James S. Dunlop
16 lectures

Astrophysics 4/MPhys 4/5
2012/13

Radiation and Matter

I. Physical Introduction

1 Review of Classical Radiation

This section reviews the properties of the electromagnetic fields in free space and shows

(1) that they are derivable from four potentials (ϕ, A_x, A_y, A_z),

(2) that the potentials can be changed (gauge transformations) without changing the physics i.e. *the resulting fields,*

(3) that the equations governing the potentials (and by implication Maxwell's field equations) are Lorentz covariant,

(4) that at distances large compared with source size and wavelength (the wave zone) the potentials and fields can be expressed as transverse waves with $E = cB$ and $\mathbf{E}$ perpendicular to $\mathbf{B}$,

(5) that the fields due to a point charge lead – via the Poynting vector – to Larmor's formula for classical radiation (power proportional to acceleration squared),

(6) that radiation in an otherwise empty cubic box can be expressed as the sum of plane standing or travelling waves with an infinite but discrete set of frequencies and related wave vectors and polarization directions; the total energy is the sum of the energies of each of the modes.

(7) that the dynamical similarity between the real component of the amplitude of the vector potential and the displacement of a simple harmonic oscillator leads to a classical formalism called the 'oscillator expansion' which will be exploited in the section on quantization

1.1 Maxwell's equations, field energy and potentials

1.1.1 Maxwell's equations

Maxwell's equations *in vacuo* are

$$\nabla_\wedge \mathbf{E} = -\partial \mathbf{B}/\partial t \qquad \nabla . \mathbf{B} = 0$$

$$\nabla_\wedge \mathbf{B} = \mu_0(\mathbf{j} + \partial(\epsilon_0 \mathbf{E})/\partial t) \qquad \nabla . \mathbf{E} = \rho/\epsilon_0 .$$

1.1.2 Field energy and Poynting vector

By manipulating these equations (see appendix) we can arrive at the result

$$\frac{\partial}{\partial t}\int(\frac{\epsilon_0}{2}\mathbf{E}^2+\frac{1}{2\mu_0}\mathbf{B}^2)\,dV=-\int\mathbf{j}\,.\,\mathbf{E}\,dV-\int_{\text{surface}}\mathbf{E}_\wedge\mathbf{B}/\mu_0\;.\;d\mathbf{s}.$$

The first term on the right hand side is dissipation of energy thro'out the volume (*Ohmic dissipation*) so in a situation where the second term is zero (*e.g.* boundary surface at infinity where fields are negligible), this is the total rate of loss of energy, which implies (a) that the integral on the left side is the field energy, and (b) that the surface integral also represents energy loss, and that the vector

$$\mathbf{S}=\mathbf{E}_\wedge\mathbf{B}/\mu_0,$$

known as the Poynting vector, may be regarded as the energy flux (note that only the *surface integral* is defined, and there are situations - not encountered here - in which this simple identification of $\mathbf{S}$ with energy flux is suspect).

1.1.3 Electromagnetic potentials

By using the first pair of Maxwell's equations, which evidently express constraints on the fields rather than saying how they arise, the fields $\mathbf{E},\mathbf{B}$ can be seen to be derivatives of potential functions $\mathbf{A}$ and ϕ:

$$\mathbf{B}=\nabla_\wedge\mathbf{A}\qquad\qquad\mathbf{E}=-\partial\mathbf{A}/\partial t-\nabla\phi.$$

The potentials $\mathbf{A},\phi$ can be changed by the gradient of an arbitrary function $\chi(\mathbf{r},t)$, without changing the values of $\mathbf{E},\mathbf{B}$. This is called *gauge invariance*.

$$\mathbf{A}\rightarrow\mathbf{A}'=\mathbf{A}+\nabla\chi,\qquad\qquad\phi\rightarrow\phi'=\phi-\partial\chi/\partial t.$$

Substituting the potentials in the second pair of Maxwell's equations, re-arranging, and re-labelling the constant $(\mu_0\epsilon_0)^{-1}=c^2$, we see the wave magic in Maxwell's equations – due to the insertion by Maxwell of the $\partial\mathbf{E}/\partial t$ term –

$$\nabla^2\mathbf{A}-\frac{\partial^2\mathbf{A}}{c^2\partial t^2}-\nabla\left(\nabla.\mathbf{A}+\frac{\partial\phi}{c^2\partial t}\right)=-\mu_0\mathbf{j}$$

and

$$\nabla^2\phi-\frac{\partial^2\phi}{c^2\partial t^2}+\partial/\partial t\left(\nabla.\mathbf{A}+\frac{\partial\phi}{c^2\partial t}\right)=-\rho/\epsilon_0.$$

These are *nearly* wave equations, and can be made so using the gauge invariance of the fields: It is possible to choose χ so that the transformed potentials satisfy $\nabla.\mathbf{A}'+\partial\phi'/c^2\partial t=0$, and then the potentials (drop the primes) satisfy *wave equations*

$$\Box^2 \mathbf{A} = -\mu_0 \mathbf{j} \qquad \Box^2 \phi = -\rho/\epsilon_0$$

where $\Box^2$ is shorthand for $\nabla^2 - (\partial^2/c^2\partial t^2)$. Note that this operator is Lorentz invariant. Furthermore, ρ transforms like the time-component of a Lorentz 4-vector since $dq = \rho dx\, dy\, dz$, the charge element, is invariant as is $dt\, dx\, dy\, dz$. It is straightforward to show that the rest of the 4-vector is $\mathbf{j}$. This carries the implication that if we write $j^\mu = (\mathbf{j}, \rho c)$ and $A^\mu = (\mathbf{A}, \phi/c)$, the equations for the potentials take the Lorentz covariant form

$$\Box^2 A^\mu = -\mu_0 j^\mu .$$

1.1.4 Solution of the wave equation

At this point it has been shown that Maxwell's equations can be written as a well-posed problem in potential theory: four equations in four unknowns with four source functions. Moreover, it is also clear that Maxwell's theory is relativistic (*i.e.* Lorentz covariant). This is not pursued further here.

Returning to the potential equations, we shall consider the vector potential equation

$$\Box^2 \mathbf{A} = -\mu_0 \mathbf{j},$$

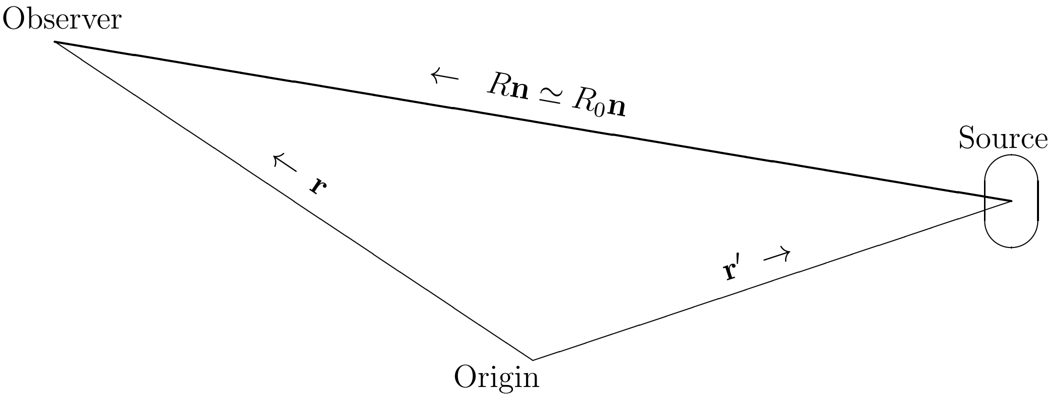
whose solution (see appendix) is

$$\mathbf{A}(\mathbf{r}, t) = \left(\frac{\mu_0}{4\pi}\right) \int \frac{\mathbf{j}(\mathbf{r}', t_{ret})}{|\mathbf{r} - \mathbf{r}'|}\, d^3\mathbf{r}'$$

where t_{ret}, the 'retarded time', is given by

$$t_{ret} = t - |\mathbf{r} - \mathbf{r}'|/c.$$

1.2 Potentials in the wave zone – radiant energy



By reference to the diagram, $|\mathbf{r} - \mathbf{r}'|/c$ may be written $\mathbf{r}.\mathbf{n}/c - \mathbf{r}'.\mathbf{n}/c$ and so

$$t_{ret} = t - \mathbf{r}.\mathbf{n}/c + \mathbf{r}'.\mathbf{n}/c.$$

Note that the distance from source to observer is very great, so that $|\mathbf{r}-\mathbf{r}'|$ is approximately constant *in the denominator*, and will be approximated by R_0, the distance from the observer to a standard point in the source. In the function $\mathbf{j}(\mathbf{r}', t_{ret})$ which is now the integrand, the variable $\mathbf{r}'$ is the (dummy) integration variable on integration over the source, and so the result is a function of the non-integrated parameters, $t - \mathbf{r}.\mathbf{n}/c$:

$$\begin{aligned}\mathbf{A}(\mathbf{r},t) &\simeq \frac{\mu_0}{4\pi R_0}\int \mathbf{j}(\mathbf{r}', t_{ret})\, d^3\mathbf{r}' \\ &\simeq \frac{\mu_0}{4\pi R_0}\int \mathbf{j}(\mathbf{r}', [t - \mathbf{r}.\mathbf{n}/c] + \mathbf{r}'.\mathbf{n}/c)\, d^3\mathbf{r}' \\ &= \mathbf{F}(t - \mathbf{r}.\mathbf{n}/c)\,/\,R_0, \text{ say.}\end{aligned}$$

The ‘0’ is dropped from the R after this. This expression for $\mathbf{A}$ is that of an outgoing wave, amplitude decreasing as $1/R$, travelling with velocity c[1]. We have uncovered the constancy of the velocity of light.

We can use (see appendix) the following relations for a vector $\mathbf{F}$ (or scalar f) whose argument is of the form $t - \mathbf{n}.\mathbf{r}/c$:

$$\nabla.\,\mathbf{F} = -\frac{\mathbf{n}}{c}\,.\,\frac{\partial}{\partial t}\mathbf{F}, \qquad \nabla_\wedge\mathbf{F} = -\frac{\mathbf{n}}{c}{}_\wedge\frac{\partial}{\partial t}\mathbf{F}, \qquad \nabla f = -\frac{\mathbf{n}}{c}\frac{\partial}{\partial t}f$$

and since in this, the wave zone, variations in the denominator R are negligible, these results can be applied to $\mathbf{A}$, and to $\mathbf{B} = \nabla_\wedge\mathbf{A}$; this allows us to write the first and third Maxwell equations as

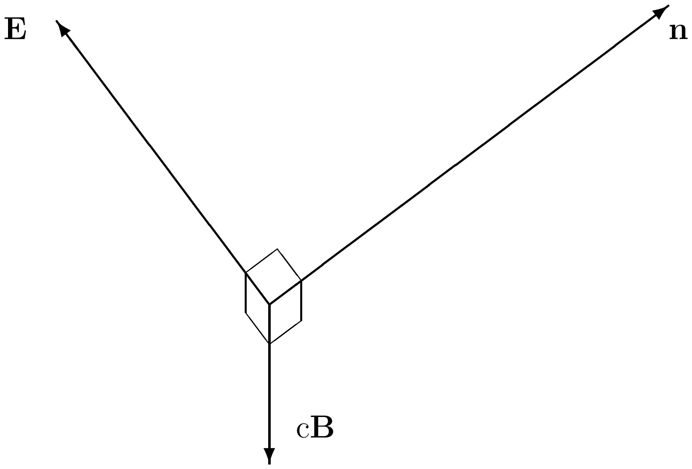
$$\mathbf{n}_\wedge\mathbf{E} = c\mathbf{B}, \qquad\qquad c\mathbf{n}_\wedge\mathbf{B} = -\mathbf{E}.$$

---

[1] Note that a static ($1/R$, Coulomb) potential may be part of the solution; we disregard it in what follows; its gradient falls off as R^{-2} and it carries no radiation.

(Use $\mathbf{j} = 0$ in the wave zone; if $\mathbf{n}_\wedge\dot{\mathbf{E}} = c\dot{\mathbf{B}}$, then $\mathbf{n}_\wedge\mathbf{E} = c\mathbf{B}$, etc. You can also get these results directly from the source-free Maxwell equations).

Thus **E**, **B**, **n** form a right-handed triad (like x, y, z), and $E = cB$.



The energy transported by the field is determined by the Poynting vector **S** which in the wave zone becomes

$$\mathbf{S} = -(\mathbf{n}_\wedge\mathbf{B}c)_\wedge(\mathbf{B}/\mu_0) = B^2c\mathbf{n}/\mu_0 = \epsilon_0 c\dot{\mathbf{A}}_\perp^2\mathbf{n}$$

and the form derived in the wave zone for $\mathbf{A} \propto R^{-1}\mathbf{F}(t - \mathbf{r}.\mathbf{n}/c)$ can be substituted. There is a simple and very important case which can be explicitly solved – the point charge.

1.3 Point charges, the Liénard-Wiechert potentials, the Larmor radiation formula

The current density of a point charge is

$$\mathbf{j}(\mathbf{r}', t) = q\mathbf{v}\delta(\mathbf{r}' - \mathbf{x}(t))$$

where $\mathbf{x}(t)$ is the given trajectory of the point charge ($\mathbf{v} = d\mathbf{x}/dt$) and so the vector potential is (previous section)

$$\mathbf{A}(\mathbf{r}, t) \simeq \frac{\mu_0}{4\pi R}\int \mathbf{j}(\mathbf{r}', t_{ret})\, d^3\mathbf{r}' = \frac{\mu_0 q}{4\pi R}\int \mathbf{v}(t_{ret})\delta(\mathbf{r}' - \mathbf{x}(t_{ret}))\, d^3\mathbf{r}'.$$

Since the delta function depends on $\mathbf{r}'$ through t_{ret} we cannot just collapse the integral on the delta function at $\mathbf{r}' = \mathbf{x}$. We can deal with this most simply by using the neat trick of writing

$$\mathbf{j}(\mathbf{r}', t_{ret}) = \int \mathbf{j}(\mathbf{r}', \tau)\delta(\tau - t_{ret})\, d\tau$$

getting

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0 q}{4\pi R}\int\int \mathbf{v}(\tau)\delta(\mathbf{r}' - \mathbf{x}(\tau))\delta(\tau - t_{ret})\, d\tau\, d^3\mathbf{r}'.$$

As τ is independent of $\mathbf{r}'$ we can now integrate over space by collapsing the space delta function onto $\mathbf{x}$, noting that t_{ret} then becomes $t - |\mathbf{r} - \mathbf{x}(\tau)|/c$, to get

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0 q}{4\pi R}\int \mathbf{v}(\tau)\delta(\tau - t_{ret})\, d\tau.$$

Now note that (a) for Dirac delta functions

$$\int f(x)\delta(g(x))\, dx \equiv \left(f(x) \,/\, dg/dx\right)_{g=0},$$

and (b) $\partial(\tau - t_{ret})/\partial\tau = (1 - \mathbf{v}.\mathbf{n}/c)$, so that finally (writing $\mathbf{v}/c = \boldsymbol{\beta}$)

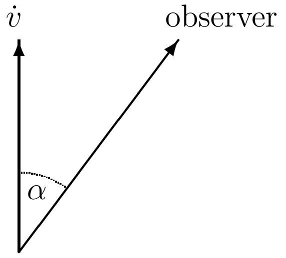
$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0 q\mathbf{v}}{4\pi R}\frac{1}{1 - \boldsymbol{\beta}.\mathbf{n}}$$

and an analogous equation for ϕ ($\mu_0 \to \epsilon_0^{-1}$, $q\mathbf{v} \to q$).
These are the Liénard-Wiechert potentials *in the wave zone*. We drop the relativistic factor for now, and so

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0 q\mathbf{v}}{4\pi R},$$

and from section 1.2

$$\mathbf{S} = \frac{1}{4\pi R^2}\frac{q^2/c^3}{4\pi\epsilon_0}\dot{v}^2 \sin^2\alpha$$



Since $d\Omega = \sin\alpha\, d\alpha\, d\phi$ and $\int \sin^3\alpha\, d\alpha\, d\phi = 8\pi/3$,

$$P = \int \mathbf{S}.d\mathbf{s} = \int SR^2\, d\Omega = \frac{2}{3}\frac{q^2/c^3}{4\pi\epsilon_0}\dot{v}^2,$$

which is Larmor's radiation formula.

1.4 The radiation field as a set of oscillators

We specify the electromagnetic field, free of sources, by its vector potential $\mathbf{A}$, satisfying

$$\nabla^2\mathbf{A} - \ddot{\mathbf{A}}/c^2 = 0.$$

We choose a gauge such that $\phi = 0$. Further, since $\rho = 0$,

$$\nabla .\, \mathbf{E} = \nabla .\, \left(-\frac{\partial \mathbf{A}}{\partial t} \right) = 0,$$

or $\nabla .\, \mathbf{A} = const.$ Now regauge with gauge function χ independent of t; it doesn't change $\phi = 0$, but can be used to transform $\mathbf{A}$ to give also $\nabla .\, \mathbf{A} = 0$. We now have all of the information about the electromagnetic field in just two components of the vector potential (one suppressed by requiring $\nabla .\, \mathbf{A} = 0$), instead of the four components of the complete set.

In order to deal with each wave separately, use the old trick of putting the radiation in a big box with boundary conditions, so that the states become denumerable, analogously to the transverse vibrations of a string held stationary at either end.

We enclose the electromagnetic field in a very large cubic box with side L. Then the potential is expressible as a triple Fourier series with amplitudes $\mathbf{X}_\mathbf{k}$:

$$\mathbf{A}(\mathbf{r}, t) = \sum_\mathbf{k} \mathbf{X}_\mathbf{k}(t) e^{i\mathbf{k}.\mathbf{r}}.$$

The wave vector $\mathbf{k} = (2\pi/\lambda)\mathbf{n} = (k_x, k_y, k_z)$, and $\mathbf{k}$ takes only the discrete values $(2\pi/L)(n_x, n_y, n_z)$, with the n's positive or negative integers. The summation is over these integers.

The condition $\nabla .\, \mathbf{A} = 0$ leads to $\mathbf{k} \,.\, \mathbf{X}_\mathbf{k}(t) = 0$, ie the waves are transverse. Furthermore, the wave equation gives

$$\ddot{\mathbf{X}}_\mathbf{k}(t) + \omega^2 \mathbf{X}_\mathbf{k}(t) = 0$$

where $\omega = ck$, and so a solution is $\mathbf{X}_\mathbf{k}(t) = \mathbf{X}^{(1)}_\mathbf{k} e^{-i\omega t} + \mathbf{X}^{(2)}_\mathbf{k} e^{+i\omega t}$ with $\mathbf{X}^{(1)}_\mathbf{k}, \mathbf{X}^{(2)}_\mathbf{k}$ arbitrary time-independent vectors. Putting this in the expression for $\mathbf{A}(\mathbf{r}, t)$ and reversing the signs of k in all the terms in the second sum (doesn't affect the result!),

$$\mathbf{A}(\mathbf{r}, t) = \sum_\mathbf{k} \mathbf{X}^{(1)}_\mathbf{k} e^{-i(\omega t - \mathbf{k}.\mathbf{r})} + \sum_\mathbf{k} \mathbf{X}^{(2)}_{-\mathbf{k}} e^{+i(\omega t - \mathbf{k}.\mathbf{r})}.$$

Then note that if we put $\mathbf{X}^{(2)}_{-\mathbf{k}} = \left(\mathbf{X}^{(1)}_\mathbf{k}\right)^*$ that this expression is *real*, as required.

The direction of vibration $\mathbf{X}_\mathbf{k}$, perpendicular to $\mathbf{k}$, can be expressed in terms of two directions, represented by the unit vectors $\mathbf{e}_{\mathbf{k},\alpha}$ where $\mathbf{e}_{\mathbf{k},1}$ is *e.g.* in the 'vertical' plane through $\mathbf{k}$, and $\mathbf{e}_{\mathbf{k},2}$ completes a right-handed triad with $\mathbf{n} = \mathbf{k}/k$ and $\mathbf{e}_{\mathbf{k},1}$.

Writing $\mathbf{X}^{(1)}_\mathbf{k} = \sqrt{(\hbar/(2\epsilon_0 V\omega)}\, \mathbf{e}_{\mathbf{k},\alpha} a_{\mathbf{k},\alpha}$ then gives

$$\mathbf{A}(\mathbf{r}, t) = \sum_{\mathbf{k},\alpha} \sqrt{\frac{\hbar}{2\epsilon_0 V\omega}}\, \mathbf{e}_{\mathbf{k},\alpha} \left(a_{\mathbf{k},\alpha}(t) e^{i\mathbf{k}.\mathbf{r}} \;+\; a^*_{\mathbf{k},\alpha}(t) e^{-i\mathbf{k}.\mathbf{r}} \right).$$

The constant square root in this equation is chosen with an ulterior motive to which we return in due course – it obviously multiplies the amplitudes $a_{\mathbf{k},\alpha}$, $a^*_{\mathbf{k},\alpha}$. Notice that the factor e^{-iwt} has been absorbed into $a_{\mathbf{k},\alpha}$ and correspondingly e^{iwt} into $a^*_{\mathbf{k},\alpha}$.

Any electromagnetic field in the box can be expressed as a sum of such waves (since they are the only ones satisfying the boundary conditions, they form a complete set). Thus the field can be expressed as a denumerable set of independent amplitudes $a_{\mathbf{k},\alpha}$.

Let us review what has been done. The classical source-free electromagnetic field is expressed in terms of its potentials from which all other properties are derivable. These potentials (except for the static potentials) can be written as an infinite set of transverse waves, if the gauge is chosen so that $\nabla.\mathbf{A} = 0$. The field is assumed to be contained within a volume so large that edge effects are negligible, except that the boundary conditions ensure that the wave numbers take discrete values (very closely spaced, so that sums over these can be approximated by integrals when convenient; the larger the assumed box, the closer the spacing). We refer to the state of vibration of the wave in one of these discrete states as a *mode*. Then

$$\mathbf{A}(\mathbf{r}, t) = \sum_n (\mathbf{A}_n a_n + \mathbf{A}^*_n a^*_n)$$

where n is an index labelling the modes.[2] In the case of plane waves, $n \equiv (\mathbf{k}, \alpha)$.

1.5 The oscillator energies

The total field energy $H = \int \frac{1}{2}(\epsilon_0 E^2 + B^2/\mu_0)\, dV$ can be expressed first in terms of the vector potential, expanded into linear or spherical modes as above, which after a LOT of grinding down gives a sum over states of independent terms

$$H_n = \frac{1}{2}(a_n a^*_n + a^*_n a_n)\, \hbar\omega,$$

where we haven't commuted the a, a^* for future convenience.

What is the linear *momentum density* of the plane wave field? Consider a pulse of radiation in a given direction containing energy E, and for which special relativity requires $E = pc$ as it travels with the speed of light. By measuring the flow of this pulse across a surface we find $SA\Delta t = E$ where S is the magnitude of the Poynting vector, A is the cross section of the pulse, $\Delta t = l/c$ is the time for the pulse of length l to pass. Hence $S = c^2 p/(Al)$, and is therefore c^2 times the momentum density in the pulse.

A calculation similar to that for energy then gives for total momentum

$$\mathbf{K} = c^{-2} \int \mathbf{S}\, dV = \sum \frac{1}{2}(a_{\mathbf{k},\alpha} a^*_{\mathbf{k},\alpha} + a^*_{\mathbf{k},\alpha} a_{\mathbf{k},\alpha})\, \hbar\mathbf{k}.$$

Now introduce real linear combinations of a_n, a^*_n

$$Q_n = \sqrt{\hbar/2\omega}(a_n + a^*_n), \quad P_n = \dot{Q}_n = -i\omega\sqrt{\hbar/2\omega}(a_n - a^*_n).$$

---

[2] It is possible to use the fundamental vibrations of other complete sets instead: for instance, spherical waves of given angular momentum, instead of plane waves

Then the energy per mode can be written

$$H_n = \frac{1}{2}(P_n^2 + \omega^2 Q_n^2).$$

Furthermore P_n and Q_n obey

$$\partial H_n/\partial Q_n = -\dot{P}_n, \qquad \partial H_n/\partial P_n = \dot{Q}_n,$$

so behave as dynamically conjugate momentum and coordinate respectively, with respect to the *Hamiltonian* H; these are Hamilton's equations.

These equations are identical to the equations of motion of a harmonic oscillator of unit mass and natural frequency $\omega = kc$ (see next section). Each state is formally the same as such an oscillator, each independent of the others. The energy of the electromagnetic field in the box is identical with that of a set of independent oscillators, one for each mode $(\mathbf{k}, \alpha)$. This was the start of Dirac's breakthrough to quantum electrodynamics.

2 Quantum Radiation

This section
(a) Reviews the quantum mechanics of SHM, using the 'number representation' and ladder operators

(b) Completes the analogy between radiation oscillations in a box and SHM, and carries across the quantization procedure to the waves, thus inventing photons or radiation quanta.

2.1 Quantized SHM

First of all the quantum mechanics of the simple harmonic oscillator will be reviewed from the point of view of annihilation and creation operators. Then the connection noted above with modes of the radiation field is used to quantize the radiation field (*Dirac's second step to quantum electrodynamics, soon after the invention of proper quantum mechanics*).

Assume for simplicity that the oscillating mass of the oscillator is unity. Then the energy equation is

$$H = \frac{1}{2}(p^2 + \omega^2 q^2),$$

where q is the displacement and p is the corresponding canonical momentum (in this case it is v) with respect to the Hamiltonian H. These obey Hamilton's equations

$$\partial H/\partial q = -\dot{p} \qquad \partial H/\partial p = \dot{q}.$$

To make this a quantum calculation p, q and therefore H become operators and p, q obey

$$[q, p] = i\hbar.$$

Now the clever bit. New operators c, $c^\dagger$ are defined by (*cf* a, a^* in previous section)

$$c = (2\hbar\omega)^{-1/2}(\omega q + ip) \quad c^\dagger = (2\hbar\omega)^{-1/2}(\omega q - ip).$$

so that

$$[c,\, c^\dagger] = 1 \quad \text{and} \quad H = \frac{\hbar\omega}{2}(cc^\dagger + c^\dagger c).$$

These two results give also the relations

$$cc^\dagger = (H/\hbar\omega) + \tfrac{1}{2}, \qquad c^\dagger c = (H/\hbar\omega) - \tfrac{1}{2}$$

which are used below.

Here is Dirac's argument for the states of a harmonic oscillator.

Let $|E\rangle$ be an eigenstate of H with eigenvalue E (*i.e.* $H|E\rangle = E|E\rangle$). The 'squared length' (it would be $\int \psi^* \psi dV$ in wave mechanical representation) of such an eigenstate is $\langle E|E\rangle$ and similarly for the state $c|E\rangle$ it is $\langle E|c^\dagger c|E\rangle$. From the operator relations above,

$$\langle E|c^\dagger c|E\rangle = (E/\hbar\omega - \tfrac{1}{2})\langle E|E\rangle.$$

where $\langle E|E\rangle$ must be positive and $\langle E|c^\dagger c|E\rangle$ must be positive or zero. Thus $(E/\hbar\omega - \frac{1}{2}) \geq 0$. The equality only occurs for some state $|E_0\rangle$ if $c|E_0\rangle = 0$; denote such a state as $|0\rangle$, *i.e.* $c|0\rangle = 0$.

Next,

$$Hc|E\rangle = (cH - \hbar\omega c)|E\rangle = (E - \hbar\omega)c|E\rangle$$

and so if $|E\rangle$ is an eigenstate of H with eigenvalue E, $c|E\rangle$ is an eigenstate with eigenvalue $E - \hbar\omega$. By repeated application of operator c we can form eigenstates with eigenvalues $E - n\hbar\omega$, with n an integer as large as we please, creating energy eigenvalues arbitrarily negative, *unless* at some point in the sequence we create the eigenstate $|0\rangle$ introduced above. This therefore must happen. Thus possible states are $|0\rangle$ with energy eigenvalue $\frac{1}{2}\hbar\omega$ and states with eigenvalues $n\hbar\omega$ greater than that, n being any positive integer, $E_n = (n + \frac{1}{2})\hbar\omega$.

These states can be made from $|0\rangle$ using the operator $c^\dagger$ since

$$Hc^\dagger|E\rangle = (E + \hbar\omega)c^\dagger|E\rangle$$

i.e. $c^\dagger|E\rangle$ is the eigenstate 'one up' from $|E\rangle$. Name the 'nth up from zero' eigenstate, after normalization, $|n\rangle$. It will have eigenvalue $(n + \frac{1}{2})\hbar\omega$.

To normalize the eigenstates suppose $|0\rangle$ is defined so that $\langle 0|0\rangle = 1$. Then proceed by induction. If the nth eigenstate is normalized,

$$\langle n|cc^\dagger|n\rangle = \langle n|H/\hbar\omega + \tfrac{1}{2}|n\rangle = (n+1)\langle n|n\rangle = n + 1.$$

Now the $(n+1)$th eigenstate is created, un-normalized, by $c^\dagger|n\rangle$. Put

$$c^\dagger|n\rangle = \alpha|n+1\rangle, \text{ so that } \langle n|cc^\dagger|n\rangle = |\alpha|^2\langle n+1|n+1\rangle = |\alpha|^2$$

where α is a constant and $|n+1\rangle$ is the normalized eigenstate. By comparison with the previous result, $\alpha = \sqrt{n+1}$ and so

$$c^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle.$$

Conversely

$$\langle n|c^\dagger c|n\rangle = n\langle n|n\rangle = n$$

and so

$$c\,|n\rangle = \sqrt{n}\,|n-1\rangle.$$

The operator $c^\dagger$ increases the energy by one quantum (creates a quantum) and amplifies the state by $\sqrt{n+1}$ and the operator c decreases the energy (annihilates a quantum) and amplifies the state by $\sqrt{n}$.

Using these operators all the properties of SHM oscillators can be deduced.

2.2 Quantizing the radiation oscillators

The radiation oscillators are quantized identically to SHM. The classical values defined in section 1.4 become non-commuting operators:

$$[Q_n, P_n] = i\hbar, \qquad [a_n, a_n^\dagger] = 1$$

and clearly the amplitudes $a_n, a_n^\dagger$ (which have the same relation to P_n, Q_n as c, $c^\dagger$ to p, q), have become the annihilation and creation operators for quanta in the state n, which, analogously with the harmonic oscillator case, can be labelled $|N_n\rangle$. Note the use of superscript '†' instead of '*' to denote the Hermitian conjugate of an operator, as opposed to the complex conjugate of a number.

The energy in state $|N_n\rangle$ is $(N_n + \frac{1}{2})\hbar\omega$ but the $\frac{1}{2}$ is dropped for reasons of expediency (it can't be measured ...). A transition caused by the operator a_n from $|N_n\rangle$ gives $\sqrt{N_n}|N_n - 1\rangle$. The creation operator produces state $\sqrt{N_n + 1}|N_n + 1\rangle$.

Photons It is convenient to talk of the energy quanta in these states as *photons*, with *energy* $\hbar\omega$ and *linear momentum* $\hbar\mathbf{k}$ in the case of plane waves, or *angular momentum* $l\hbar$ in the case of spherical waves, as specified by the state. The same physical system is being described whether we specify the numbers of photons in each momentum and polarization state, or specify the energy (quantized) of all oscillators, identified by their momentum and polarization.

(The state functions multiplying the operators we choose to normalize to 'one photon per state in the box'.)

Notice, however, that **photons** only exist in the context of the specified states! All sorts of paradoxes and conundrums ensue if this is ignored.

3 Quantum interaction of radiation with matter

The previous section sketched the quantum nature of free radiation. This section
(a) introduces Fermi's Golden rule as one expression of time dependent perturbation theory, giving the rate of transition between two states under the action of a perturbation

(b) sketches a formal way in which the interaction between radiation and matter can be expressed as a perturbation on free particles and fields

(c) derives the form of the transition rate

(d) introduces the dipole approximation

(e) shows that the photon occupation numbers in thermodynamic equilibrium are those of the black body distribution

3.1 Fermi's Golden Rule

Time dependent perturbation of a stationary system results in a state function that varies in time in a non-trivial way (*i.e.* not just by a term $e^{iEt/\hbar}$). The initial (stationary) state function Ψ would be represented as a linear superposition of the eigenstates of the system, the coefficients b_j being the amplitudes to be in each state. With the perturbation switched on, these amplitudes change :

$$\Psi = \sum b_j(t) u_j(\mathbf{r})\, e^{-iE_j t/\hbar}.$$

Schrödinger's equation is

$$i\hbar\, \partial\Psi/\partial t = H\Psi$$

where the Hamiltonian $H = H_0 + H'$ and H' is the perturbing Hamiltonian. If the initial state is 'i' this reduces to

$$i\hbar\,(db_f/dt) = H'_{fi}\, e^{-i(E_i - E_f)t/\hbar}$$

since $\int u_f^* u_i\, dV = \delta_{if}$ and H'_{fi} is defined as $\int u_f^* H' u_i\, dV$.

This leads, if H' is time independent, and with the usual set of approximations, to

$$w = |b_f|^2/t \quad \xrightarrow{t\to\infty} \quad \frac{2\pi}{\hbar}|H'_{fi}|^2\, \delta(E_f - E_i)$$

This is the rate of transition from a specified state 'i' to a specified final state 'f'. If there are many possible final states 'f', then the probabilities of transition add. If these matrix elements all have virtually identical values, and the final states are closely-spaced in energy with a density of $(dN/dE)_f$ states per unit energy interval, then the the result of the sum of rates is Fermi's Golden Rule :

$$w = \frac{2\pi}{\hbar}|H'_{fi}|^2\, (dN/dE)_f, \quad E_f = E_i$$

3.2 The interaction Hamiltonian from gauge invariance

Since observables are bilinear functions of the wave function ($\psi^*\psi$ for example), the same physics is described by ψ and $e^{i\alpha}\psi$ (gauge invariance). However, the Schrödinger equation turns out *not* to be invariant to this change. The free particle equation is

$$i\hbar\frac{\partial\psi}{\partial t} = \frac{\mathbf{p}^2}{2m}\psi = -\frac{\hbar^2}{2m}\nabla^2\psi$$

while inserting $e^{i\alpha}\psi$ for ψ in the equation gives

$$i\hbar(\frac{\partial}{\partial t} + i\dot{\alpha})\psi = -\frac{\hbar^2}{2m}(\nabla + i\nabla\alpha)^2\psi$$

and the solution will depend on $\alpha(\mathbf{r}, t)$. A way to sort this is to include, in addition to ψ, fields which transform under change of gauge so as to eliminate the ephemeral α. We can do this by adding $f(\mathbf{r}, t)$ to the operator $\partial/\partial t$, and $\mathbf{F}(\mathbf{r}, t)$ to ∇ in the original equation, and arranging that they transform by the subtraction of $i\dot{\alpha}$ and $i\nabla\alpha$ respectively, and the job is done. The doctored Schrödinger equation is

$$i\hbar(\frac{\partial}{\partial t} + f)\psi = -\frac{\hbar^2}{2m}(\nabla + \mathbf{F})^2\psi$$

which is gauge invariant at the expense of the introduction of the new fields $f, \mathbf{F}$ and their transformations under $\psi \to e^{i\alpha}\psi$. On examination, these fields have the same gauge properties as we saw already in electromagnetism, and can be treated as the potentials $\mathbf{A}, \phi$ on multiplication by a factor. The gauge invariant Schrödinger equation can then be written

$$i\hbar\frac{\partial\psi}{\partial t} = \frac{(\mathbf{p} - q\mathbf{A})^2}{2m}\psi + q\phi\psi$$

This substitution (relativistically $p^\mu \to p^\mu - qA^\mu$) is known traditionally as *minimal coupling.*

As an approximation, we use $(\mathbf{p} - q\mathbf{A})^2 = \mathbf{p}^2 - q(\mathbf{p} \,.\, \mathbf{A} + \mathbf{A} \,.\, \mathbf{p})$, neglecting the $\mathbf{A}^2$ term. Also, because $\nabla.\mathbf{A} = 0$, we can convert[3] $\mathbf{p}\,.\mathbf{A}$ to $\mathbf{A}\,.\mathbf{p}$, so in this approximation, Schrödinger's equation becomes (remember we have $\phi = 0$)

$$i\hbar\frac{\partial\psi}{\partial t} = \frac{\mathbf{p}^2}{2m}\psi - \frac{q}{m}\mathbf{A} \,.\, \mathbf{p} \equiv H_0 + H_{\text{int}}.$$

The interaction, or perturbation, Hamiltonian is thus

$$H' = -\frac{q}{m}\mathbf{A} \,.\, \mathbf{p}$$

3.3 Radiative transitions

The chosen initial state of the system is the *product* of an unperturbed atomic state, say $|X\rangle$, and unperturbed radiation state specified by all the photon occupation numbers, say $|N_{\mathbf{k},\alpha}\rangle$ (this is shorthand for $|N_{\mathbf{k}_1,\alpha_1}, N_{\mathbf{k}_2,\alpha_2}, \text{...to infinity}\rangle$). The states are products because the atomic states are independent of those of the radiation. The chosen final state of the atomic system is $|Y\rangle$, so the final state is $|Y\rangle|N'_{\mathbf{k},\alpha}\rangle$.

The matrix element of H' between an initial state and the final state is therefore

$$\langle Y|\langle N'_{\mathbf{k},\alpha}|\; H' \;|X\rangle|N_{\mathbf{k},\alpha}\rangle$$

The radiation factor $\langle N'_{\mathbf{k},\alpha}|\mathbf{A}|N_{\mathbf{k},\alpha}\rangle$ can be simplified by using a little physical intuition (although the maths of Fermi's Golden rule will ensure we get it right even without, but at the expense of a page of calculation). First, as we derived in section 1.4,

$$\mathbf{A} = \sum_{\mathbf{k},\alpha} \sqrt{\frac{\hbar}{2\epsilon_0\omega V}}\, \mathbf{e}_{\mathbf{k},\alpha} \left(a_{\mathbf{k},\alpha}(t)e^{i\mathbf{k}.\mathbf{r}} \;+\; a^\dagger_{\mathbf{k},\alpha}(t)e^{-i\mathbf{k}.\mathbf{r}}\right)$$

and we can see that $\mathbf{A}$ as an operator comprises a linear combination of $a_{\mathbf{k},\alpha}$ and $a^\dagger_{\mathbf{k},\alpha}$ operators.

---

[3] Most easily seen in wave function representation: $\nabla \,.\, \mathbf{A}\psi = (\nabla \,.\, \mathbf{A})\,\psi + \mathbf{A} \,.\, \nabla\psi = \mathbf{A} \,.\, \nabla\psi$.

If we choose atomic state Y so that $E_Y > E_X$, then to conserve energy (enforced by Fermi's Golden rule), since the atomic system gains energy the radiation must lose it. We need only consider states for which $\hbar|\mathbf{k}|c = E_Y - E_X$, and only annihilation (not creation) operators. There will be a huge number of possible final states of the radiation field (*i.e.* of the atomic system + radiation field), so that requirement in Fermi's Golden rule is met, and we need only do the calculation for one of them. The radiation matrix element therefore contains only one of the terms in the sum above,

$$\langle N'_{\mathbf{k},\alpha}|\sqrt{\frac{\hbar}{2\epsilon_0\omega V}}\,\mathbf{e}_{\mathbf{k},\alpha}e^{i\mathbf{k}.\mathbf{r}}a_{\mathbf{k},\alpha}|N_{\mathbf{k},\alpha}\rangle = \sqrt{\frac{\hbar}{2\epsilon_0\omega V}}\,\mathbf{e}_{\mathbf{k},\alpha}e^{i\mathbf{k}.\mathbf{r}}\langle N'_{\mathbf{k},\alpha}|a_{\mathbf{k},\alpha}|N_{\mathbf{k},\alpha}\rangle.$$

The matrix element has non-zero value only when the set of occupation numbers N' is identical to the set N, except that $N'_{\mathbf{k},\alpha} = N_{\mathbf{k},\alpha} - 1$ for the particular term $(\mathbf{k}, \alpha)$ that we chose. This is because each of the oscillators is independent, and the states are orthonormal, *i.e.*, in the bra-ket they produce a factor 1 if the Ns are the same, zero otherwise; and $a_{\mathbf{k},\alpha}|N_{\mathbf{k},\alpha}\rangle = \sqrt{N_{\mathbf{k},\alpha}}|N_{\mathbf{k},\alpha} - 1\rangle$. Thus the radiation matrix element becomes

$$\sqrt{\frac{\hbar}{2\epsilon_0\omega V}}\,\mathbf{e}_{\mathbf{k},\alpha}e^{i\mathbf{k}.\mathbf{r}}\sqrt{N_{\mathbf{k},\alpha}}.$$

We can now write down the matrix element for the absorption of a photon by an atomic system :

$$H'_{fi} = \frac{q}{m}\sqrt{\frac{\hbar}{2\epsilon_0\omega V}}\sqrt{N_{\mathbf{k},\alpha}}\,M_{YX}(\mathbf{k},\alpha), \qquad \text{where } M_{YX}(\mathbf{k},\alpha) = \mathbf{e}_{\mathbf{k},\alpha}\,.\,\int \psi_Y^* e^{i\mathbf{k}.\mathbf{r}}\mathbf{p}\psi_X\,dV.$$

Emission : The corresponding term, if Y had lower energy than X and a photon were emitted, would be the nearly the same : $\sqrt{N_{\mathbf{k},\alpha}}$ would become $\sqrt{N_{\mathbf{k},\alpha} + 1}$ and $e^{i\mathbf{k}.\mathbf{r}}$ would become $e^{-i\mathbf{k}.\mathbf{r}}$.

Density of final states : If this matrix element is put into Fermi's Golden Rule formula, the density of final states is that of the photons times that of the atomic final states. The number of photon states near the one specified is given by

$$dN^{\text{photon}} = (dN/dE)\,dE = \frac{d^3\mathbf{k}}{(2\pi/L)^3} = (2\pi)^{-3}Vk^2dk\,d\Omega.$$

Note that $dN/dE \propto d\Omega$ so therefore is 'w' – write it as $w(\Omega)_{abs}\,d\Omega$. Since $E = \hbar kc$, the density of states (per energy) $(dN/dE) = (2\pi)^{-3}Vk^2\,d\Omega\,/\hbar c$.

This then gives (putting $q = -e$ for the electron and recalling that $kc = \omega$) the transition rate for absorption

$$w(\Omega)_{abs}d\Omega = \frac{\omega d\Omega}{2\pi\hbar c^3 m^2}\,N_{\mathbf{k},\alpha}\,\frac{e^2}{4\pi\epsilon_0}\,|M_{YX}(\mathbf{k},\alpha)|^2$$

This is the probability per second that a photon of polarization α*, frequency* $\omega = |\mathbf{k}|c$ *and direction within* $\mathrm{d}\Omega$ *of* $\mathbf{k}$ *will be absorbed, and the term got by replacing* $N_{\mathbf{k},\alpha}$ *by* $N_{\mathbf{k},\alpha} + 1$ *is the corresponding emission probability* $w(\Omega)_{em}d\Omega$.

3.4 Black body radiation

From the above the emission transition rates for photons in $d\Omega$ about $\mathbf{k}$ with polarization α

$$w(\Omega)_{em}/w(\Omega)_{abs} = (N_{\mathbf{k},\alpha} + 1)/N_{\mathbf{k},\alpha}.$$

In thermodynamic equilibrium in a two state atomic system, (i) $n_u = n_l e^{-x}$, with $x = (E_u - E_l)/kT = \hbar\omega/kT$, and (ii) $n_l w_{abs} = n_u w_{em}$, *i.e.* number going up equals number going down per second. Hence the equilibrium occupancy is

$$N_{\mathbf{k},\alpha} = 1/(e^x - 1).$$

The density of photon states (in the box again!) in directions within $d\Omega$ is $2V\ d\Omega\ \nu^2\ d\nu/c^3$, using $k = 2\pi\nu/c$. The factor 2 is because we are counting both polarization states, which have the same k and therefore the same energy.

The energy density of radiation in thermodynamic equilibrium (it's isotropic so we can integrate over all angles by setting $d\Omega \to 4\pi$) is therefore

$$u(\nu) = \frac{8\pi\nu^2}{c^3}\frac{h\nu}{e^{h\nu/kT} - 1}$$

and the *specific intensity* of black body radiation, $B_\nu(T)$, is $c/4\pi$ times this, i.e.

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

(Specific intensity is discussed below).

3.5 Electric dipole transitions

3.5.1 The dipole transition rate

The dipole approximation consists of putting $e^{i\mathbf{k}.\mathbf{r}} = 1$ in the expression for $M_{YX}(\mathbf{k}, \alpha)$, which is the same as saying that the perturbing 'potential' doesn't vary over the extent of the atomic system (i.e. no tidal or quadrupole and higher effects) and is justified if the wavelength of the radiation is much greater than the atomic dimensions.

One further simplification is required: in the non-relativistic limit $\mathbf{p} = m\mathbf{v}$. The operator $\mathbf{v}$ is related to operator $\mathbf{r}$ by requiring that the matrix elements of $\mathbf{v}$ are the time derivatives of the matrix elements of $\mathbf{r}$:

$$\mathbf{v}_{YX} = \frac{\mathrm{d}}{\mathrm{d}t}\int \psi_Y^* \mathbf{r} \psi_X\ \mathrm{d}V = \int \frac{\partial \psi_Y^*}{\partial t}\mathbf{r}\psi_X\ \mathrm{d}V + \int \psi_Y^* \mathbf{r}\frac{\partial \psi_X}{\partial t}\ \mathrm{d}V.$$

Using Schrödinger's equation to replace the partial derivatives with the eigenstate energies we get

$$\mathbf{v}_{YX} = (1/i\hbar)(-E_Y \mathbf{r}_{YX} + E_X \mathbf{r}_{YX}) = i\omega \mathbf{r}_{YX}$$

Thus we can replace $\mathbf{p}$ by $im\omega\mathbf{r}$ in the matrix element, which gives an extra factor ω^2 in the rate.

In this approximation therefore,

$$w(\Omega)_{abs}d\Omega = \frac{\omega^3 d\Omega}{2\pi\hbar c^3}\, N_{\mathbf{k},\alpha}\, \frac{e^2}{4\pi\epsilon_0}\, \left|\mathbf{r}_{YX}.\mathbf{e}_{\mathbf{k},\alpha}\right|^2, \qquad \text{where } \mathbf{r}_{YX} = \int \psi_Y \mathbf{r} \psi_X \, dV$$

The corresponding value for emission is identical with the exception that $N_{\mathbf{k},\alpha}$ is replaced by $N_{\mathbf{k},\alpha} + 1$.

Consider the *spontaneous* emission rate (*i.e.* reduce $N_{\mathbf{k},\alpha} + 1$ to 1). Sum over the polarization directions α and integrate over all emission directions[4]. We get the rate of spontaneous emission of a photon of any polarization in any direction:

$$w_{spon} = \frac{4\omega^3}{3\hbar c^3}\, \frac{e^2}{4\pi\epsilon_0}\, \left|\mathbf{r}_{YX}\right|^2.$$

This is the quantum mechanical counterpart (when multiplied by $\hbar\omega$) of the Larmor formula.

If the radiation is isotropic, *i.e.* $N_{\mathbf{k},\alpha}$ is the same for states with wave vectors of magnitude k in every direction, or if we are averaging over randomly oriented atomic systems, we have as corresponding rates

$$w_{abs} = N_{\mathbf{k},\alpha} w_{spon}, \qquad w_{em} = (N_{\mathbf{k},\alpha} + 1) w_{spon}$$

$$w_{em} = w_{abs} + w_{spon}.$$

3.5.2 Dipole selection rules

w depends on the matrix element $\int \psi_Y^* \mathbf{r} \psi_X \, dV$. Typically, these single particle wave functions ψ can be expanded – as is the case for the hydrogen atom – in products of spherical harmonics $Y(\theta, \phi)$ (for a non-symmetric atomic system sums of several such products may be necessary). The space integral then factors into a radial integral, and an angle integral.

As an example, the z-component of $\mathbf{r}$ is $r\cos\theta$. The integration over $dV = r^2 dr\, d\Omega$ is sepaerable and the resulting (θ, ϕ) integral in the matrix element is

$$\int Y^*_{l_Y, m_Y}(\theta, \phi)\ \cos\theta\ Y_{l_X, m_X}(\theta, \phi)\ d\Omega.$$

This integral is only non-zero if $l_Y = l_X \pm 1$ (because of the orthogonality properties of the $Y(\theta, \phi)$, and in general if state Y is of opposite parity to state X). The selection rules are, for reference:

Electric dipole : $J_Y = J_X - 1, J_X, J_X + 1$ (J is total angular momentum quantum number)
parity change
In the case of LS coupling, also
no spin change
same rules for orbital AM L as for total AM J.

---

[4] $|\mathbf{e}_{\mathbf{k},1}.\mathbf{r}|^2 + |\mathbf{e}_{\mathbf{k},2}.\mathbf{r}|^2 = |\mathbf{r}|^2 - |(\mathbf{k}/k).\mathbf{r}|^2 = |\mathbf{r}|^2 \sin^2\theta$ where θ is the angle between $\mathbf{k}$ and $\mathbf{r}$, since $\mathbf{e}_{\mathbf{k},1}, \mathbf{e}_{\mathbf{k},2}$ and $\mathbf{k}$ are all mutually perpendicular. Integrate $\sin^2\theta$ over all directions to get $8\pi/3$. This leads to the result.

3.6 Electric quadrupole and magnetic dipole transitions

The above selection rules often forbid astronomically interesting transitions, and we so have to improve on the dipole approximation.

3.6.1 Transition rates

The electric dipole approximation was $e^{i\mathbf{k}.\mathbf{r}}\mathbf{v} \simeq \mathbf{v} \rightarrow i\omega\mathbf{r}$, which was justified by the smallness of $\mathbf{k}.\mathbf{r}$ when $\mathbf{r}$ ranged over atomic dimensions, so that $e^{i\mathbf{k}.\mathbf{r}} \simeq 1$. But if the resulting term vanishes, we must examine the next term in the expansion of $e^{i\mathbf{k}.\mathbf{r}} = 1 + i\mathbf{k} \,.\, \mathbf{r} + ...$ This gives the matrix element (*cf* $M_{YX}(\mathbf{k}, \alpha)$ in section 3.3)

$$\mathbf{e}_{\mathbf{k},\alpha}. \int \psi_Y^* \; i\mathbf{k}.\mathbf{r}\mathbf{p}\psi_X \; dV = im\mathbf{k}. \int \psi_Y^* \; \mathbf{r}\mathbf{v}\psi_X \; dV \;.\; \mathbf{e}_{\mathbf{k},\alpha},$$

which is evidently smaller than the dipole term by the factor of order k times the size of the atomic system, so that the transition rate is less by roughly $(a_0/\lambda)^2$.

We shall write the operator as the sum of symmetric and antisymmetric parts:

$$\mathbf{r}\mathbf{v} = \frac{1}{2}(\mathbf{r}\mathbf{v} + \mathbf{v}\mathbf{r}) + \frac{1}{2}(\mathbf{r}\mathbf{v} - \mathbf{v}\mathbf{r}).$$

The first – **quadrupole** – term is $\frac{1}{2}d(\mathbf{r}\mathbf{r})/dt$, and just as for $\mathbf{v} = d(\mathbf{r})/dt$, for which the matrix element was $i\omega\mathbf{r}_{YX}$, so we find the matrix element of the first term is $\frac{1}{2}i\omega(\mathbf{r}\mathbf{r})_{YX}$. It couples the ellipsoidal (quadrupole) charge distribution to the electric field (like tides in gravity). Transitions via this process are important for some nuclear state transitions, and for ground state transitions in molecular hydrogen, for example.

In the second term put $\mathbf{v} = \mathbf{p}/m$, and the expression for this part of the matrix element becomes (remember $\mathbf{k} = k\mathbf{n} = (\omega/c)\mathbf{n}$)

$$\frac{1}{2m} i\mathbf{k}.(\mathbf{r}\mathbf{p} - \mathbf{p}\mathbf{r}) \;.\; \mathbf{e}_{\mathbf{k},\alpha} = -\frac{1}{2m} i\mathbf{k}_\wedge\mathbf{L} \;.\; \mathbf{e}_{\mathbf{k},\alpha} = \frac{i\omega}{2mc}\mathbf{L}_\wedge\mathbf{n} \;.\; \mathbf{e}_{\mathbf{k},\alpha} = \frac{i\omega}{2mc}\mathbf{L}.(\mathbf{e}_{\mathbf{k},\alpha\wedge}\mathbf{n})$$

where we've used $\mathbf{k}.(\mathbf{r}\mathbf{p} - \mathbf{p}\mathbf{r}) = -\mathbf{k}_\wedge(\mathbf{r}_\wedge\mathbf{p}) = -\mathbf{k}_\wedge\mathbf{L}$, $\mathbf{L}$ being the orbital angular momentum operator. Note that $(\mathbf{e}_{\mathbf{k},\alpha\wedge}\mathbf{n}) = \mathbf{e}_{\mathbf{k},\text{other }\alpha}$ This term looks similar to the 'component of electric dipole' operator derived previously, $i\omega\mathbf{r} \,.\, \mathbf{e}_{\mathbf{k},\alpha}$, with operator $e\mathbf{r}$ (magnitude of order ea_0) replaced by **magnetic dipole** operator $e\mathbf{L}/2mc$ (magnitude of order $e\hbar/2mc$, the Bohr magneton, times an angular momentum quantum number), and $\mathbf{e}_{\mathbf{k},\alpha}$ (chosen to specify the direction of $\mathbf{A}$ and hence $\mathbf{E}$) replaced by the polarization vector at right angles, in the $\mathbf{B}$ direction; *i.e.* this term couples the electromagnetic $\mathbf{B}$ field to the magnetic moment of the electric current associated with orbital motion in the atomic system, much like a current loop aerial (classically, the electric dipole energy is $\mathbf{E}.\mathbf{d}$ while the magnetic dipole energy is $\mathbf{B}.\boldsymbol{\mu}$).

The non-relativistic expression we used for the current density, $\mathbf{v}$, neglects electron spin angular momentum $\mathbf{s}$, and a more detailed treatment gives $\mathbf{L} \rightarrow \mathbf{L} + \mathbf{2s}$. Thus the transition rate for **the magnetic dipole** radiation is the same as for the electric dipole, but with electric dipole moment replaced by magnetic dipole moment: $e\mathbf{r} \rightarrow e(\mathbf{L} + \mathbf{2s})/2mc$.

3.6.2 Electric quadrupole and magnetic dipole selection rules

The selection rules for these types of transition – which evidently exclude each other at this level of approximation – are :

Electric quadrupole : $J_Y = (J_X \pm 2)\ \mathit{or}\ J_X\ \ (J = 0 \not\leftrightarrow\ J = 0)$
no parity change;
and for LS coupling also
no orbital AM change, no spin change.

Magnetic dipole (incl. spin) : $J_Y = J_X - 1, J_X, J_X + 1$
no parity change;
and for LS coupling: no spin change, so orbital AM change, only J may (and must) change, representing fine structure transitions in a single term.

4 Radiative Transfer

This section
(a) Introduces specific intensity as a full classical description of radiation, and its properties

(b) relates the emission coefficient and opacity – required to derive the specific intensity – to the transition rate derived previously

(c) describes line profiles

4.1 Specific intensity and the equation of radiative transfer

Specific Intensity is the nearly complete classical description of the radiation field (it misses polarization, which requires another three closely related distributions – the other Stokes parameters, and it doesn't describe the fluctuation spectrum). It is the "energy per everything" of the radiation field at a given point and in a given direction.

In order to measure it, one therefore requires a device that:
1. presents a given area
2. chooses a particular direction and defines a solid angle in that direction
3. limits the radiation accepted to a given frequency range
4. allows radiation to pass for a given time
5. measures the energy (or counts the photons) in such a measurement.

Such a device could be called a specific intensitometer. It is in fact an astronomical telescope (of given aperture, equipped with sky aperture in the focal plane, filter, shutter and photometer).

Specific intensity and photon occupancy I_ν is related to $N_{\mathbf{k},\alpha}$ in a straightforward manner. The energy flux is $I_\nu d\Omega\, d\nu$ and this is equal to the number of photon states *per unit volume* in $d^3\mathbf{k}$ ($= k^2dkd\Omega/(2\pi)^3$) times $h\nu$ times c times occupancy of these photon states (NB $k = 2\pi\nu/c$), *i.e.*

$$I_\nu d\Omega\, d\nu = (k^2dkd\Omega/(2\pi)^3) \times h\nu c \times (N_{\mathbf{k},1} + N_{\mathbf{k},2}).$$

Hence (supposing that the radiation is unpolarized: $N_{\mathbf{k},1} = N_{\mathbf{k},2}$)

$$I_\nu = (2h\nu^3/c^2)\, N_{\mathbf{k},\alpha}$$

where the factor 2 is because both polarizations are assumed equally populated.

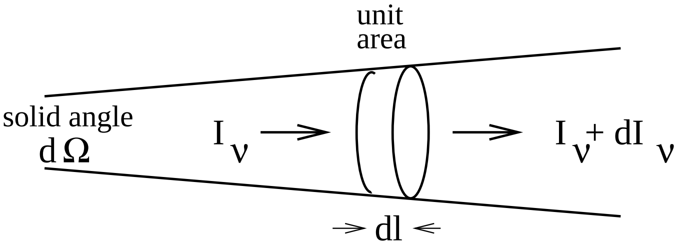
Radiative transfer Consider for concreteness the transfer of radiation due to a particular bound/bound transition ($l \to u$, say, from *lower* to *upper*) of randomly oriented atoms. The number of absorbers in the elemental volume shown below is $n_l dl$ and the number of emitters is $n_u dl$. The photon occupancy is $N_{\mathbf{k},\alpha}$. The number of 'up' transitions per second due to photons in the beam $d\Omega$ is

$$n_l dl\; w_{abs}(d\Omega/4\pi) = n_l dl\; w_{spon} N_{\mathbf{k},\alpha}(d\Omega/4\pi)$$

while the number per second of 'down' transitions is

$$n_u dl\, w_{spon}(N_{\mathbf{k},\alpha} + 1)(d\Omega/4\pi).$$

$(I_\nu d\Omega/h\nu)$ photons arrive at the volume per second, $(n_l - n_u)\, dl\, w_{spon} N_{\mathbf{k},\alpha}(d\Omega/4\pi)$ are removed and $n_u\, dl\, w_{spon}(d\Omega/4\pi)$ are added – all in the beam $d\Omega$.



Before continuing, we must allow for the fact that absorbtion can occur at a range of frequencies near ν_{YX}, due to *e.g.* Doppler shifts from thermal motion. The total absorbtion rate is now spread over frequency proportional to ϕ_ν and $\int \phi_\nu d\nu = 1$.

Then the change in flux at a given frequency as the beam passes through the volume can be written

$$dI_\nu d\Omega/h\nu = [-(n_l - n_u) dl\, N_{\mathbf{k},\alpha} w_{spon}(d\Omega/4\pi) + n_u dl\, w_{spon}(d\Omega/4\pi)]\, \phi_\nu$$

or by substituting $I_\nu = (2h\nu^3/c^2)\, N_{\mathbf{k},\alpha}$,

$$\frac{dI_\nu}{dl} = -(n_l - n_u)\frac{w_{spon}\lambda^2}{8\pi} I_\nu \phi_\nu + n_u \frac{w_{spon} h\nu}{4\pi}\phi_\nu.$$

This can be written

$$\frac{dI_\nu}{dl} = -\kappa_\nu I_\nu + \mathcal{E}_\nu$$

which is known as the **equation of radiative transfer**. Here

$$\kappa_\nu = (n_l - n_u) w_{spon} \lambda^2 \phi_\nu / 8\pi$$

is the *opacity* (units m^{-1}) and

$$\mathcal{E}_\nu = n_u w_{spon} h\nu \phi_\nu / 4\pi$$

(units are W m^{-3} sr^{-1}) is the *emissivity*. Each of those is calculable from local conditions in the material.

The opacity can also be represented in terms of a cross section σ_ν for absorption per lower state

$$\kappa_\nu = n_l\, \sigma_\nu (1 - n_u/n_l).$$

This result follows from the definitions of opacity and cross section.

Correction for stimulated emission — The factor $(n_l - n_u)$ introduced above is very important as the correction for stimulated emission. As we saw in section 3.3 stimulated absorption and emission between two states proceeds at equal rates. The *net* rate of absorption per unit volume is is $w_{abs}(n_l - n_u) = w_{abs} * n_l(1 - n_u/n_l)$.

The factor $(1 - n_u/n_l)$ can have a dominating effect. For example, if the states are Boltzmann populated (thermodynamic equilibrium), and if (Rayleigh-Jeans limit) $h\nu << kT$, the factor becomes $h\nu/kT$, a very small quantity. The states are so nearly equally populated that almost as much stimulated emission as absorption occurs, and the material appears almost transparent!

In certain cases, a non-thermal effect can tip the balance, and $n_u > n_l$ in which case the opacity is negative and the specific intensity *grows*! This is maser emission (Microwave – it usually is – Amplification by Stimulated Emission of Radiation).

Modification for multiple states per level. Often one requires the emissivity and opacity for transitions between two *energy levels* L and U – each level containing several (g_L and g_U) states u and l respectively. The rate w^{ul}_{spon} is found for each possible pair of states in upper and lower levels, and the individual opacities and emissivities are added to give the totals denoted by the superscripts LU and UL:

$$\kappa_\nu^{LU} = \left(\sum_l n_l \sum_u w^{ul}_{spon} - \sum_u n_u \sum_l w^{ul}_{spon} \right) (\lambda^2/8\pi)\, \phi_\nu$$

$$\mathcal{E}_\nu{}^{UL} = \sum_u n_u \sum_l w^{ul}_{spon}\, (h\nu/4\pi)\, \phi_\nu$$

If the individual states in a level are equally populated (*e.g.* by collisions) then we can use $n_u = \mathcal{N}_U/g_U$ and $n_l = \mathcal{N}_L/g_L$, whence

$$\kappa_\nu^{LU} = (\mathcal{N}_L/g_L - \mathcal{N}_U/g_U)\; S\; (\lambda^2/8\pi)\, \phi_\nu$$

$$\mathcal{E}_\nu{}^{UL} = (\mathcal{N}_U/g_U)\; S\; (h\nu/4\pi)\, \phi_\nu$$

where $S = \sum_l \sum_u w^{ul}_{spon}$. In this case, the modification is to use the original number densities per state, and change w_{spon} to S.

4.2 Optical depth and a solution to the equation of radiative transfer

Before considering the detailed behaviour of specific intensity, note that in empty space opacity and emissivity are zero. As a consequence, dI_ν/dl is zero, therefore and *specific intensity is constant along a line of sight (or ray) in empty space*.

This constancy is the crux of observational astronomy! The source imprints on I_ν its properties, via emissivity and opacity. The the radiation travels through empty space carrying this fixed value of I_ν, which is measured on earth with a telescope. From this the properties of the source can be unravelled. We shall see how this is done.

Define **optical depth** τ_ν by $d\tau_\nu = \kappa_\nu\, dl$ to get

$$\frac{dI_\nu}{d\tau_\nu} + I_\nu = \frac{\mathcal{E}_\nu}{\kappa_\nu} = S_\nu$$

where the latter equality *defines* the **source function** S_ν.

If κ_ν, S_ν are independent of position, the equation integrates to

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

If there is no background radiation, then

$$I_\nu = S_\nu(1 - e^{-\tau_\nu})$$

and two extreme cases can be seen.

(1) $\tau_\nu << 1$ — optically thin case :
expand $\exp(-\tau_\nu)$ to two terms and get

$$I_\nu = S_\nu \tau_\nu = \mathcal{E}_\nu l$$

(2) $\tau_\nu >> 1$ — optically thick case :
$\exp(-\tau_\nu) \to 0$ and

$$I_\nu = S_\nu = \mathcal{E}_\nu / \kappa_\nu = \mathcal{E}_\nu \times \text{(mean free path)}$$

4.3 Line profiles

The transitions $u \to l$ appear to involve exact energy differences. However, the perturbation causing the transition spoils this. The uncertainty principle can be stated in terms of energy and time as

$$\Delta E \, t \geq \hbar$$

where t is the time available to measure the energy, and ΔE is the uncertainty in the measurement. Since $t \simeq 1/w_{spon}$, the energy uncertainty is $\hbar w_{spon}$, giving a finite width to the line resulting from the transition equal in frequency to w_{spon}. This is very narrow, but becomes significant as the 'damping wings' in very strong lines. The profile of such a line is the *Lorentz* profile, proportional to $1/[(w_{spon}/2)^2 + (\nu - \nu_{ul})^2]$. An effect normally much greater, at least in emission lines, is Doppler, or thermal, broadening. Each atomic system in a source in which kinetic energy is in thermal equilibrium at temperature T is moving with a velocity associated with the energy distribution. In one dimension (along the line of sight) the velocities are distributed as $\exp(-E/kT) = \exp(-v^2/(2kT/m))$. The $1/e$ velocity of this distribution is $\sqrt{2}v_{iso}$ where $v_{iso} = \sqrt{kT/m}$. The frequency of the line is altered by $\delta\nu/\nu = v/c$, and the line has a Gaussian profile with $1/e$ frequency width $\sqrt{2}v_{iso}/\lambda$.

To express this frequency spread we modify the emissivity and opacity coefficients by multiplying each by a normalized profile (or frequency probability distribution). Put $\mathcal{E}_\nu = \mathcal{E}_{ul}\phi_\nu$ and $\kappa_\nu = \kappa_{lu}\tilde{\phi}_\nu$ with ϕ_ν, $\tilde{\phi}_\nu$ each satisfying $\int \phi_\nu d\nu = 1$. Then each, integrated over the frequency range, gives the value consistent with the total transition rate.

In the example of Doppler broadening the expression for ϕ_ν is

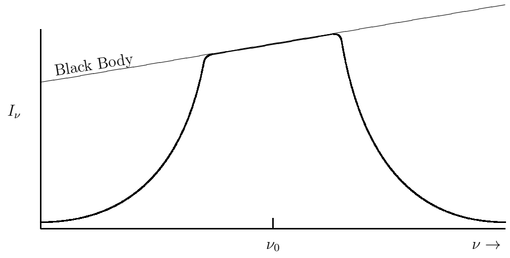
$$\phi_\nu = \frac{1}{\sqrt{\pi}b} e^{(\nu-\nu_0)^2/b^2}, \quad \text{where } b = \sqrt{2}v_{iso}/\lambda.$$

b is called the Doppler broadening parameter. Note that the peak value of ϕ_ν is $1/(\sqrt{\pi}b)$, expressing the approximate truth that since area equals unity by normalization, height = 1/width.

In the frequently met case that the emissivity and the opacity have the same origin, the source function is independent of the profile shape since it is the ratio $\mathcal{E}_\nu/\kappa_\nu$. This will occur for instance if the emission and absorption processes are dominated by the same Doppler broadening, or generally are part of the same process.

For example, consider the emission line due to the dipole transition from rotational state $J = 1$ to $J = 0$ of the CO molecule. Clouds with small optical depths at the centre of the line will be observed to have the Doppler broadening profile. However, when the optical depth at the centre of the line approaches and exceeds one, the optically thick solution applies and the specific intensity there cannot exceed the (at most slowly varying in frequency) source function. As the column density of the cloud is increased, more of the line reaches this limit, and the specific intensity of the line doesn't become any higher and flattens at the value of the source function.

If the levels in the molecule are determined entirely by collisions (from molecular hydrogen) in a gas which is thermalized, then the level population will be given by the Boltzmann distribution, and the ratio $\mathcal{E}_\nu/\kappa_\nu = S_\nu = B_\nu(T)$. The profile will be a Gaussian profile chopped off at the black body specific intensity at the kinetic temperature of the cloud (see figure).

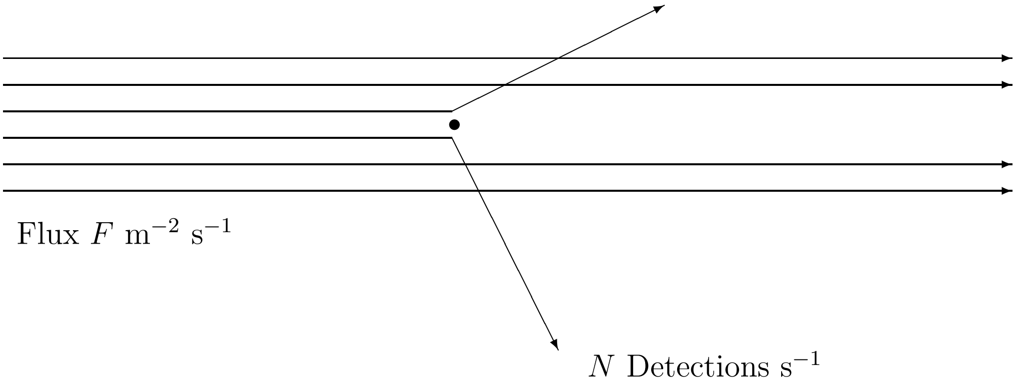


5 Collisions

This section
(a) describes cross section

(b) derives reciprocity between collisions causing upward and downward transitions

(c) defines critical density for a pair of levels



5.1 Cross section and collision rate

The cross section for a process is defined above. The rate of detection of the process occurring, N s^{-1}, is proportional to the flux F. The constant of proportionality, σ, has dimensions m^2 and is called the **cross-section**.

In Maxwellian distributions, collisions between species '1' and '2' proceed at a rate

$$n_1 n_2 \langle \sigma v \rangle \quad \mathrm{m^{-3}\ s^{-1}}$$

where n_1, n_2 are particle number densities (m^{-3}) and $\langle \sigma v \rangle$ is the Maxwell average of the product σv. (see appendix).

This also leads gives the definition of **mean free path** (the mean distance travelled by particle '2' before "hitting" a particle '1' — a "hit" is specified as an elastic collision, an excitation, a reaction or whatever, and the cross section σ for that process can be calculated or measured)

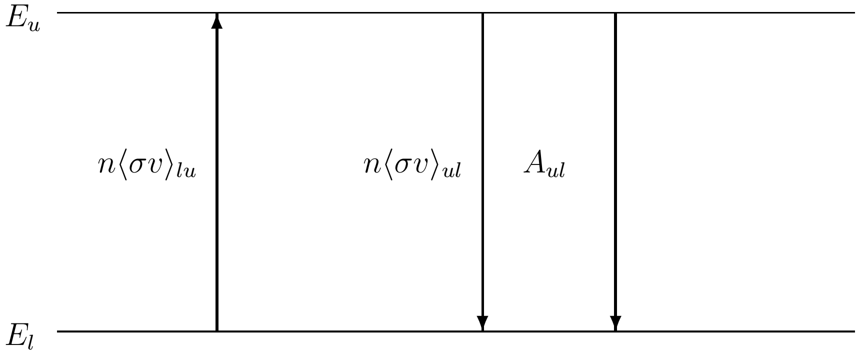
$$\mathrm{MFP} = \frac{1}{n_1 \sigma}.$$

The **collision time** t_{coll} is the mean time for such a hit, and is evidently

$$t_{coll} = \frac{1}{n_1 \langle \sigma v \rangle}.$$

5.2 Critical density

A useful concept in interpreting level populations is **critical density**, defined in a two-level system:



The radiation field is presumed to be negligible, and collisions cause up and down transitions. The particles colliding with the 2-level system have number density n. In statistical equilibrium (rate of ups equals rate of downs in an ensemble of such systems)

$$n_l\, n\langle\sigma v\rangle_{lu} = n_u(n\langle\sigma v\rangle_{ul} + g_l w_{spon}).$$

In the case where n is extremely large, we (a) reach the thermodynamic equilibrium rate, and (b)

$$\frac{n_u}{n_l} = \frac{\langle\sigma v\rangle_{lu}}{\langle\sigma v\rangle_{ul}}$$

Hence, applying the Boltzmann distribution $n_u/g_u = \exp(-\Delta E/kT)$, we get a very important **reciprocity**

$$\langle\sigma v\rangle_{lu} = (g_u/g_l)\langle\sigma v\rangle_{ul} e^{-\Delta E/kT}.$$

Now rearranging the original equation for n_u and n_l, using the reciprocity relation between σ_{lu} and σ_{ul}, and **defining**

$$n_{\text{crit}} = g_l w_{spon}/\langle\sigma v\rangle_{ul},$$

we get

$$\frac{n_u}{n_l} = \frac{(g_u/g_l)e^{-x}}{1 + n_{\text{crit}}/n}, \qquad x = \frac{E_u - E_l}{kT}.$$

There are two limiting cases.

1. When $n \gg n_{\text{crit}}$ the denominator is nearly unity, and the levels are Boltzmann populated. The emissivity is then independent of collider density n and is given by

$$\mathcal{E}_\nu = \frac{h\nu}{4\pi} n_u g_l w_{spon} \phi_\nu$$

where

$$n_u = n_{\rm total} g_u e^{-E_u/kT}/Q(T)$$

and $Q(T) = \sum g_n e^{-E_n/kT}$ is the *partition function.* In the thermodynamic equilibrium case, other things being equal, the emissivity is proportional to density.

2. When $n \ll n_{\rm crit}$ the downward collision rate is dwarfed by radiative transitions, and rate of collisional transitions upward equals rate of radiative transitions downward:

$$\frac{n_u}{n_l} = \left(\frac{n}{n_{\rm crit}}\right)\left(\frac{n_u}{n_l}\right)_{Boltzmann}.$$

The emissivity is independent of A_{ul} but depends on the collision rate $\langle\sigma v\rangle_{ul}$ and is

$$\mathcal{E}_\nu = \frac{h\nu}{4\pi} n n_l \frac{g_u}{g_l} \langle\sigma v\rangle_{ul}\, e^{-x}\phi_\nu,$$

proportional to density squared.