Chapter 2

Radiation Fundamentals

2.1 Brightness and Flux Density

Astronomers study an astronomical source by measuring the strength of its radiation as a function of direction on the sky (by mapping or imaging) and frequency (spectroscopy), plus other quantities (time, polarization) to be considered later. Clear and quantitative definitions are needed to describe the strength of radiation and how it varies with direction, frequency, and distance between the source and the observer. The concepts of brightness and flux density are deceptively simple, but they regularly trip up even experienced astronomers. It is very important to understand them clearly because they are so fundamental.

Consider the simplest possible case of radiation traveling from a source through empty space, so there is no **absorption** (destruction of photons), **scattering** (changing the direction of photons), or **emission** (creation of new photons) along the path to an observer. In the **ray-optics approximation**, radiated energy flows in straight lines. This approximation is valid only for systems much larger than the wavelength λ of the radiation, a criterion easily met by astronomical sources. You may find it helpful to visualize electromagnetic radiation as a stream of light particles (photons), essentially bullets that travel in straight lines at the speed of light. To motivate the following mathematical definitions, imagine you are looking at the Sun. The "brightness" of the Sun appears to be about the same over most of the Sun's surface, which looks like a nearly uniform disk even though it is actually a sphere. This means that a photograph of the Sun would be uniformly exposed across the Sun's disk. It also turns out that the exposure would not change if photographs were made at different distances from the Sun, from points near Mars, the Earth, and Venus, for example (Figure 2.1).
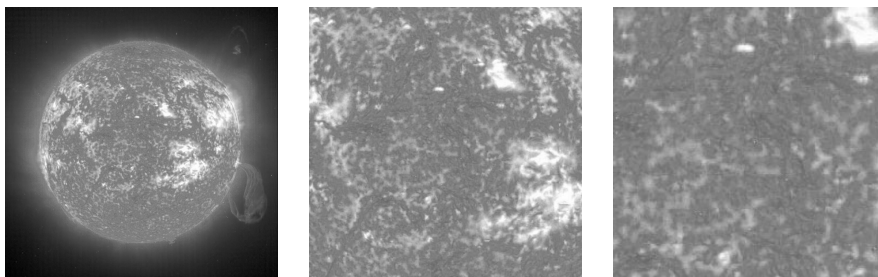

Figure 2.1: The Sun as it would appear in three photos taken with the same camera from long (left), medium (center), and short distances (right) would have constant brightness but increasing angular size. Image credit: SOHO/EIT Consortium (ESA & NASA).

The angular size of the Sun depends on the distance between the Sun and the camera, but the number of photons falling on the detector per unit area per unit time *per unit solid angle* does not. The photo taken from near Venus would not be overexposed, and the one from near Mars would not be underexposed. The total number of solar photons from all directions reaching the camera per unit area per unit time (or the total energy absorbed per unit area per unit time) does decrease with increasing distance, but *only* because the solid angle subtended by the Sun decreases. Thus we distinguish between the **brightness** or **intensity** of the Sun's radiation, which does not depend on distance, and the apparent **flux**, which does. Some authors [16, 20] prefer to use *brightness* for the power per unit area per unit solid angle emitted at the source and *specific intensity* for the power per unit area per unit solid angle along the path to the detector; others [107, 116] do not distinguish between them. The two are identical if there is no absorption or emission between the source and the detector, and we will use these terms interchangeably.

Note also that the number of photons per unit area hitting the detector is proportional to $\cos\theta$ if the normal to the detector is tilted by an angle θ from the direction of the incoming rays (Figure). This is just the same projection effect that reduces the amount of water collected by a tilted rain gauge by $\cos\theta$. Likewise at the source, such as the spherical Sun, the projected area perpendicular to the line of sight is proportional to $\cos\theta$, where θ is the angle between the line of sight and the normal to the Sun's surface.

The total brightness is contributed by photons of all frequencies. The brightness per unit frequency is called the **specific intensity** (also **spectral intensity** or **spectral brightness**). The notation for specific intensity is I_ν, where the subscript ν is used to indicate "per unit frequency." In the ray-optics approximation, specific intensity can be defined quantitatively in terms of

- $d\sigma$ = an infinitesimal surface area (e.g., of a detector);
- θ = the angle between a "ray" of radiation and the normal to the surface;
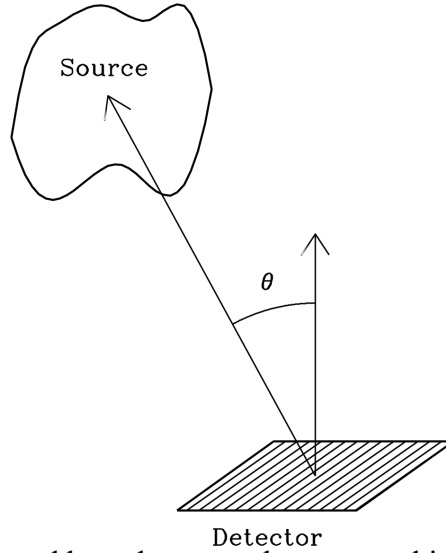- $d\Omega$ = an infinitesimal solid angle measured from the observer's location.

Figure 2.2: Specific intensity measured by a detector whose normal is at an angle θ from the line of sight.

The surface containing $d\sigma$ can be any surface, real or imaginary; that is, it could be the physical surface of the detector, the source, or an imaginary surface anywhere along the ray. If energy dE from within the solid angle $d\Omega$ flows through the projected area $\cos\theta\,d\sigma$ in time dt and in a narrow frequency band of width $d\nu$, then

$$dE = I_\nu \cos\theta\,d\sigma\,d\Omega\,dt\,d\nu. \tag{2.1}$$

Power is defined as the flow of energy per unit time, so the corresponding power dP is

$$dP = \frac{dE}{dt} = I_\nu\,(\cos\theta\,d\sigma\,d\Omega\,d\nu)\,.$$
$$(\text{watts}) \qquad\qquad (\text{m}^2\ \text{sr}\ \text{Hz})$$

Thus the quantitative definition of **specific intensity** or **spectral brightness** is

$$\boxed{I_\nu \equiv \frac{dP}{(\cos\theta\,d\sigma)\ \ d\nu\,d\Omega}} \tag{2.2}$$

and the MKS units of I_ν are $\text{W}\ \text{m}^{-2}\ \text{Hz}^{-1}\ \text{sr}^{-1}$.

Radio astronomers almost always measure frequencies, but observers in other wavebands normally measure wavelengths rather than frequencies, so the brightness per unit wavelength

$$I_\lambda \equiv \frac{dP}{(\cos\theta\,d\sigma)\ d\lambda\,d\Omega} \qquad (2.3)$$

is also widely used. The MKS units of I_λ are $\text{W m}^{-3}\ \text{sr}^{-1}$. The relation between the intensity *per unit frequency* I_ν and the intensity *per unit wavelength* I_λ can be derived from the requirement that the power dP in the frequency interval ν to $\nu + d\nu$ must equal the power in the corresponding wavelength interval λ to $\lambda + d\lambda$:

$$|I_\nu d\nu| = |I_\lambda d\lambda|\,. \qquad (2.4)$$

Thus

$$\frac{I_\lambda}{I_\nu} = \left|\frac{d\nu}{d\lambda}\right| = \frac{c}{\lambda^2} = \frac{\nu^2}{c}. \qquad (2.5)$$

The reasons for specifying the brightness in an infinitesimal frequency range $d\nu$ or wavelength range $d\lambda$ are (1) the detailed spectra of sources carry astrophysically important information, (2) source properties (e.g., opacity) may vary with frequency, and (3) most general theorems about radiation are true for *all* narrow frequency ranges (e.g., specific intensity is conserved along a ray path in empty space), so they are also true for any wider frequency range. Thus the conservation of specific intensity (Equation 2.7) implies that **total intensity** defined as

$$I \equiv \int_0^\infty I_\nu\,(\nu)\,d\nu = \int_0^\infty I_\lambda\,(\lambda)\,d\lambda \qquad (2.6)$$

is also conserved.

Theorem. Specific intensity is conserved (is constant) along any ray in empty space.

This follows directly from geometry. Let $d\sigma_1$ and $d\sigma_2$ be two infinitesimal surfaces along a ray of length r, as shown in Figure 2.3.
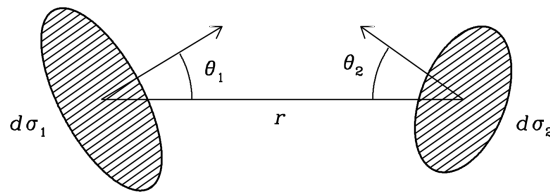

Figure 2.3: Specific intensity conserved along a ray of length r.

Let $d\Omega_1 \ll 1$ rad be the solid angle subtended by $d\sigma_2$ as seen from the center of the surface $d\sigma_1$ and $d\Omega_2 \ll 1$ rad be the solid angle subtended by $d\sigma_1$ as seen from the center of the surface $d\sigma_2$. Then

$$dB_1 = \frac{\cos\theta_2\,d\sigma_2}{r^2},$$
$$dB_2 = \frac{\cos\theta_1\,d\sigma_1}{r^2}.$$

The power dP_1 in the frequency range ν to $\nu + d\nu$ flowing through the area $d\sigma_1$ in solid angle $d\Omega_1$ is

$$dP_1 = \frac{dE_1}{dt} = (I_\nu)_1\,\cos\theta_1\,d\Omega_1\,d\sigma_1\,d\nu$$

$$= (I_\nu)_1 \cos\theta_1 \left(\frac{\cos\theta_2 \, d\sigma_2}{r^2} \right) d\sigma_1 \, d\nu$$

$$= (I_\nu)_1 \left(\frac{\cos\theta_1 \cos\theta_2 \, d\sigma_1 \, d\sigma_2}{r^2} \right) d\nu.$$

Likewise

$$dP_2 = \frac{dE_2}{dt} = (I_\nu)_2 \cos\theta_2 \left(\frac{\cos\theta_1 \, d\sigma_1}{r^2} \right) d\sigma_2 \, d\nu$$

$$= (I_\nu)_2 \left(\frac{\cos\theta_1 \cos\theta_2 \, d\sigma_1 \, d\sigma_2}{r^2} \right) d\nu.$$

Radiation energy is conserved in free space (where there is no absorption or emission), so $dE_1 = dE_2$ and

$$(I_\nu)_1 = (I_\nu)_2. \qquad \text{QED} \tag{2.7}$$

The conservation of specific intensity has two important consequences:

1. **Brightness is independent of distance.** Thus the camera setting for a good exposure of the Sun would be the same, regardless of whether the photograph was taken close to the Sun (from near Venus, for example) or far away from the Sun (from near Mars, for example), so long as the Sun is resolved in the photograph.
2. **Brightness is the same at the source and at the detector.** Thus you can think of brightness in terms of energy flowing out of the source or as energy flowing into the detector (Figure 2.4).
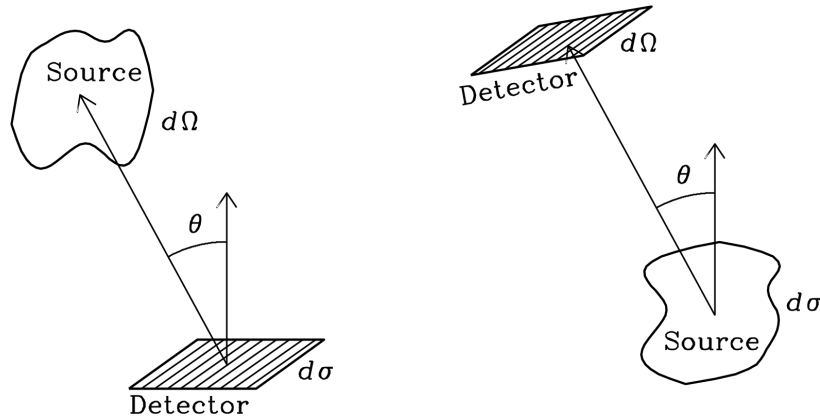

Figure 2.4: Two ways of looking at brightness.

The conservation of brightness applies to any lossless optical system (a system of lenses and mirrors, for example) that can change the direction of a ray. **No passive optical system can increase the specific intensity or total intensity of radiation**. If you look at the Moon through a large telescope, the Moon will appear bigger (in angular size) but not brighter. Many people are disappointed when they see a large galaxy through a telescope because it looks so dim; they expected to see a brilliantly glowing disk of stars, as in the photograph of Andromeda in Figure 2.5. The difference is not in the telescope; it is in the detector—the photograph appears brighter only because the photograph has accumulated more light over a long exposure time.

Figure 2.5: No passive optical system (e.g., a telescope) can increase the specific intensity of an extended source. The Andromeda galaxy (M31) appears much brighter in this photograph than it does to the eye, either with or without the aid of a telescope, only because a long photographic exposure accumulates more light. Image credit: Robert Gendler.

If a source is discrete, meaning that it subtends a well-defined solid angle, the spectral power received by a detector of unit projected area (Figure 2.6) is called the **flux density** S_ν of the source. Equation (2.2) implies

$$\frac{dP}{d\sigma d\nu} = I_\nu \cos\theta \, d\Omega, \tag{2.8}$$

so integrating over the solid angle subtended by the source yields

$$\boxed{S_\nu \equiv \int_{\text{source}} I_\nu\,(\theta, \phi) \cos\theta \, d\Omega.} \tag{2.9}$$

If the source angular size is $\ll 1$ rad, $\cos\theta \approx 1$ and the expression for flux density is much simpler:

$$\boxed{S_\nu \approx \int_{\text{source}} I_\nu\,(\theta, \phi) \, d\Omega.} \tag{2.10}$$

This is usually the case for astronomical sources, and astronomers rarely use flux densities to describe sources so extended that the $\cos\theta$ factor must be retained (e.g., the diffuse emission from our Galaxy).

In practice, when should spectral brightness and when should flux density be used to describe a source? If a source is **unresolved**, meaning that it is much smaller in angular size than the point-source response of the eye or telescope observing it, its flux density can be measured but its spectral brightness cannot. To the naked eye, the unresolved red giant star Betelgeuse appears to be one of the brightest stars in the sky. Yet calling it a "bright star" is misleading because the total intensity of this relatively cool star is lower than the total intensity of every hotter but more distant star that is scarcely visible to the eye. Betelgeuse appears "brighter" than most other stars only because it subtends a much larger solid angle and therefore its flux is

higher. If a source is much larger than the point-source response, its spectral brightness at any position on the source can be measured directly, but its flux density must be calculated by integrating the observed spectral brightnesses over the source solid angle. Consequently, flux densities are normally used to describe only relatively compact sources.
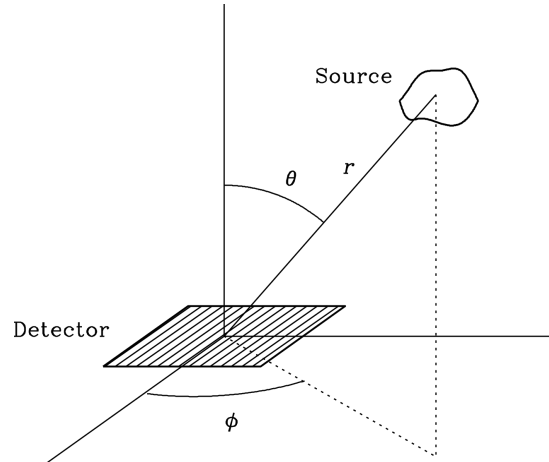

Figure 2.6: An illustration of the definition of flux density.

The MKS units of flux density, $\mathrm{W\ m^{-2}\ Hz^{-1}}$, are much too big for practical astronomical use, so astronomers use smaller ones:

$$1 \text{ jansky} = 1 \text{ Jy} \equiv 10^{-26} \mathrm{\ W\ m^{-2}\ Hz^{-1}}, \tag{2.11}$$

1 millijansky = 1 mJy $\equiv 10^{-3}$ Jy, and 1 microjansky = 1 μJy $\equiv 10^{-6}$ Jy. Optical astronomers often express flux densities as **AB magnitudes** defined in terms of Jy by

$$\mathrm{AB\ magnitude} \equiv -2.5 \log_{10}\left(\frac{S_\nu}{3631 \mathrm{\ Jy}}\right). \tag{2.12}$$

Unlike brightness, flux density depends on source distance d. Because $\int_{\text{source}} d\Omega \propto 1/d^2$ and brightness is conserved, Equation 2.10 implies the **inverse square law**:

$$S_\nu \propto d^{-2}. \tag{2.13}$$

The specific intensity or brightness is an intrinsic property of a source, while the flux density of a source also depends on the distance between the source and the observer.

The **total flux** or **flux** S from a source is the integral over frequency of flux density:

$$S \equiv \int_0^\infty S_\nu d\nu. \tag{2.14}$$

Its dimensions are power divided by area, so its MKS units are $\mathrm{W\ m^{-2}}$. Total flux is a rarely used quantity in observational radio astronomy, so radio astronomers often delete the subscript from S_ν and use the symbol S to indicate flux density. This is convenient but potentially confusing. Likewise, the word "flux" is sometimes used as a shorthand for "flux density" in the literature, even though it is formally incorrect.

The **spectral luminosity** L_ν of a source is defined as the total power per unit bandwidth radiated by the source at frequency ν; its MKS units are $\mathrm{W\ Hz^{-1}}$. The area of a sphere of radius d is $4\pi d^2$, so the relation between the spectral luminosity and the flux density of an *isotropic* source radiating in free space is

$$L_\nu = 4\pi d^2 \, S_\nu, \tag{2.15}$$

where the distance d between the source and the observer is much larger than the dimensions of the source itself. Beware that some radio sources emit anisotropically, relativistically beamed quasars (Section 5.6.1) for example. Unfortunately, Equation 2.15 cannot be used to calculate the total (integrated over 4π sr) spectral luminosity of a beamed quasar from a flux-density measurement made from just one direction. Spectral luminosity is an intrinsic property of the source because it does not depend on the distance d between the source and the observer—the d^2 in Equation. 2.15 cancels the d^{-2} dependence of S_ν. The **luminosity** or **total luminosity** L of a source is defined as the integral over all frequencies of the spectral luminosity:

$$L \equiv \int_0^\infty L_\nu \, d\nu. \tag{2.16}$$

Astronomers sometimes call L the **bolometric luminosity** because a bolometer is a broadband detector that measures the heating power of radiation at all frequencies.

2.2 Radiative Transfer

In free space, the specific intensity I_ν of radiation is conserved along a ray:

$$\frac{dI_\nu}{ds} = 0, \tag{2.17}$$

where s is the coordinate along the ray between the source and the detector. What happens if there is an intervening medium between s_{in} and s_{out} (Figure 2.7)?
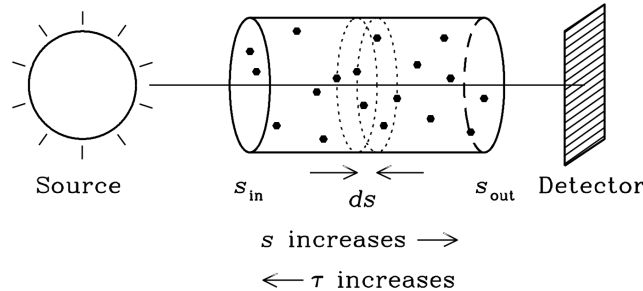

Figure 2.7: Absorption between a source and a detector. The coordinate along the ray increases from s_{in} at the input end of the absorber to s_{out} at the output end of the absorber. The optical depth τ is measured in the opposite direction, starting from $\tau = 0$ at s_{out} and increasing as s decreases.

2.2.1 Absorption

Think of the ray as a beam of photons, or light particles, some of which may be absorbed by the medium and vanish. The infinitesimal probability dP of a photon being absorbed (e.g., by hitting an absorbing particle) in a thin slab of thickness ds is directly proportional to ds: $dP = \kappa \, ds$, where the constant of proportionality

$$\kappa \equiv \frac{dP}{ds} \tag{2.18}$$

is called the **linear absorption coefficient**, and its dimension is inverse length. A value $\kappa = 1 \mathrm{~m}^{-1}$ means that over some small distance $\Delta s \ll \kappa^{-1}$ along the ray, $\Delta s = 10^{-3}$ m for example, the small fraction

$\kappa \Delta s = 1 \text{ m}^{-1} \times 10^{-3} \text{ m} = 10^{-3}$ of the photons will be absorbed. Why consider such thin layers rather than the whole absorber at once? Only if the layer is thin enough that dP is infinitesimal does the probability of absorption within the layer increase linearly with thickness ds so that κ is a constant independent of ds. In a layer thick enough to absorb a significant fraction of the photons, fewer photons are absorbed near the far end of the layer simply because fewer photons have survived to be absorbed there, so the probability of absorption increases nonlinearly with thickness.

This simple *macroscopic* model doesn't depend on the microscopic physical processes by which photons are absorbed. It is like the ideal gas model, which relates the macroscopic properties of a gas (temperature, pressure, and density) without considering the microscopic details about how individual gas molecules behave.

The fraction of the specific intensity lost to absorption in the infinitesimal distance ds along the ray is

$$\frac{dI_\nu}{I_\nu} = -\kappa \, ds \quad \text{(absorption only)}. \tag{2.19}$$

Integrating both sides of Equation 2.19 along the absorbing path gives the output specific intensity as a fraction of the input specific intensity:

$$\int_{s_{\text{in}}}^{s_{\text{out}}} \frac{dI_\nu}{I_\nu} = -\int_{s_{\text{in}}}^{s_{\text{out}}} \kappa \, (s') \; ds' = \ln I_\nu \Big|_{s_{\text{in}}}^{s_{\text{out}}}, \tag{2.20}$$

$$\ln[I_\nu \, (s_{\text{out}})] - \ln[I_\nu \, (s_{\text{in}})] = -\int_{s_{\text{in}}}^{s_{\text{out}}} \kappa \, (s') \; ds', \tag{2.21}$$

$$\frac{I_\nu \, (s_{\text{out}})}{I_\nu \, (s_{\text{in}})} = \exp\left[-\int_{s_{\text{in}}}^{s_{\text{out}}} \kappa \, (s') \; ds'\right]. \tag{2.22}$$

The dimensionless quantity

$$\boxed{\tau \equiv -\int_{s_{\text{out}}}^{s_{\text{in}}} \kappa \, (s') \; ds'} \tag{2.23}$$

is called the **optical depth** or **opacity** of the absorber. Note that $d\tau = -\kappa \, ds$. The "backward" direction of the path integration along the line of sight was chosen in Equation 2.23 to make $\tau > 0$ and increasing as you look deeper into an absorber. Thus

$$\frac{I_\nu \, (s_{\text{out}})}{I_\nu \, (s_{\text{in}})} = \exp(-\tau) \quad \text{(absorption only)}. \tag{2.24}$$

If $\tau \ll 1$, the absorber is said to be **optically thin**; if $\tau \gg 1$, it is **optically thick**.

2.2.2 Emission

The intervening medium may also emit photons, again by some unspecified microscopic process. In any infinitesimal volume $(ds \, d\sigma)$ of thickness ds and cross section $d\sigma$, the probability per unit time that an

isotropic source will emit a photon into the solid angle $d\Omega$ is directly proportional to the volume and solid angle:

$$\dot{P}_{\text{em}} \propto ds \, d\sigma \, d\Omega. \qquad (2.25)$$

The **emission coefficient** j_ν is defined so that

$$\boxed{j_\nu \equiv \frac{dI_\nu}{ds}} \qquad (2.26)$$

if there is no absorption. The dimensions of j_ν follow from its definition; they are power per unit volume per unit frequency per unit solid angle, and the corresponding MKS units are W m^{-3} Hz^{-1} sr^{-1}. Combining the effects of absorption (Equation 2.19) and emission (Equation 2.26) yields the **equation of radiative transfer**

$$\boxed{\frac{dI_\nu}{ds} = -\kappa I_\nu + j_\nu.} \qquad (2.27)$$

Because the absorption and emission processes have not been specified, κ and j_ν seem to be independent. However, they are not independent in full **thermodynamic equilibrium (TE)**. In TE, matter and radiation are in equilibrium at the same temperature T. **Cavity radiation** from a container bounded by opaque walls pierced by a small hole (Figure 2.8) is equilibrium radiation. The intensity and spectrum of the radiation emerging from the hole is independent of the wall material (e.g., wood painted green, shiny copper, gray concrete, etc.) and any absorbing material (e.g., gas, dust, fog, etc.) that may be inside the cavity.
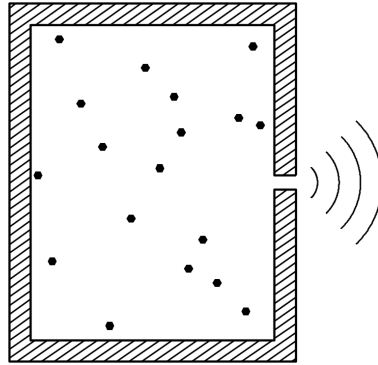

Figure 2.8: Thermal radiation from a cavity depends only on its temperature and is independent of both the material used to make the cavity walls and the contents of the cavity.
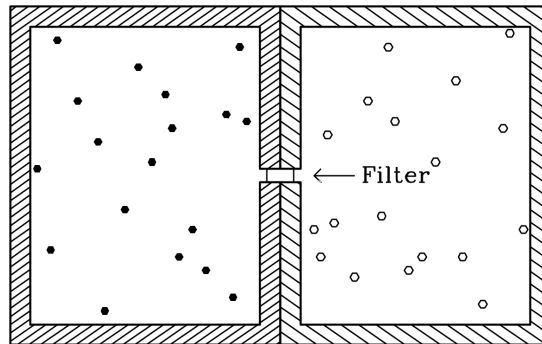

Figure 2.9: Kirchhoff's thought experiment invokes two cavities in thermodynamic equilibrium connected through a filter that passes radiation in the narrow frequency range ν to $\nu + d\nu$. The cavities may be made of different materials and contain different emitting/absorbing particles.

Kirchhoff derived the relation between κ and j_ν in TE using the thought experiment illustrated in Figure 2.9. Two cavities made of different materials and containing different absorbers are connected through a passive (meaning, no energy is needed for it to work) filter transparent only in the narrow

frequency range ν to $\nu + d\nu$. In equilibrium at any temperature T, radiation can transfer no *net* power from one cavity to the other, lest one cavity cool down and the other heat up. That would violate the second law of thermodynamics because the two cavities at different temperatures could be used to drive a heat engine.

In (full) thermodynamic equilibrium (TE) at temperature T,

$$\frac{dI_\nu}{ds} = 0 \quad \text{and} \quad I_\nu = B_\nu(T), \tag{2.28}$$

where $B_\nu(T)$ is the spectrum of equilibrium radiation at temperature T. The symbol $B_\nu(T)$ is used because equilibrium cavity radiation is the same as the **blackbody radiation** for a perfect absorber at temperature T, even if it is not in a cavity. Rearranging the equation of radiative transfer (Equation 2.27)

$$\frac{dI_\nu}{ds} = 0 = -\kappa B_\nu(T) + j_\nu \tag{2.29}$$

yields **Kirchhoff's law** for a system in TE:

$$\boxed{\frac{j_\nu(T)}{\kappa(T)} = B_\nu(T),} \tag{2.30}$$

which is valid at any frequency ν. Equation 2.30 is remarkable because it connects the properties $j_\nu(T)$ and $\kappa(T)$ of any kind of *matter* to the single universal spectrum $B_\nu(T)$ of equilibrium *radiation*. Note also that Kirchhoff's argument is independent of the shapes of the cavities, so cavity radiation must be isotropic.

Although Kirchhoff's law was derived for a system in thermodynamic equilibrium, its applicability is not limited to radiation in full thermodynamic equilibrium with its material environment. Kirchhoff's law also applies whenever the radiating/absorbing *material* is in thermal equilibrium, in *any* radiation field. If the emitting/absorbing material is in thermal equilibrium at a well-defined temperature T, it is said to be in **local thermodynamic equilibrium (LTE)** even if it is not in equilibrium with the radiation field. For example, gas molecules in the Earth's lower atmosphere have a Maxwellian speed distribution (Equation 4.34), so the gas is in LTE. The gas has a well-defined **kinetic temperature** $T \sim 300$ K measurable with an ordinary thermometer, but it is in a distinctly nonequilibrium radiation field (anisotropic $T \sim 5800$ K sunlight during the day, the cold dark sky at night, plus anisotropic emission from the ground).

Kirchhoff's law applies in LTE as well as in TE. To show this, recall that $B_\nu(T)$ is independent of the properties of the radiating/absorbing material. In contrast, both $j_\nu(T)$ and $\kappa(T)$ depend *only* on the materials in the cavity (e.g., whether the walls are made of copper or concrete) and on the temperature of that material; they do not depend on the ambient radiation field or its spectrum.

This generalized version of Kirchhoff's law is an exceptionally valuable tool for calculating the emission coefficient from the absorption coefficient or vice versa. For example, in Section 4.3 Kirchhoff's law gives $\kappa(\nu)$ immediately following a calculation of $j_\nu(\nu)$ for free–free emission by an HII region (a cloud of ionized hydrogen) in LTE.

The blackbody spectrum $B_\nu(T)$ (Equation 2.86) falls exponentially at frequencies above the peak, but only as a power law at lower frequencies (solid curve in Figure 2.20). Kirchhoff's law is not intuitively obvious from experience because room-temperature ($T \sim 300$ K) objects in our environment are much too cold ($h\nu/kT \gg 1$) to emit detectable amounts of visible light. A glass of water might absorb 10% of the sunlight passing through it, but we do not see any emitted light because 10% (or even 100%) of the blackbody radiation emitted by a room-temperature object is very nearly zero. One familiar example of Kirchhoff's law is a charcoal fire with flames and glowing coals. The infrared (heat) radiation from barely glowing black coals in LTE is much more intense than that from the hotter but nearly transparent visible flames, as you can verify by using a shield to cover either the coals or the flames. In contrast, the radio

emission implied by Kirchhoff's law is always significant because $h\nu/kT \ll 1$ when $T \sim 300$ K and the blackbody spectrum below the peak falls off only as ν^{-2}. This point is illustrated by radio emission and absorption in the Earth's atmosphere (Section 2.2.3).

2.2.3 Emission and Absorption of Radio Waves in the Earth's Atmosphere

At radio frequencies higher than $\nu \sim 1$ GHz, absorption by the Earth's atmosphere may be large enough to affect the accuracy of flux-density measurements and atmospheric emission can increase noise errors. Radio astronomers can determine the amount of **atmospheric absorption** by measuring the amount of **atmospheric emission** as a function of **zenith angle**, or angle from the vertical, and using Kirchhoff's law to calculate the zenith opacity of the (roughly isothermal) atmosphere at frequency ν from its kinetic (thermometer) temperature $T_{\text{atm}} \sim 300$ K and its radio brightness. The celestial sky above the atmosphere is much colder ($T \sim 3$ K) at high frequencies, so the "background" emission above the atmosphere can usually be ignored.
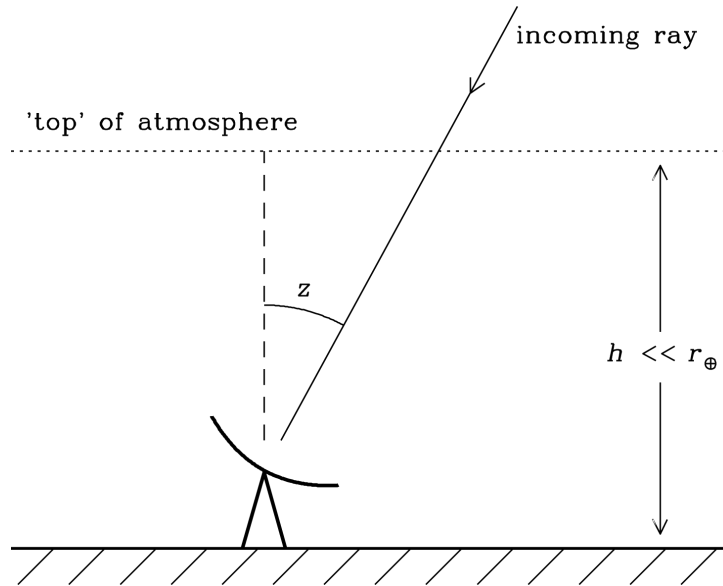

Figure 2.10: Most atmospheric emission and absorption occurs in a layer whose height h is only a few km, much smaller than the radius of the Earth $r_\oplus$, so the plane-parallel approximation for the spherical Earth shown here is accurate.

This is done by tilting the radio telescope and measuring I_ν as a function of the zenith angle z as shown in Figure 2.10. There are two practical complications: (1) The atmospheric signal is noise indistinguishable from other noise sources, especially noise generated in the receiver itself. The output voltage of the receiver is proportional to the sum of these input noise powers. The amount of receiver noise power may not be well known, so only the *change* in I_ν with zenith angle z can be measured, not the absolute value of I_ν at any zenith angle. (2) The power gain of the receiver system is difficult to calculate from first principles, so most measurements are made relative to some calibration source. For example, the receiver might be calibrated by having the feed look alternately at two absorbing plates with different known temperatures.

Because B_ν is directly proportional to T in the Rayleigh–Jeans low-frequency approximation

$$B_\nu \approx \frac{2kT\nu^2}{c^2}, \tag{2.31}$$

radio astronomers often find it convenient to specify the spectral brightness I_ν, even if $I_\nu \neq B_\nu$, in terms of the equivalent Rayleigh–Jeans **brightness temperature** T_{b} *defined* by the equation

$$I_\nu = \frac{2kT_{\text{b}}\nu^2}{c^2}. \tag{2.32}$$

Thus for *any* I_ν,

$$T_{\mathrm{b}}\left(\nu\right) \equiv \frac{I_\nu c^2}{2k\nu^2}.\tag{2.33}$$

Equation 2.33 explicitly indicates that T_{b} can vary with frequency. Brightness temperature is just another way to specify power per unit solid angle per unit bandwidth in terms of the Rayleigh–Jeans approximation. It is convenient because radio telescopes are often calibrated by absorbers or "loads" of known temperature and because the temperature of a radio source is frequently a quantity of physical interest. Beware that brightness temperature is *not* the same as physical temperature. Nonthermal sources have frequency-dependent brightness temperatures but they do not have well-defined physical temperatures. Thermal sources have brightness temperatures lower than their physical temperatures if they are semitransparent (e.g., the atmosphere) or partially reflecting (e.g., the Moon). Even true blackbody radiators have $T_{\mathrm{b}} = T$ only in the low-frequency limit $h\nu \ll kT$.

The observed output from the radio telescope might vary with z as shown in Figure 2.11.
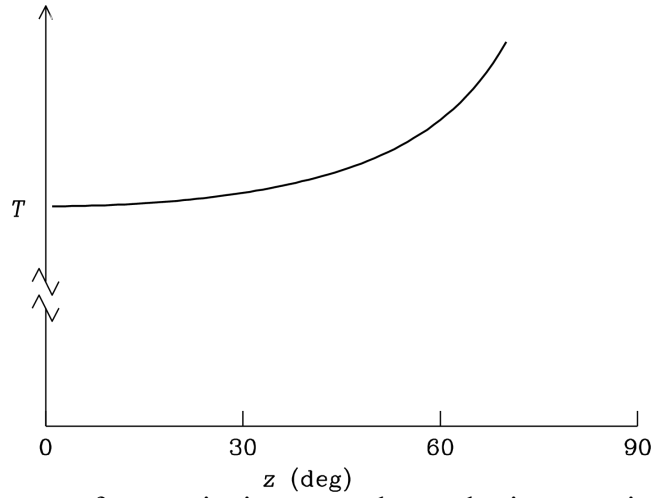

Figure 2.11: The receiver output from a tipping scan shows the increase in sky brightness with zenith angle, but the zero point of the sky brightness temperature is not well determined.

To understand and interpret these observations, start with the radiative transfer equation (2.27),

$$\frac{dI_\nu}{ds} = -\kappa I_\nu + j_\nu.\tag{2.34}$$

Kirchhoff's law (Equation 2.30) can eliminate the unknown j_ν because the atmosphere is in LTE:

$$j_\nu = \kappa B_\nu\left(T_{\mathrm{atm}}\right),\tag{2.35}$$

where T_{atm} is the kinetic temperature of the atmosphere as measured by an ordinary thermometer. Dividing by κ_ν yields

$$\frac{1}{\kappa}\frac{dI_\nu}{ds} = \frac{-dI_\nu}{d\tau} = -I_\nu + B_\nu\left(T_{\mathrm{atm}}\right).\tag{2.36}$$

Multiply both sides of this differential equation by $\exp(-\tau)$ and integrate along the ray in the telescope beam from the top of the atmosphere to the ground. Let $\tau_{\mathrm{A}}\left(z\right)$ be the total optical depth of the atmosphere along the ray at zenith angle z. Then

$$\int_0^{\tau_A} e^{-\tau} \frac{dI_\nu}{d\tau} d\tau = \int_0^{\tau_A} [I_\nu - B_\nu (T_{\mathrm{atm}})] e^{-\tau} d\tau. \tag{2.37}$$
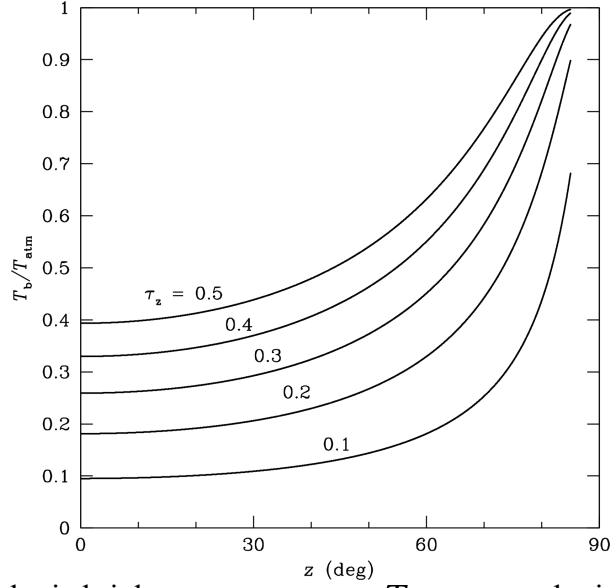

Figure 2.12: Ratios of atmospheric brightness temperature T_b to atmospheric kinetic temperature T_{atm} for different zenith opacities τ_z as functions of zenith angle z.

Next integrate the left side by parts and take $B_\nu (T_{\mathrm{atm}})$ outside the integral (i.e., make the approximation that the atmosphere below altitude h is nearly isothermal):

$$e^{-\tau} I_\nu \big|_0^{\tau_A} - \int_0^{\tau_A} -e^{-\tau} I_\nu \, d\tau = \int_0^{\tau_A} I_\nu \, e^{-\tau} d\tau - B_\nu (T_{\mathrm{atm}}) \int_0^{\tau_A} e^{-\tau} d\tau, \tag{2.38}$$

$$I_\nu (\tau = \tau_A) e^{-\tau_A} - I_\nu (\tau = 0) = B_\nu (T_{\mathrm{atm}}) \left(e^{-\tau_A} - 1 \right). \tag{2.39}$$

The term $I_\nu (\tau = \tau_A) \approx 0$ can be neglected because the brightness of emission above the atmosphere is small and doesn't depend on zenith angle at high radio frequencies, so that

$$I_\nu (\tau = 0) = \left(1 - e^{-\tau_A} \right) B_\nu (T_{\mathrm{atm}}). \tag{2.40}$$

The path length through a plane-parallel atmosphere is proportional to sec z so at any zenith angle z,

$$\tau_A \approx \tau_Z \sec z, \tag{2.41}$$

where $\tau_Z \equiv \tau_A (z = 0)$ is the zenith opacity of the atmosphere above the radio telescope. Inserting the Rayleigh–Jeans approximation for $B_\nu (T_{\mathrm{atm}})$ gives

$$I_\nu = [1 - \exp(-\tau_Z \sec z)] \frac{2k T_{\mathrm{atm}} \nu^2}{c^2}. \tag{2.42}$$

Thus the brightness temperature of the atmospheric emission as a function of zenith angle is

$$T_{\rm b} = \frac{I_\nu c^2}{2k\nu^2} = T_{\rm atm} \left[1 - \exp(-\tau_{\rm Z} \sec z)\right]. \qquad (2.43)$$

The variation of $T_{\rm b}$ with zenith angle is shown for different zenith opacities in Figure 2.12. By fitting observed data to such curves, radio astronomers can estimate the zenith opacity and correct the measured flux densities of celestial sources for atmospheric absorption.

If (1) the zenith opacity is low ($\tau_{\rm z} \ll 1$), (2) the frequency is high enough ($\nu > 1$ GHz) that the radio-source background is weak, and (3) there is little pickup of "spillover" radiation from the ground, then a good approximation to the observed system noise temperature $T_{\rm s}$ (Equation 3.150) is

$$T_{\rm s} \approx [T_{\rm r} + T_{\rm cmb}] + T_{\rm atm}\tau_{\rm z} \sec(z), \qquad (2.44)$$

where $T_{\rm r}$ is the noise temperature of the receiver alone and $T_{\rm cmb} \approx 2.7$ K is the brightness temperature of the cosmic microwave background. The quantity in square brackets is independent of z and $T_{\rm atm}$ is close to the surface air temperature, which can easily be measured with a thermometer. Equation 2.44 implies that the zenith opacity

$$\tau_{\rm z} \approx \frac{\Delta T_{\rm s}/T_{\rm atm}}{\Delta \sec(z)} \qquad (2.45)$$

can be read directly from the slope of the linear tipping-curve plot of $T_{\rm s}/T_{\rm atm}$ versus $\sec(z)$. The solid line in Figure 2.13 shows an example in which $\tau_{\rm z} \approx 0.04$. Extrapolating the data to (the geometrically impossible) $\sec(z) = 0$ yields $[T_{\rm r} + T_{\rm cmb}]/T_{\rm atm}$. If $T_{\rm atm} = 280$ K, the example $[T_{\rm r} + T_{\rm cmb}]/T_{\rm atm} = 0.14$ implies $[T_{\rm r} + T_{\rm cmb}] \approx 39.2$ K and thus $T_{\rm r} \approx 36.5$ K. Alternatively, if $T_{\rm r}$ is measured in the laboratory, the tipping curve can be used to determine $T_{\rm cmb}$, the temperature of the cosmic microwave background, corrected for atmospheric emission. This is essentially how Penzias and Wilson [81] discovered the $T_{\rm cmb} \approx 3$ K CMB at $\nu \approx 4$ GHz using the low-spillover Bell Labs horn antenna (Figure 3.19).
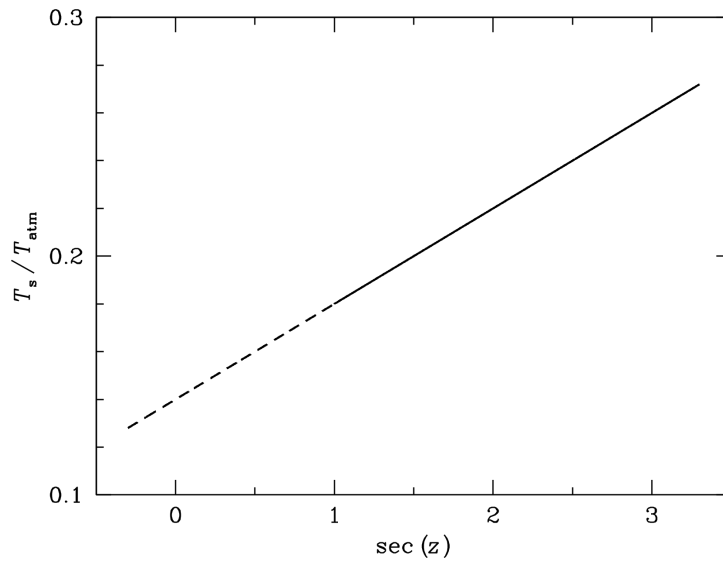

Figure 2.13: If the zenith opacity is low, a plot of the observed ratio ($T_{\rm s}/T_{\rm atm}$) versus $\sec(z)$ is nearly linear (solid line), and its slope equals the zenith opacity; $\tau_{\rm z} \approx 0.04$ in this case. Extrapolating the solid line to $\sec(z) = 0$ (dashed line) yields the ratio $[T_{\rm r} + T_{\rm cmb}]/T_{\rm atm} \approx 0.14$.

2.2.4 Emission, Absorption, and Reflection from an Opaque Body

A body is **opaque** if its opacity is so high that photons cannot pass through it. The **absorption coefficient** $a(\nu)$ of an opaque body is the probability that an incident photon of frequency ν will be absorbed, and the

reflection coefficient $r(\nu)$ is the probability that it will be reflected. Thus $a = 1$ for a blackbody and $r = 1$ for a perfect reflector. Each photon is either absorbed or reflected by any opaque body so

$$a(\nu) + r(\nu) = 1. \tag{2.46}$$

The **emission coefficient** $e(\nu)$ of an opaque body is defined as the ratio of the spectral power it emits per unit area at frequency ν to that of a blackbody of the same temperature.

The values of $a(\nu)$ and $e(\nu)$ for any opaque body in local thermodynamic equilibrium are not independent. Imagine an opaque body surrounded by a filter passing frequencies in the range ν to $\nu + d\nu$ inside a cavity (Figure 2.14) in thermodynamic equilibrium at some temperature T. In equilibrium, the spectral power absorbed by the opaque body must equal the spectral power emitted, so

$$\boxed{e(\nu) = a(\nu) = 1 - r(\nu).} \tag{2.47}$$

This is **Kirchhoff's law for opaque bodies**. It implies that the brightness temperature T_b of an opaque body in LTE at physical temperature T is given by

$$T_b(\nu) = a(\nu)\,T = [1 - r(\nu)]\,T. \tag{2.48}$$

Because $a \leq 1$, then $T_b \leq T$. A perfect reflector ($r = 1$) will always have $T_b = 0$; that is, it does not emit.
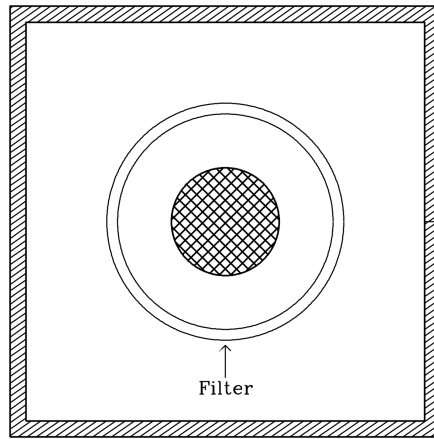

Figure 2.14: Thought experiment involving an opaque body surrounded by a filter passing frequencies ν to $\nu + d\nu$ inside a cavity in thermodynamic equilibrium.

Example. Temperature differences across the GBT reflector (Figure 8.1) caused by differential solar heating can deform the surface and degrade its performance. A special paint that is white at visible wavelengths, black in the mid-infrared, and transparent at radio wavelengths keeps the surface cool and does not harm performance at radio wavelengths by absorbing incoming radio waves or emitting radio noise. The special paint exploits Kirchhoff's law to perform three separate functions simultaneously: It is opaque and white in the visible portion of the spectrum to reflect sunlight ($T \approx 5800$ K). It is black in the mid-infrared so that the GBT ($T \approx 300$ K) can cool itself efficiently by reradiation. It is transparent at radio wavelengths so that it neither absorbs incoming radio waves nor emits thermal noise at radio wavelengths.
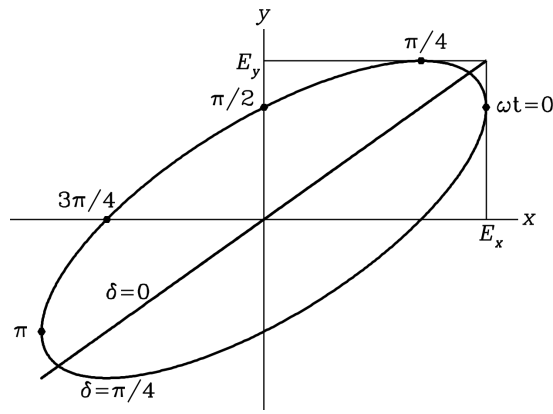
2.3 Polarization

Figure 2.15: The electric field vector of any monochromatic wave traveling in the $\hat{z}$-direction pointing up out of the page traces an ellipse that can be written in the form $\vec{E} = [\hat{x}E_x \exp(i\phi_x) + \hat{y}E_y \exp(i\phi_y)] \exp[i(\vec{k} \cdot \hat{z} - \omega t)]$. The ellipse shown has $\delta = +\pi/4$ and $E_x/E_y = 1.4$. If $\delta \equiv \phi_x - \phi_y = 0$, the ellipse becomes a line, while $\delta = \pm\pi/2$, $E_x = E_y$ would make it a circle. When $\delta > 0$, the tip of the vector rotates clockwise as viewed from the source below the page and counterclockwise as seen by an observer above the page, as shown by the time samples at $\omega t = 0, \ \pi/4, \ \pi/2, \dots$.

The instantaneous transverse electric field $\vec{E}$ of a monochromatic electromagnetic wave traveling in the $\hat{z}$-direction can be projected onto orthogonal $\hat{x}$- and $\hat{y}$- (e.g., horizontal and vertical) directions:

$$\vec{E} = [\hat{x}E_x \exp(i\phi_x) + \hat{y}E_y \exp(i\phi_y)] \ \exp[i(\vec{k} \cdot \hat{z} - \omega t)], \tag{2.49}$$

where $k \equiv 2\pi/\lambda$ is the magnitude of the **wave vector** $\vec{k}$ pointing in the direction of wave travel,

$$\omega \equiv 2\pi\nu \tag{2.50}$$

is the **angular frequency**,

$$\delta \equiv \phi_x - \phi_y \tag{2.51}$$

is the **phase difference** between the orthogonal fields E_x and E_y, and

$$E^2 = |\vec{E}|^2 = E_x^2 + E_y^2. \tag{2.52}$$

Any time-independent combination of phases and amplitudes yields an **elliptically polarized** wave (Figure 2.15) whose electric field vector traces out an ellipse in the (x, y) plane. If the phase difference δ is zero, the electric field vector does not rotate and the wave is **linearly polarized**. If $E_x = E_y$ and $|\delta| = \pi/2$, the electric field vector rotates with angular frequency ω and traces out a circle; such radiation is said to be **circularly polarized**. The **sign convention for circular polarization** used by the Institute of Electrical and Electronics Engineers (IEEE) and by the International Astronomical Union (IAU) is to call the polarization right-handed or left-handed depending on whether the rotation is clockwise ($\delta > 0$) or counterclockwise ($\delta < 0$), respectively, as viewed *from the source toward the observer*. An observer looking toward the source sees the electric-field vector from right-handed polarization rotating counterclockwise, as shown by the different time samples $\omega t = 0, \pi/4, \pi/2, \dots$ in Figure 2.15. Beware that some optics textbooks use the opposite sign convention.

The radiation from astronomical sources is wideband noise whose electric field vector varies rapidly and randomly in amplitude and direction. If radiation in a unit frequency range $\Delta\nu = \Delta\omega/(2\pi)$ is averaged over timescales $\tau \gg (\Delta\omega)^{-1}$, its average polarization can be characterized by the four **Stokes parameters**,

$$I = \langle E_x^2 + E_y^2 \rangle / R_0, \tag{2.53}$$
$$Q = \langle E_x^2 - E_y^2 \rangle / R_0, \tag{2.54}$$
$$U = \langle 2 E_x E_y \cos \delta \rangle / R_0, \tag{2.55}$$
$$V = \langle 2 E_x E_y \sin \delta \rangle / R_0, \tag{2.56}$$

where the brackets indicate time averages, R_0 is the radiation resistance of free space (Equation 3.29), and I is the total flux density, regardless of polarization. The **polarized flux density** is

$$I_{\rm p} = (Q^2 + U^2 + V^2)^{1/2} \tag{2.57}$$

and the **degree of polarization** is defined as

$$p \equiv \frac{I_{\rm p}}{I}. \tag{2.58}$$

If $\langle E_x \rangle$ and $\langle E_y \rangle$ have equal amplitudes and their phases are completely uncorrelated, then $Q = U = V = 0$, $I_{\rm p} = 0$, $p = 0$, and the wave is said to be **unpolarized**. For example, blackbody radiation is unpolarized. An antenna sensitive only to one polarization (e.g., a dipole antenna oriented so it is sensitive only to the $\hat{x}$ component of linear polarization or a helical antenna sensitive only to the left-handed component of circular polarization) will detect only half the power radiated by an unpolarized source. Two orthogonally polarized antennas (e.g., dipoles aligned in the $\hat{x}$- and $\hat{y}$-directions or left-handed and right-handed helical antennas) are needed to collect all of the unpolarized power.

Many astronomical sources are **partially polarized** with $0 < p < 1$. The quantity $(Q^2 + U^2)^{1/2}$ measures the linearly polarized component of flux, and the ratio Q/U depends on the linear polarization position angle. The circularly polarized flux is given by $|V|$, with $V > 0$ indicating right-handed and $V < 0$ indicating left-handed circular polarization.

2.4 Blackbody Radiation

A perfect absorber has $a(\nu) = 1$ at all frequencies and is called a **blackbody**. Kirchhoff's law (Equation 2.47) requires that every blackbody has emission coefficient $e(\nu) = 1$ at all frequencies as well, so the radiated spectrum $B_\nu(T)$ of any blackbody at temperature T is the same as the spectrum of radiation in thermodynamic equilibrium inside a cavity of temperature T, even if the interior walls of the cavity are not black. Thus the intensity and spectrum of blackbody radiation depend *only* on the temperature of the blackbody or cavity. The most fundamental feature of blackbody radiation is that it is **equilibrium radiation**; the only reason for discussing **cavity radiation** is that the cavity traps radiation long enough for it to come into equilibrium. The radiation escaping from a small hole in the cavity is also **blackbody radiation** because radiation entering the hole from the outside has almost no chance of escaping. A small hole in the side of a very cold cavity looks black. The energy density and brightness spectrum of *equilibrium radiation* will be derived in Sections 2.4.1 and 2.4.2, even though it will be called *blackbody radiation*.

The same is true for the electrical noise generated by a warm **resistor**, a device that completely absorbs electrical energy, and which plays an important role in radio astronomy. The standard derivations of the Rayleigh–Jeans and Planck radiation equations are worth repeating because blackbody radiation is so fundamental and because their one-dimensional analogs yield the spectrum of electrical noise generated by a resistor in equilibrium at temperature T.

2.4.1 The Rayleigh–Jeans Approximation

Consider a large (side length $a \gg \lambda$, where λ is the longest wavelength of interest) cubical **cavity** filled with radiation in thermodynamic equilibrium. The purpose of the cavity is to generate radiation and to confine the radiation long enough for it to reach equilibrium. The radiation must be generated by thermal

accelerations of charged particles in the walls of any cavity with $T > 0$. The walls have must have nonzero conductivity because walls of zero conductivity would be transparent—they would generate no currents in response to the electric fields of incoming radiation and the radiation would simply escape. The equilibrium radiation is otherwise independent of the wall material. For walls having nonzero conductivity, the transverse electric field strength at the walls is $E = 0$ in equilibrium because fields with $E \neq 0$ induce currents in the walls and lose energy. Only those standing waves with $E = 0$ at the walls will persist after some time $t \gg a/c$.

All possible standing-wave **modes** (a mode is a field pattern that oscillates sinusoidally with a fixed frequency and phase) in the cavity can be enumerated. For example, consider all standing waves whose wave vectors point in the x-direction (Figure 2.16). The boundary conditions $E = 0$ at $x = 0$ and at $x = a$ mean that only those waves having the discrete wavelengths

$$\frac{\lambda_x}{2} = a, \quad \frac{2\lambda_x}{2} = a, \quad \frac{3\lambda_x}{2} = a, \quad \ldots \tag{2.59}$$

can have nonzero amplitudes. They all satisfy

$$\frac{n_x \lambda_x}{2} = a, \tag{2.60}$$

where $n = 1, 2, 3, \ldots$. Likewise,

$$\frac{n_y \lambda_y}{2} = \frac{n_z \lambda_z}{2} = a \tag{2.61}$$

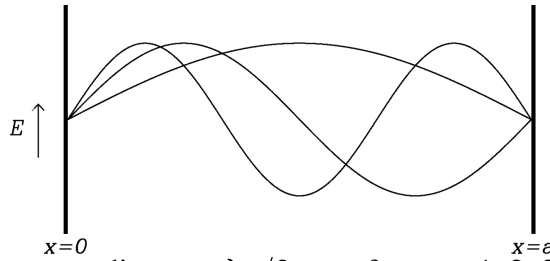for wave vectors pointing in the y- and z-directions, respectively.

Figure 2.16: Standing waves corresponding to $n\lambda_x/2 = a$ for $n = 1, 2, 3$. Abscissa: x-axis of the cavity, bounded by walls at $x = 0$ and $x = a$. Ordinate: electric field strengths of the three longest standing waves satisfying the boundary condition $E = 0$ at the walls.

What about a wave vector pointing in some arbitrary direction? Let α, β, γ be the angles between the wave vector and the x-, y-, z-axes, respectively. From Figure 2.18 it is clear that

$$\lambda = \lambda_x \cos \alpha, \tag{2.62}$$

where λ is the wavelength measured in the direction of the wave vector and $\lambda_x \geq \lambda$ is the component of λ projected onto the x-axis. Thus

$$\lambda_x = \frac{\lambda}{\cos \alpha}. \tag{2.63}$$

Notice that $\lambda_x > \lambda$; it is λ divided by $\cos \alpha$, not multiplied by $\cos \alpha$. Simultaneously, the standing waves must also satisfy

$$\lambda_y = \frac{\lambda}{\cos\,\beta} \text{ and } \lambda_z = \frac{\lambda}{\cos\,\gamma}. \tag{2.64}$$

Equations 2.60 through 2.64 imply that standing waves of *all* orientations must satisfy

$$n_x = \frac{2a}{\lambda_x}, \quad n_y = \frac{2a}{\lambda_y}, \quad n_z = \frac{2a}{\lambda_z}, \tag{2.65}$$

$$n_x = \frac{2a\,\cos\,\alpha}{\lambda}, \quad n_y = \frac{2a\,\cos\,\beta}{\lambda}, \quad n_z = \frac{2a\,\cos\,\gamma}{\lambda}. \tag{2.66}$$

Squaring and summing the three parts of Equation 2.66 gives

$$n_x^2 + n_y^2 + n_z^2 = \left(\frac{2a}{\lambda}\right)^2 \left(\cos^2\,\alpha + \cos^2\,\beta + \cos^2\,\gamma\right). \tag{2.67}$$

The Pythagorean theorem implies $\left(\cos^2\,\alpha + \cos^2\,\beta + \cos^2\,\gamma\right) = 1$ so all standing waves must satisfy

$$n_x^2 + n_y^2 + n_z^2 = \left(\frac{2a}{\lambda}\right)^2. \tag{2.68}$$

The permitted frequencies $\nu = c/\lambda$ of the standing waves are

$$\nu = \frac{c}{2a}\sqrt{n_x^2 + n_y^2 + n_z^2} \tag{2.69}$$

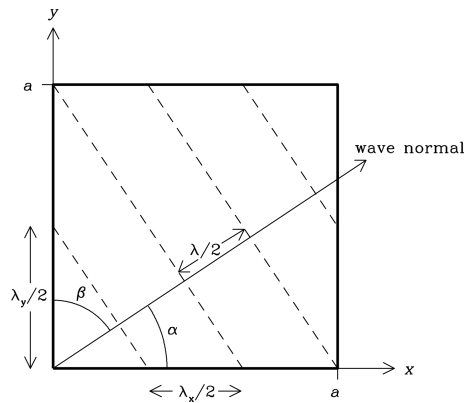for all positive integers n_x, n_y, and n_z.

Figure 2.17: This two-dimensional figure illustrates standing waves propagating in a cavity with wave vectors at angles α and β from the x- and y-axes, respectively. Examples of wave nodes, where $|E| = 0$, are indicated by dashed lines for the case $n_x = 3$, $n_y = 2$.
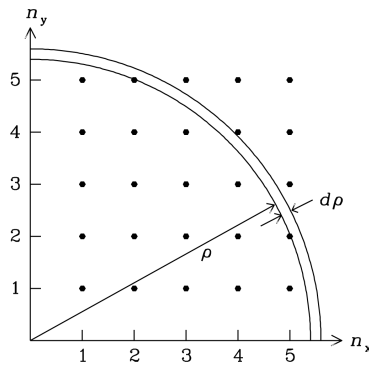
Figure 2.18: The (x, y) plane in the space whose axes are $(n_x > 0, \ n_y > 0, \ n_z > 0)$. Dots at all positive integers mark standing waves in the $(n_x, \ n_y)$ plane. In this space, frequency is proportional to the radial distance ρ from the origin.

The permitted standing waves can be represented as a lattice in the positive octant of the space whose axes are $n_x > 0$, $n_y > 0$, and $n_z > 0$ (Figure 2.18). Each point of the lattice represents one possible mode of equilibrium cavity radiation. The space density of points in this lattice is unity, so the average number of points in any volume equals that volume.

Let ρ be the radial coordinate in $(n_x, \ n_y, \ n_z)$-space. Then

$$\rho^2 = n_x^2 + n_y^2 + n_z^2 \tag{2.70}$$

and

$$\nu = \frac{c}{\lambda} = \left(\frac{c}{2a}\right)\rho. \tag{2.71}$$

The number $N_\nu(\nu)\, d\nu$ of independent modes having frequencies in the range ν to $\nu + d\nu$ equals the volume of the spherical octant shell between ρ and $\rho + d\rho$, multiplied by 2 to account for the two independent orthogonal polarizations (Section 2.3) of electromagnetic waves:

$$N_\nu(\nu)\, d\nu = \frac{4\pi\rho^2 d\rho}{8} \times 2 = \pi\left(\frac{2a\nu}{c}\right)^2 \frac{2a}{c} d\nu. \tag{2.72}$$

Classically, each radiation mode can have any energy $E > 0$. In thermodynamic equilibrium at temperature T, the probability $P(E)$ that any mode has energy E is given by the continuous **Boltzmann probability distribution**

$$P(E) \propto \exp\left(-\frac{E}{kT}\right). \tag{2.73}$$

Then the average energy per mode is

$$\langle E \rangle = \frac{\int_0^\infty E P(E)\, dE}{\int_0^\infty P(E)\, dE} = kT. \tag{2.74}$$

Thus each mode has the same average energy $\langle E \rangle = kT$. The spectral energy density $u_\nu(T)$ of cavity radiation at frequency ν is the total energy of all N_ν modes (Equation 2.72) at that frequency divided by the

volume a^3 of the cavity:

$$u_\nu(T) = \frac{N_\nu(\nu)\ kT}{a^3} = \frac{8\pi a^3}{a^3}\frac{\nu^2}{c^3}kT = 8\pi kT\frac{\nu^2}{c^3}. \tag{2.75}$$

The **spectral energy density** of radiation (blackbody or not) is spectral energy per unit volume. It equals the total flow of spectral power per unit area divided by the flow speed c:

$$\boxed{u_\nu = \frac{1}{c}\int I_\nu d\Omega.} \tag{2.76}$$

Calling the specific intensity of blackbody radiation B_ν and making use of the fact the blackbody radiation is isotropic yields, for blackbody radiation,

$$u_\nu = \frac{1}{c}\int_{4\pi} B_\nu d\Omega = \frac{4\pi}{c}B_\nu. \tag{2.77}$$

Combining Equations 2.75 and 2.77 gives

$$\frac{4\pi}{c}B_\nu = \frac{8\pi kT\nu^2}{c^3}. \tag{2.78}$$

Solving for B_ν yields the **Rayleigh–Jeans approximation** for the spectral brightness of blackbody radiation

$$\boxed{B_\nu = \frac{2kT\nu^2}{c^2} = \frac{2kT}{\lambda^2}} \tag{2.79}$$

that is valid only in the low-frequency limit $h\nu \ll kT$.

Notice that

1. the spectral brightness B_ν is proportional to frequency squared because the volume of a spherical shell in three dimensions is proportional to frequency squared;
2. all modes have $\langle E \rangle = kT$, the classical assumption that breaks down at high frequencies;
3. the Rayleigh–Jeans approximation implies that B_ν diverges at high ν; this is called the **ultraviolet catastrophe**;
4. B_ν is independent of direction; blackbody radiation is **isotropic**;
5. blackbody radiation is **unpolarized**; you can show this with a thought experiment involving two cavities connected by a passive polarizing filter, as in Figure 2.9;
6. this result derived for a cubical cavity applies to equilibrium radiation in a cavity of any shape or (large) size, which you can demonstrate by a thought experiment in which a cubical cavity is connected through a small filter window to the other cavity.

2.4.2 The Planck Radiation Law

The only flaw in the derivation of the Rayleigh–Jeans approximation is the classical assumption that each radiation mode can have any energy $E > 0$. To eliminate the ultraviolet catastrophe, Planck postulated that the possible mode energies are not continuously distributed, but rather they are quantized and must satisfy the new constraint

$$E = nh\nu, \quad n = 1, 2, 3, \ldots, \tag{2.80}$$

where $h \approx 6.63 \times 10^{-27}$ erg s is Planck's constant and n is the number of **photons**, or discrete particles of light each having energy

$$\boxed{E = h\nu.} \tag{2.81}$$

Then

$$P(E) = P(nh\nu) \propto \exp\left(-\frac{nh\nu}{kT}\right) \tag{2.82}$$

and the average energy per mode is calculated by summing over only the discrete energies permitted instead of integrating over all energies. Planck's quantized replacement for Equation 2.74 is

$$\langle E \rangle = \frac{\displaystyle\sum_{n=0}^{\infty} nh\nu P(nh\nu)}{\displaystyle\sum_{n=0}^{\infty} P(nh\nu)} = \frac{\displaystyle\sum_{n=0}^{\infty} nh\nu \exp\left(-\frac{nh\nu}{kT}\right)}{\displaystyle\sum_{n=0}^{\infty} \exp\left(-\frac{nh\nu}{kT}\right)}. \tag{2.83}$$

Planck's sum is evaluated in Appendix B.1; it is

$$\langle E \rangle = \frac{h\nu}{\exp\left(\frac{h\nu}{kT}\right) - 1} = kT \left[\frac{\frac{h\nu}{kT}}{\exp\left(\frac{h\nu}{kT}\right) - 1} \right]. \tag{2.84}$$

The quantity kT is the classical mean energy per mode, and the quantity in square brackets is the quantum correction factor. In the limit $h\nu \ll kT$ the quantum correction factor is unity, but in the limit $h\nu \gg kT$ it falls exponentially to zero. That is, the high-frequency or "ultraviolet" modes may contain few or even zero photons, thereby avoiding the "ultraviolet catastrophe."

Thus the correct blackbody radiation law becomes

$$B_\nu = \frac{2kT\nu^2}{c^2} \left[\frac{\frac{h\nu}{kT}}{\exp\left(\frac{h\nu}{kT}\right) - 1} \right], \tag{2.85}$$

where the first factor is the Rayleigh–Jeans approximation and the quantity in square brackets is the quantum correction factor. **Planck's equation** for the spectral brightness B_ν of blackbody radiation is usually written in the simpler form

$$B_\nu(\nu, T) = \frac{2h\nu^3}{c^2} \frac{1}{\exp\left(\frac{h\nu}{kT}\right) - 1}. \qquad (2.86)$$

The corresponding brightness per unit wavelength B_λ follows from Equation 2.5; it can be written either as a function of frequency:

$$B_\lambda(\nu, T) = \frac{\nu^2}{c} B_\nu(\nu, T) = \frac{2h\nu^5}{c^3} \frac{1}{\exp\left(\frac{h\nu}{kT}\right) - 1} \qquad (2.87)$$

or as a function of wavelength:

$$B_\lambda(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{\exp\left(\frac{hc}{\lambda kT}\right) - 1}. \qquad (2.88)$$

Integrating Planck's law over all frequencies (see Appendix B.2) gives the **Stefan–Boltzmann law** specifying a finite **integrated brightness** for a blackbody radiator at temperature T:

$$B(T) \equiv \int_0^\infty B_\nu(T)\, d\nu = \frac{\sigma T^4}{\pi}. \qquad (2.89)$$

The quantity

$$\sigma \equiv \frac{2\pi^5 k^4}{15 c^2 h^3} \approx 5.67 \times 10^{-5} \frac{\text{erg}}{\text{cm}^2 \ \text{s} \ \text{K}^4 \ (\text{sr})} \qquad (2.90)$$

is called the **Stefan–Boltzmann constant** and is also derived in Appendix B.2. The unit sr is dimensionless because solid angle has dimensions angle^2 = (length/length)^2. A dimensionless unit can be inserted or deleted without changing the dimensions of an equation—hence the parentheses around sr in Equation 2.90. For clarity it should be used where appropriate. For example, it is needed in the blackbody brightness (erg cm^{-2} s^{-1} K^{-4} sr^{-1}) Equation 2.89 but not in the radiation energy density (erg cm^{-3}) Equation 2.93 or in the flux (erg cm^{-2} s^{-1} K^{-4}) Equation 2.111.

For isotropic radiation, the spectral energy density (Equation 2.76) is simply

$$u_\nu = \frac{4\pi I_\nu}{c}, \qquad (2.91)$$

so the total **radiation energy density**

$$u \equiv \int_0^\infty u_\nu\, d\nu = \frac{4\pi I}{c} \qquad (2.92)$$

of blackbody radiation is

$$u = \frac{4\sigma T^4}{c} = aT^4, \tag{2.93}$$

where the quantity

$$a \equiv 4\sigma/c \approx 7.56577 \times 10^{-15} \ \text{erg} \ \text{cm}^{-3} \ \text{K}^{-4} \tag{2.94}$$

is called the **radiation constant**.

   In any narrow frequency range, the number density of photons is their spectral energy density divided by the energy per photon, $E = h\nu$. The total **photon number density** n_γ is

$$n_\gamma = \int_0^\infty \frac{u_\nu}{h\nu} d\nu. \tag{2.95}$$

The **photon number density of blackbody radiation** at temperature T is

$$n_\gamma = \frac{4\pi}{ch} \int_0^\infty \frac{B_\nu(T)}{\nu} d\nu \tag{2.96}$$

$$= \frac{8\pi}{c^3} \int_0^\infty \frac{\nu^2 d\nu}{\exp\left(\frac{h\nu}{kT}\right) - 1} \tag{2.97}$$

$$= \frac{8\pi}{c^3} \left(\frac{kT}{h}\right)^3 \int_0^\infty \frac{x^2 dx}{e^x - 1}, \tag{2.98}$$

where the integral

$$\int_0^\infty \frac{x^2 dx}{e^x - 1} \approx 2.404 \tag{2.99}$$

is evaluated in Appendix B.2. Numerically, the photon number density of blackbody radiation is

$$\left(\frac{n_\gamma}{\text{cm}^{-3}}\right) \approx 20.3 \left(\frac{T}{\text{K}}\right)^3. \tag{2.100}$$

   The **mean photon energy of blackbody radiation** is

$$\langle E_\gamma \rangle = \frac{u}{n_\gamma} = \frac{4\sigma T^4}{cn_\gamma} = \frac{\pi^4 kT}{15} \left(\int_0^\infty \frac{x^2 dx}{e^x - 1}\right)^{-1} \approx 2.70 \, kT. \tag{2.101}$$

The frequency $\langle \nu \rangle$ corresponding to this mean photon energy is

$$\left(\frac{\langle \nu \rangle}{\text{GHz}} \right) = \frac{\langle E_\gamma \rangle}{h} \approx 56 \left(\frac{T}{\text{K}} \right). \tag{2.102}$$

It is close to the frequency ν_{max} at which B_ν, the brightness *per unit frequency* of a blackbody, is maximum. Setting

$$\frac{\partial B_\nu (\nu)}{\partial \nu} = 0 \tag{2.103}$$

yields

$$\boxed{\left(\frac{\nu_{\text{max}}}{\text{GHz}} \right) \approx 59 \left(\frac{T}{\text{K}} \right).} \tag{2.104}$$

The wavelength λ_{max} at which

$$\frac{\partial B_\lambda (\lambda)}{\partial \lambda} = 0 \tag{2.105}$$

maximizes B_λ, the brightness *per unit wavelength* of a blackbody. It is

$$\boxed{\left(\frac{\lambda_{\text{max}}}{\text{cm}} \right) \approx 0.29 \left(\frac{T}{\text{K}} \right)^{-1}.} \tag{2.106}$$

Equation 2.106 is the familiar form of **Wien's displacement law** used by optical astronomers, whose spectrometers measure wavelengths instead of frequencies. Note that the peak frequency $c/\lambda_{\text{max}} \approx 103 \ \text{GHz} \ T (\text{K})$ is much higher than $\nu_{\text{max}} \approx 59 \ \text{GHz} \ T (\text{K})$.

The power emitted per unit area per unit frequency by any opaque isotropic radiator of brightness I_ν is the spectral flux density

$$F_\nu = \int I_\nu \cos \theta \, d\Omega, \tag{2.107}$$

where the integration covers the hemisphere above the unit area:

$$F_\nu = \int_0^{\pi/2} I_\nu \cos \theta \left(\int_0^{2\pi} \sin \theta \, d\phi \right) d\theta = 2\pi I_\nu \int_0^{\pi/2} \cos \theta \sin \theta \, d\theta, \tag{2.108}$$

$$\boxed{F_\nu = \pi I_\nu.} \tag{2.109}$$

For a blackbody at temperature T, the spectral flux density is

$$F_\nu(T) = \pi B_\nu(T).\tag{2.110}$$

The corresponding total power per unit area integrated over all frequencies is

$$\boxed{F(T) = \pi B(T) = \sigma T^4.}\tag{2.111}$$

2.5 Noise Generated by a Warm Resistor

A **resistor** is a passive electronic component that absorbs the electrical power applied to it and converts that power into heat; it is the "blackbody" of electronic circuits. Just as motions of charged particles in the walls of a warm cavity generate photons, motions of charged particles in a resistor at any temperature $T > 0$ K generate electrical noise. The frequency spectrum of this noise depends only on the temperature of an ideal resistor and is independent of the material in the resistor.

An **antenna** is a passive device that converts electromagnetic radiation into electrical currents in wires (when it is used as a receiving antenna) or vice versa (when it is used as a transmitting antenna). The noise generated by a resistor is indistinguishable from the noise coming from a receiving antenna surrounded by blackbody radiation of the same temperature. Warm resistors are useful in radio astronomy as standards for calibrating antennas and receivers, so the noise power per unit bandwidth received by a radio telescope is often described in terms of the Rayleigh–Jeans **antenna temperature**. The antenna temperature of a receiving antenna is defined as the temperature of an ideal resistor that would generate the same Rayleigh–Jeans noise power per unit bandwidth as appears at the antenna output. Like brightness temperature, antenna temperature is not a physical temperature. The sensitivity and gain of a radio receiver can be calibrated by connecting its input alternately to hot and cold resistors (sometimes called hot and cold "loads") having known temperatures, and the amount of noise generated in a **radiometer** (a radio receiver that measures noise power) can be described by the **radiometer noise temperature**, the temperature of a resistor at the input of an imaginary noiseless radiometer having the same gain as the actual receiver that would generate the same noise power output. Like brightness temperature, radiometer noise temperature is *not* a physical temperature.

The derivation of the electrical power per unit bandwidth P_ν generated by current in a resistor is the one-dimensional analog of the the three-dimensional derivation of the blackbody spectrum [75, 9]. At low radio frequencies, $h\nu \ll kT$ and the Rayleigh–Jeans approximation is accurate. Recall that the Rayleigh–Jeans derivation of B_ν starts with a large cube of side length $a \gg \lambda$ containing standing waves of thermal radiation. The classical average energy in each standing-wave mode is $\langle E \rangle = kT$, and the number of modes with frequency ν to $\nu + d\nu$ is proportional to ν^2, so $B_\nu \propto \nu^2$.

Consider two identical resistors at temperature T connected by a lossless transmission line (e.g., a pair of parallel wires) of length a much larger than the longest wavelength of interest (Figure 2.20). Standing waves on the line must satisfy

$$a = \frac{n\lambda}{2}, \qquad n = 1, 2, 3, \ldots,\tag{2.112}$$

where λ is the wavelength. Electrical signals do not travel at exactly the speed of light on a transmission line, but at some slightly lower velocity $v < c$ so $\nu = v/\lambda$ and

$$n = \frac{2a\nu}{v}.\tag{2.113}$$

For $a > \lambda$, the number of modes per unit frequency is

$$N_\nu = \frac{2a}{v}. \tag{2.114}$$

The classical Boltzmann law says that each mode has average energy $\langle E \rangle = kT$ in equilibrium, so the average energy per unit frequency E_ν in the transmission line is

$$E_\nu = N_\nu kT = \frac{2akT}{v}. \tag{2.115}$$

This energy takes a time $\Delta t = a/v$ to flow from one end of the transmission line to the other, so the classical power (energy per unit time) per unit frequency flowing on the transmission line is

$$P_\nu = \frac{E_\nu}{\Delta t} = 2kT, \tag{2.116}$$

independent of v. Thus the total spectral power P_ν generated by the two identical resistors must be $2kT$ and, by symmetry, the spectral power generated by each resistor is

$$\boxed{P_\nu = kT} \tag{2.117}$$

in the limit $h\nu \ll kT$. This equation is called the **Nyquist approximation** and is the electrical equivalent of the Rayleigh–Jeans equation for radiation. Because the "space" of the transmission line has only one dimension instead of three, $P_\nu \propto \nu^0$ instead of $P_\nu \propto \nu^2$. The dimensions of P_ν are power per unit frequency (e.g., W Hz^{-1}) and the dimensions of kT are energy (e.g., joules), which appear to be different at first glance. However, Hz^{-1} = s, so power per unit frequency W Hz^{-1} = W s = joule also has dimensions of energy. The latter is simpler, but the former is used because it is conceptually more appropriate for noise power per unit bandwidth.
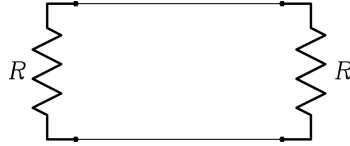
Figure 2.19: Two resistors connected by a lossless transmission line of length $a \gg \lambda$, the longest wavelength of interest. In equilibrium, the transmission line can support only those standing waves having zero voltages at the ends; other modes are suppressed by the lossy resistors.
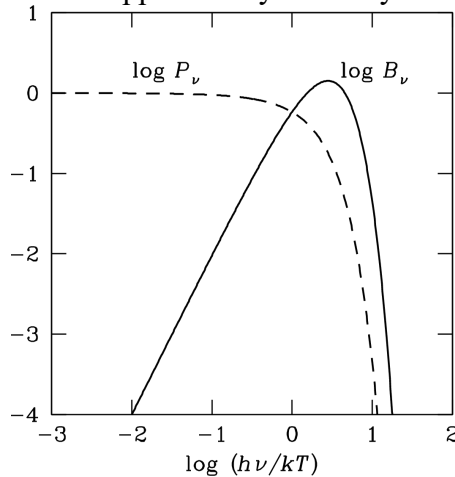
Figure 2.20: Comparison of the spectrum B_ν of blackbody radiation with the spectrum P_ν of noise generated by a resistor.

Like the Rayleigh–Jeans approximation, the Nyquist approximation implies an "ultraviolet catastrophe" that can be cured by Planck's quantization rule. If the electrical energy in each mode is restricted to an

integer multiple of $h\nu$, the Nyquist formula becomes

$$P_\nu = kT \left[\frac{\frac{h\nu}{kT}}{\exp\left(\frac{h\nu}{kT}\right) - 1} \right],$$

(2.118)

where the quantity in square brackets equals the quantization correction for blackbody radiation (Equation 2.85). The exact **Nyquist formula** is usually written in the form

$$\boxed{P_\nu = \frac{h\nu}{\exp\left(\frac{h\nu}{kT}\right) - 1}.}$$

(2.119)

Figure 2.20 compares the Nyquist and Planck spectra. At low frequencies $\nu \ll kT/h$, the specific intensity B_ν of blackbody radiation in three dimensions (solid curve) is proportional to ν^2. Its one-dimensional analog, the spectral power density of noise generated by a resistor (dashed curve), is proportional to ν^0. Quantization causes the sharp exponential cutoffs of both curves at high frequencies.

2.6 Cosmic Microwave Background Radiation

The universe is filled with blackbody radiation whose temperature now is $T_0 = 2.725 \pm 0.002$ K. This **cosmic microwave background** (CMB) was produced by the hot "big bang" marking the birth of the universe. This section reviews the expanding universe, derives the evolution of blackbody radiation during the expansion, and indicates how the CMB carries cosmological information.

2.6.1 The Expanding Universe

The expanding universe is nearly homogeneous and isotropic on all scales much larger than clusters of galaxies. Consequently the recession velocity v of a nearby galaxy at rest in its local environment is proportional to its (independently measured) distance d. The constant of proportionality is called the **Hubble parameter**

$$H \equiv \lim_{d \to 0} \left(\frac{v}{d} \right).$$

(2.120)

$H(t)$ is everywhere the same at any particular time t after the big bang, but it changes with time. The present value of the Hubble parameter is called the **Hubble constant** H_0, and its measured value is [83]

$$H_0 \approx 67.8 \pm 0.9 \ \text{km} \ \text{s}^{-1} \ \text{Mpc}^{-1}.$$

(2.121)

Substituting 3.0856×10^{19} km ≈ 1 Mpc shows that the dimension of H is inverse time and the Hubble constant can be written as $H_0 \approx (4.55 \times 10^{17} \ \text{s})^{-1}$. The **Hubble time** defined by

$$t_{\text{H}} \equiv H_0^{-1} \approx 4.55 \times 10^{17} \ \text{s} \, / \, 10^{7.5} \ \text{s} \ \text{yr}^{-1} \approx 1.44 \times 10^{10} \ \text{yr}$$

(2.122)

would be the present age of the universe had it started as a point and expanded at a constant rate thereafter. (The useful approximation $10^{7.5}$ s ≈ 1 yr used in Equation 2.122 is easy to remember and accurate to 0.2%.)

The radius of the observable universe is roughly the **Hubble distance** at which the Hubble parameter indicates a recession velocity equal to the speed of light:

$$d_{\mathrm{H}} \equiv \frac{c}{H}.$$

(2.123)

It is currently $d_{\mathrm{H}} = c/H_0 \approx 1.36 \times 10^{28}$ cm $\approx 4.4 \times 10^9$ pc $\approx 1.4 \times 10^{10}$ light years.

Although general relativity is needed to describe the universe as a whole, the homogeneity of the universe means that many of its global properties can be derived by extrapolating Newtonian results obtained from regions of size $d \ll d_{\mathrm{H}}$. For example, the **critical density** ρ_{c} is defined as the density at which the kinetic energy of expansion equals the gravitational energy decelerating the expansion. Consider a spherical region of radius d centered on an observer and a unit test mass on its surface, which is receding with radial velocity $v = Hd$. If the average density inside the sphere is ρ_{c}, the sum of the kinetic and gravitational potential energies per unit mass will be zero:

$$\frac{v^2}{2} - \frac{GM}{d} = 0,$$

(2.124)

$$\frac{H^2 d^2}{2} = \frac{G}{d}\left(\frac{4\pi d^3 \rho_{\mathrm{c}}}{3}\right),$$

(2.125)

so the critical density today is

$$\boxed{\rho_{\mathrm{c}} = \frac{3H_0^2}{8\pi G} \approx 8.6 \times 10^{-30} \text{ g cm}^{-3}}$$

(2.126)

for $H_0 = 67.8$ km s^{-1} Mpc^{-1}. Ordinary **baryonic matter** consists by mass primarily of protons and neutrons, whose masses are $m \approx 1.67 \times 10^{-24}$ g, so a mean number density $n_{\mathrm{b}} \approx 5.2 \times 10^{-6}$ cm^{-3} of protons and neutrons would be needed to supply the critical density. However, in a matter-dominated universe, mutual gravitational attraction decelerates the expansion. The expansion age of a low-density ($\rho \ll \rho_{\mathrm{c}}$) universe is nearly t_{H}, but a critical-density matter-dominated universe should be only $(2/3)\,H_0^{-1} \approx 9.6$ Gyr old, which is less than the ages of the oldest known stars. **Dark energy** actually dominates the mass-energy of the universe today, and it has the peculiar property that it *accelerates* the expansion enough that the actual age of the universe is closer to t_{H}, eliminating the problem of stars older than the universe.

2.6.2 Blackbody Radiation in the Expanding Universe

How does the blackbody CMB evolve as the universe expands with time t since the big bang? The universe is spatially homogeneous and isotropic on large scales, so the properties of the CMB can be calculated by considering what happens to blackbody radiation in a small imaginary cube whose side length slowly grows along with the universe. Here "small" only means much smaller than d_{H} so that relativistic effects can be ignored; the side length might be the distance between two galaxies that are currently separated by ~ 100 Mpc. Although the imaginary cube has no walls, CMB radiation escaping from the cube is balanced by CMB radiation entering the cube from similar adjacent cubes. Consequently the energy density and spectrum of CMB radiation within the cube is exactly the same as that in a cubical cavity having mirrored walls and undergoing a slow adiabatic (meaning, there is no heat transfer to or from the cube)

expansion. The radiation remains in equilibrium during the slow expansion, so it is always blackbody radiation.

The expansion of both the universe and the imaginary cube can be described by a single dimensionless **expansion scale factor** $a(t)$, where $a(t_0) \equiv 1$ today. The photon number density n_γ in the cube falls as a^{-3} because the total number of photons in the expanding cube doesn't change as a grows. The photon number density of blackbody radiation is proportional to T^3 (Equation 2.100), so the temperature of the CMB declines as $T = a^{-1}T_0$, where $T_0 \approx 2.725$ K. Also, the mean energy per photon of blackbody radiation is proportional to T. Thus the wavelength λ of any individual photon is proportional to a, even if the photon is from a source which does not have a blackbody spectrum.

Astronomers frequently use the term **redshift** defined by

$$z \equiv \frac{\lambda_{\mathrm{o}} - \lambda_{\mathrm{e}}}{\lambda_{\mathrm{e}}} = \frac{\lambda_{\mathrm{o}}}{\lambda_{\mathrm{e}}} - 1 = \frac{\nu_{\mathrm{e}}}{\nu_{\mathrm{o}}} - 1, \qquad (2.127)$$

where λ_{e} and ν_{e} are the wavelength and frequency emitted by a source at redshift z, and λ_{o} and ν_{o} are the observed wavelength and frequency at $z = 0$. The redshift z and expansion factor a are related by

$$(1 + z) = a^{-1}. \qquad (2.128)$$

Thus the CMB temperature at redshift z is

$$T = T_0 (1 + z). \qquad (2.129)$$

The CMB radiation is ubiquitous, so interstellar matter (e.g., interstellar dust or gas) is heated by the CMB to at least the temperature given by Equation 2.129. The interstellar dust temperature in nearby galaxies like our own is ~ 20 K, which is well above $T_0 \sim 3$ K but comparable with the CMB temperature felt by galaxies at $z \sim 6$. The CMB will influence the spectral line and continuum emission visible from high-redshift galaxies.

Prior to the time corresponding to the redshift $z_\star = 1091 \pm 1$, the CMB temperature was $T > 3000$ K and the radiation ionized enough of the hydrogen atoms filling the universe to keep the universe opaque. This "wall" beyond which photons cannot escape is called the **surface of last scattering**, and it defines the limit of the visible universe. During the **recombination era** when the age of the universe was only $t_\star = (379 \pm 5) \times 10^3$ yr, almost all of the free protons and electrons combined to form neutral hydrogen atoms and the universe became transparent to the CMB. (The term *re*combination is misleading because the protons and electrons were combining for the first time.) The CMB photons received today were last scattered at that time and have been traveling in straight lines ever since, so an image of the CMB (Figure 8.16) made now ($z = 0$) shows its temperature distribution as it was when $z_\star \approx 1091$.

2.6.3 Prediction and Discovery of the CMB

Well before the CMB was observed, George Gamow and his students calculated the relative hydrogen and helium abundances (Figure 2.21) produced by nucleosynthesis a few minutes after a hot big bang, when the temperature of the CMB was $T > 10^9$ K [2]. The abundance of deuterium in particular depends sensitively on the relative number densities n_{b} of baryons (primarily protons plus neutrons by mass) and of photons n_γ. The ratio $\eta \equiv n_{\mathrm{b}}/n_\gamma \approx 6 \times 10^{-10}$ has remained constant since that time because both n_{b} and and n_γ are proportional to a^{-3}. Equation 2.100 gives the photon number density of the $T_0 \approx 2.725$ K CMB; it is $n_\gamma \approx 411$ cm^{-3}. If $\eta \approx 6 \times 10^{-10}$, then the baryon number density is $n_{\mathrm{b}} \approx 2.5 \times 10^{-7}$ cm^{-3}, which is only about 5% of the density needed to close the universe (Equation 2.126). Alpher and Herman [3] presented the first estimate $T_0 \approx 5$ K for the present temperature of the CMB based on the chemical

composition of the universe. They also showed that only light elements can be formed by big-bang nucleosynthesis because the free neutrons needed to assemble heavier elements decay with a half life of about 10 minutes. Stars produce the heavier elements.
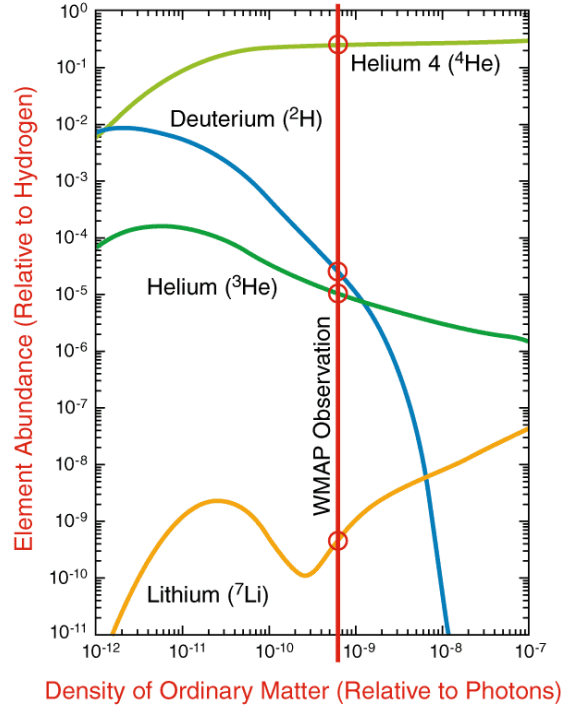

Figure 2.21: The energy density of the early universe was dominated by hot blackbody radiation. By $t \approx 1$ s, the temperature fell to $T \approx 10^{10}$ K, and neutrons decayed into protons and electrons or combined with protons to make deuterium nuclei, most of which combined to make helium nuclei during the first three minutes. The chemical composition of the early universe is a function of the baryon/photon ratio $\eta \approx 6 \times 10^{-10}$ by number. "Ordinary" baryonic matter now accounts for about 4.6% of the energy density of the universe. Image credit: NASA/WMAP Science Team.

Although 5 K is a very strong signal by radio-astronomy standards, the CMB is hard to distinguish from other noise sources in receivers and radio telescopes because it is very nearly isotropic. Unlike emission from a compact radio source, the CMB cannot be isolated by a differential measurement comparing two adjacent patches of sky. It was accidentally discovered in 1964 by Arno Penzias and Robert Wilson [81], who reported painstaking measurements indicating an isotropic "excess antenna temperature" coming from the Bell Labs horn antenna at 4080 MHz. Subsequent measurements, especially by the COBE[1] (COsmic Background Explorer) satellite, have demonstrated that this "excess noise" matches a blackbody spectrum within 50 parts per million well into the far infrared.

2.6.4 Dipole Anisotropy

After contamination by foreground sources near the Galactic plane has been removed, the largest observed anisotropy in the CMB is the **dipole anisotropy** with amplitude

$$\Delta T \approx 3.37 \cos \theta \ \text{mK}, \tag{2.130}$$

where θ is the angle from the direction (Galactic latitude $b = +48°$, Galactic longitude $l = 264°$) of the Earth's motion relative to the universe as a whole. The speed $v_\oplus$ of the Earth is much smaller than the speed of light, so the nonrelativistic Doppler equation can be used to estimate the redshift induced by the Earth's motion:

$$\frac{v_\oplus}{c} \approx \frac{\Delta \lambda}{\lambda} \approx z \approx \frac{\Delta T}{T} \approx \frac{3.37 \times 10^{-3} \ \text{K}}{2.725 \ \text{K}} \approx 1.24 \times 10^{-3}, \tag{2.131}$$

or $v_\oplus \approx 370 \;\mathrm{km}\;\mathrm{s}^{-1}$. This speed is consistent with the vector sum of independently measured motions of the Earth around the Sun, the Sun around the Galactic center, and the peculiar motion of our Galaxy in the local universe.

2.6.5 Intrinsic Anisotropy

The age $t_\star$ of the universe at recombination can be calculated because the physics of the homogeneous and isotropic early universe was so simple. During that time, small inhomogeneities in the density of dark matter were amplified by gravity. Baryons started falling into the gravitational potential wells, and the CMB was coupled to this flow by Thomson scattering (Equation 5.33) off the numerous free electrons present prior to $t_\star$. Radiation pressure resisted this infall, setting up **baryon acoustic oscillations** analogous to sound vibrations in air, but with a "sound" speed $c_\mathrm{s} \approx c/\sqrt{3}$. The strongest oscillations in the CMB brightness had wavelengths equal to the diameter of the sound horizon at the time of recombination:

$$\lambda_\star \approx 2c_\mathrm{s}t_\star \approx \frac{2ct_\star}{\sqrt{3}} \approx \frac{2 \cdot 379 \times 10^3}{\sqrt{3}} \;\mathrm{ly} \approx 438 \times 10^3 \;\mathrm{ly}, \tag{2.132}$$

which was about $134 \;\mathrm{kpc}$. After recombination, the photons decoupled and this "standard ruler" grew by a factor $(1 + z_\star) \approx 1092$ to its present size of $146 \;\mathrm{Mpc}$. Figure 8.16 shows these CMB brightness fluctuations, projected onto a sphere [106]. The angular power spectrum of CMB fluctuations (Figure 1.16) observed by the **Wilkinson Microwave Anisotropy Probe (WMAP)** and by the **Planck Mission** reveal that the angular size of this ruler is $\theta_\star = 0.010388 \pm 0.000027 \;\mathrm{rad}$, so the comoving distance to the surface of last scattering is $D_\mathrm{A} = (1 + z_\star)\,\lambda_\star/\theta_\star \approx 14.1 \pm 0.16 \;\mathrm{Gpc}$. The **radius of the observable universe** (including gravitational radiation and neutrinos from redshifts $z \gg z_\star$) is only about 2% larger. If we live in a flat ΛCDM universe (Λ Cold Dark Matter universe, where Λ stands for the cosmological constant associated with dark energy), its present age is $t = 13.75 \pm 0.13 \;\mathrm{Gyr}$.

The intrinsic rms of the CMB over the whole sky is $\sigma \approx 18\,\mu\mathrm{K}$, which is only about $\sim 10^{-5}$ of the average $T \approx 2.725\;\mathrm{K}$. This leads to the **horizon problem**: Why should the CMB temperature be the same at points separated by more than $\theta_\star \approx 0.01 \;\mathrm{rad}$, which were never in causal contact by the time of recombination? The simple big-bang model also has a **flatness problem**. The universe is nearly "flat" today so its density is close to the critical density, and it must have been *much* closer to the critical density when it was much smaller (e.g., during the nucleosynthesis era when $z \sim 10^9$). What "fine tuned" the density of the universe?

The preferred solution for both problems is that the early universe underwent a brief period ($\sim 10^{-32}\;\mathrm{s}$) of **inflation** during which the scale factor a grew exponentially by a factor $> 10^{26}$. The observable universe is only an infinitesimal and causal part of the whole universe, nearly all of which is outside our horizon. The **vacuum energy** associated with inflation ensures that the total density contributed by both matter and energy approaches the critical density.

2.7 Radiation from an Accelerated Charge

Maxwell's equations imply that *all classical electromagnetic radiation is generated by accelerating electrical charges*. It is possible to derive the intensity and angular distribution of the radiation from a point charge (a charged particle) subject to an arbitrary but small acceleration $\Delta v/\Delta t$ via Maxwell's equations, but the complicated math obscures the physical interpretation that remains clear in J. J. Thomson's illuminating derivation (see [69]).

If a particle with electrical charge q is at rest or moving with a constant velocity, its electric field lines are purely radial: $|\vec{E}| = E_r$. Suppose a charged particle initially at rest is accelerated to a small velocity $\Delta v \ll c$ in a short time Δt. This disturbs the lines of force, and the disturbance travels outward at the speed of light c. Figure 2.22 shows that at time t after the acceleration, the disturbance will have propagated to $r = ct$ and the perpendicular component of the electric field will have magnitude

$$\frac{E_\perp}{E_r} = \frac{\Delta v\, t\, \sin\theta}{c\Delta t}, \tag{2.133}$$

where θ is the angle between the acceleration vector and the line of sight connecting the charge to the observer. The web applet `https://phet.colorado.edu/en/simulation/radiating-charge` provides a very clear and interactive demonstration of this effect.
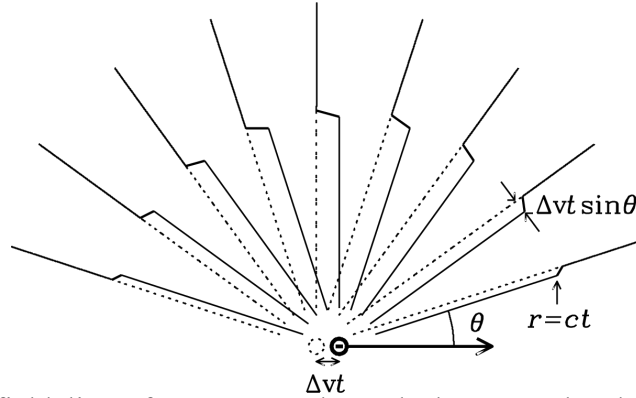

Figure 2.22: The electric field lines from an accelerated electron. The dotted circle shows the initial position of the electron, and the dotted lines are the radial lines of force emanating from that position. At time t after a small acceleration $\Delta v/\Delta t$, the electron position has moved by $\Delta v\, t$ and its lines of force have shifted transversely by $\Delta v\, t \sin\theta$.

Coulomb's law for the radial component E_r of the **electric field** (electric force per unit charge) a distance r from a stationary charge q is

$$\boxed{E_r = \frac{q}{r^2}} \tag{2.134}$$

in Gaussian CGS units. Substituting r/c for t in Equation 2.133 and using Coulomb's law to eliminate E_r gives

$$E_\perp = \frac{q}{r^2}\left(\frac{\Delta v}{\Delta t}\right)\frac{r\sin\theta}{c^2} \tag{2.135}$$

and

$$\boxed{E_\perp = \frac{q\dot{v}\sin\theta}{rc^2}}, \tag{2.136}$$

where

$$\dot{v} \equiv \lim_{\Delta t \to 0}\left(\frac{\Delta v}{\Delta t}\right). \tag{2.137}$$

Equation 2.136 is valid for *any* small acceleration, not just a sinusoidal acceleration at a single frequency. The transverse field $E_\perp$ at $r = ct$ mimics the acceleration at $t = 0$. Thus a sinusoidal acceleration would result in a sinusoidal variation of $E_\perp$ with the same frequency.

Notice that $E_\perp \propto r^{-1}$ (Equation 2.136) falls more slowly with distance r than $E_r \propto r^{-2}$ (Equation 2.134). Far from the charged particle, only $E_\perp$ will contribute significantly to the observed electric field. From a distant observer's point of view, only the "visible" component of acceleration perpendicular to the line of sight ($\dot{v}\sin\theta$) contributes to the radiated electric field; the "invisible" component of acceleration parallel to the line of sight ($\dot{v}\cos\theta$) does not appear to radiate—what you see is what you get. Likewise, the radiated electric field is linearly polarized in the direction parallel to the component of acceleration perpendicular to the line of sight.

How much power is radiated in each direction? The **Poynting flux**, or power per unit area (e.g., erg s^{-1} cm^{-2}), is

$$\vec{S} = \frac{c}{4\pi}\vec{E}\times\vec{B}. \tag{2.138}$$

In CGS units $|\vec{E}| = |\vec{B}|$ so

$$\boxed{|\vec{S}| = \frac{c}{4\pi}E^2.} \tag{2.139}$$

Inserting Equation 2.136 into Equation 2.139 gives

$$|\vec{S}| = \frac{c}{4\pi}\left(\frac{q\dot{v}\sin\theta}{rc^2}\right)^2 = \left(\frac{q^2\dot{v}^2}{4\pi c^3}\right)\frac{\sin^2\theta}{r^2} \tag{2.140}$$

at large distances r. The accelerated charge radiates with a dipolar power pattern $\propto \sin^2\theta$ shaped like a doughnut whose axis is parallel to the acceleration $\dot{v}$ (Figure 2.23).
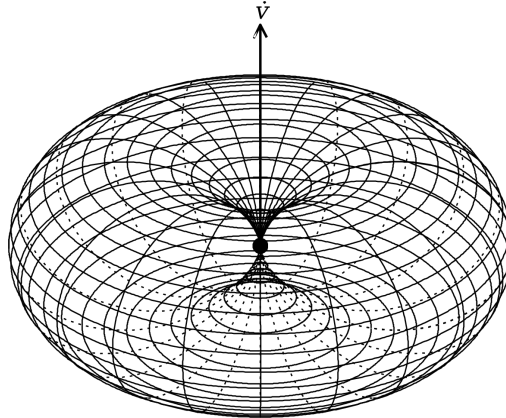

Figure 2.23: The power pattern of Larmor radiation from a charged particle shown for an acceleration vector $\dot{v}$ tilted 60 degrees from the line of sight. The power received in any direction is proportional to the component of $\dot{v}$ perpendicular to the line of sight.

The total power emitted is the integral of $|\vec{S}|$ over the spherical surface of radius r:

$$P = \int_{\text{sphere}}|\vec{S}|\,dA = \frac{q^2\dot{v}^2}{4\pi c^3}\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\frac{\sin^2\theta}{r^2}r\sin\theta\,d\theta\,r\,d\phi \tag{2.141}$$

$$= \frac{q^2\dot{v}^2}{2c^3}\int_{\theta=0}^{\pi}\sin^3\theta\,d\theta. \tag{2.142}$$

The integral $\int_{\theta=0}^{\pi}\sin^3\theta\,d\theta = 4/3$, so the total power emitted by the accelerated charged particle is

$$P = \frac{2}{3} \frac{q^2 \dot{v}^2}{c^3}.$$

(2.143)

This result is called **Larmor's equation**. It implies that any charged particle radiates when accelerated and that the total radiated power is proportional to the square of the acceleration. The largest astrophysical accelerations are usually produced by electromagnetic forces, so the acceleration is proportional to the charge/mass ratio of the particle. In celestial sources, radiation from electrons is typically ($m_\mathrm{p}/m_\mathrm{e})^2 \approx 4 \times 10^6$ times stronger than radiation from protons.

Larmor's extremely useful equation will be the basis for our derivations of radiation from a short dipole antenna as well as for free–free and synchrotron emission from astrophysical sources. Beware that *Larmor's formula is nonrelativistic*; it is valid only in frames moving at velocities $v \ll c$ with respect to the radiating particle. To treat particles moving at nearly the speed of light in the observer's frame, we must first use Larmor's equation to calculate the radiation in the particle's rest frame and then transform the result to the observer's frame in a relativistically correct way. Also, *Larmor's formula does not incorporate the constraints of quantum mechanics*, so it should be applied only with great caution to microscopic systems such as atoms. For example, Larmor's equation incorrectly predicts that the electron orbiting in the lowest energy level of a hydrogen atom will quickly radiate away all of its kinetic energy and fall into the nucleus. On the other hand, it correctly predicts the radio power emitted by an electron orbiting in a very high energy level of a hydrogen atom.

2.8 Dust Emission at Radio Wavelengths

All small solid particles in space are called **dust grains** by astronomers. **Interstellar dust** was first recognized because it scatters or absorbs ultraviolet, visible, and near-infrared photons from stars. **Extinction** is defined as the total dimming of the light coming directly from a point source caused by both photon scattering and photon absorption, and it is usually expressed in units of $\mathrm{mag} = 0.4 \, \mathrm{dex}$. The amount of extinction is inversely proportional to wavelength for $0.1 \, \mu\mathrm{m} \leq \lambda \leq 7 \, \mu\mathrm{m}$, indicating that the grains responsible have sizes $a \leq \lambda/(2\pi) \leq 1 \, \mu\mathrm{m}$. Interstellar dust grains are much smaller than common terrestrial dust particles visible to the naked eye, and they are much smaller than the wavelengths $\lambda > 300 \, \mu\mathrm{m}$ (frequencies $\nu < 1000$ GHz) accessible to ground-based radio astronomy (Figure 1.3).

Dust scattering occurs because the oscillating electric field of incident radiation forces electrons within dust grains to oscillate and hence reradiate at the same frequency in all directions, as shown in Section 2.7. In the radio limit $\lambda \gg a$ these radiators are very inefficient, so their scattering cross section is proportional to λ^{-4} (Rayleigh's law). Rayleigh's law applies to all small grains, regardless of composition. Scattering in the limit $a \ll \lambda$ is called **Rayleigh scattering**, and Rayleigh scattering of sunlight by air molecules is why the sky is so blue and the setting Sun is so red. Although photon scattering by dust *in front of* a star makes the star appear both dimmer and redder in color, scattering by dust *within* an extended object such as an external galaxy affects neither its total brightness nor its color. Only the absorption component of its internal dust extinction can redden or dim an external galaxy.

Some of the incident photons are absorbed, and their energy heats the dust grains. The energy carried by a single ultraviolet photon can significantly raise the temperature of a very small ($a \ll 1 \, \mu\mathrm{m}$) dust grain, so the smallest grains are not in thermodynamic equilibrium with the local interstellar radiation field. Larger interstellar dust grains have enough heat capacity to come into equilibrium at well-defined temperatures $20 < T_\mathrm{d} \, (\mathrm{K}) < 200$ such that the power absorbed is balanced by power reemitted primarily at far-infrared (FIR) wavelengths $\lambda \sim 100 \, \mu\mathrm{m}$. The nearly equal UV/optical and FIR peaks in the cosmic background (Figure 1.4) imply that about half of all UV/optical starlight escapes and half is absorbed within its parent galaxy. This result is an average over all galaxies, but individual galaxies can range from nearly transparent (e.g., elliptical galaxies with little interstellar dust) to nearly opaque (e.g., the compact ultraluminous starburst galaxy Arp 220).

The dust absorption cross section varies as λ^{-2} when $a \ll \lambda$, so at radio wavelengths (1) dust absorption ($\propto \lambda^{-2}$) contributes much more than dust scattering ($\propto \lambda^{-4}$) to the total extinction, and (2) all galaxies are nearly transparent. Thus radio (and far-infrared $\lambda > 30 \, \mu\mathrm{m}$) observations can see into dusty

starburst galaxies, and the radio/FIR luminosity is nearly proportional to the recent rate of star formation, unaffected by dust obscuration.
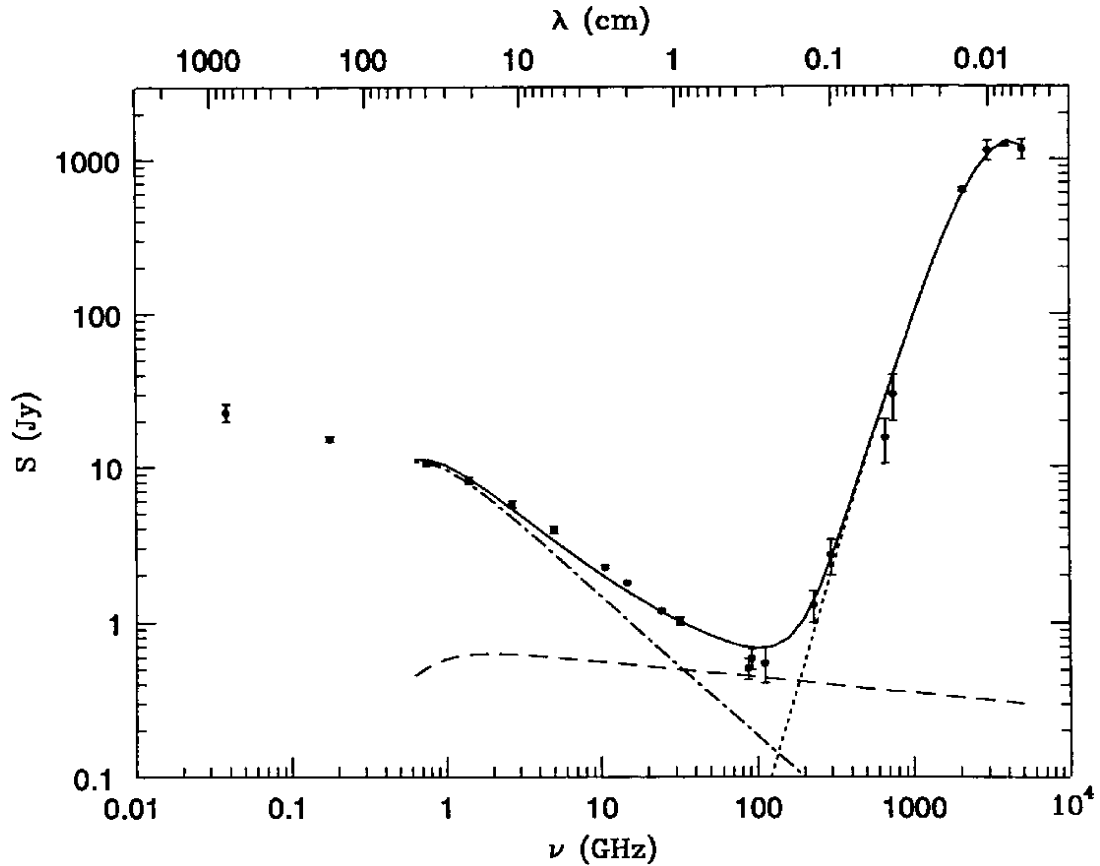

Figure 2.24: The radio and far-infrared spectrum of the nearby starburst galaxy M82 [30]. The contribution of free–free emission is indicated by the nearly horizontal dashed line. Synchrotron radiation (dot-dash line) and thermal dust emission (dots) dominate at low and high frequencies, respectively. Free–free absorption from Hɪɪ regions distributed throughout the galaxy absorbs some of the synchrotron radiation and flattens the overall spectrum at the lowest frequencies. Figure 8.13 shows a radio image of M82.

Obeying Kirchhoff's law for opaque bodies (Equation 2.47), small dust grains are also inefficient emitters. Their absorption and emission coefficients both scale with frequency as $\epsilon(\nu) \propto \nu^{\beta}$, where $1 < \beta < 2$ and $\beta \to 2$ in the Rayleigh limit. Thus dusty radio sources do not have Rayleigh–Jeans radio spectra $S_{\nu} \propto \nu^{2}$; they have relatively "blue" spectra rising as $S_{\nu} \propto \nu^{2+\beta}$. The dotted curve in Figure 2.24 showing dust emission from the galaxy M82 indicates $S_{\nu} \propto \nu^{3}$ to ν^{4} for $\nu < 1000$ GHz.

A striking consequence of such a steeply rising spectrum is that the flux density of a galaxy observed through the $\lambda \sim 1$ mm ($\nu \sim 300$ GHz) atmospheric window (Figure 1.3) is nearly independent of its redshift in the range $1 < z < 10$. The strongest submm galaxies tend to be the most luminous, and radio surveys at wavelengths near $\lambda = 1$ mm are effective at tracing the rate of star formation over the entire redshift range during which most stars formed [10].

Astronomical dust is also present in **protoplanetary disks**. These "dust" particles are much larger than interstellar dust grains, so they can be strong radio emitters with nearly blackbody spectra at $\lambda \sim 1$ mm. The spectacular ALMA (Figure 8.5) image of the dust disk surrounding the young star HL Tau (Figure 8.9) has sufficient angular resolution (≈ 0.035 arcsec) to show dark rings carved out by young planets [1].