharf, 1994: *From Electrostatics to Optics: A Concise Electrodynamics Course*, Springer, p. 157).

For a good general reference on radiometry and photometry see Robert W. Boyd, 1983: *Radiation and the Detection of Optical Radiation*, John Wiley & Sons.

For a nonmathematical discussion of solid angle, radiance, irradiance, luminance, brightness and examples of the subjectivity of brightness see Craig F. Bohren, 1991: *What Light Through Yonder Window Breaks*?, John Wiley & Sons, Ch. 15.

Poynting's theorem is derived in treatises on electromagnetic theory, for example, Julius Adams Stratton, 1941: *Electromagnetic Theory*, McGraw-Hill, pp. 131-7; John David Jackson, 1975: *Classical Electrodynamics*, 2^{nd} ed., John Wiley & Sons, pp. 236–7. Poynting's original paper is reprinted in *Collected Scientific Papers by John Henry Poynting*, Cambridge University Press (1920).

We owe the insight that solid angle is a measure to Rudolph W. Preisendorfer, 1965: *Radiative Transfer on Discrete Spaces*, Pergamon, p. 23.

For an illuminating schematic diagram showing how mass density [Eq. (4.17)] depends on volume see Ludwig Prandtl and Oskar G. Tietjens, 1957: *Fundamentals of Hydro- and Aeromechanics*. Dover, Fig. 1, p. 9. A similar diagram is in George Keith Batchelor, 1967: *An Introduction to Fluid Mechanics*. Cambridge University Press. Fig. 1.2.1, p. 5. This figure is fundamental to understanding all densities, that is, limits of quotients, whether explicitly called densities or not.

For more on how the human perceptual system imposes constancy of size, shape, color, and brightness see treatises on perception, for example Herschel W. Leibowitz, 1965: *Visual Perception*, MacMillan; Tom Cornsweet, 1970: *Visual Perception*, Harcourt Brace Jovanovich; Robert Sekuler and Randolph Blake, 1985: *Perception*. Alfred A. Knopf.

For a discussion and photograph of the subsun see Robert Greenler, 1980: *Rainbows, Halos, and Glories*. Cambridge University Press, pp. 73-4, and Alistair B. Fraser, David K. Lynch, and Stanley D. Gedzelman, 1994: Subsuns, Bottlinger's ring and elliptical halos, *Applied Optics*, Vol. 33, pp. 4580-6.

The invariance principle $L/n^2 = \mathrm{const.}$ [Eq. (4.43)] is in a review article by E. A Milne, 1930: Thermodynamics of the stars. *Handbuch der Astrophysik*, Vol. 3, Part I, p. 74 (reprinted in the collection edited by Donald H. Menzel, 1966: *Selected Papers on the Transfer of Radiation*. Dover) but goes back even earlier to Max Planck's *Theory of Heat Radiation* (p. 35), the second German edition of which was published in 1913, its 1914 English translation reprinted by Dover in 1959.

For a discussion of the luminous efficiency curve and the Purkinje effect see Yves Le Grand, 1957: *Light, Colour and Vision*, John Wiley & Sons, Chs. 4 & 6. You can open this book almost at random and find something about vision worth reading.

References relevant to fluorescence are given at the end of Chapter 1.

R. C. Hilborn's amusing, but sadly true, quip about radiometry is in *American Journal of Physics* (Vol. 52, 1984, p. 668).

For a humorous but biting criticism of the sloppy use of intensity see J. M. Palmer, 1993: Getting intense on intensity. *Metrologia*, Vol. 30, 371–2.

The quotation in the first paragraph in Section 4.3 is our translation of Santiago Ramón y Cajal, 1943: *El Mundo Visto a los Ochenta Años*, 5^{th} ed. Espas-Calpe Argentina, pp. 22–23.

To find out what Newton really wrote about colors (or anything) we suggest the radical step of reading his own words. Newton's *Optiks* is readily available as a Dover edition published in 1952.

Keynes's essay, Newton, The man is reprinted in Robert Karplus, 1970: *Physics and Man*, W. A. Benjamin, pp. 22–9.

For an ethnolinguistic view of color words see Brent Berlin and Paul Kay, 1969: *Basic Color Terms: Their Universality and Evolution*, University of California Press. For recent statistical evidence in support of the hypothesis in this book of "a total universal inventory of exactly 11 basic color categories... from which the 11 or fewer basic color terms of any given language are always drawn" see Paul Kay and Terry Regler, 2003: Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences*, Vol. 100, pp. 9085–9. A subsequent paper Terry Regler, Paul Kay, and Richard S. Cook, 2005: Focal colors are universal after all. *Proceedings of the National Academy of Sciences*, Vol. 102, pp. 8386–91 presents evidence that the most likely set of terms are the six corresponding to English white, black, red, green, yellow, and blue.

Books written by committees are rarely worth reading. A notable exception is *The Science of Color* (Optical Society of America, 1963), written by the Committee on Colorimetry of the Optical Society of America.

For a figure (p. 50) showing the CIE diagram with color names assigned to regions in this diagram see Fred W. Billmeyer, Jr. and Max Saltzman, 1981: *Principles of Color Technology*, 2^{nd} ed. John Wiley & Sons.

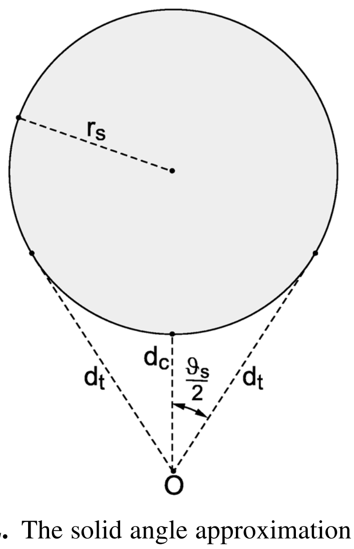
An encyclopedic treatise on colorimetry, chock full of data, is Günter Wyszecki and W. S. Stiles, 1982: *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2^{nd} ed. John Wiley & Sons.

For a detailed discussion of the MacAdam ellipses see David L. MacAdam, 1985: *Color Measurement: Theme and Variations*, 2^{nd} ed. Springer.

For a recent review of colorimetry, which begins with the assertion that "Since perceived color is a property of the human eye and brain, and not a property of physics…colorimetry is inextricably linked" to understanding the biology of vision, see Andrew Stockman, 2004: Colorimetry in *The Optics Encyclopedia*, Vol. 1, Thomas G. Brown, Katherine Creath, Herwig Kogelink, Michael A. Kriss, Joanna Schmit, Marvin J. Weber, Eds. Wiley-VCH, pp. 207-26.

Problems

4.1. What is the solid angle at O of the sphere in the following diagram as a function of the distances d_{c} and r_{s}?



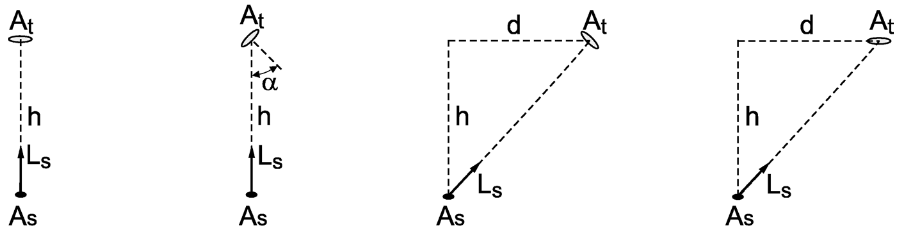
4.2. The solid angle approximation of Eq. (4.6) is equivalent to $\pi\{d_{\mathrm{t}}\vartheta_{\mathrm{s}}/2\}^2/d_{\mathrm{t}}^2$ in the notation of Problem 4.1. Three other approximations are $\pi r_{\mathrm{s}}^2/(d_{\mathrm{c}}+r_{\mathrm{s}})^2$, $\pi\{d_{\mathrm{t}}\sin(\vartheta_{\mathrm{s}}/2)\}^2/d_{\mathrm{t}}^2$, and $\pi\{d_{\mathrm{t}}\sin(\vartheta_{\mathrm{s}}/2)\}^2/\{d_{\mathrm{t}}\cos(\vartheta_{\mathrm{s}}/2)\}^2$. All of these approximate expressions produce similar (and accurate) values if $d_{\mathrm{c}} \gg r_{\mathrm{s}}$. If d_{c} is not much greater than r_{s} which approximate expression is most accurate? To explore this question find d_{c} (in terms of r_{s}) that leads to a 10% error in $\pi r_{\mathrm{s}}^2/(d_{\mathrm{c}}+r_{\mathrm{s}})^2$ relative to the exact solid angle. For this value of d_{c} rank the performance of the four approximate expressions. Using a figure similar to the one in Problem 4.1, give geometrical interpretations of the four approximations and use them to provide plausible explanations for the rankings obtained.

4.3. In the derivation of Eq. (4.5) we implicitly assume that Earth is a point. By how much does the solid angle subtended by the sun at Earth vary when we account for its finiteness? Consider only daylight locations on Earth. Take the radius of the sun to be 6.6×10^8 m, that of Earth to be 6.4×10^6 m, and the distance between the centers of Earth and sun to be 1.5×10^{11} m.

4.4. If you are inside a hollow sphere what is its solid angle? What is the solid angle of the sphere if you are just outside its surface? What is the solid angle of an infinite plane if you are 1 m from it? What if you are 10^{11} m from it? If you are inside a hollow cube what is its solid angle? What about just outside the cube and located at the center of one of its faces? Just outside the cube and located at the center of one of its edges? And what about just outside the cube at one of its corners?

4.5. A disc of radius r_{s} is a source of uniform and isotropic radiance L_{s}. A target disc of radius r_{t} is at a distance d_{st} from the source disc. The centers of these two discs lie on a common axis perpendicular to both with $r_{\mathrm{s}}, r_{\mathrm{t}} \ll d_{\mathrm{st}}$. Show that the radiant power incident on the target disc is $L_{\mathrm{s}}(2\pi^2 r_{\mathrm{s}}^2)(1-\cos\vartheta_{\mathrm{st}})$, where $\vartheta_{\mathrm{st}} = \sin^{-1}\{r_{\mathrm{t}}/\sqrt{r_{\mathrm{t}}^2+d_{\mathrm{st}}^2}\}$, and that a good approximation to the radiant power is $L_{\mathrm{s}}\pi^2 r_{\mathrm{s}}^2 r_{\mathrm{t}}^2/d_{\mathrm{st}}^2$.

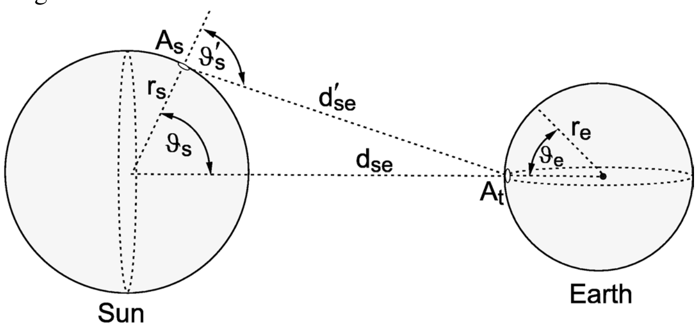
4.6. In the accompanying diagrams the isotropic radiance from disc A_{s} is L_{s}. Derive expressions for the radiant power incident on disc A_{t} in terms of A_{s}, A_{t}, h and d assuming that h is much greater than the dimensions of A_{t} and A_{s}.



4.7. A disc is a uniform source of isotropic radiance. What is the total power through a target disc? The two discs are parallel to each other, and a common axis goes through their centers. The distance h between them is small compared with the target disc radius r_{t} but large compared with the source disc radius r_{s}. Is the irradiance over the target disc uniform? Why or why not? Now change this problem. Find the irradiance at the center of the target disc for arbitrary source disc radius and distance between the two discs. Check your solution against limiting cases, which is always good advice when you have solved any problem.

4.8. Assume the sun to be an isotropic emitter with uniform radiance L_{s}. Show that the expression for the irradiance at A_{t} (with $\vartheta_{\mathrm{e}} = 0°$) is $L_{\mathrm{s}}\pi r_{\mathrm{s}}^2/d_{\mathrm{se}}^2$. There are three ways to solve this problem. First, use the irradiance at the surface of the sun and spherical symmetry;

radiance is not needed. In the second approach use only Eq. (4.49) and appropriately adjusted limits of integration to separate those angles for which L_s is zero and not zero. Finally, solve this problem the hard way. This approach requires both setting up a double integral with the proper limits over that part of the sun that emits to Earth and using the angle ϑ_s' and the distance d_se'. This integral can be evaluated analytically for L_s uniform and isotropic, but would have to be evaluated numerically if the variation of L_s over the sun were taken into account. If evaluating the exact integral proves to be too difficult, try evaluating it approximately for $r_\mathrm{s}/d_\mathrm{se} \ll 1$. How would the solar irradiance at A_t change if it were located at latitude ϑ_e in the diagram?



4.9. We are so accustomed to writing the solar irradiance as $L_\mathrm{s}\Omega_\mathrm{s}$, where L_s is the (assumed uniform and isotropic) radiance of the sun and Ω_s its solid angle subtended at Earth, that it is easy to forget that this expression is approximate if we use the exact expression, Eq. (4.5), for the solid angle. How much error do we make when we use the exact solid angle to determine the solar irradiance?

4.10. To our eyes the clear night sky is black even with a full moon. But film can respond to light our eyes cannot by increasing the exposure time. Thus one should be able to photograph a blue moonlit sky by attaching a camera to a tripod, pointing the camera at a patch of moonlit sky, opening the shutter and leaving it open. The only question is, How long? To answer this question, we pointed a camera at a patch of blue sunlit sky and noted that, for the film and aperture (f-stop) chosen, the exposure time selected by the camera's light meter was 1/250 s. Estimate the exposure time in moonlight (for the same film and aperture). Assume a full moon and take the reflectivity of the moon for visible light to be 10%.

4.11. Several years ago the rules for hunting in Pennsylvania were changed. Previously, hunters were not allowed to fire a rifle within 100 yards ($\approx$100 m) of a residence; now the minimum distance is 150 yards ($\approx$150 m). Assume that there are open fields near a house and that a single hunter is standing at the minimum distance. How has the new regulation changed the probability that a resident of the house would be hit accidentally by a bullet fired by this hunter? Needless to say, the hunter is not trying to shoot the resident. If trees surround the house, a bullet might hit a branch before reaching the house. Assume that the probability a bullet travels 100 yards without hitting a branch is 0.5. What is the probability that it travels 150 yards without hitting a branch? State all your assumptions. NOTE: This problem is not about hunting.

4.12. Suppose that you have been given the task of measuring the irradiance from a snowpack illuminated by direct sunlight. Take the detector to be a circular disc of diameter D. This detector necessarily casts a shadow on the snow, and the shadow region does not contribute much to the irradiance measured by the detector. Thus the presence of the detector introduces an error in the measurement of the upward irradiance. Estimate the height h above the snowpack that you must place the detector so that this shadow error is a few percent. Take the sun to be directly overhead and a parallel source of light. Don't worry about minor details. Give the quickest and dirtiest – but still correct – answer you can come up with.

4.13. We once received a message from a woman in Michigan who wondered "why water in lakes and streams appears much darker, often black in the winter". What is a possible explanation for what she observed?

HINT: This explanation makes no use of any detailed knowledge about lakes and streams in Michigan.

4.14. The following was taken from an advertisement for an "ultra bright laser pointer": "The 650 nm (nanometer, wavelength) laser light is five times brighter than 670 nm laser because the 650 nm is closer to the 633 nm of red light." Discuss.

4.15. A *pyranometer* measures the time-varying downward solar irradiance during daylight hours. On days with broken clouds the irradiance sometimes is observed to be greater than it would be if the sky were clear. Explain. What is the upper limit of the factor by which the measured irradiance can be greater when broken clouds are present than when the sky is clear, all else being equal? For simplicity, you may take the sun to be directly overhead.

4.16. Why does the temperature of an electric stove coil not increase without limit after it is turned on? Estimate the distance (relative to the size of the coil) from the coil at which the rate you are radiatively heated by it is comparable with the rate you are radiatively heated by the surroundings. You may assume a coil temperature of 1500 K. Make any reasonable assumptions. Elaborate derivations and detailed calculations are neither necessary nor desirable.

4.17. Some people still believe that seasonal temperature variations in midlatitudes are a consequence of the eccentricity of the Earth's orbit around the sun (i.e., the distance between the Earth and sun varies over the year by about 3%). What simple but convincing quantitative arguments can you make to show that variations in the distance between Earth and sun are not the cause of the seasons? What simple but convincing quantitative arguments can you make to show that the true cause of seasonal temperature variations is the obliquity of the Earth's axis (i.e., the rotational axis is inclined by about 23 degrees to the plane of the Earth–sun orbit)?

HINT: The easiest way to do this problem is to base your arguments on surface temperatures on the moon, which has no atmosphere or oceans to complicate matters.

4.18. We note in Section 1.6 that Earth, from the point of view of an observer on the moon, is an infrared sun with an effective blackbody temperature of around 255 K. What is the corresponding terrestrial irradiance at the moon? The distance from Earth to moon is about 60 Earth radii.

4.19. The albedo of Earth was known before there were artificial satellites (the moon is a *natural* satellite). Given that the luminance of the full moon is 9300 times that of the new moon, estimate Earth's albedo. This is revisited in Problem 5.23.

Determining Earth's albedo by measuring the *earthshine* (faint moonlight resulting from sunlight reflected by Earth) has a long and interesting history and is a topic of considerable current interest. For more on this see André Danjon, 1954: Albedo, color, and polarization of the Earth, in *The Earth as a Planet*, Gerard P. Kuiper, Ed., University of Chicago Press, pp. 726–38; Donald R. Huffman, Charles Weidman, and Sean Twomey, 1990: Repetition of Danjon earthshine measurements for determination of long term trends in the Earth's albedo. *Colloque André Danjon*, N. Capitaine and S. Debarbat, Eds., Observatoire de Paris, Journees 1990, Systemes de Reference Spatio-Temporels, Paris, 28-29-30 Mai, pp. 111–6; P. R., Goode, J, Qiu, V. Yurchyshyn, J. Hickey, M-C. Chu, E. Kolbe, C. T. Brown, and S. E., Koonin, 2001: Earthshine observations of the Earth's reflectance. *Geophysical Research Letters*, Vol. 28, pp. 1671–4.

4.20. Estimate the ratio of the total visible radiance from black paper (with reflectivity 10%, say) illuminated by bright sunlight to the radiance of white paper (with reflectivity 90%, say) illuminated by a 100 W light bulb at the kind of distances you would be from a reading lamp. What does this ratio tell you about the meaning of the terms black and white?

4.21. Estimate how much the solar irradiance (over a small area) can be increased with an ordinary magnifying glass (e.g., the kind used to read maps). All you need is a magnifying glass and ruler to answer this. Then estimate the maximum temperature (again, over a small area) to which an object might be raised by placing it at the focal point of a magnifying glass exposed to direct sunlight. Be sure that your answer makes sense.

4.22. Estimate by how much subsuns are brighter than clouds.

4.23. Using only your eyes investigate the degree to which an ordinary blackboard or white paper is a diffuse (isotropic) reflector of visible radiation.

4.24. Is the solid angle subtended by the sun at Earth the same as the solid angle subtended by Earth at the sun?

4.25. A colleague forwarded us the following email message (spelling and grammatical errors have not been corrected) that he, and many other people, received:

> "Everyone should mark their calendars this month for the 'Last Lunar Harrah' of the Millennium: This year will be the first full moon to occur on the winter solstice, Dec. 22... Since a full moon on the winter solstice occurred in conjunction with a lunar perigee (point in the moon's orbit that is closest to Earth), the moon will appear about 14% larger than it does at apogee (the point in its elliptical orbit that is farthest from the Earth). Since the Earth is also several million miles closer to the sun at this time of the year than in the summer, sunlight striking the moon is about 7% stronger, making it brighter. Also, this will be the closest perigee of the Moon of the year since the moon's orbit is constantly deforming. If the weather is clear and there is a snow cover where you live... it is believed that even car headlights will be superfluous.
>
> In layman's terms it will be a super bright full moon, much more than the usual AND it hasn't happened this way for 133 years! Our ancestors 133 years ago saw this. Our descendants 100 or so years from now will see this again."

Quantitatively evaluate the assertion that the moon will be "super bright" to the extent that driving in moonlight without headlights would be possible.

HINT: Determine by how much the lunar irradiance at Earth's surface increases.

4.26. What is the total amount of light scattered by a particle with scattering cross section C_{sca} at a height h above an infinite plane surface at every point of which the isotropic radiance is L? Now suppose that this surface is a diffuse reflector with reflectivity 1 illuminated normally by sunlight. What is the ratio of the total scattering of light from the surface to total scattering of direct sunlight?

HINT: The first part of this problem is done most easily in a cylindrical polar coordinate system with the z-axis passing through the particle and normal to the surface.

4.27. What would be the range of wavelengths (frequencies) over which the human eye is sensitive to electromagnetic radiation if the number of octaves were the same as that for the range of frequencies over which the human ear is sensitive to acoustic radiation (sound)? Take the center of this range of wavelengths to be such that the number of octaves above the middle of the visible spectrum equals the number below. To answer this question you will have to learn the range of audible sound frequencies and what is meant by an octave (how many octaves on a piano?).

4.28. Can you think of reasons why the human eye would not have evolved to respond to the range of wavelengths calculated in the previous problem?

HINT: Figure 2.25 may help.

4.29. In *Optical Treatise on the Gradation of Light*, published in 1760 (an earlier version appeared in 1729), Pierre Bouguer asserted that "at a depth of about 311 ft in sea water, the light of the sun becomes equal to that of the full moon seen at the surface of the earth." Verify Bouguer's assertion. Make any reasonable assumptions. All that is wanted is a rough check. Was Bouguer at least approximately correct (more than 240 years ago) or was he hopelessly wrong? Figure 5.12 should be helpful.

4.30. A man once told us that at a summer dog show in Arizona he was advised by a judge to never wet dogs to cool them. Water drops on the fur act as little lenses that greatly increase the rate of solar heating of the dog, and hence the dog's temperature would rise. What do you think of this advice? How would you give different advice in a way that could be understood by someone with little scientific training? Can you think of any convincing experiments or demonstrations to support your advice?

4.31. Show that combining any two visible sources S_1 and S_2 in any proportion yields colorimetric coordinates (x, y) that lie along the straight line between the respective colorimetric coordinates (x_1, y_1) and (x_2, y_2) of the sources.

HINT: For this and related problems the safest approach is to begin with X, Y, and Z.

4.32. Use the results of Problem 4.31 to show that the perceived color of any source of light, except a strictly monochromatic source (i.e., one with chromaticity coordinates lying on the spectrum locus), can be obtained in an infinite number of ways.

4.33. Use the results of Problem 4.31 to show that any spectral color can be obtained by adding white light to a monochromatic source.

4.34. Use the results of Problem 4.31 to show that any nonspectral color can be obtained by subtracting from white light a monochromatic source with wavelength equal to that of the complementary color of the nonspectral color.

4.35. Derive an expression for the excitation purity as a function of the colorimetric coordinates of the achromatic point, of the spectrum of interest, and the dominant wavelength. Derive a similar expression for the chromatic purity. Finally, determine the relation between excitation and chromatic purity. At what points are the two purities the same? Where is the difference between them greatest?

4.36. Can the perceived color we call blue be obtained by superposing sources of light that contain *no* blue light in the sense of light with wavelengths between 450 and 490 nm? If so, how?

HINT: Use the CIE diagram.

4.37. We stated in Section 4.3 that the light from three different (real) lamps cannot be added so as to match all visible wavelengths. This is equivalent to saying that three lamps cannot yield all points of color space. What about adding three laser beams of different wavelengths? If three won't do the job, what about four? If not four, what about five, and so on?

HINT: Use the CIE chromaticity diagram. Laser beams are represented by points on the spectrum locus. An analytical proof is not needed (perhaps not possible) here. You can answer these questions with crude sketches.

4.38. This problem is related to the previous one. For what shape of color space (chromaticity diagram) could all visible wavelengths be matched by combining three laser beams?

4.39. First do Problem 4.37, then answer the following questions. Would it be possible using various (finite) combinations of inks or dyes to obtain the full gamut of colors in color space? Is truly full-color photographic film possible (i.e., film that exactly reproduces the colors seen by humans)? What about truly full-color television?

4.40. Derive an expression for transforming from (x, y) color space to (u', v') color space. Is the excitation purity, defined as a ratio of distances on the CIE 1931 chromaticity diagram, the same as an excitation purity defined in a similar way for the CIE 1976 UCS chromaticity diagram?

4.41. G. I. Taylor is well known to meteorologists because of his many contributions to meteorology. What is less well known is that his first piece of research, done while he was a graduate student, was, in effect, photography photon by photon for very long times. That is, Taylor adjusted the irradiance of his light source so that only one photon at a time occupied the volume of his apparatus. For sake of argument, assume that this volume was that of a box 10 cm on a side (the approximate volume of a camera) with the source being one side of the box. Estimate the irradiance of Taylor's light source such that only one photon at a time occupies this volume.

HINT: See Problem 1.1 in Chapter 1.

4.42. Estimate the brightness temperature of the full moon at, say, the middle of the visible spectrum. You may take the visible reflectivity of the moon to be 10%. Estimate the color temperature of the moon. If these two temperatures are different, explain. Describe the color of a blackbody at the brightness temperature you obtain for the full moon. Explain any difference between this color and that of the moon.

4.43. You can buy at grocery stores Food Colors & Egg Dyes, four small (29 ml) vials containing different dyes (red, yellow, green, and blue). The color associated with each vial is indicated by the color of its cap. But the vial with the yellow cap presents us with a puzzle. The dye in this vial is red, and yet a few drops added to a glass of water turns the water yellow. Explain.

4.44. This problem is related to the previous one. A drop or two of the yellow dye added to a glass of water colors the water yellow. And a drop or two of the blue dye added to a glass of water colors the water blue. But if you look at the glass of yellow water through the glass of blue water (or vice versa), you see what is unquestionably green. Explain. This has a practical application: plastic re-sealable zipper bags with seals composed of two parts, a grooved yellow strip, and its blue mate. When the two are snapped together tightly the result is a green strip.

4.45. *Fahrenheit 451*, by the science-fiction writer Ray Bradbury, is an anti-utopian novel about a future world in which firemen, instead of putting out fires, burn books. The title comes from the *ignition temperature* of paper, the temperature to which it must be raised in order to burst into flames. What is the *minimum* (net) irradiance of direct sunlight illuminating paper in order for it to reach this temperature? Show that even allowing for a high reflectivity of paper for solar radiation, attenuation by the atmosphere, reflection and absorption by a magnifying lens, such a lens can focus solar radiation to yield irradiances well in excess of this minimum. To do this problem you'll have to get your hands on a magnifying lens. And once you have it, do some paper-burning experiments.

4.46. No one should measure any quantity without knowing beforehand its expected range of values. Estimate the maximum luminance of a diffusely-reflecting object illuminated by daylight.

HINT: Figure 1.5 should be helpful. All that is wanted here is a rough estimate.

4.47. This problem is related to the previous one. Estimate the maximum luminance of the (full) moon. You may take the visible reflectivity of the moon to be 10%. Then estimate the maximum luminance of a snowpack in moonlight. What do these values and that obtained in the previous problem tell you about the dynamic range of the human eye to different magnitudes of luminance?

4.48. This problem is related to the previous one. Estimate the luminance of a snowpack illuminated by a crescent moon (just before the new moon). To do this problem requires a moon calendar showing phases of the moon.

4.49. According to electromagnetic theory, which underlies all of optics, the fundamental measureable quantity (other than the separate electric and magnetic fields) is the Poynting vector, which is essentially an irradiance. Yet according to radiometry, the fundamental quantity is radiance, from which irradiance can be obtained by integration. This seems contradictory. Explain.

4.50. In order to check our result for Problem 4.47, we measured the luminance of the (nearly) full moon on a fairly clear night. But a problem we faced is that the field of view of our instrument is one degree whereas the moon subtends half a degree. How did we account for this?

HINT: The previous problem should be of help.

4.51. What is the relation between the spectral (or total) energy density u of an isotropic radiation field (e.g., the radiation field inside a cavity in equilibrium) and the spectral (or total) radiance?

4.52. This problem is related to the previous one. We stated without proof in Problem 1.40, that the pressure of a photon gas (for cavity radiation) is one-third its (total) energy density. Prove this. You will need the result from Problem 1.23.

HINT: Determine the total pressure on a specularly reflecting wall with 100% reflectivity.

4.53. This problem requires working the previous two. What must be the total radiance of equilibrium radiation (within a cavity) such that its pressure is one atmosphere? How does this radiance compare with that of the sun?

4.54. The *Bond albedo* for a spherical body is defined as the ratio of the total amount of reflected radiation to the total amount of incident (monodirectional) radiation. Derive an expression for the Bond albedo for a spherical body with a reflectivity R that depends on the angle of incidence. Assume azimuthal symmetry and uniformity of the body (reflectivity does not depend on position). Show that this albedo is the same as that for an isotropic source illuminating the body at any point.

4.55. We showed at the end of Section 4.2.1 that total irradiance is conserved across an air–water interface. In so doing we assumed that the transmissivity from air to water is close to 1 (reflectivity negligible). The justification for this was that the reflected (isotropic) irradiance is about 6% of the incident (isotropic) irradiance. Now let's turn this problem around. Suppose that a source of isotropic radiation is in the water. Is total irradiance still conserved across the interface using logic similar to Eqs. (4.62)–(4.64)? If not, what needs to be done to ensure that it is conserved?

HINT: Divide the angles of incidence into those less than the critical angle and those greater, and make a simple approximation.

4.56. A factor is missing from each of the terms in Eq. (1.65), but this makes no difference because it is common to all of them. What is this factor and why do you think that we omitted it?

4.57. Find the radiance from an infinite, diffusely reflecting plane when illuminated by a uniform and isotropic source disc at a distance z from the plane and parallel to it. The dimensions of the area A of the disc are $\ll z$. Express your answer as a function of the angle between the normal to the source disc and the line connecting the center of the source disc to a point on the plane.

4.58. Consider an infinite planar source, isotropic and uniform, above which and parallel to it is an infinite planar isotropic reflector with reflectivity R. The two planes are displaced relative to each other by a distance d along the z-axis. The source plane occupies the region in the xy-plane $-\infty < x < \infty$, $0 < y < \infty$, whereas the reflector occupies the region $-\infty < x < \infty$, $-\infty < y < 0$. Ignoring multiple reflections between the two planes find the reflected radiance at every point of the reflector. Your answer can be expressed succinctly as a function of an angle and checked using the known result for a source plane $-\infty < x < \infty$, $-\infty < y < \infty$.

4.59. To avoid wallowing in terminology, the bane of radiometry and photometry, we do not define *hemispherical emissivity* in Chapter 1, although we do note that emissivity depends on direction, in general. Hemispherical emissivity is the ratio of emitted irradiance to the Planck irradiance, whereas directional emissivity is the ratio of emitted radiance to the Planck radiance. Show how the hemispherical emissivity can be obtained from the directional emissivity. Assume azimuthal symmetry, that is, emission depends only on the angle between the direction of emission and the normal to the emitting surface. First guess what the relation might be, then obtain it carefully. If there is a discrepancy between your guess and the result of a careful analysis, ask yourself why.

4.60. On page 92 of *Mathematics: The Loss of Certainty* (1980, Oxford University Press) Morris Kline adduces the following example that "the notion of equality cannot be applied automatically to experience": "... colors a and b may seem to be the same as do colors b and c but a and c can be distinguished." Explain. We note that Kline wrote many excellent books on mathematics worth reading.

4.61. On page 131 of Dudley Towne's *Wave Phenomena* (cited at the end of Ch. 3) we find the following problem: "White light is normally incident from air upon the surface of a medium of high dispersion, the index of refraction for blue light being greater than that for red. Describe the color of the reflected light." Answer this question.

HINT: You will need to bring together ideas from Chapter 3, this chapter, and possibly even refractive index data for real materials. We suspect that our answer to Towne's question is not that expected by him. This is a simple question with a complex answer.

4.62. Consider a white disc of radius a. The disc is a uniform source of isotropic radiation of radiance L_0. Scattering and absorption by the medium between the disc and a point of observation is negligible. As the disc is moved away from this point, in what sense does the radiance of the disc change? This should become obvious if you sketch the radiance as a function of all relevant variables. How does your answer depend on the point of observation?

HINT: This is not a difficult problem but it does require recognizing that radiance is a function. Begin this problem by first understanding the second part of Problem 4.7and then move the point in the second part of Problem 4.7 away from the center of the target disc.

4.63. Consider a cylindrical tube of radius a, the inside walls of which are black. A uniform, isotropic source with radiance L_0 is placed at one end of the tube ($z = 0$). What is the radiance at any depth z in the tube? Is the radiance function at a given z the same at all radial points $0 \leq r \leq a$? What about the irradiance? Obtain an approximate expression for the (average) irradiance as a function of z for $z \gg a$. Try to devise a simple experiment to test your result, at least qualitatively.

4.64. We considered one explanation for the yellow of fog lamps in Problem 3.19. Here is another: the human eye is more sensitive to yellow light. Discuss. In particular, suppose that we were to place a yellow filter over a white headlight. Would the luminance change and, if so, in what direction? Under what simple condition could you be reasonably certain that the luminance of a yellow headlight was greater than that of a white headlight? Why did we say "reasonably" rather than absolutely? For more on fog lamps see J. H. Nelson, 1938: Optics of headlights. *Journal of Scientific Instruments*, Vol. 15, pp. 317–22. You might want to revisit this problem one more time after reading Sections 8.2 and 8.4.1.

4.65. A pinhole camera is, in essence, an opaque screen with a tiny hole in it. The hole serves the same purpose as a lens in that it results in a one-to-one transformation between object and image. The hole must be small for a sharp image. How does the radiance of the sun's image seen on the image plane (assumed to be a diffuse reflector) of the pinhole camera depend on the distance from the pinhole to this plane and other relevant physical variables? What does this result tell you about the limitations of a pinhole camera? You can do a simple experiment to verify your result. Poke a hole in a sheet of aluminum foil with a pin, image the sun onto a sheet of paper, and vary its distance from the hole.

4.66. The ASA number on film indicates how "fast" it is: the greater the ASA number, the shorter the exposure time (for a given light source and aperture size). Typically, one might use ASA 100 or 200 for outdoor photography. What ASA would be required for a pinhole camera to be able to duplicate what an ordinary camera (with a lens) does? You'll have to make some rough measurements to answer this question.

4.67. A lightning flash is an intense but very brief source of radiation. Approximate a vertical flash as a cylindrical source of radius R and length H, where $H \gg R$. Assume that emission by the flash is uniform in space over H and uniform in time over a time interval Δt. During this time N_{p} photons in the frequency interval $\Delta\nu$ centered on ν ($\Delta\nu \ll \nu$) are emitted isotropically. What is the irradiance of the flash? What is the corresponding radiance? Suppose that an irradiance detector is at a distance D ($R \ll D \ll H$) from the center of the flash, in a horizontal plane that intersects the flash at its midpoint. In clear air (scattering and absorption negligible), what is the time interval over which the detector will receive radiation from the flash? What is the average irradiance measured by the detector during this time interval? Describe qualitatively how this time interval and average irradiance change if the flash occurs in a cloud, in which the detector is also embedded, for which $\beta D \gg 1$, where β is the scattering coefficient (absorption assumed negligible). Also describe how the directionality of the radiation field at the detector changes because of the cloud.

4.68. Suppose we assume that there exists a spectrum we call "physicist's white", namely a perfectly flat spectrum over the visible. We still have to specify what we mean by flat, that is, for what variable (frequency, wavelength, etc.) is the spectrum flat? Suppose that it is flat as a function of frequency. What is the corresponding wavelength spectrum? Suppose that the spectrum is flat as a function of wavelength. What is the corresponding frequency spectrum?