This course covers high energy astrophysics, defined loosely as the astrophysics of objects that emit radiation in the X-ray and γ-ray parts of the electromagnetic spectrum. The delineation between high energy astrophysics and astrophysics at other wavelengths is largely driven by instrumentation: the detectors that one uses to measure radiation at X-ray and γ-ray energies are fundamentally different than the ones used at lower energies, mainly because these very energetic photons penetrate matter easily, and thus cannot be handled using the same sorts of optical devices used for IR through UV wavelengths, nor the sorts of receives used at radio wavelengths. However, the divisions between the fields are also driven partly by the types of sources to which they are sensitive, and the processes by which those sources produce their radiation. The processes that occur in the highly ionised plasma around a black hole or in a galaxy cluster are different in nature from those that occur in the mostly-neutral atmosphere of a star like the Sun.

The first half of this course will focus on this last point: how do astrophysical objects produce high energy radiation? To address this question, we will need to develop some general theory of radiation and its interaction with matter, focusing on the interaction processes that are most relevant at high energies. The goal in this lecture and the next is to begin building up a statistical theory for the propagation of radiation and its interaction with matter. We will then complement this statistical theory with a theory formulated starting from Maxwell's equations and electrodynamics that addresses the microphysics of how radiation is produced. Combining these two will give us the tools we need to begin studying astrophysical sources and objects.

I. Statistical description of radiation fields

A. The specific intensity

Let us consider radiation propagating through vacuum, or through a homogenous medium. Imagine placing a slab of near-zero temperature material with infinitesimal area dA in the radiation field, and leaving it there for an infinitesimal time dt. As the slab absorbs radiation it will heat up, and its thermal energy will increase from nearly zero at the start to some value dE. We can use the amount of energy absorbed to define a basic quantity the radiation flux. Specifically, we define

$$F = \frac{dE}{dA\,dt}. \tag{1}$$

The motivation for dividing out dA and dt is obvious: the longer we expose the slab the more it will heat up, and the larger the slab, the more radiation it will absorb. Dividing out dA and dt removes these factors, so that we are left with a quantity that is particular to the radiation field, rather than to the details of our experiment. The radiation flux has units of energy per time per area – in cgs units, this would be erg s^{-1} cm^{-2}.

While this is a useful description, we can imagine making more precise ones. For example, we have posited that our experimental slab absorbs all the radiation, but we know that radiation can be broken up into a spectrum of different frequencies. We can imagine that our experimental slab absorbs radiation only over a narrow range of frequencies $d\nu$, and lets all other radiation pass unaffected. In this case the amount of energy absorbed will be decreased, and it makes sense to divide out the frequency range $d\nu$ as well, since if we double the frequency range we will also double the amount of energy absorbed. This motivates us to define the specific flux

$$F_\nu = \frac{dE}{dA\,dt\,d\nu}. \tag{2}$$

Finally, we can imagine making a further refinement of our detector. As we have set up our experiment, our slab absorbs radiation coming from any direction. We can imagine, however, making our detector directional, so that it only absorbs radiation coming from some particular direction – after all, this is what real telescopes do, since they are set up so that only radiation arriving from the direction in which the telescope is pointed will be bounced to the focal point and thence onward to the detector. Let us suppose that our slab only accepts radiation arriving from some range of solid angles $d\Omega$ centred on some particular direction $\hat{\mathbf{n}}$. As before, it makes sense to divide out $d\Omega$, since if the detector absorbs from twice as large a range of solid angle, it will absorb twice as much radiation. We define the quantity derived in this way as the *specific intensity*,

$$I_\nu = \frac{dE}{dA\,dt\,d\nu\,d\Omega}. \tag{3}$$

This quantity describes how much energy the radiation field delivers per unit detector area, per unit time, per unit radiation frequency, per unit solid angle. It has dimensions of energy per time per area per frequency per solid angle – in cgs, erg s^{-1} cm^{-2} Hz^{-1} sr^{-1}.

The intensity is a function of both frequency and direction: that is, at any given point in space, the intensity of the radiation field can be different in different directions, and at different frequencies. As a trivial example of this, consider a point inside the beam of a laser pointer. The intensity will be quite low at any frequency other than that of the laser beam, and the intensity will be high along the direction toward the laser point, and low in all other directions.

B. The transfer equation in vacuum

A crucial fact about radiation in vacuum far from sources of electromagnetic fields is that it is composed of photons / waves that propagate in straight lines at a fixed speed c, and conserve their frequency and direction as they do so. We will give precise meaning to the statement "far enough" in a few lectures, where we will also demonstrate the remainder of these assertions for classical fields. For now, however, we will simply accept these statements as given.

The statements above imply that the intensity is that it is conserved along rays. To see this, consider any straight line, and pick any two points along that line separated by a distance s. We will denote these points as points 1 and 2. Now let dA_1 and dA_2 be two infinitesimal areas normal to this line at the two points. Furthermore, let $d\Omega_1$ be the angle subtended by dA_2 as viewed from point 1, and similarly $d\Omega_2$ is the angle subtended by dA_1 at point 2. Clearly we have

$$d\Omega_1 = \frac{dA_2}{s^2} \qquad d\Omega_2 = \frac{dA_1}{s^2}. \tag{4}$$

Now consider that any photon that is travelling in a straight line through point 1 in a direction within solid angle $d\Omega_1$ of the line must pass through dA_2, and similarly for photons passing through point 2. Thus it follows that the energy passing through the two points must be the same:

$$dE_1 = dE_2 = I_{\nu,1}\, dA_1\, dt\, d\Omega_1\, d\nu_1 = I_{\nu_2} dA_2\, dt\, d\Omega_2\, d\nu_2. \tag{5}$$

From this we can immediately deduce that $I_{\nu,1} = I_{\nu,2}$, i.e., the intensity is constant along rays. Mathematically, we can express this as

$$\frac{dI_\nu}{ds} = 0 \tag{6}$$

in steady state and in free space, i.e., with no matter to absorb or emit radiation. Here differentiation with respect to s means taking the derivative along the direction $\hat{\mathbf{n}}$ for which we are writing down the intensity – again, remember that the intensity is a function of direction, so there are different intensities in different directions.

If we relax the assumption of steady state, we can generalise this by noting that light travels at speed c, so the intensity passing through point 2 must be equal to the intensity that passed through point 1 a time $t = s/c$ ago. Again, we an express this as a differential equation by considering two points with infinitesimal separation ds. In this case we must have

$$\frac{\partial I_\nu}{\partial t} = -c\frac{\partial I_\nu}{\partial s}. \tag{7}$$

This just asserts that the rate at which the intensity at a given position changes must be equal to the rate of change of the intensity along a ray at fixed time, multiplied by the rate c at which the light moves. The minus sign in the equation is because ds along the direction the light is going, and we care about what is happening upstream rather than downstream of the point in question.

Intuitively, an easy way to understand Equation 7 is to imagine what would happen if the radiation intensity upstream of the current location were much higher than the radiation intensity at the location of interest. Clearly in this case the radiation intensity should rise. In this case we would have $dI_\nu/ds \ll 0$ (since the radiation intensity is falling as one moves along the ray), and thus $dI_\nu/dt \gg 0$. The sharper the gradient in intensity along the ray (the larger $|dI_\nu/ds|$ is), the more rapidly the intensity will rise (the larger dI_ν/dt will be).

An alternative way of writing this, which avoids introducing the length s along a specified ray, is

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \hat{\mathbf{n}} \cdot \nabla I_\nu = 0. \tag{8}$$

Here $\hat{\mathbf{n}}$ is a unit vector that specifies which direction of radiation we're interested in, and this equation applies equally and independently to each unit vectors.

C. Moments of the intensity

From the intensity we can now define a series of more familiar quantities. One the most basic is the energy density in the radiation field. Imagine taking an infinitesimal volume dV, and trying to estimate the amount of energy within it carried by photons of frequency ν. We do not care about which directions these photons are going, we just want to add up their energy.

Since the photons and the energy they carry move at speed c, if an energy I_ν crosses the volume every unit time in direction $\hat{\mathbf{n}}$ at frequency ν, the amount of energy within the volume at any given time must be I_ν/c. To understand this argument, it is perhaps more intuitive to think about something like cars. Suppose cars are going down a one-lane road at 100 km / h, and there is one car every 2 km. How many will go past a given sign on the road in one hour? The answer is easy: we say that in $t = 1$ hr, the cars will travel $d = vt = 100$ km, and the number of cars within the 100 km line that will pass the road sign is $d \times n = (100 \text{ km})(1 \text{ car}/2 \text{ km}) = 50$ cars. Thus the flux of cars is $F = (d \times n)/t = nv$, and the density of cars can be computed from the flux as F/v. Our expression is exactly this.

We are only interested in the total energy density, not the direction, so to get this we must next integrate over all possible directions. This gives

$$E_\nu = \frac{1}{c}\int I_\nu \, d\Omega \tag{9}$$

as our final expression for the energy density at a given frequency. We could of course integrate over all frequencies if we wanted just the total energy independent of frequency.

Suppose instead of the energy we are interested in the flux? This is just the amount of energy that crosses a given surface of infinitesimal area dA per unit time, at a given frequency. The idea is pretty similar to the energy: we need to integrate over all possible directions photons could be travelling that would bring them across the surface. Consider a surface that lies in the xy plane, so that photons travelling in direction $\hat{\mathbf{n}}$ are travelling at an angle $\hat{\mathbf{n}} \cdot \hat{\mathbf{z}}$ to the surface. The velocity of the photons normal to the surface is $c\hat{\mathbf{n}} \cdot \hat{\mathbf{z}}$, so the total flux across the surface at frequency ν is

$$F_\nu = \int I_\nu \hat{\mathbf{n}} \cdot \hat{\mathbf{z}} \, d\Omega, \tag{10}$$

where the integral goes over all possible photon directions $\hat{\mathbf{n}}$. The factor $\hat{\mathbf{n}} \cdot \hat{\mathbf{z}}$ accounts for the fact that photons that are moving directly normal to the area ($\hat{\mathbf{n}} \cdot \hat{\mathbf{z}} = 1$) will contribute more to the flux than photons that are close to parallel to the surface ($\hat{\mathbf{n}} \cdot \hat{\mathbf{z}} \approx 0$).

Of course placing the surface in the xy plane was an arbitrary source. In general we can define the vector flux

$$\mathbf{F}_\nu = \int I_\nu \hat{\mathbf{n}} \, d\Omega, \tag{11}$$

so that in order to get the flux crossing a surface with a normal vector pointing in any particular direction we can just take $\mathbf{F}_\nu \cdot \hat{\mathbf{k}}$, where $\hat{\mathbf{k}}$ is the normal vector to the surface.

The quantities E_ν and $\mathbf{F}_\nu$ are referred to as *moments* of the radiation intensity. That is, they involve integrating the radiation field over direction, after multiplying by some number of powers of the direction vector $\hat{\mathbf{n}}$. An arbitrary number of moments can be defined. For example, the second moment is

$$\mathbb{P}_\nu = \frac{1}{c}\int I_\nu \hat{\mathbf{n}} \otimes \hat{\mathbf{n}} \, d\Omega, \tag{12}$$

is the radiation pressure tensor, where $\otimes$ indicates the outer product. In Cartesian terms, the components of $\mathbb{P}$ describe the rate at which radiation momentum in a particular direction is being transported in some other direction; for example, $(\mathbb{P})_{xz}$ is the rate at which the x component of momentum is being transported in the z direction by the radiation field.

II. Matter-radiation interaction

The transfer equation in vacuum is supremely boring – it is

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \hat{\mathbf{n}} \cdot \nabla I_\nu = 0, \tag{13}$$

which just says that photons keep going the direction they are already going, and do so at constant speed c. The interesting part is in how the photons interact with matter. We therefore now introduce source terms to sit on the right hand side of this equation. We can think of the terms we are introducing here as placeholders, in the sense that we will describe ways that radiation can interact with matter, but we will not yet specify the microphysics of exactly what is going on, or how to calculate the terms that describe these processes from the properties of the matter that is doing the interaction. We will fill in the missing microphysical information over the next few weeks.

A. Emission

Radiation has to come from somewhere. We define the emission coefficient j_ν for as the energy emitted per unit time per unit solid angle per unit frequency per unit volume by some matter.[1] The total power radiated per unit frequency is just

$$P_\nu = \int j_\nu \, d\Omega. \tag{14}$$

For an isotropic radiation source (usually the case for astrophysical sources, at least when viewed in the correct reference frame), $P_\nu = 4\pi j_\nu$.

Consider a collection of photons that travel a distance ds through a region containing material with an emission coefficient j_ν. The beam of photons has cross-sectional area dA, and thus the volume traversed by the beam is $dV = dA\, ds$. Since the intensity added per unit volume is $j_\nu \, dV$, it follows that the intensity added to the beam as it moves a distance ds through the emitting material is $j_\nu \, ds$, so $\partial I_\nu / ds = j_\nu$.

B. Absorption

Just as matter can emit radiation, it has to be able to absorb it – it not, we would not be able to build detectors. Consider radiation passing through a medium containing particles capable of absorbing it. Each absorbed has a finite probability of absorbing any photon that goes past it. We can describe this microphysically in terms of a cross section: suppose that the density of absorbers is n. As the photons in a beam of cross-sectional area dA move a distance ds, they pass $n\, dA\, ds$ absorbers.

Further suppose that each absorber has a cross section σ_ν – we can think of this as a real, physical cross-sectional area, except that it can depend on frequency. The total absorbing area within the beam of area dA, over a distance ds, is $n\sigma_\nu \, dA\, ds$. The total area of the beam is dA, so the fraction of the area that is blocked by the absorbers is $(n\sigma_\nu \, dA\, ds)/dA$. Thus the amount of energy absorbed is $I_\nu \, (n\sigma_\nu \, ds)$. Thus we can describe the absorption process by

$$\partial I_\nu / \partial s = -n\sigma_\nu I_\nu. \tag{15}$$

[1] Some authors also define the emissivity $\epsilon_\nu = 4\pi j_\nu / \rho$, where ρ is the matter density, as the energy emitted per unit mass per unit time per unit frequency by an isotropic emitter. For maximum confusion, however, some authors *also* define j_ν as the emissivity rather than the emission coefficient. Unfortunately the notation and definitions are not consistent from one sub-field to another, even within astronomy. Just be careful when reading the literature to see which convention is being used, and check the units to be sure.

We define the combination

$$n\sigma_\nu \equiv \alpha_\nu \tag{16}$$

as the absorption coefficient of the matter. We can also define the opacity of the matter as the absorption coefficient per unit mass, $\kappa_\nu = \alpha_\nu/\rho$.

C. The transfer equation with emission and absorption; steady state formal solution

Putting the emission and absorption terms together, we arrive at the transfer equation for pure emission and absorption:

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \hat{\mathbf{n}} \cdot \nabla I_\nu = j_\nu - \alpha_\nu I_\nu. \tag{17}$$

If we are in steady state, or if we are simply in a situation where the radiation field reaches equilibrium very quickly compared to the matter (very often the case), we can drop the time derivative term, and we have the steady state transfer equation:

$$\hat{\mathbf{n}} \cdot \nabla I_\nu = j_\nu - \alpha_\nu I_\nu. \tag{18}$$

Equivalently, along any given ray

$$\frac{dI_\nu}{ds} = j_\nu - \alpha_\nu I_\nu, \tag{19}$$

where s measures distance along the ray.

If either j_ν or α_ν is zero, the solutions to this equation are trivial. If $\alpha_\nu = 0$, the solution is

$$I_\nu(s) = I_\nu(s_0) + \int_{s_0}^{s} j_\nu(s')\, ds', \tag{20}$$

which just asserts that the intensity at a point s along the ray is equal to the intensity at a point s_0 further upstream plus the integral of the emission coefficient along all the intervening distance.

If $j_\nu = 0$, the solution is

$$I_\nu(s) = I_\nu(s_0) \exp\left[-\int_{s_0}^{s} \alpha_\nu(s')\, ds'\right] \equiv I_\nu(s_0) e^{-\tau_\nu}, \tag{21}$$

which just says that the initial intensity at any point s_0 upstream along the ray decreases exponentially, with the exponent determined by the integral of the absorption coefficient along the ray. We define this integral as τ_ν, the optical depth.

When both j_ν and α_ν are non-zero, we can still write down a formal solution as follows. The first step is to make a change of variable from s to τ_ν. We define

$$d\tau_\nu = \alpha_\nu\, ds, \tag{22}$$

and dividing both sides of Equation 19 by α_ν we have

$$\frac{dI_\nu}{d\tau_\nu} = S_\nu - I_\nu, \tag{23}$$

where $S_\nu = j_\nu/\alpha_\nu$ is called the source function. We next move the I_ν term on the right hand side to the left and multiply by e^{τ_ν} on both sides, which gives

$$
\begin{aligned}
\frac{dI_\nu}{d\tau_\nu} e^{\tau_\nu} + I_\nu e^{\tau_\nu} &= S_\nu e^{\tau_\nu} & (24) \\
\frac{d}{d\tau_\nu}\left(e^{\tau_\nu} I_\nu\right) &= S_\nu e^{\tau_\nu} & (25) \\
\left(e^{\tau_\nu} I_\nu\right)_{\tau_\nu} - \left(e^{\tau_\nu} I_\nu\right)_{\tau_\nu=0} &= \int_0^{\tau_\nu} e^{\tau_\nu'} S_\nu(\tau_\nu')\, d\tau' & (26) \\
I_\nu(\tau_\nu) &= I_\nu(0) e^{-\tau_\nu} + \int_0^{\tau_\nu'} e^{-(\tau_\nu - \tau_\nu')} S_\nu(\tau_\nu')\, d\tau' & (27)
\end{aligned}
$$

This result has an easy physical interpretation. The intensity at a given optical depth τ_ν along the ray has two terms. The first, $I_\nu(0)e^{-\tau_\nu}$, is the intensity that entered the ray at optical depth zero (wherever we choose to place that – it is arbitrary), reduced exponentially by the optical depth along the path. The second is that we take the emission at each optical depth along the ray, decrease it by exp of the optical depth between that point and the point where we're calculating, and integrate all these contributions along the ray.

If the source function S_ν is constant, then the solution is again trivial, and we have

$$
I_\nu(\tau_\nu) = I_\nu(0)e^{-\tau_\nu} + S_\nu\left(1 - e^{-\tau_\nu}\right) = S_\nu + e^{-\tau_\nu}\left(I_\nu(0) - S_\nu\right). \tag{28}
$$

Thus as $\tau_\nu \to \infty$, the intensity approaches the source function, and the "initial condition" at $\tau_\nu = 0$ is forgotten.

III. Thermodynamics of radiation fields

We now proceed to derive on the more important properties of radiation fields and the source function: their functional form when the radiation in question is thermal, meaning that the radiation field and the matter with which it interacts are in thermal equilibrium.

A. Kirchhoff's Law of thermal radiation

We start this discussion with a derivation of Kirchhoff's Law of thermal radiation. Consider two closed boxes held at the same temperature T. We allow each one to sit for a long time, so that the radiation intensity I_ν inside the box comes into equilibrium with the box walls. Now we cut a small hole in each box and place the two holes next to one another, allowing radiation to flow out of one box and into the other. We also place a filter between the two holes, which allows only radiation at some specific frequency ν, or some narrow range of frequencies around ν, to pass. If one box or the other has a higher value of I_ν, then there will be a net flow of radiation from the box with the higher I_ν into the one with the lower I_ν.

However, since the two boxes are both in equilibrium at the same temperature, the net flow of energy from one to the other must be zero regardless of which filter we put between them. If it were not, that would violate the second law of thermodynamics, because we could catch some of the photons flowing from one box into the other (say with a photovoltaic cell) and use it to do work – and doing work with zero heat transfer is not allowed by the second law of thermodynamics. Thus regardless of any other properties of the boxes – shape, size, contents, etc., simply the fact that they are in equilibrium at the same temperature tells us that I_ν inside each box is the same.

However, notice that this is true no matter what filter we use. It is also true regardless of where we put the holes in the boxes, or how we orient them. Thus we learn that, in thermal equilibrium, I_ν must be a function of temperature alone. It does not depend on position, direction, or any other property of the matter producing the radiation. There is a universal function $B_\nu(T)$ that describes the frequency distribution of all radiation produced by matter in thermal equilibrium.

This has an immediate implication with regard to our formal solution above. Recall that we showed that, after propagating through an optical depth τ_ν through a medium with a uniform source function S_ν, the radiation intensity along any ray approaches S_ν. It immediately follows that, if the material is in thermal equilibrium, the source function must be $S_\nu = B_\nu(T)$, where T is the material temperature. Thus for any matter in thermal equilibrium, it immediately follows that

$$j_\nu = \alpha_\nu B_\nu(T), \tag{29}$$

where T is the matter temperature. We will make use of this result throughout this class, because it can save us a great deal of work. If we're considering material in thermal equilibrium, we need not calculate how it emits and absorbs radiation separately. As soon as we know one, we immediately know the other from Kirchhoff's Law.

B. The Planck spectrum

Kirchhoff's Law tells us that there is a function $B_\nu(T)$ that describes thermal radiation, but does not tell us what it is. Indeed, Kirchhoff, who died in 1887, before the quantum revolution, did not live to see the problem solved. The solution is fundamentally quantum mechanical, and was derived by Planck in 1900. The first part of the argument, however, is purely classical.

Consider a cubical box with sides of length L, with conducting walls; these assumptions will not alter our conclusion, since we have already shown that $B_\nu(T)$ is universal and does not depend on the properties of the box. In a classical picture, light consists of propagating electromagnetic waves, and as we will show next week, the electric field $\mathbf{E}$ of these waves obeys the wave equation

$$\frac{\partial^2 \mathbf{E}}{\partial^2 t} - c^2 \nabla^2 \mathbf{E} = 0, \tag{30}$$

where $\mathbf{E}$ is the electric field vector, and valid solutions of this equation plus its magnetic equivalent have $\mathbf{E}$ perpendicular to the direction of wave propagation. Since the walls of the box are conducting, $\mathbf{E}$ must be identically zero in the directions tangent to the box walls at any point on the walls – if it were not, a current would flow to create a cancelling electric field.

If we place the lower left corner of the box at the origin, then one can show by direct substitution that the following is a valid solution to the system:

$$\mathbf{E} = \mathbf{E}_0 e^{i\omega t} \sin k_x x \, \sin k_y y \, \sin k_z z \equiv \mathbf{E}_{(n_x, n_y, n_z)} \tag{31}$$

where $k_x = \pi n_x / L$ where $n_x = 1, 2, 3, \ldots$ and similarly for y and z. The vector $\mathbf{E}_0$ is a (possibly complex) constant vector amplitude subject to the constraint $\mathbf{E} \cdot \mathbf{k} = 0$, where $\mathbf{k} = (k_x, k_y, k_z)$. Finally, the angular frequency $\omega = c/|\mathbf{k}|$. One can show with a bit more effort (appealing to Sturm-Liouville Theory) that these solutions with different n_x, n_y, and n_z are the eigenfunctions of the equation, in the sense that any valid solution to the equation can be written out as a weighted sum of the modes, i.e., as

$$\mathbf{E} = \sum_{n_x=1}^{\infty} \sum_{n_y=1}^{\infty} \sum_{n_z=1}^{\infty} w(n_x, n_y, n_z)\mathbf{E}(n_x, n_y, n_z). \tag{32}$$

The vector amplitude $\mathbf{E}_0$ for each mode has three components,[2] but the number of degrees of freedom is reduced by one by the constraint equation $\mathbf{E} \cdot \mathbf{k} = 0$, so there are only two independent components. These correspond to polarisation modes. We can choose to describe these in various ways, which we will discuss in detail next week, but for now we simply note that, for each (n_x, n_y, n_z), there are two distinct orthogonal modes.

We can now count how many distinct modes there are up to angular frequency ω. Since $\omega = c/|\mathbf{k}|$, a given mode (n_x, n_y, n_z) has angular frequency ω or less if and only if

$$\sqrt{n_x^2 + n_y^2 + n_z^2} \leq \frac{\omega L}{\pi c}. \tag{33}$$

If we ignore discreteness (i.e., if we assume $n_x, n_y, n_z \gg 1$), we can find the number of (n_x, n_y, n_z) triplets satisfying this condition just by thinking of this problem as equivalent to asking the volume of 1/8 of a sphere of radius $\omega L/\pi c$ – 1/8 of a sphere because we are only interested in the octant where n_x, n_y, and n_z are all positive. In this case we have

$$N(< \omega) = 2 \cdot \frac{1}{8} \left[\frac{4}{3}\pi \left(\frac{\omega L}{\pi c} \right)^3 \right] = \frac{\omega^3 L^3}{3\pi^2 c^3} = \frac{8\pi \nu^3 L^3}{3c^3}, \tag{34}$$

where in the last step we used $\omega = 2\pi\nu$. Note that the factor of 2 in the front is due to the two distinct polarisation modes. Taking the derivative of this gives the density of modes per unit frequency

$$\frac{dN}{d\nu} = \frac{8\pi\nu^2 L^3}{c^3} \tag{35}$$

The next step is the quantum mechanical one. The energy in each distinct mode is not allowed to take on any value. Instead the energies are quantised so that the energy of a mode can only by $E = vh\nu$, where $v = 0, 1, 2, \ldots$. The average energy of this mode in a system with temperature T can then be derived from the Boltzmann distribution. The probability that a particular mode has quantum number v is

$$p(v) = \frac{e^{-vh\nu/k_B T}}{\sum_{v=0}^{\infty} e^{-vh\nu/k_B T}}, \tag{36}$$

and the corresponding mean energy in the mode is

$$\langle E \rangle = \sum_{v=0}^{\infty} vh\nu p(v) = h\nu \sum_{v=0}^{\infty} \frac{ve^{-vh\nu/k_B T}}{\sum_{v'=0}^{\infty} e^{-v'h\nu/k_B T}}. \tag{37}$$

If we let $x = e^{-h\nu/k_B T}$, this becomes

$$\langle E \rangle = h\nu \sum_{v=0}^{\infty} \frac{vx^v}{\sum_{v'=0}^{\infty} x^{v'}} = h\nu \frac{x}{1-x} = \frac{h\nu}{e^{h\nu/k_B T} - 1}, \tag{38}$$

[2] In fact, if we're being really pedantic it has six, since each amplitidue can be complex valued. However, we will see next week that the complex part just represents a phase shift in the wave, which is equivalent to shifting the zero of time, and thus does not really count as a separate mode.

where we used the formula for the sum of a geometric series.

Putting this all together, the total energy in the radiation field at frequency ν is

$$\frac{dN}{d\nu}\langle E\rangle = \frac{8\pi h\nu^3 L^3/c^3}{e^{h\nu/k_BT}-1}. \tag{39}$$

Since we have already convinced ourselves based on Kirchhoff's arguments that the energy density must be uniform within the box, the energy density is therefore this divided by the volume of the box L^3:

$$E_\nu = \frac{8\pi h\nu^3/c^3}{e^{h\nu/k_BT}-1}. \tag{40}$$

However, we demonstrated earlier that energy density and intensity are related by

$$E_\nu = \frac{1}{c}\int I_\nu d\Omega. \tag{41}$$

Since $I_\nu = B_\nu(T)$ is isotropic for a thermal radiation field, it follows that

$$E_\nu = \frac{4\pi}{c}B_\nu(T), \tag{42}$$

and therefore

$$B_\nu(T) = \frac{2h\nu^3/c^2}{e^{h\nu/k_BT}-1}. \tag{43}$$

We have therefore found the functional form for $B_\nu(T)$. This is known as the Planck function.