Chapter 1

The Description of Radiation

In this chapter we'll consider how to describe radiation and also consider the special characteristics of radiation in thermodynamic equilibrium.

Levels of Description

Depending on the properties we wish to emphasize and the level of accuracy and approximation that we are willing to accept, we can describe a radiation field in one of several ways:

1. As a quantum-mechanical field. This is useful when we wish to consider the interaction of radiation with matter at the microscopic level – the emission, absorption, or scattering of individual photons by individual atoms, molecules, or electrons. This description is the most accurate, but it is often not well suited to describing macroscopic phenomena.

2. As a classical electromagnetic field. This description is familiar from undergraduate physics, but it is not especially useful in stellar atmospheres, as wave phenomena are only important for far-infrared and radio waves, which account for a negligible fraction of the total luminosity. Still, it is worth keeping in mind the possibility of wave phenomena when considering other applications of radiation transfer.

3. As a semi-classical gas of photons. Again, this description is familiar from undergraduate physics. We consider the radiation to consist of a gas of photons traveling in straight lines at speed c and only being destroyed or created by discrete interactions with matter. This is useful as a bridge between the quantum mechanical and thermodynamic descriptions.

4. As a flow of energy. This thermodynamic description is unlikely to be familiar from undergraduate physics, as undergraduate classical thermodynamics courses typically deal only with matter. However, considering radiation in this manner is extremely useful when considering the macroscopic thermodynamics of the atmosphere, such as the restriction that the atmosphere must be in thermal equilibrium.

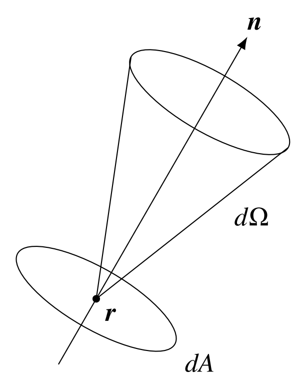

Figure 1.1: The geometry used in the definition of the specific intensity.

We will use both the photon gas and energy flow descriptions in stellar atmospheres. In reality, the two are closely linked, as photons carry energy $h\nu$ at a speed c, so converting from one to the other often involves little more than multiplication or division by factors of $h\nu$ and c.

The Specific Intensity

In radiation transfer we most commonly work with a somewhat unusual macroscopic thermodynamic quantity called the specific intensity. The reasons for this should soon become apparent.

Definition

Consider Figure 1.1. The specific intensity $I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t)$ at a position $\boldsymbol{r}$, in a direction $\boldsymbol{n}$, at a frequency ν, and at a time t is such that the energy dE transported by radiation across an area dA, centered on $\boldsymbol{r}$ and perpendicular to $\boldsymbol{n}$, into a solid angle $d\Omega$ about $\boldsymbol{n}$, in a frequency interval $(\nu, \nu + d\nu)$, in a time interval $(t, t + dt)$ is given by

$$dE = I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t)\, dA\, d\Omega\, d\nu\, dt. \tag{1.1}$$

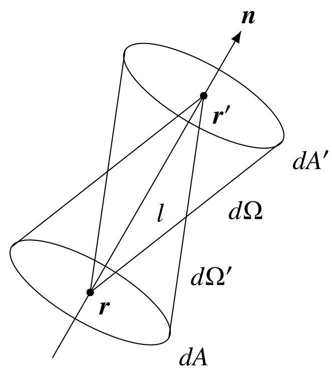

Figure 1.2: The geometry used in the proof of the conservation of specific intensity.

The conventional units of I_ν are erg s^{-1} cm^{-2} Hz^{-1} sr^{-1}. The specific intensity is sometimes called the brightness.

Another form of the specific intensity is I_λ, the specific intensity per unit wavelength instead of per unit frequency. The definition of I_λ is analogous to that of I_ν, but λ and $d\lambda$ replace ν and $d\nu$. The conventional units of I_λ are erg s^{-1} cm^{-2} Å^{-1} sr^{-1} or erg s^{-1} cm^{-2} μm^{-1} sr^{-1}. Since the energy dE transferred is the same, regardless of whether we consider I_ν or I_λ, we have

$$I_\nu\, d\nu = I_\lambda\, d\lambda, \tag{1.2}$$

where $d\nu$ and $d\lambda$ refer to the same range of photons, or

$$I_\lambda = I_\nu \left|\frac{d\nu}{d\lambda}\right| = I_\nu \frac{c}{\lambda^2}. \tag{1.3}$$

When converting between I_ν and I_λ, we have to be careful with the units of λ. For example, if I_ν is in the conventional units of erg s^{-1} cm^{-2} Hz^{-1} sr^{-1} and we want I_λ in the conventional units of erg s^{-1} cm^{-2} Å^{-1} sr^{-1}, we must multiply by c in cm s^{-1} and then divide once by λ in cm and then again by λ in Å.

The quantities I_ν and I_λ are also known as the monochromatic specific intensities, to contrast them with the total or integrated specific intensity I, defined by

$$I \equiv \int_0^\infty d\nu\, I_\nu. \tag{1.4}$$

These conventions, using a subscript of λ instead of ν to specify a monochromatic quantity per unit wavelength instead of per unit frequency and dropping the subscript to specify a total or integrated quantity, apply to all quantities derived from the specific intensity.

The Conservation of Specific Intensity

An important property of the specific intensity is that in free space it is conserved along a ray. Consider Figure 1.2, which shows two points on a ray, $\boldsymbol{r}$ and $\boldsymbol{r}' = \boldsymbol{r} + l\boldsymbol{n}$ and two areas at those points, normal to the ray, dA and dA'. We will consider all the photons in the frequency interval $(\nu, \nu + d\nu)$ that pass through dA in the time interval $(t, t + dt)$ and later pass through dA' in the interval $(t', t' + dt)$. Clearly, the photons pass through dA' a time l/c after they pass through dA, so $t' = t + l/c$.

In free space, and ignoring the gravitational redshift (see Problem 1.1), gravitational lensing, and very rare photon-photon scatterings, the energy carried by these photons will be conserved. More precisely, the energy dE that passes through dA in the time interval $(t, t + dt)$ in the frequency interval $(\nu, \nu + d\nu)$, at angles that will take it through dA' will be equal to the energy dE' that passes through dA' in the time interval $(t', t' + dt)$ in the frequency interval $(\nu, \nu + d\nu)$ at angles that took it through dA. From the definition of specific intensity we have

$$dE = I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t)\, dA\, d\Omega\, d\nu\, dt \tag{1.5}$$

and

$$dE' = I_\nu(\boldsymbol{r}', \boldsymbol{n}, \nu, t')\, dA'\, d\Omega'\, d\nu\, dt \tag{1.6}$$

where $d\Omega$ is the solid angle subtended by dA' at $\boldsymbol{r}$ and $d\Omega'$ is the solid angle subtended by dA at point $\boldsymbol{r}'$. Since, $dE = dE'$, we have

$$I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t)\, dA\, d\Omega = I_\nu(\boldsymbol{r}', \boldsymbol{n}, \nu, t')\, dA'\, d\Omega'. \tag{1.7}$$

However, $d\Omega = dA'/l^2$ and $d\Omega' = dA/l^2$, and so

$$I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t)\, dA\, dA'/l^2 = I_\nu(\boldsymbol{r}', \boldsymbol{n}, \nu, t')\, dA'\, dA/l^2. \tag{1.8}$$

and

$$I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t) = I_\nu(\boldsymbol{r}', \boldsymbol{n}, \nu, t'), \tag{1.9}$$

or

$$I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t) = I_\nu(\boldsymbol{r} + l\boldsymbol{n}, \boldsymbol{n}, \nu, t + l/c). \tag{1.10}$$

This shows that in the absence of interaction with matter, the specific intensity is conserved as radiation streams along its path. A few moments of consideration show that if the radiation field is in steady state, then

$$I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t) = I_\nu(\boldsymbol{r} + l\boldsymbol{n}, \boldsymbol{n}, \nu, t), \tag{1.11}$$

that is, the specific intensity is constant along straight lines.

Why Use the Specific Intensity?

Why do we use the specific intensity rather than some more physically obvious quantity such as the density of photons or the density of energy?

First, we need a quantity that encapsulates all of the information about radiation. The specific intensity contains all of the information required to describe unpolarized light completely, as it specifies the quantity of radiation at a given point, in a given direction, at a given frequency, and at a given time.

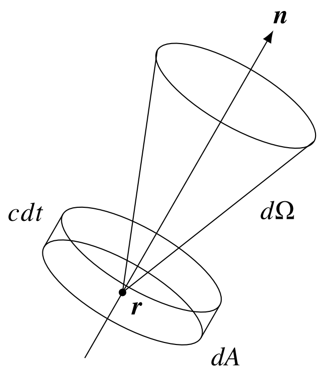

Figure 1.3: The geometry used in the determination of the relation between the photon distribution function f and the specific intensity I_ν.

Second, the conservation of the specific intensity along a ray in the absence of interaction with matter means that in the transfer of specific intensity we need only consider changes due to interaction with matter and ignore changes due to geometry. For example, the density of photons emitted by a point source into free space decreases as $1/r^2$ simply because of geometric dilution, but the specific intensity remains the same. This is an enormous simplification.

The Photon Distribution Function

The equivalent of the specific intensity I_ν in the photon-gas description is the photon distribution function f, the phase space density of photons. Phase space is a six-dimensional space specified by the three components of position and the three components of momentum. The photon distribution function $f(\boldsymbol{r},\boldsymbol{p},t)$ is such that the number of photons dN in a volume of real space $d\boldsymbol{r}$ and volume of momentum space $d\boldsymbol{p}$ centered on position $\boldsymbol{r}$ and momentum $\boldsymbol{p}$ is given by

$$dN = f(\boldsymbol{r},\boldsymbol{p},t)\, d\boldsymbol{r}\, d\boldsymbol{p}. \tag{1.12}$$

We'll now develop the relation between f to I_ν. Consider Figure 1.3. We know from the definition of the specific intensity that the energy dE transported by radiation across an area dA, centered on $\boldsymbol{r}$ and perpendicular to $\boldsymbol{n}$, into a solid angle $d\Omega$ about $\boldsymbol{n}$, in a frequency interval $(\nu, \nu + d\nu)$, in a time interval $(t, t + dt)$ is given by

$$dE = I_\nu(\boldsymbol{r},\boldsymbol{n},\nu,t)\, dA\, d\Omega\, d\nu\, dt. \tag{1.13}$$

In terms of the photon distribution function,

$$dE = h\nu f\, d\boldsymbol{r}\, d\boldsymbol{p}, \tag{1.14}$$

where the volume of phase space $d\boldsymbol{r}\, d\boldsymbol{p}$ contains the all of the photons that pass through an area dA, centered on $\boldsymbol{r}$ and perpendicular to $\boldsymbol{n}$, into a solid angle $d\Omega$ about $\boldsymbol{n}$, in a frequency interval $(\nu, \nu + d\nu)$, in a time interval $(t, t + dt)$. Since photons travel a distance $c dt$ in time dt, we have

$$d\boldsymbol{r} = c\, dt\, dA. \tag{1.15}$$

Furthermore, since the momentum of a photon is $p = h\nu/c$, the photons traveling into a solid angle $d\Omega$ in a frequency interval $(\nu, \nu+d\nu)$ are those in the partial shell in momentum space between $(p, p + dp)$ with directions in $d\Omega$. That is,

$$d\boldsymbol{p} = p^2\, dp\, d\Omega = \frac{h^3\nu^2}{c^3}\, d\nu\, d\Omega. \tag{1.16}$$

Thus,

$$dE = I_\nu(\boldsymbol{r},\boldsymbol{n},\nu,t)\, dA\, d\Omega\, d\nu\, dt = \frac{h^4\nu^3}{c^2} f\, dA\, d\Omega\, d\nu\, dt, \tag{1.17}$$

and so

$$I_\nu(\boldsymbol{r},\boldsymbol{n},\nu,t) = \frac{h^4\nu^3}{c^2} f. \tag{1.18}$$

Like the specific intensity, the photon distribution function also completely describes unpolarized light and also is conserved along a ray. However, the specific intensity is often preferred because it is more directly related to macroscopic thermodynamic quantities, such as energy and temperature.

Distribution functions are commonly used in physics and astrophysics, for example, in the statistical mechanics of gases and in the dynamics of stellar systems; they are all defined in terms of phase space densities. A standard result for distribution functions is that they are conserved along a path. Thus, we can think of the conservation of I_ν along a ray as coming from the conservation of f (and ν) along the same ray. (Similarly, we can consider the equation of radiation transfer that we will derive in chapter 2 as a collisional Boltzmann equation for f.)

The Moments of the Specific Intensity

The moments of the specific intensity with respect to the direction vector $\boldsymbol{n}$ are doubly useful, as they have physical meaning (such as the energy, the energy flux, and the radiation pressure) and are used in the development of solutions for the equations of radiative transfer. In dyadic form, the m-th moment is defined as

$$M_m(I_\nu) \equiv \frac{1}{4\pi}\int_{4\pi} d\Omega\, I_\nu(\boldsymbol{r},\boldsymbol{n},\nu,t)\boldsymbol{n}^m. \tag{1.19}$$

In general the zeroth moment is a scalar, the first moment is a vector, and the second moment is a symmetric tensor of rank 2. In component form, these moments are

$$M_0(I_\nu) \equiv \frac{1}{4\pi}\int_{4\pi} d\Omega\, I_\nu(\boldsymbol{r},\boldsymbol{n},\nu,t), \tag{1.20}$$

$$M_1(I_\nu)_i \equiv \frac{1}{4\pi}\int_{4\pi} d\Omega\, I_\nu(\boldsymbol{r},\boldsymbol{n},\nu,t)n_i, \tag{1.21}$$

and

$$M_2(I_\nu)_{ij} \equiv \frac{1}{4\pi} \int_{4\pi} d\Omega \, I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t) n_i n_j. \tag{1.22}$$

Here, n_i and n_j are Cartesian components of $\boldsymbol{n}$.

We often work in a plane-parallel symmetry, which allows us to simplify the first and second moments considerably. If we take the z axis to be the vertical axis of the plane-parallel symmetry, then we have reflectional symmetry about the planes $x = 0$ and $y = 0$ and rotational symmetry about the z axis.

Consider the x component of the first moment, $M_1(I_\nu)_x$. This is given by

$$M_1(I_\nu)_x = \frac{1}{4\pi} \int_{4\pi} d\Omega \, I_\nu(n_x, n_y) n_x, \tag{1.23}$$

where we indicate the explicit dependence on the components n_x and n_y of $\boldsymbol{n}$. There is no corresponding dependence on n_z, since n_z is given in terms of n_x and n_y by $n_x^2 + n_y^2 + n_z^2 = 1$. We omit the implicit dependence on $\boldsymbol{r}$, ν, and t for conciseness. We split the integral into two hemispheres,

$$M_1(I_\nu)_x = \frac{1}{4\pi} \left[\int_{n_x > 0} d\Omega \, I_\nu(n_x, n_y) n_x + \int_{n_x < 0} d\Omega \, I_\nu(n_x, n_y) n_x \right]. \tag{1.24}$$

We then make the substitution $n'_x = -n_x$ in the second integral.

$$M_1(I_\nu)_x = \frac{1}{4\pi} \left[\int_{n_x > 0} d\Omega \, I_\nu(n_x, n_y) n_x - \int_{n'_x > 0} d\Omega' \, I_\nu(-n'_x, n_y) n'_x \right]. \tag{1.25}$$

We then note that the reflectional symmetry through the plane $n_x = 0$ implies that

$$I_\nu(n_x, n_y) = I_\nu(-n_x, n_y), \tag{1.26}$$

and so the two integrals cancel, leaving

$$M_1(I_\nu)_x = 0. \tag{1.27}$$

We can make a similar argument using the reflectional symmetry through the plane $n_y = 0$ to show that the y component $M_1(I_\nu)_y$ is also zero. However, in plane-parallel symmetry we do not in general have symmetry through the plane $z = 0$, and so in general the z component $M_1(I_\nu)_z$ is not zero. Thus, we can fully specify the vector first moment by its z component, $M_1(I_\nu)_z$, the only component that is not identically zero. That is, we have,

$$M_1(I_\nu)_x = M_1(I_\nu)_y = 0 \tag{1.28}$$

and

$$M_1(I_\nu)_z = \frac{1}{4\pi} \int_{4\pi} d\Omega \, I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t) n_z. \tag{1.29}$$

More intuitively, we could argue that in plane-parallel symmetry, any vector must be either parallel or anti-parallel to the z axis. Thus, the x and y components of the first moment must be zero.

Similarly, consideration of the second moment shows that the off-diagonal components must be zero. Rotational symmetry about the z axis further implies that the xx and yy components must be equal. Finally, we note that $\boldsymbol{n}$ is a unit vector, so $n_x^2 + n_y^2 + n_z^2 = 1$, and so the trace of the second moment satisfies $M_2(I_\nu)_{xx} + M_2(I_\nu)_{yy} + M_2(I_\nu)_{zz} = M_0(I_\nu)$. Thus, we can fully specify the tensor second moment by $M_2(I_\nu)_{zz}$ and $M_0(I_\nu)$ as

$$M_2(I_\nu)_{xx} = M_2(I_\nu)_{yy} = \frac{1}{2}(M_0(I_\nu) - M_2(I_\nu)_{zz}), \tag{1.30}$$

and

$$M_2(I_\nu)_{zz} = \frac{1}{4\pi} \int_{4\pi} d\Omega \, I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t) n_z n_z. \tag{1.31}$$

To summarize, in plane-parallel symmetry the first three moments are fully-specified by the values of $M_0(I_\nu)$, $M_1(I_\nu)_z$, and $M_2(I_\nu)_{zz}$. We normally write these as J_ν, H_ν, and K_ν, where

$$J_\nu \equiv M_0(I_\nu) \equiv \frac{1}{4\pi} \int_{4\pi} d\Omega \, I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t), \tag{1.32}$$

$$H_\nu \equiv M_1(I_\nu)_z \equiv \frac{1}{4\pi} \int_{4\pi} d\Omega \, I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t) n_z, \tag{1.33}$$

and

$$K_\nu \equiv M_2(I_\nu)_{zz} \equiv \frac{1}{4\pi} \int_{4\pi} d\Omega \, I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t) n_z^2. \tag{1.34}$$

In spherical symmetry, we can construct a local cartesian coordinate system with the z axis parallel to the local radial axis. This local cartesian coordinate system then has reflectional symmetry about the planes $x = 0$ and $y = 0$ and rotational symmetry about the z axis. These were precisely the symmetries we used above in plane-parallel symmetry to show that we could reduce the first three moments to three components. We can see, therefore, that in spherical symmetry we can do the same.

To integrate over solid angle, we change variables to the polar angle θ between $\boldsymbol{n}$ and the z axis and the azimuthal angle ϕ (see Figure 1.4), and so have

$$\frac{1}{4\pi} \int_{4\pi} d\Omega = \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \, \sin\theta. \tag{1.35}$$

In plane-parallel symmetry, we have rotational symmetry about the z axis, which allows us to integrate over ϕ directly, giving

$$\frac{1}{4\pi} \int_{4\pi} d\Omega = \frac{1}{2} \int_0^{\pi} d\theta \, \sin\theta. \tag{1.36}$$

The substitution $\mu \equiv \cos\theta$ allows this integral to be expressed in an especially simple form. Instead of θ varying from 0 to π, we then have μ varying from $+1$ to -1, with $\mu = +1$ being parallel to the symmetry axis ($\theta = 0$), $\mu = 0$ being perpendicular to the

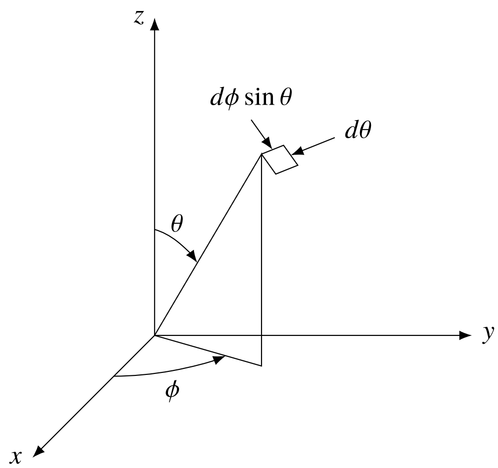

Figure 1.4: The element of solid angle $d\Omega$ is equal to $\sin\theta\, d\theta\, d\phi$.

symmetry axis ($\theta = \pi/2$), and $\mu = -1$ being anti-parallel to the symmetry axis ($\theta = \pi$). Since $d\mu = \sin\theta\, d\theta$, we obtain

$$\frac{1}{4\pi}\int_{4\pi} d\Omega = \frac{1}{2}\int_{-1}^{+1} d\mu. \tag{1.37}$$

This integration over ϕ and substitution of μ for θ is a very common trick when we have plane-parallel geometry.

Returning to the moments, we have $n_z = \cos\theta = \mu$, so

$$J_\nu = \frac{1}{2}\int_{-1}^{+1} d\mu\, I_\nu(\mu), \tag{1.38}$$

$$H_\nu = \frac{1}{2}\int_{-1}^{+1} d\mu\, I_\nu(\mu)\mu, \tag{1.39}$$

and

$$K_\nu = \frac{1}{2}\int_{-1}^{+1} d\mu\, I_\nu(\mu)\mu^2. \tag{1.40}$$

That is, we can interpret J_ν, H_ν, and K_ν as the first three moments of I_ν with respect to μ. We now consider physical interpretations of these moments.

Zeroth Moments

The Mean Intensity

The zeroth moment J_ν or mean intensity is defined above as

$$J_\nu = \frac{1}{2}\int_{-1}^{+1} d\mu\, I_\nu. \tag{1.41}$$

Physically, J_ν is just the angular mean of the specific intensity. If I_ν is isotropic, then $J_\nu = I_\nu$.

The Photon Density

The monochromatic density N_ν of photons per unit frequency is given by

$$N_\nu = \frac{4\pi J_\nu}{h\nu c}. \tag{1.42}$$

and the total photon density N is given by

$$N = \frac{4\pi}{hc}\int_0^\infty d\nu\, \frac{J_\nu}{\nu}. \tag{1.43}$$

These expressions are derived in Problem 1.3.

The Energy Density

The monochromatic energy density u_ν per unit frequency is given by

$$u_\nu = \frac{4\pi J_\nu}{c}. \tag{1.44}$$

and the total energy density E is given by

$$u = \frac{4\pi J}{c}. \tag{1.45}$$

These expressions are derived in Problem 1.3.

First Moments

The Eddington Flux

The first moment H_ν or Eddington flux is defined above as

$$H_\nu = \frac{1}{2}\int_{-1}^{+1} d\mu\, \mu I_\nu. \tag{1.46}$$

The Eddington flux has no direct physical interpretation, but it is widely used in solutions to the radiation transfer problem. If I_ν is isotropic, then $H_\nu = 0$.

The Energy Flux

Consider Figure 1.5. We define the energy flux F_ν such that the energy dE that flows through area dA' perpendicular to the z-axis, into all solid angles, in frequency interval $(\nu, \nu + d\nu)$, in time interval $(t, t + dt)$ is

$$dE \equiv F_\nu(\boldsymbol{r}, \nu, t) dA' d\nu dt. \tag{1.47}$$

From the definition of specific intensity, we have

$$dE = \int_{4\pi} d\Omega\, I_\nu(\boldsymbol{r}, \nu, t) dA d\nu dt, \tag{1.48}$$

in which dA is the projection of dA' onto a plane perpendicular to the $\boldsymbol{n}$. The two areas are related by

$$dA = dA' \cos\theta, \tag{1.49}$$

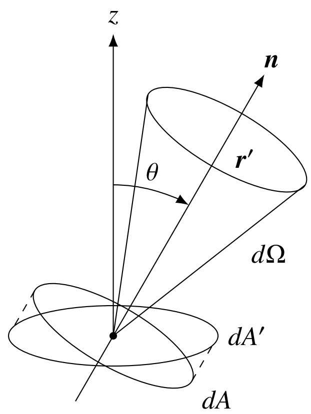

Figure 1.5: The geometry used in the derivation of the flux.

where θ is the angle between $\boldsymbol{n}$ and the z-axis. Thus, equating the two expressions for dE, we have

$$F_\nu = \int_{4\pi} d\Omega\, I_\nu(\boldsymbol{r}, \boldsymbol{n}, \nu, t) \cos\theta. \tag{1.50}$$

This is proportional to the first-moment, and so we have

$$F_\nu = 4\pi H_\nu = 2\pi \int_{-1}^{+1} d\mu\, \mu I_\nu. \tag{1.51}$$

The total energy flux $F = \int_0^\infty d\nu F_\nu = 4\pi H$ is one of the most important quantities in an atmosphere. We use it so much that we often refer to it simply as "the" flux.

If I_ν is isotropic, then $F_\nu = 0$. This result simply states that if the specific intensity is isotropic, then there is no net flux of energy. This is to expected on symmetry grounds; an isotropic specific intensity has no preferred direction in which energy might flow.

If I_ν is isotropic over the outward hemisphere and zero over the inward hemisphere, for example, if a flat surface of constant brightness is emitting into free space, then we have

$$I_\nu(\mu) = \begin{cases} I_\nu(1) & \text{for } \mu > 0, \\ 0 & \text{for } \mu < 0, \end{cases} \tag{1.52}$$

and the flux is given by

$$F_\nu = \pi I_\nu(1). \tag{1.53}$$

This expression is derived in Problem 1.4. This expression for the flux leads to the use of the quantity F_ν/π, which is called the "astrophysical flux" in contrast to the "physical flux" F_ν. The astrophysical flux would be of largely historical interest were it not that the emergent flux of a model atmospheres is most often tabulated in terms of the astrophysical flux F_ν/π rather than the energy flux F_ν. For example, Kurucz distributes tables of astrophysical fluxes for his model atmospheres. Note that some authors, in particular Chandrasekhar (1960), Mihalas (1978), Gray (1992), use the symbol F_ν for the astrophysical flux and either πF_ν or the symbol $\mathcal{F}_\nu$ for the physical flux, but here we follow Milne (1930) and Rybicki & Lightman (1979) in using F_ν for the physical flux.

The Photon Flux

We can generalize the definition of a flux so that the "flux of X" such that the quantity of X dE that flows through area dA' perpendicular to the z-axis, into all solid angles, in frequency interval $(\nu, \nu + d\nu)$, in time interval $(t, t + dt)$ is

$$dX \equiv F_X(\boldsymbol{r}, \nu, t) dA' d\nu dt. \tag{1.54}$$

A simple application is to define the photon flux as the number of photons that flow through an area perpendicular to the z axis, in a frequency interval, and in a time interval. Since energy is carried by photons and since each photon has an energy $h\nu$, the photon flux is related to the energy flux and is given by

$$\frac{1}{h\nu} F_\nu. \tag{1.55}$$

Second Moments

The Eddington Pressure

The second moment K_ν or Eddington pressure is defined above as

$$K_\nu = \frac{1}{2} \int_{-1}^{+1} d\mu\, I_\nu \mu^2. \tag{1.56}$$

Like the Eddington flux, the Eddington pressure has no direct physical interpretation, but it is widely used in solutions to the radiation transfer problem.

The ratio

$$f \equiv \frac{K_\nu}{J_\nu} \tag{1.57}$$

is known as the Eddington factor. (Although the same symbol is used for both the Eddington factor and the photon distribution function, they are used in very different contexts and no real ambiguity should arise.) If I_ν is isotropic, then $K_\nu = I_\nu/3$, and the Eddington factor is 1/3. For a non-isotropic radiation field, the Eddington factor will not in general be 1/3, but will always satisfy in $0 \leq f \leq 1$.

The Radiation Pressure

The radiation pressure is defined as the flux of momentum in the radiation field. The flux of momentum is the flux of a vector quantity and so in general will be a second-rank tensor. However, as we saw above, if we have rotational symmetry, the tensor can be essentially reduced to a scalar quantity: the flux parallel to the symmetry axis (i.e., upwards or outwards) of the component of momentum parallel to the symmetry axis.

We normally consider pressure to be the force per unit area exerted on the walls of a vessel by the fluid contained within the vessel rather than the flux of momentum. The two quantities have the same units and are clearly related, because the force on the walls is derived from the change in momentum of the particles when the impinge on the walls. In a gas consisting of atoms or molecules, both quantities are normally identical. However, the interaction of a photon with a wall can be more complex – it can be reflected, absorbed, emitted, or transmitted – and in gases of photons the flux of momentum and force per unit area are in general not the same.

To derive the radiation pressure, consider Figure 1.5 again, and recall that the energy carried by radiation across dA' is $I_\nu \cos\theta dA d\Omega d\nu dt$. The z-component of momentum carried by radiation across the same area is $I_\nu \cos^2\theta dA d\Omega d\nu dt/c$, where the factor of $1/c$ relates the momentum of a photon to its energy and the additional factor of $\cos\theta$ comes from considering just the component of the momentum of a photon parallel to the z-axis. From this, we can see that the radiation pressure p_ν is given by

$$p_\nu = \frac{1}{c}\int_{4\pi} d\Omega\, I_\nu \cos^2\theta. \tag{1.58}$$

Again, we see that this is related to the second moment of I_ν, and we have

$$p_\nu = \frac{4\pi}{c} K_\nu = \frac{2\pi}{c}\int_{-1}^{+1} d\mu\, \mu^2 I_\nu. \tag{1.59}$$

If I_ν is isotropic, then $p_\nu = 4\pi J_\nu/3c$.

We can interpret the Eddington factor f as the ratio of the monochromatic pressure to the monochromatic energy density p_ν/u_ν, because p_ν and u_ν are simply K_ν and J_ν multiplied by the same factor of $4\pi/c$. From classical thermodynamics, we know that the ratio of pressure to internal energy is $\gamma - 1$, in which γ is the ratio of the specific heats c_P/c_V. Thus, an isotropic radiation field has $\gamma = 4/3$.

However, isotropic radiation is a special case. In general, the Eddington factor and hence the ratio of the pressure to the internal energy depend on the degree of anisotropy in the radiation field and can range from 0 to 1. This has an analogy in gases of massive particles: an ideal gas of massive particles has an isotropic distribution of velocities when at rest and the ratio of pressure to internal thermal energy is 2/3; however, when the gas is has a bulk motion, the velocity distribution is not isotropic and the ratio of ram pressure to bulk kinetic energy varies from 0 to 1/2 depending on the direction of the flow.

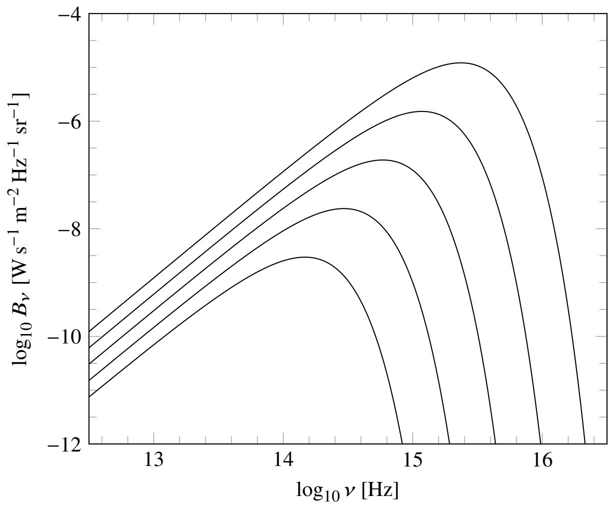

Figure 1.6: The Planck function B_ν at temperatures of (bottom to top) 2500 K, 5000 K, 10,000 K, 20,000 K, and 40,000 K. A frequency of $10^{14.5}$ Hz corresponds to a wavelength of about 1 μm.

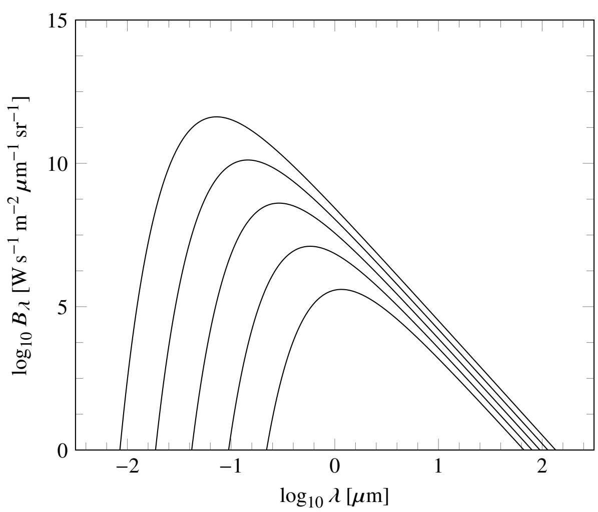

Figure 1.7: The Planck function B_λ at temperatures of (bottom to top) 2500 K, 5000 K, 10000 K, 20000 K, and 40000 K.

Black-Body Radiation

In thermodynamic equilibrium at a temperature T, the radiation field is uniform, time-independent, and has a frequency distribution given by the Planck functions, $I_\nu^* = B_\nu$ and $I_\lambda^* = B_\lambda$ where

$$B_\nu \equiv \frac{2h\nu^3}{c^2}(e^{h\nu/kT} - 1)^{-1}, \tag{1.60}$$

and

$$B_\lambda \equiv \frac{2hc^2}{\lambda^5}(e^{hc/kT\lambda} - 1)^{-1}. \tag{1.61}$$

Such radiation is known as "black-body" radiation. It is conventional to use the superscript * to denote the value of physical quantities in equilibrium.

Figures 1.6 and 1.7 show the Planck functions B_ν and B_λ at temperatures, frequencies, and wavelengths of relevance for stellar atmospheres. The figures illustrate the monotonic increase in the Planck functions with temperature at a given frequency or wavelength. They also show the two important limiting cases of the Planck functions, the low-frequency Rayleigh-Jeans tail for $h\nu \ll kT$, which has

$$B_\nu \approx \frac{2kT}{c^2}\nu^2, \tag{1.62}$$

and the high-frequency Wien tail for $h\nu \gg kT$, which has

$$B_\nu \approx \frac{2h\nu^3}{c^2}e^{-h\nu/kT}. \tag{1.63}$$

In the Rayleigh-Jeans tail, the Planck function changes only linearly with temperature at a given frequency, whereas in the Wien tail the change with temperature is much more dramatic. We will see, however, that the dominant radiation in an atmosphere often has $h\nu \sim kT$, close to the peak of B_ν, in which case we must use the exact expression for B_ν.

Since black-body radiation is isotropic we have

$$J_\nu^* = I_\nu^* = B_\nu, \tag{1.64}$$

$$F_\nu^* = 0, \tag{1.65}$$

and

$$p_\nu^* = \frac{4\pi}{3c}I_\nu = \frac{4\pi}{3c}B_\nu. \tag{1.66}$$

The total mean intensity is

$$J^* = B = \int_0^\infty B_\nu \, d\nu \tag{1.67}$$

If substitute for B_ν and then use $x \equiv h\nu/kT$, we find

$$B = \left(\frac{2k^4}{c^2h^3}\right)T^4\int_0^\infty dx\, x^3(e^x - 1)^{-1}. \tag{1.68}$$

The integral is a pure number, and can be shown to have the value $\pi^4/15$. Thus

$$B = \left(\frac{2\pi^4k^4}{15c^2h^3}\right)T^4. \tag{1.69}$$

The total energy density is

$$u^* = \frac{4\pi}{c}J^* = \frac{4\pi}{c}B = aT^4, \tag{1.70}$$

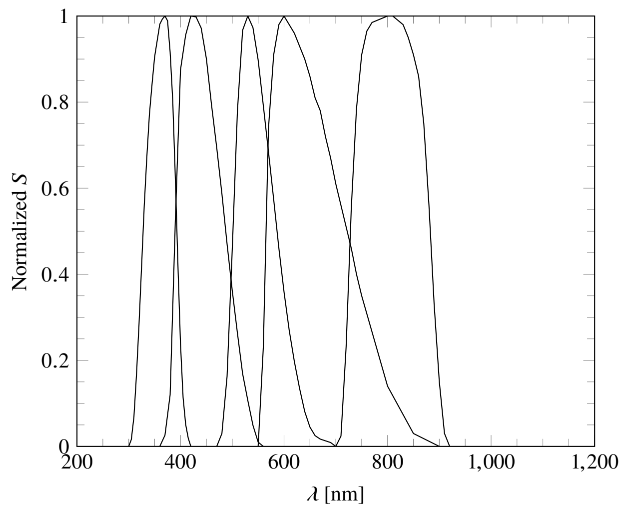

Figure 1.8: The normalized system response S of the Johnson-Cousins $UBVRI$ filters (Bessell 1990), excluding the atmosphere.

in which $a \equiv 8\pi^5k^4/15c^3h^3$ is the radiation constant. This result can also be obtained by consideration of the thermodynamics of black-body radiation (Rybicki & Lightman 1979, pp. 17–18).

If a surface emits as a black body, i.e., has $I_\nu = B_\nu$ over the outward hemisphere and has $I_\nu = 0$ over the inward hemisphere, then by equation 1.53 the flux is $F_\nu = \pi B_\nu$, and the total flux is

$$F = \pi B = \sigma T^4, \tag{1.71}$$

in which $\sigma \equiv ac/4$ is the Stefan-Boltzmann constant, $\sigma \equiv (2\pi^5k^4)/(15c^2h^3) = 5.670 \times 10^{-5}$erg s^{-1} cm^{-2} K^{-4}.

The Effective Temperature

We commonly use the *effective temperature* T_{eff} as a surrogate for the total flux in an atmosphere. The effective temperature is defined in terms of the total flux F by

$$F \equiv \sigma T_{\text{eff}}^4. \tag{1.72}$$

The motivation for this definition is that, as we saw in equation 1.71, the total flux from a surface that emits as a black body with a temperature T is σT^4. Nevertheless, we need to be clear that specifying the effective temperature is simply a different means to give the total flux; the effective temperature is not a real thermodynamic temperature and does not imply that the matter in the atmosphere is in thermodynamic equilibrium.

Magnitudes and Colors

We measure the magnitude of a star by measuring the flux after passing its spectrum through a filter. Until recently, the most common filters were the Johnson-Cousins filters; UBV were Johnson's

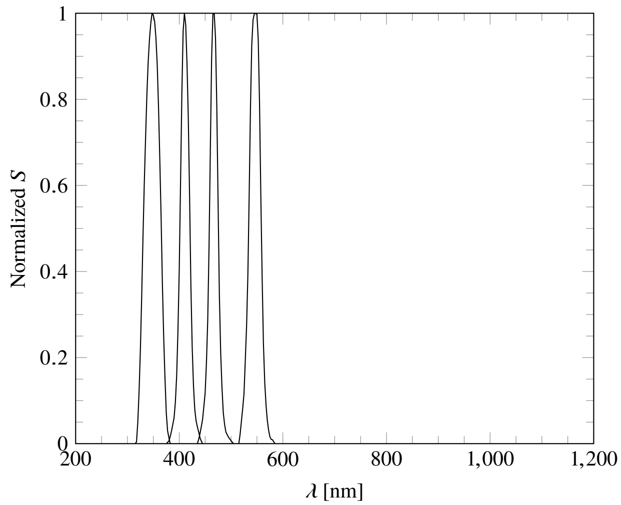

Figure 1.9: The normalized system response S of the Strömgren *uvby* filters (Bessell 2011), excluding the atmosphere.

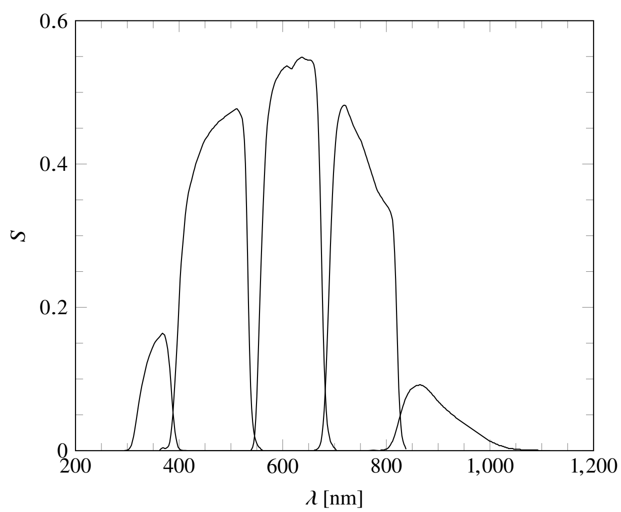

Figure 1.10: The absolute system response S of the SDSS $u'g'r'i'z'$ filters at an airmass of 1.3 (Doi et al. 2010).

Table 1.1: The Johnson-Cousins, Strömgren-Crawford, and SDSS Filter Systems

System	Filter	$\bar{\lambda}/\mu$m	$\Delta\lambda/\bar{\lambda}$
Johnson-Cousins	U	0.365	0.18
	B	0.445	0.21
	V	0.551	0.16
	R	0.658	0.21
	I	0.806	0.18
	J	1.220	0.17
	H	1.630	0.19
	K	2.190	0.18
	L	3.450	0.14
	M	4.750	0.10
Strömgren-Crawford	u	0.349	0.09
	v	0.411	0.05
	b	0.467	0.04
	y	0.547	0.04
	β_n	0.489	0.01
	β_w	0.489	0.03
SDSS	u'	0.356	0.17
	g'	0.483	0.29
	r'	0.626	0.22
	i'	0.767	0.20
	z'	0.910	0.15

original filters; the rest were added later. These bands range from the atmospheric cut off at about 3200 Å into thermal infrared beyond 2.2 μm. The Strömgren-Crawford system is also used quite often in precision stellar photometry, as they are narrower (with band widths of less than 10% compared to around 20% for the Johnson filters), and so measure the spectrum of a star more finely. More recently, the SDSS filters have become more popular. The mean wavelengths and bandwidths of these filters are given in Table 1.1 and responses are given in Figures 1.8, 1.9, and 1.10. Other systems include the Geneva, DDO, Vilnius, Thuan-Gunn, Hipparcos, Tycho, and Pan-STARRS systems.

The apparent magnitude m_X in a band X is defined as

$$m_X = -2.5 \log_{10} F_X + C_X \tag{1.73}$$

where F_X is the mean flux in the band and C_X is a constant. The mean flux is given by

$$F_X = \frac{\int d\lambda\, \lambda S F_\lambda}{\int d\lambda\, \lambda S}, \tag{1.74}$$

in which S is the response of the system: the relative probability that a photon of wavelength λ will be detected by the system (atmosphere, telescope, filter, and detector). The additional factors of λ are let over from multiplying by λ/hc to convert F_λ from energy per unit wavelength to number of photons per unit wavelength. We often make the crude assumption that S is a δ-function at λ_X, and so

$$F_X \approx F_\lambda(\lambda_X). \tag{1.75}$$

In practice, magnitudes are measured by comparing the flux of a star to the flux of standard stars whose magnitudes are specified by the system. We often use X instead of m_X.

For the Johnson-Cousins and Strömgen systems, the constants C_X were originally chosen so that Vega would have magnitude 0 in all bands. However, as the calibration was refined and effectively defied by fainter standards, Vega was found to have $V = 0.03$ and to have slightly different magnitudes in other bands.

For the SDSS system, the magnitudes are on the so-called AB scale, so where the constants C_X are chosen so that a star with a constant monochromatic flux of 3631 Jy (that is, 3.631×10^{-20}erg s^{-1} cm^{-2} Hz) would have magnitude 0 in all bands.

The absolute magnitude M_X of a star in a band X is the apparent magnitude it would have if it were at a distance of 10 pc instead of its real distance d. From the inverse square law, we can see that

$$M_X = m_X - 5\log_{10}(d/10\ \mathrm{pc}). \tag{1.76}$$

We form color indices by subtracting a pair of magnitudes. It doesn't matter whether we subtract two apparent or two absolute magnitudes, as color indices are independent of distance. Traditionally, the redder magnitude is subtracted from the bluer magnitude, e.g., $U - B$, $B - V$, and $V - R$. Thus, a more positive color means a redder spectrum and a more negative color means a bluer spectrum. In the Johnson system, A0V stars by definition have zero color. Thus, if a star has a positive color, it is redder than an A0V star, and if it has a negative color, it is bluer than an A0V star. In the AB system, stars that have constant flux density F_ν by definition have zero color.

Notes and Further Reading

Specific Intensity

The specific intensity and its moments are discussed by Mihalas (1978, pp. 2–18), Rybicki & Lightman (1979, pp. 2–8), Boehm-Vitense (1989, ch. 4 and 5), Shu (1991, pp. 3–8 and pp. 11–12), Gray (1992, ch. 5), and Rutten (2003, pp. 9–12). The vector and tensor forms of the flux and radiation pressure in general geometries are discussed by Mihalas (1978, pp. 9–19).

Milne (1930, pp. 74–75) and Mihalas (1978, p. 4) discuss a more general conservation property of the specific intensity: that I/n^2 is conserved along a path, where n is the refractive index, provided the coefficient of reflection at each interface is zero.

Polarized Light

For polarized light, in addition to I_ν we need to also describe the degree and angle of linear polarization and the degree of circular polarization. The Stokes parameters Q_ν, U_ν, and V_ν, are commonly used to specify these additional three qualities. For definitions, see Chandrasekhar (1960, pp. 24–35), Rybicki & Lightman (1979, pp. 62–69), Shu (1991, ch. 12), and Rutten (2003, pp. 135–137). Hecht (1998, pp. 366–367) gives an operational definition of the Stokes parameters in terms of a polarizing filter (and using the notation s_0, s_1, s_2, and s_3 instead of I, Q, U, and V).

Black-Body Radiation

Black-body radiation is discussed by Rybicki & Lightman (1979, pp. 15–27), Mihalas (1978, pp. 6–7), and Gray (1992, ch. 6).

Magnitudes

The zero-points and effective bandpasses of the Johnson-Cousins filters are given by Bessell & Brett (1988) and Bessell & Murphy (2012). The zero-points and effective bandpasses of the Strömgren filters are given by Bessell (2011). The effective bandpasses of the SDSS filters are given by Fukugita et al. (1996) and Doi et al. (2010)