Astronomy 271B Lecture Notes, Spring 2016
©Edward L. Wright, 2016

This is a course about the Fundamentals of Astrophysics: in particular electromagnetic radiation. As astronomers we are limited to observing what happens in the Universe – we are not able to reach out and touch something. Our observations are almost entirely done using electromagnetic radiation. There are a very few exceptions: meteorites, charged cosmic rays and neutrinos. Now that advanced LIGO has seen sources, we must add gravitational waves to the list of observed radiation. But almost everything we know about the Universe comes from light and non-visible EM radiation that reaches the Solar system.

I will try to get these LaTeX'd notes out in a timely fashion. They will not include all the diagrams that I draw on the board, so you still need to take notes during class, but they should help you remember what went on and also get the equations right.

A NOTE ABOUT UNITS: In this course we will use electrostatic cgs units because most astrophysicists use them. Those of you who studied E&M in MKS units will suffer the slings and arrows of outrageous unit conversion, but can probably fix a toaster. Just remember that 1 picofarad is about 1 cm and that 1 gauss (that's 10^{-4} Tesla) is the same as 1 statvolt/cm which is 300 V/cm. All the best E&M books use cgs units: Purcell's E&M in the Berkeley Physics course, and Jackson in early editions.

The electrostatic force between two charges is $F = k_1 q_1 q_2 / r^2$. The magnetic force between two wires separated by distance r and carrying currents I_1 and I_2 per unit length λ is $dF/d\lambda = 2k_2 I_1 I_2 / r$. Note that the speed of light is always given by $c^2 = k_1/k_2$, the ratio of electrostatic to magnetic force strengths. For MKS units $k_2 = 10^{-7}$ and thus $k_1 = 10^{-7} c^2$, while for electrostatic cgs units $k_1 = 1$ and $k_2 = c^{-2}$. In the electrostatic cgs system $\epsilon_\circ = 1/4\pi$ and $\mu_\circ = 4\pi/c^2$. The electrostatic cgs unit for charge is the electrostatic unit or esu, and 1 Coulomb is 2.99792458×10^9 esu. That ratio is exactly $c/10$ cm/sec, and since the potential energy is $U = qV$, the unit of electric potential in the cgs system is 299.792458 Volts which is known as a statvolt. The unit of the electric field is statvolt/cm which is exactly the same as the unit for the magnetic field which is the gauss. This fortunate coincidence arises because the Lorentz force law, which is

$$F = q(\vec{E} + \vec{v} \times \vec{B}) \tag{1}$$

in MKS units, is changed to

$$F = q(\vec{E} + \vec{\beta} \times \vec{B}) \qquad \text{with} \qquad \beta = v/c \tag{2}$$

in cgs units, so $\vec{E}$ and $\vec{B}$ have the same units. Since we will be combining $\vec{E}$ and $\vec{B}$ into a single electromagnetic field tensor $F_{\mu\nu}$ it is very useful that $\vec{E}$ and $\vec{B}$ have the same units. The unit for capacitance is esu/statvolt which is the centimeter so 1 Farad is $(2.99792458)^2 \times 10^{11}$ cm.

Hence 1 cm is 1.1 pF. Since the RC time constant is still in seconds, $(2.99792458)^2 \times 10^{11}$ Ohms is 1 sec/cm.

Since the resistance is given by $R = \rho L/A$ where ρ is the resistivity, L is the length and A is the area, the units for the resistivity ρ are *seconds*. This is approximately the order of magnitude for the length of time needed to discharge a spherical charge separation by electrical conduction. For elongated objects like needles the discharge time is much longer than the resistivity. To return to the toaster, the resistivity of nichrome is 110×10^{-8} Ohm-meters in MKS, or 110 $\mu\Omega$-cm in the mixed units normally used. Since 1 Ohm is 1.1×10^{-12} sec/cm, the resistivity of nichrome is 1.2×10^{-16} sec. The discharge time is significant for wavelengths $\lambda \approx 2\pi\rho c \approx 0.2$ μm for solid spheres and much longer wavelengths for needle-shaped or porous dust grains.

1. Dust Absorption in the ISM

not in R&L

Typically in the ISM 1% of the mass is in dust. The heavy elements (C to U) make up about 2% of the mass of the gas currently, but CO forms until one runs out of C or O, and CO stays gaseous down to very low temperatures.

The "typical" dust grain has a radius a of about 0.1 μm $= 10^{-5}$ cm. This value is probably more set by $\lambda/2\pi$ where λ is the wavelength of the light used to study the dust. Because the dust grains are small, there is more absorption and scattering for short wavelengths of light than for longer wavelengths. The sum of absorption and scattering is known as *extinction*. The extinction *vs.* wavelength is denoted be A_λ. By comparing the color of stars to the colors expected for the temperatures implies by their spectral types, the color excess E can be determined. Most commonly one talks about $E(B-V) = A_B - A_V$. If the color excess, or *reddening* is known, one can determine the total extinction A_V through the ratio of total to selective extinction, $R = A_V/E(B-V) \approx 3.1$. Note that if $A_\lambda \propto \nu$, then R would be $R = 0.55^{-1}/(0.44^{-1} - 0.55^{-1}) = 4$ so the extinction curve is less steep than $1/\lambda$ in the visible, but in the near infrared the extinction declines more rapidly than $1/\lambda$ with increasing wavelength.

The typical dust grain has a mass $m_g = 4\pi a^3 \rho/3 \approx 10^{-14}$ gm which corresponds to about 10^{10} H atoms. Thus the number density of grains is about 10^{-12} of the number density of hydrogen. In the typical ISM density $n_H = 1$ cm^{-3}, the distance between dust grains is about 10^4 cm. Thus the ISM is very dilute and the dust grains act independently.

The dielectric constant of the ISM is given by $\epsilon = 1 + 4\pi\chi$ where $\chi = P/E$ and P is the

dipole density. P is given by $n_g\alpha_g E$ where α_g is the *polarizability* of the dust grain. Thus

$$\epsilon = 1 + 4\pi n_g \alpha_g \tag{3}$$

But what is the polarizability of a dust grain? It has the dimensions of $(\text{length})^3$ in cgs units. For a perfectly conducting sphere, assume that there is a surface charge density $\sigma = \sigma_\circ \cos\theta$, where θ is the angle between the incident E field and the point on the surface. This charge density makes an E field at the center of the sphere given by

$$\vec{E}_{ind} = \hat{z} \int \sigma_\circ \cos^2 \theta d\Omega = \frac{4\pi\sigma_\circ}{3} \tag{4}$$

Thus, $\sigma_\circ = \frac{3}{4\pi}E_{ext}$ for the induced field to cancel the external field in the center of the sphere. This surface charge density has a dipole moment given by

$$d = \int a \cos\theta (\sigma_\circ \cos\theta a^2) d\Omega = \frac{4\pi}{3} a^3 \sigma_\circ = a^3 E_{ext} \tag{5}$$

so the polarizability of a conducting sphere is

$$\alpha_g = a^3 \tag{6}$$

For a dielectric sphere, the dipole density P inside the sphere is constant, as is the total electric field E. The surface charge density is $\sigma = P \cos\theta$. Thus the induced field is

$$\vec{E}_{ind} = -\frac{4\pi}{3}\vec{P} \tag{7}$$

But the dipole density is given by

$$\vec{P} = \frac{(\epsilon_g - 1)\vec{E}}{4\pi} \tag{8}$$

and the interior electric field is given by

$$\vec{E} = \vec{E}_{ext} + \vec{E}_{ind} \tag{9}$$

Since all the electric fields are along the z axis and the direction of the induced field is opposite to the applied field, we can find the magnitude of the induced field by solving

$$\frac{3}{4\pi}E_{ind} = \frac{(\epsilon_g - 1)}{4\pi}(E_{ext} - E_{ind}) \tag{10}$$

so

$$E_{ind}\left(\frac{3}{4\pi} + \frac{(\epsilon_g - 1)}{4\pi}\right) = \frac{(\epsilon_g - 1)}{4\pi}E_{ext} \tag{11}$$

and finally

$$E_{ind} = \frac{\epsilon_g - 1}{\epsilon_g + 2}E_{ext} \tag{12}$$

3

The dipole moment is obviously

$$d = \frac{4\pi}{3}a^3 P = a^3 E_{ind} = \frac{\epsilon_g - 1}{\epsilon_g + 2}a^3 E_{ext} \tag{13}$$

and the polarizability of a dielectric *sphere* is

$$\alpha_g = \frac{\epsilon_g - 1}{\epsilon_g + 2}a^3 \tag{14}$$

For a conductor, $\epsilon = 1 - \omega_p^2/\omega^2 \to -\infty$ so $\alpha_g = a^3$ as found earlier. These formula are derived using electrostatics ($\omega \to 0$) but are OK for $\lambda >> a$.

The imaginary part of the wavevector $k = \omega\sqrt{\epsilon}/c$ is given by

$$\Im(k) = \frac{\omega}{c}(1 + 2\pi n_g \alpha_g) \tag{15}$$

Now $\Im(k) \times D$ gives the fractional reduction in the amplitude of $|E|$, so the reduction in the power is twice this great. But this reduction in power is also given by

$$D \times n_g \times \sigma_{ext} = 2 \times \Im(k) \times D \tag{16}$$

Thus the extinction cross section of dust grain is given by

$$\sigma_{ext} = \frac{4\pi\omega\Im(\alpha_g)}{c} \tag{17}$$

For a *sphere*,

$$\sigma_{ext} = 4\pi a^3 \frac{\omega}{c}\Im\left(\frac{\epsilon_g - 1}{\epsilon_g + 2}\right) \tag{18}$$

1.1.   Scattering

The oscillating dipole moment of the dust grain emits power given by

$$P = \frac{2e^2 a^2}{3c^3} = \frac{2d^2\omega^4}{3c^3} \tag{19}$$

The Poynting flux times the scattering cross-section equals this power so

$$\sigma_{scat} = \frac{8\pi}{3}\left(\frac{\omega}{c}\right)^4 a^6 \left|\frac{\epsilon_g - 1}{\epsilon_g + 2}\right|^2 \tag{20}$$

Since the scattering goes down like ω^4, the long wavelength extinction is dominated by absorption.

4

The Kramers-Kronig relation for dust grains gives:

$$\alpha_{DC} = \frac{1}{4\pi^3} \int_0^\infty \sigma_{ext} d\lambda \tag{21}$$

Note that σ can't approach a constant greater than zero as $\lambda \to \infty$ because that would require an infinite DC polarizability. An extinction going like $1/\lambda$ is also ruler out, although only logarithmicly.

To have a large σ_{ext} at long wavelengths, one needs to have a large α_{DC}. But for spheres, α_{DC} is proportional to the volume, and we can't have a large volume without requiring too much mass in the dust grains. To get around this limit, we need to find grains with a large ratio α_{DC}/V where V is the volume.

This is possible for conducting needles. Consider a needle with length L and diameter D in an external electric field oriented parallel to the long axis of the needle. A linear charge density $\lambda(z)$ will be induced on the needle. Assume that $\lambda(z) \propto z$ with the origin at the middle of the needle. The voltage of a linear charge density is given by $V = 2\lambda \ln r$. Taking the potential drop between $r = D/2$ and $r = L/2$ and setting this equal to the external voltage Ez gives

$$\lambda(z) = \frac{Ez}{2\ln(L/D)} \tag{22}$$

so the dipole moment is

$$d = \int \lambda(z) z \, dz = \frac{E}{2\ln(L/D)} \int_{-L/2}^{L/2} z^2 dz = \frac{L^3}{24\ln(L/D)} E \tag{23}$$

Thus the polarizability is proportional to the (longest dimension)^3 instead of the volume. For these formula to work it is necessary that the induced voltage cancel the applied voltage, which will happen if the grain is made out of a conducting material. Thus elongated conducting dust grains can be very efficient long wavelength absorbers.

The absorption cross-section can be found by comparing the ohmic loss in the grain to the Poynting flux. The ohmic loss is

$$\text{Power} = \int jE \, dV = \int \sigma_e E^2 dV \tag{24}$$

where σ_e is the electrical conductivity of the material. Thus the absorption cross-section is given by

$$\sigma_{abs} = \frac{4\pi\sigma_e V}{c} \tag{25}$$

Note that in cgs units the electrical conductivity σ_e has the units sec^{-1} and is equal to the usual value in mho/cm multiplied by $2.99792458^2 \times 10^{11}$. For a typical amorphous carbon $\sigma_e = 100$ mho/cm $= 9 \times 10^{13}$ sec^{-1}, the absorption cross-section is about $30,000$ cm^2/cc or

approximately 10^4 cm^2/gm of dust and 100 cm^2/gm of total ISM. The visual extinction of 1 per kpc is 200 cm^2/gm of total ISM, so conducting needles can have absorptions as high as the visual extinction extending into the far-IR and sub-mm. For randomly oriented needles, these values should be divided by 3 since only 1 of the 3 axes is along the long axis of the needle.

This absorption cross-section assumes that current is flowing freely in the incident field, but for long wavelengths the dipole moment gets established and the net field goes to zero, and so does the absorbed power. The time constant for setting up the dipole moment can be estimated by calculating the resistance and capacitance of the needle. The resistance R is $R = L/(\pi D^2 \sigma_e/4)$ and has units of sec/cm. The capacitance can be estimated from the voltage across the needle, $V = 4\lambda(L/2)\ln(L/D)$, and the charge on one "plate" of the needle, $Q = \int_0^{L/2} \lambda(z)dz = (L/4)\lambda(L/2)$. Thus $C = Q/V = L/(16\ln(L/D))$ and has units of cm. The time constant is

$$\tau = RC = \frac{(L/D)^2}{4\pi \ln(L/D)\sigma_e} \tag{26}$$

For $L/D = 100$ and $\sigma_e = 10^{14}$ sec^{-1}, this is $\tau = 1.7 \times 10^{-12}$ sec. Thus the cutoff at $\omega\tau = 1$ occurs at $\omega = 6 \times 10^{11}$, $\nu = 9 \times 10^{10}$ Hz, or $\lambda = 3$ mm.

1.2. Fractal Dust Grains

Dust grains in the ISM are probably severely radiation damaged, so the crystal structure will be very imperfect, and the shapes are likely to be extremely diverse. Neither spheres nor needles offer much diversity, but fractal shapes produced by random aggregation offer an example of possible shapes. Computing the absorption cross-sections of odd shapes can be done using the discrete-dipole approximation (Purcell & Pennypacker, 1973, ApJ, 186, 705) which divides a grain into polarizable spheres, and computes their dipoles in the field due to both the external field and their mutual interactions.

In Wright (1987, ApJ, 320, 818) the RC network theory for the absorption cross-section of arbitrarily shaped dust grains in developed. This is a simplification of the DDA in the limit applicable to the conducting needle considered earlier. In this theory, the potential applied by the external electric field causes currents to flow, leading to V^2/R losses and absorption. A dust grain is modeled by a collection of spheres of radius r which all interact capacitively, along with resistive interactions between spheres in contact.

The applied field $\vec{E}$ generates a potential

$$V_{\circ,i} = \vec{E} \cdot \vec{r}_i \tag{27}$$

on the ith sphere. The total potential depends on the charges on all the balls, Q_i on the ith
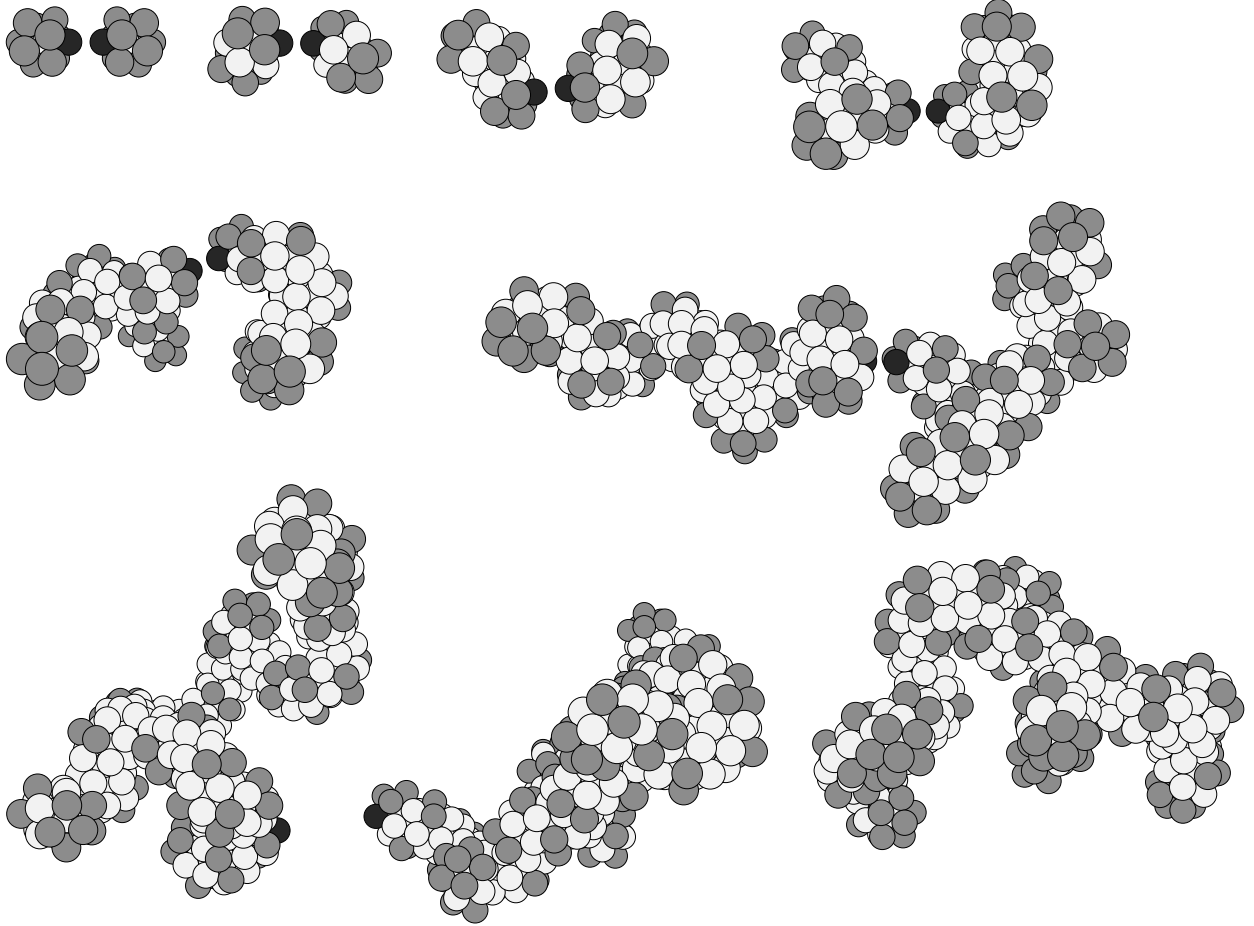
Fig. 1.— Fractal shapes produced by random aggregation processes.

sphere. One gets

$$V_i = \sum G_{ij} Q_j + V_{\circ,i} \tag{28}$$

where the Green's function G_{ij} is $1/r$ if $i = j$, or $1/r_{ij}$ if $i \neq j$. Note that G has units of inverse capacitance.

The current flow between the spheres is given by

$$\frac{dQ_i}{dt} = \sum Y_{ij} V_j \tag{29}$$

The matrix element Y_{ij} contains $1/R$ if there is a link with resistance R between the ith sphere and the jth sphere, or 0 otherwise. The diagonal element are defined as $Y_{ii} = -\sum_{i \neq j} Y_{ij}$, so if V_i is constant there is no current flow. Y has units of inverse resistance, with SI unit the siemens (used to be "mho") and cgs units of cm/sec.

If we assume a $\exp(-i\omega t)$ time dependence, we get an equation for the charges on the spheres:

$$-i\omega \mathbf{Q} = \mathbf{YGQ} + \mathbf{YV}_\circ \tag{30}$$

7

This has the solution

$$\mathbf{Q} = -(i\omega\mathbf{I} + \mathbf{YG})^{-1}\mathbf{YV}_{\circ} \tag{31}$$

The power dissipated in the resistive links can be calculated as

$$P = -\mathbf{V}^{T*}\mathbf{YV} \tag{32}$$

which is analogous to $P = V^2/R$. The minus sign comes from the definition of Y with a negative diagonal. The voltage is

$$\mathbf{V} = -i\omega\mathbf{Y}^{-1}\mathbf{Q} \tag{33}$$

which gives the dissipated power as

$$
\begin{aligned}
P &= -\omega^2\mathbf{V}_{\circ}{}^T\mathbf{Y}\left(\omega^2\mathbf{Y} + \mathbf{YGYGY}\right)^{-1}\mathbf{V}_{\circ} \\
&= -\omega^2\mathbf{V}_{\circ}{}^T\mathbf{Y}\left(\omega^2\mathbf{I} + \mathbf{GYGY}\right)^{-1}\mathbf{V}_{\circ}
\end{aligned} \tag{34}
$$

Since $G \sim 1/C$ and $Y \sim 1/R$, this formula is schematically $P = (V^2/R)[(\omega\tau)^2/(1 + (\omega\tau)^2)]$, the dissipation in a series RC circuit. But with a complicated shape there are many different time constants. The dissipated power is also given by the cross-section times the Poynting flux, giving

$$P = \sigma_a S = \sigma_a\frac{c\vec{E}^2}{4\pi} \tag{35}$$

We can calculate the absorption cross-section in various orientations of the grain with respect to the incident electric field. If the z axis is chosen to be axis of largest moment of inertia of the grain, and σ_z is the absorption when $\vec{E}$ is along the z axis, then the average cross-section is

$$\sigma = (\sigma_x + \sigma_y + \sigma_x)/3 \tag{36}$$

and the polarization in perfect spinning alignment with all the spin axes aligned along z is

$$\text{Polarization} = \frac{\sigma_x + \sigma_y - 2\sigma_z}{\sigma_x + \sigma_y + 2\sigma_z} \tag{37}$$

The absorption cross-section for fractal dust grains varies like $\sigma \propto \nu^2$ for low ν, then $\sigma \propto \nu^\beta$ for intermediate ν, and $\nu \propto$ const for high ν. For a cluster aggregation (CL) fractal construction method, $\beta \approx 0.6$. The CL process gives a fractal dimension $D \approx 1.7$. Grains constructed by diffusion limited aggregation (DLA) have fractal dimension $D \approx 2.5$ and give $\beta \approx 1$. The intermediate ν range corresponds to the range of time constants in a grain with a complicated shape.

2. Fourier Transforms

The Fourier transform is defined by R&L as

$$\hat{f}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(t)e^{i\omega t}dt \tag{38}$$

8

Other authors put the "2π" in other places so be careful when mixing formulae from different sources. We can easily find the FT of a normalized Gaussian:

$$
\begin{aligned}
g(t) &= \frac{1}{\tau\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{t^2}{\tau^2}\right) \\
\hat{g}(\omega) &= \frac{1}{\tau(2\pi)^{3/2}}\int \exp\left(-\frac{1}{2}\frac{t^2}{\tau^2}+i\omega t\right)dt \\
&= \frac{1}{\tau(2\pi)^{3/2}}\int \exp\left(-\frac{1}{2}\left[\frac{t}{\tau}-i\omega\tau\right]^2-\frac{1}{2}\omega^2\tau^2\right)dt \\
&= \frac{1}{2\pi}\exp\left(-\frac{1}{2}\omega^2\tau^2\right)
\end{aligned}
\tag{39}
$$

In the limit as $\tau \to 0$, $g(t) \to \delta(t)$ and clearly $\hat{g}(\omega) \to (2\pi)^{-1}$.

We can use this result to find the FT of an unnormalized Gaussian:

$$
\begin{aligned}
h(t) &= \tau\sqrt{2\pi}g(t) = \exp\left(-\frac{1}{2}\frac{t^2}{\tau^2}\right) \\
\hat{h}(\omega) &= \tau\sqrt{2\pi}\hat{g}(\omega) = \frac{\tau}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\omega^2\tau^2\right)
\end{aligned}
\tag{40}
$$

Note that $\hat{h}(\omega)$ is a normalized Gaussian in ω space so its integral is unity for any τ. As $\tau \to \infty$, $h(t) \to 1$ and $\hat{h}(\omega) \to \delta(\omega)$. This is a very important and useful result:

$$
\delta(x) = \frac{1}{2\pi}\int e^{ixy}dy
\tag{41}
$$

2.1.   Inverse Fourier Transform

We can now prove the correctness of the inverse Fourier transform:

$$
f(t) = \int_{-\infty}^{+\infty} \hat{f}(\omega)e^{-i\omega t}d\omega
\tag{42}
$$

We evaluate

$$
\begin{aligned}
\int_{-\infty}^{+\infty} \hat{f}(\omega)e^{-i\omega t}d\omega &= \int \frac{1}{2\pi}\int f(t')e^{i\omega t'}dt'e^{-i\omega t}d\omega \\
&= \int f(t')\left(\frac{1}{2\pi}\int e^{i\omega(t'-t)}d\omega\right)dt' \\
&= \int f(t')\Big(\delta(t'-t)\Big)dt' = f(t)
\end{aligned}
\tag{43}
$$

9

2.2. Convolution Theorem

Now consider the convolution theorem: for

$$
\begin{aligned}
h(t) &= \int f(t')g(t-t')dt' \\
&= \int\int \hat{f}(\omega')e^{-i\omega't'}d\omega' \int \hat{g}(\omega'')e^{-i\omega''(t-t')}d\omega''dt' \\
\hat{h}(\omega) &= \frac{1}{2\pi}\int\int\int \hat{f}(\omega')e^{-i\omega't'}d\omega' \int \hat{g}(\omega'')e^{-i\omega''(t-t')}d\omega''dt'e^{i\omega t}dt \\
&= \frac{1}{2\pi}\int\int \hat{f}(\omega')\hat{g}(\omega'')\left(\int e^{i(\omega''-\omega')t'}dt'\right)\left[\int e^{i(\omega-\omega'')t}dt\right]d\omega'd\omega'' \\
&= \frac{1}{2\pi}\int\int \hat{f}(\omega')\hat{g}(\omega'')\Big(2\pi\delta(\omega''-\omega')\Big)\Big[2\pi\delta(\omega-\omega'')\Big]d\omega'd\omega'' \\
&= 2\pi \hat{f}(\omega)\hat{g}(\omega)
\end{aligned}
\tag{44}
$$

Thus the Fourier transform of the convolution of two functions is the product of the two Fourier transforms.

2.3. Parseval's Theorem

Now let $g(t) = f(-t)$. Then $\hat{g}(\omega) = \hat{f}(\omega)^*$, the complex conjugate of $\hat{f}$. This special case of the convolution theorem gives

$$
\begin{aligned}
h(t) &= \int f(t')f(t'-t)dt' \\
\hat{h}(\omega) &= 2\pi \hat{f}(\omega)\hat{f}(\omega)^* = 2\pi|\hat{f}(\omega)|^2
\end{aligned}
\tag{45}
$$

The inverse Fourier transform of $\hat{h}(\omega)$ evaluated at $t = 0$ then gives

$$
\int f(t')f(t')dt' = 2\pi \int |\hat{f}(\omega)|^2 d\omega
\tag{46}
$$

which is Parseval's theorem.

2.4. Simple Properties of the Fourier Transform

1. Linearity: For $h(t) = \alpha f(t) + \beta g(t)$, $\hat{h}(\omega) = \alpha \hat{f}(\omega) + \beta \hat{g}(\omega)$

2. For real $f(t)$, $\hat{f}(-\omega) = \hat{f}(\omega)^*$.

3. For $g(t) = f(\alpha t)$, $\hat{g}(\omega) = \alpha^{-1}\hat{f}(\alpha^{-1}\omega)$

4. For $g(t) = f(t + T)$, $\hat{g}(\omega) = e^{-i\omega T} \hat{f}(\omega)$

The converse of the convolution theorem says that the Fourier transform of a product, $h(t) = f(t)g(t)$, is a convolution:

$$
\begin{aligned}
\hat{h}(\omega) &= \frac{1}{2\pi} \int f(t)g(t)e^{i\omega t} dt \\
&= \frac{1}{2\pi} \int \left(\int \hat{f}(\omega')e^{-i\omega' t} d\omega' \right) \left(\int \hat{g}(\omega'')e^{-i\omega'' t} d\omega'' \right) e^{i\omega t} dt \\
&= \frac{1}{2\pi} \int \int \int \exp(-i\omega' t - i\omega'' t + i\omega t) dt \, \hat{f}(\omega')\hat{g}(\omega'') d\omega' d\omega'' \\
&= \int \delta(\omega - \omega' - \omega'') \hat{f}(\omega')\hat{g}(\omega'') d\omega' d\omega'' \\
&= \int \hat{f}(\omega - \omega'')\hat{g}(\omega'') d\omega'' \quad (47)
\end{aligned}
$$

2.5.   Periodic Functions

The Fourier transform of a periodic function is a series of delta functions: Consider $f(t)$ such that $f(t + T) = f(t)$. Then $f(t) - f(t + T) = 0$ and the Fourier transform of zero is zero. Thus

$$
\begin{aligned}
0 &= \frac{1}{2\pi} \int f(t)e^{i\omega t} dt - \frac{1}{2\pi} \int f(t+T)e^{i\omega t} dt \\
&= \frac{1}{2\pi} \int f(t)e^{i\omega t} \left(1 - e^{-i\omega T} \right) dt \\
&= \hat{f}(\omega) \left(1 - e^{-i\omega T} \right) \quad (48)
\end{aligned}
$$

Therefore $\hat{f}(\omega) = 0$ unless $\omega = 2\pi n/T$ for integer n. Since $\hat{f}(\omega)$ is non-zero only at discrete points, the values at these must be infinite in order to have any affect, so $\hat{f}(\omega)$ is a sequence of delta functions at evenly spaced frequencies, $\omega = 2\pi n/T$.

Conversely, if a function is a series of evenly spaced delta functions, it has a periodic Fourier transform:

$$
\begin{aligned}
f(t) &= \sum f_n \delta(t - nT) \\
\hat{f}(\omega) &= \frac{1}{2\pi} \sum f_n e^{i\omega nT} \\
\hat{f}(\omega + \frac{2\pi}{T}) &= \frac{1}{2\pi} \sum f_n e^{i\omega nT} e^{2\pi ni} \\
&= \hat{f}(\omega) \quad (49)
\end{aligned}
$$

2.6. Nyquist Sampling Theorem

A delta function comb, $f(t) = \sum \delta(t - nT)$, is both a periodic function and a series of delta functions. Thus its Fourier transform has to be both a series of delta functions and periodic. The delta functions in $\hat{f}$ will be at $\omega = 2\pi n/T$ and will all have the same amplitude. To find this amplitude we only have to integrate $\hat{f}$ from $\omega = -\pi/T$ to π/T:

$$\int_{-\pi/T}^{\pi/T} \hat{f}(\omega)d\omega = \int_{-\pi/T}^{\pi/T} \frac{1}{2\pi} \sum e^{i\omega nT} d\omega \tag{50}$$

For $n \neq 0$ the integral above vanishes, but for $n = 0$ it gives $1/T$. Thus for $f(t) = \sum \delta(t - nT)$, we find the Fourier transform is $\hat{f}(\omega) = T^{-1} \sum \delta(\omega - 2\pi n/T)$. We can also show this directly by writing

$$
\begin{aligned}
f(t) &= \lim_{\epsilon \to 0} \sum (1-\epsilon)^{|n|} \delta(t - nT) \\
\hat{f}(\omega) &= \lim_{\epsilon \to 0} \frac{1}{2\pi} \left[\frac{1}{1 - (1-\epsilon)e^{i\omega T}} + \frac{1}{1 - (1-\epsilon)e^{-i\omega T}} - 1 \right] \\
&= \lim_{\epsilon \to 0} \frac{1}{2\pi} \frac{2\epsilon - \epsilon^2}{\epsilon^2 - 2(1-\epsilon)(1 - \cos(\omega T))} \\
&= \lim_{\epsilon \to 0} \frac{1}{2\pi} \frac{\frac{2}{\epsilon} - 1}{1 + \frac{4(1-\epsilon)}{\epsilon^2} \sin^2(\omega T/2)}
\end{aligned}
\tag{51}
$$

When $\sin(\omega T/2) = 0$ at $\omega = 2\pi n/T$, the limit goes to infinity. Otherwise the limit goes to zero. Near $\omega = 0$ we can use the $\sin(\omega T/2) \approx \omega T/2$ to see that

$$\hat{f}(\omega) = \lim_{\epsilon \to 0} \frac{1}{2\pi} \frac{2\epsilon^{-1}}{1 + (\omega T/\epsilon)^2} = T^{-1}\delta(\omega) \quad \text{for } \omega \approx 0 \tag{52}$$

Therefore $\hat{f}(\omega) = T^{-1} \sum \delta(\omega - 2\pi n/T)$.

Now consider the effect of multiplying a function $g(t)$ by the delta function comb. The product will be zero except at $t = nT$, so this is equivalent to sampling the function $g(t)$ at discrete times. The Fourier transform of $g_s(t) = f(t)g(t)$ is given by the converse of the convolution theorem:

$$\hat{g}_s(\omega) = \int \hat{f}(\omega - \omega'')\hat{g}(\omega'')d\omega'' = \frac{1}{T} \sum \hat{g}\left(\omega - \frac{2\pi n}{T}\right) \tag{53}$$

Thus the Fourier transform of the sampled function is an infinite sum of shifted copies of the Fourier transform of the original function. But if the original function $g(t)$ is band limited, so $\hat{g} = 0$ except for $\omega_1 < \omega < \omega_2$, and the frequency range $\omega_2 - \omega_1 < 2\pi/T$, then the shifted copies of $\hat{g}$ in the Fourier transform of g_s do not overlap. In this case one can reconstruct the Fourier transform of g from the Fourier transform of g_s. For real functions, the negative and positive frequencies must both be considered, so if spectrum of g extends up to ω_{max}

it also extends down to $-\omega_{max}$. Therefore the maximum sampling interval that allows full reconstruction of the input signal is

$$T < \frac{\pi}{\omega_{max}} \tag{54}$$

One needs at least one sample every half-cycle of the highest frequency in the signal. This is the Nyquist sampling rate.

2.7.   Random Processes

Often in astrophysics the function $f(t)$ is a *random process*. In this case, the *power spectrum*, which is the magnitude squared of the Fourier transform of f, is given by the Fourier transform of the autocorrelation of f. The autocorrelation $A(\tau)$ of a process $f(t)$ is the average value of $f(t)f(t-\tau)$. For a *stationary* random process, this only depends on $|\tau|$, and does not depend on t. We will do this average by averaging this quantity over a time interval T. We do this by multiplying $f(t)$, which typically extends over all times, by a window function that is unity for $|t| < T/2$ and zero otherwise. Call this windowed function $f_T(t)$. Note that using the average over time to get the average requires that a random process be *ergodic*. The integral one performs to evaluate the autocorrelation function as a time average is almost but not quite a convolution. The difference is that in a convolution the time variable runs backward for one of the functions being convolved. But the Fourier transform of a time-reversed real function is just the complex conjugate to the Fourier transfrom of the function, so the convolution theorem can be used, giving:

$$
\begin{aligned}
h_T(\tau) &= \int f_T(t')f_T(t'-\tau)dt' = T\langle f(t')f(t'-\tau)\rangle = TA(\tau) \\
\hat{h}_T(\omega) &= 2\pi\hat{f}_T(\omega)\hat{f}_T(\omega)^* = 2\pi|\hat{f}_T(\omega)|^2 = \frac{T}{2\pi}\int A(\tau)e^{i\omega\tau}d\tau.
\end{aligned}
\tag{55}
$$

Note that as T gets larger, $\hat{h}_T(\omega)$ gets larger. But $\hat{h}_T(\omega)$ is the distribution of pulse energy over frequency, so if we divide by T we get the distribution of power over frequency. Thus

$$\lim_{T\to\infty}\frac{\hat{h}_T(\omega)}{T} = (2\pi)^{-1}\int A(\tau)e^{i\omega\tau}d\tau = \hat{A}(\omega). \tag{56}$$

This is the *power spectrum* of the process, and it is the Fourier transform of the autocorrelation function.

3.   Impulse Response Function

Consider an atom or dust grain in an incident electric field. It polarizes and develops a dipole moment given by $d = \alpha E_{in}$. The polarizability α depends on frequency: $\alpha(\omega)$. Note

13

that the scattered power can be found using the Larmor formula by setting $qa = \ddot{d}$:

$$P_{scat} = \frac{2q^2a^2}{3c^3} = \frac{2\ddot{d}^2}{3c^3} = \frac{2\alpha(\omega)^2\omega^4}{3c^3}E_{in}^2 \tag{57}$$

Now for an arbitrary incident field $E_{in}(t)$ with FT $\hat{E}_{in}(\omega)$, the dipole moment's FT is given by $\hat{d}(\omega) = \alpha(\omega)\hat{E}_{in}(\omega)$. This is the product of two Fourier transforms, so we can write $d(t)$ as a convolution using the convolution theorem:

$$d(t) = \frac{1}{2\pi}\int E_{in}(t')\alpha(t - t')dt' \tag{58}$$

where $\alpha(t) = \int \alpha(\omega)e^{-i\omega t}d\omega$. The 2π factor appears in a different place because we started with $\alpha(\omega)$. If the incident field is an impulse (a delta function $\delta(t)$) then the dipole moment is $d(t) = (2\pi)^{-1}\alpha(t)$, so $(2\pi)^{-1}\alpha(t)$ is called the *impulse response function* for this system.

By causality the dipole moment can not react to the electric field before the electric field starts, so $\alpha(t) = 0$ for $t < 0$. Since for $t < 0$ you can close the contour for the integral $\alpha(t) = \int \alpha(\omega)e^{-i\omega t}dt$ through $\omega = +i\infty$, this means that $\alpha(\omega)$ has no poles in the upper half of the complex ω plane. For example, if

$$\alpha(t) = \begin{cases} 0, & \text{for } t < 0 \\ e^{-\Gamma t}, & \text{for } t > 0 \end{cases} \tag{59}$$

then

$$\alpha(\omega) = \frac{i}{2\pi}\frac{1}{\omega + i\Gamma} \tag{60}$$

which only has poles in the lower half plane.

3.1. Kramers-Kronig Relations

For functions without poles in upper half plane, one can deduce the imaginary part of the function from the real part and vice-versa. Consider the integral

$$\int_P \frac{\hat{f}(\omega')}{\omega' - \omega}d\omega' = \lim_{\epsilon \to 0}\left(\int_{-\infty}^{\omega-\epsilon} \frac{\hat{f}(\omega')}{\omega' - \omega}d\omega' + \int_{\omega+\epsilon}^{+\infty} \frac{\hat{f}(\omega')}{\omega' - \omega}d\omega'\right) \tag{61}$$

The "P" stands for *principal value* which is the value obtained using the given limit for the this improper integral through a singularity.

If $\hat{f}(\omega)$ has no poles in the upper half plane, then the principal value integral is the average of two contour integrals: one taking a contour that detours above the pole at ω and then closes through $\omega' = +i\infty$, and a second contour that detours below the pole at ω but

14

is otherwise identical. The first contour integral gives zero because there are no poles within the contour, while the second gives $2\pi i \hat{f}(\omega)$. The average gives

$$\pi i \hat{f}(\omega) = \int_P \frac{\hat{f}(\omega')}{\omega' - \omega} d\omega' \tag{62}$$

If we take the real part of this equation it gives

$$\Im(\hat{f}(\omega)) = -\frac{1}{\pi} \int_P \frac{\Re(\hat{f}(\omega'))}{\omega' - \omega} d\omega' \tag{63}$$

while the imaginary part gives

$$\Re(\hat{f}(\omega)) = \frac{1}{\pi} \int_P \frac{\Im(\hat{f}(\omega'))}{\omega' - \omega} d\omega' \tag{64}$$

The latter equation allows us to compute the dielectric constant of interstellar space from the interstellar extinction curve, for example.

The dielectric constant of the ISM is given by $\epsilon = 1 + 4\pi\chi$ where $\chi = P/E$ and P is the dipole density. P is given by $n_g \alpha_g E$ where n_g is the number density of dust grains and α_g is the polarizability of a dust grain. Thus

$$\epsilon = 1 + 4\pi n_g \alpha_g \tag{65}$$

The imaginary part of the wavevector $k = \omega\sqrt{\epsilon}/c$ is given by

$$\Im(k) = \frac{\omega}{c}\Im([1 + 2\pi n_g \alpha_g]) \tag{66}$$

Now $\Im(k) \times D$ gives the fractional reduction in the amplitude of $|E|$, so the reduction in the power is twice this great. But this reduction in power is also given by

$$D \times n_g \times \sigma_{ext} = 2 \times \Im(k) \times D \tag{67}$$

Thus the extinction cross section of dust grain is given by

$$\sigma_{ext} = \frac{4\pi\omega\Im(\alpha_g)}{c}. \tag{68}$$

Because the impulse response function $\alpha(t)$ is real, we know that $\Im(\alpha(\omega))$ must be an odd function of ω. Therefore the extinction cross-section must go like ν^2 for small enough frequencies. Thus the usual far-IR fit of $\sigma \propto \nu^\beta$ with $\beta \approx 1.5$ must break down for really long wavelengths.

We can evaluate the DC polarizability of a dust grain using a knowledge of its extinction curve and Equation 64 with $\omega = 0$:

$$\alpha_{DC} = \Re(\alpha(0)) = \frac{1}{\pi} \int_P \frac{\Im(\alpha(\omega'))}{\omega'} d\omega' = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{c\sigma_{ext}}{4\pi\omega'^2} d\omega' \tag{69}$$

15

and if we substitute $2\pi c/\lambda$ for ω we get

$$\alpha_{DC} = \frac{1}{4\pi^3} \int_0^\infty \sigma_{ext} d\lambda. \tag{70}$$

This is useful because we can usually evaluate the DC polarizability of a particle from elementary electrostatics, so this equation gives a useful integral constraint on the extinction curve.

4.   Blackbody Statistics

Each mode of the E&M field acts like a harmonic oscillator with energy levels 0, $h\nu$, $2h\nu$, ... These correspond to states with 0, 1, 2, ... photons in this mode. The probability distribution of the states in a mode is determined by the Boltzmann distribution, so

$$p(n) \propto e^{-E/kT} \tag{71}$$

but this needs to be normalized by dividing by the *partition function*

$$Z = \sum e^{-E/kT} = 1 + e^{-h\nu/kT} + \left(e^{-h\nu/kT}\right)^2 + \ldots = \frac{1}{1 - e^{-h\nu/kT}} \tag{72}$$

The mean energy per state can be found by taking the derivative of Z with respect to $\beta = 1/kT$:

$$\langle E \rangle = \frac{\sum E e^{-\beta E}}{\sum e^{-\beta E}} = -Z^{-1} \frac{\partial Z}{\partial \beta} \tag{73}$$

This derivative is easily evaluated giving

$$\langle E \rangle = Z^{-1} \frac{h\nu e^{-h\nu/kT}}{\left(1 - e^{-h\nu/kT}\right)^2} = \frac{h\nu}{e^{h\nu/kT} - 1} \tag{74}$$

which is a well known part of the Planck function.

The number of nodes per unit frequency can be found by considering the number of modes with frequencies less than ν_{max} in a box of volume V. In a box with periodic boundary conditions modes have to satisfy $k_x L_x = 2\pi n_x$, $k_y L_y = 2\pi n_y$, and $k_z L_z = 2\pi n_z$ where n_x etc. are integers and the wavenumber k is $2\pi\nu/c$. Thus the volume per mode in k space is $(2\pi/L_x)(2\pi/L_y)(2\pi/L_z) = (2\pi)^3/V$. The volume in k-space with frequency less than ν_{max} is $(4\pi/3)(2\pi\nu_{max}/c)^3$, so the number of modes is $N(< \nu_{max}) = g[(4\pi/3)(2\pi\nu_{max}/c)^3]/[(2\pi)^3/V] = (4\pi g V/3)(\nu_{max}/c)^3$ where the statistical weight g is 2 for the two possible polarizations of each E&M mode. Taking the derivative of this with respect to ν/c gives the number of modes in a box of volume dV in a frequency range $d\nu$ at frequency ν:

$$2 \times 4\pi \times \left(\frac{\nu}{c}\right)^2 \frac{d\nu}{c} dV \tag{75}$$

16

where the initial factor of 2 is for polarizations, and the 4π integrates over all directions. Thus the Planck energy density is

$$u_\nu = 8\pi \frac{\nu^2}{c^3} \frac{h\nu}{e^{h\nu/kT} - 1} \tag{76}$$

This energy density is $(4\pi/c)$ times the mean intensity J_ν which equals the intensity I_ν because the blackbody radiation field is isotropic. Thus the intensity in a blackbody is

$$I_\nu = B_\nu(T) = 2\frac{\nu^2}{c^2} \frac{h\nu}{e^{h\nu/kT} - 1} \tag{77}$$

This Planck function has two well-known limiting approximations. One is the Rayleigh-Jeans approximation which applies when $x = h\nu/kT << 1$. It is $I_\nu = 2kT(\nu/c)^2$. This RJ approximation rises continuously with frequency and gives an infinite total energy density, an *ultraviolet catastrophe* that motivated the development of quantum physics. The approximation that works in the UV is the Wien approximation, $I_\nu = 2h\nu(\nu/c)^2 \exp(-h\nu/kT)$. The approximation has the wrong form at low freqencies, but it does have a peak and a finite total energy density. The Wien displacement law gives the peak of the blackbody law as a function of temperature. The peak of $\nu^n B_\nu(T)$ occurs at the maximum of $x^{n+3}/(e^x - 1)$. Setting the derivative with respect to x of the logarithm of this to zero gives

$$\frac{n+3}{x} = \frac{e^x}{e^x - 1} \qquad \text{or} \qquad x_n = (n+3)(1 - e^{-x_n}). \tag{78}$$

The peak of B_ν occurs at $x_0 = 2.82$, the peak of νB_ν at $x_1 = 3.92$, and the peak of $B_\lambda \propto \nu^2 B_\nu$ at $x_2 = 4.97$. The scaling between T and λ is given by the second radiation constant $hc/k = (1.4387770 \pm 0.0000013)$ K cm.

The integral of the Planck function is easily evaluated. The bolometric intensity is

$$\begin{aligned}
\int B_\nu(T)d\nu &= \frac{2h}{c^2} \int \frac{\nu^3 d\nu}{e^{h\nu/kT} - 1} = \frac{2k^4 T^4}{h^3 c^2} \int \frac{x^3 dx}{e^x - 1} \\
\int \frac{x^n dx}{e^x - 1} &= \int x^n (e^{-x} + e^{-2x} + e^{-3x} + \ldots)dx = n!(1 + 2^{-(n+1)} + 3^{-(n+1)} + \ldots) \\
&= \Gamma(n+1)\zeta(n+1) \\
\int B_\nu(T)d\nu &= \frac{2\Gamma(4)\zeta(4)k^4}{h^3 c^2}T^4 = \frac{\sigma_{SB}}{\pi}T^4
\end{aligned} \tag{79}$$

Note the $1/\pi$: the Stefan-Boltzmann constant σ_{SB} includes a factor of π steradians. Thus

$$\sigma_{SB} = \frac{2\pi^5 k^4}{15h^3 c^2} \tag{80}$$

The energy density u is always $4\pi/c$ times the bolometric intensity so

$$u = \frac{4\sigma_{SB}}{c}T^4 = aT^4 \tag{81}$$

17

with $a = (7.565730 \pm 0.000027) \times 10^{-15}$ erg/cm^3/K^4. The Riemann zeta function $\zeta(s) = \sum_1^\infty n^{-s}$ is the subject of the famous Riemann zeta hypothesis, that if $\zeta(s) = 0$ then either $s = -2, -4, -6, \ldots$ or $\Re(s) = 1/2$. There is currently a \$1,000,000 prize offered for a proof of the Riemann zeta hypothesis. Useful values include $\zeta(2) = \pi^2/6$, $\zeta(3) = 1.2020569\ldots$, and $\zeta(4) = \pi^4/90$.

A rule of thumb for evaluating $\int F_\nu d\nu$ is based of looking at the flux per logarithmic frequency interval, νF_ν.

$$\int F_\nu d\nu = \int \nu F_\nu d\ln\nu \approx \mathcal{O}(1) \times (\nu F_\nu)_{pk}. \tag{82}$$

For the Planck function, $\nu B_\nu \propto x^4/(e^x - 1)$ and peaks at $x_1 = 3.92$, so $[x^4/(e^x - 1)]_{pk} = 4.78$ while $\int x^3 dx/(e^x - 1) = \pi^4/15 = 6.49 = 1.36[x^4/(e^x - 1)]_{pk}$, so in this case $\mathcal{O}(1) = 1.36$.

4.1. Variance

Now let us ask the next statistical question: what is the variance of the number of photons? Suppose we measure a brightness and detect N photons. What is the signal-to-noise ratio (SNR) of this measurement? In optical (and IR) astronomy the SNR will be approximately $\sqrt{N}$, where I have assumed that dark current and sky background are not a problem. But in radio astronomy the Dicke radiometer equation:

$$\sigma(T_A) = \frac{T_{sys}}{\sqrt{Bt}} \tag{83}$$

is used, where B is the bandwidth observed in Hz or cycles per second. If we assume a perfect system then all of the system temperature T_{sys} will be generated by signal and the SNR will be $\sqrt{Bt}$. This result doesn't seem to have anything to do with photons, but even radio waves are made of photons.

To address this mystery we need to compute the *variance* of the energy per mode in the blackbody, or the variance in the number of photons which is equivalent. The variance of any random variable z is the expected value of z squared, $\langle z^2 \rangle$, minus the square of the expected value of z, $\langle z \rangle^2$. Thus the variance of the photon number n is $\langle n^2 \rangle - \langle n \rangle^2$. We have already found $\langle n \rangle = 1/(e^x - 1)$ with $x = h\nu/kT$. To find $\langle n^2 \rangle$ we need to take two derivatives of $Z = \sum e^{-nx}$ with respect to x, obtaining

$$
\begin{aligned}
\langle n^2 \rangle &= Z^{-1}\frac{\partial^2 Z}{\partial x^2} = \frac{\sum n^2 e^{-nx}}{\sum e^{-nx}} \\
&= (1 - e^{-x})\frac{\partial}{\partial x}\left(\frac{-e^{-x}}{(1 - e^{-x})^2}\right) \\
&= (1 - e^{-x})\frac{e^{-x}(1 - e^{-x}) + 2e^{-2x}}{(1 - e^{-x})^3}
\end{aligned}
$$

18

$$= \frac{e^{-x}(1 + e^{-x})}{(1 - e^{-x})^2} \tag{84}$$

and the variance is

$$\langle n^2 \rangle - \langle n \rangle^2 = \frac{e^{-x}(1 + e^{-x})}{(1 - e^{-x})^2} - \frac{e^{-2x}}{(1 - e^{-x})^2} = \frac{e^{-x}}{(1 - e^{-x})^2} = \langle n \rangle^2 + \langle n \rangle \tag{85}$$

Thus the SNR for measuring the number of photons in a single mode is

$$\mathrm{SNR} = \frac{\langle n \rangle}{\sqrt{\mathrm{var}(n)}} = \frac{\langle n \rangle}{\sqrt{\langle n \rangle^2 + \langle n \rangle}} \tag{86}$$

In the optical limit where $\langle n \rangle << 1$ the n^2 term is unimportant and the SNR is just $\sqrt{\langle n \rangle}$, the square root of the number of photons is one mode. We actually observe many modes. Let m be the number of modes observed. Then the total number of photons seen is $N = m\langle n \rangle$ and the SNR on the measurement of N is $\sqrt{m\langle n \rangle} = \sqrt{N}$.

For a bolometer detector with area A, sensitive to radiation at angle θ in the projected area $A\cos\theta$, the throughput $A\Omega = \pi A$. The variance in the energy received in t seconds of integration is

$$\begin{aligned}
\mathrm{var}(E_r) &= 2t(\pi A) \int \left(\frac{\nu}{c}\right)^2 (h\nu)^2 \left(\langle n \rangle^2 + \langle n \rangle\right) d\nu \\
&= 2t\pi A \frac{k^5 T^5}{h^3 c^2} \int \frac{x^4 e^x}{(e^x - 1)^2} dx
\end{aligned} \tag{87}$$

We can do the integral by parts with $q = (e^x - 1)^{-1}$ so $dq = -e^x dx/(e^x - 1)^2$, and $p = x^4$ so $dp = 4x^3 dx$. Then

$$\int \frac{x^4 e^x}{(e^x - 1)^2} dx = 4 \int \frac{x^3}{e^x - 1} dx = 4\Gamma(4)\zeta(4) = \frac{4\pi^4}{15} \tag{88}$$

Now the NEP, or Noise Equivalent Power, is the standard deviation of the power in 0.5 seconds, since by the Nyquist theorem 0.5 second sampling gives a 1 Hertz bandwidth. Now the power is $P = E_r/t$, so the NEP is given by

$$\mathrm{NEP} = \left[\frac{16\pi^5 A k^5 T^5}{15 h^3 c^2}\right]^{1/2} = \left[8A\sigma_{SB} k T^5\right]^{1/2} \tag{89}$$

Note that the units of NEP are $\mathrm{Power}/\sqrt{Hz}$. The analysis above assumes the detectors is not radiating any photons. If instead the bolometer is operating in thermal equilibrium, there is an equivalent noise from the photons it emits, giving $\mathrm{NEP} = \sqrt{16A\sigma_{SB} k T^5}$.

In thermal equilibrium we can find the variance of the energy on a bolometer by considering $dS = dE/T$ with the entropy S given by $S = k \ln \Omega(E)$ where Ω is the number

of states. $dS/dE = kd\ln\Omega/dE = 1/T$ and taking the second derivative gives $d^2S/dE^2 = kd^2\ln\Omega/dE^2 = -T^{-2}dT/dE$. But dE/dT is the heat capacity of the detector C. Thus the variance of the energy, which is determined by $d^2\ln\Omega/dE^2$, is given by

$$\mathrm{var}(E) = \left(-\frac{d^2\ln\Omega}{dE^2}\right)^{-1} = kCT^2 \tag{90}$$

For a thermal conductance G the time constant τ is C/G, the power is given by $P = E/\tau$, and the variance of the power is

$$\mathrm{var}(P) = kCT^2(G^2/C^2) = kGT^2/\tau \tag{91}$$

For a filter with an exponential time constant the response is $\tau^{-1}\exp(-t/\tau)$ for $t \geq 0$ which has a Fourier transform $(2\pi)^{-1}/(i\omega\tau - 1)$. The response to power is the square of the magnitude of this, giving a response $\propto (1 + (\omega\tau)^2)^{-1}$. The integral of this from 0 to ∞ is $\frac{1}{2}\pi/\tau$ radians/second or $(4\tau)^{-1}$ Hertz. Thus the time constant corresponding to a 1 Hertz bandwidth is $\tau = 1/4$ seconds, so

$$\mathrm{NEP} = \sqrt{4kGT^2} \tag{92}$$

If radiative coupling is the only source of the thermal conductance then $G = 4A\sigma_{SB}T^3$, giving $\mathrm{NEP} = \sqrt{16A\sigma_{SB}kT^5}$ in accordance with the photon-based calculation.

4.2. Radio Astronomy

In the radio astronomy case we are in the Rayleigh-Jeans limit where $\langle n\rangle >> 1$. In this case the n^1 term drops out and we get a SNR for the number of photons in one mode is

$$\mathrm{SNR} = \frac{\langle n\rangle}{\sqrt{\langle n\rangle^2 + \langle n\rangle}} = \frac{1}{\sqrt{1 + \langle n\rangle^{-1}}} \approx 1 \tag{93}$$

Thus the SNR for one mode is just unity, and the overall SNR for m modes is just $\sqrt{m}$. But what is the number of modes observed by a radio telescope? The throughput $dAd\Omega$ is equal to λ^2 for a radio telescope, and only one polarization is observed. [The VLA has two receivers per dish to observe both polarizations but we are only considering one receiver.] The volume observed is $ctdA$ where t is the integration time or length of the observation. Thus the number of modes becomes

$$m = d\Omega \times \left(\frac{\nu}{c}\right)^2\frac{d\nu}{c}ctdA = (dAd\Omega)\left(\frac{\nu}{c}\right)^2 td\nu \tag{94}$$

Since $dAd\Omega = \lambda^2$ and $(\nu/c)^2 = \lambda^{-2}$, the result is $m = td\nu = Bt$. Thus the SNR for perfect radio observations is $\sqrt{Bt}$ just as given by the Dicke radiometer equation.
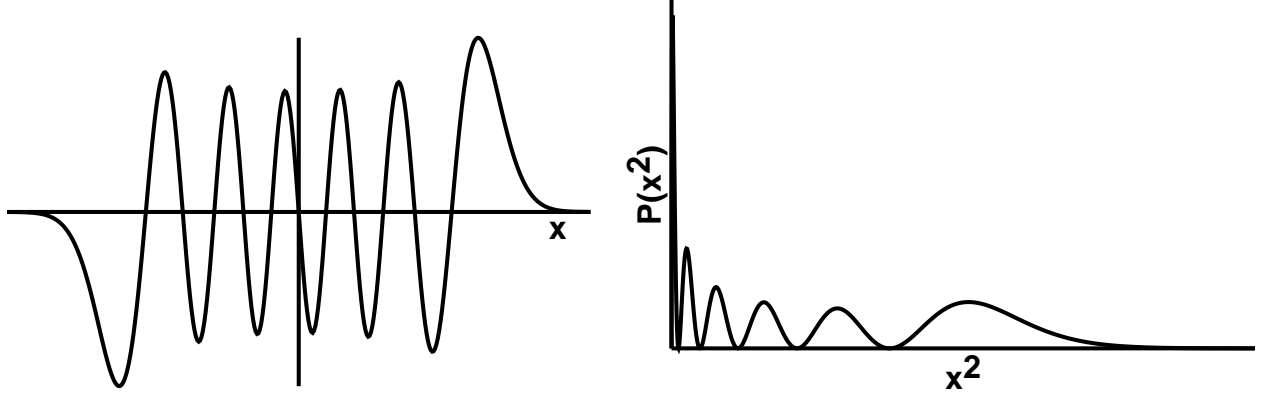
Fig. 2.— Left: The wavefunction for the $n = 11$ state of an harmonic oscillator. Right: The resulting probability density function for x^2 which is analogous to the output of a square-law detector following a coherent amplifier.

4.3. Quantum Noise Limit

But any coherent detection system is not quite perfect. There is a small amount of added noise caused by measuring the electric field of the incoming signal instead of the number of photons. The electric field does not commute with the density matrix of a thermal photon field, so an additional uncertainty is added. The electric and magnetic fields in an E&M mode are like a harmonic oscillator, so consider a simple harmonic oscillator with Hamiltonian

$$H = \frac{p^2}{2m} + \frac{kq^2}{2} \tag{95}$$

with the oscillation frequency $\omega = 2\pi\nu = \sqrt{k/m}$. The usual way to work with quantum harmonic oscillators is to form the raising and lowering operators

$$a^\dagger = (p/\sqrt{2m} + iq\sqrt{k/2})/\sqrt{h\nu}$$

and

$$a = (p/\sqrt{2m} - iq\sqrt{k/2})/\sqrt{h\nu}.$$

Then

$$h\nu \times (a^\dagger a) = \frac{p^2}{2m} + \frac{kq^2}{2} - \frac{i\sqrt{k/m}[pq - qp]}{2} = \frac{p^2}{2m} + \frac{kq^2}{2} - \frac{h\nu}{2} \tag{96}$$

Thus $H = h\nu(a^\dagger a + 0.5)$, and $n = a^\dagger a$ is the *number* operator that gives which level the harmonic oscillator is in. The product taken the other way gives $aa^\dagger = n + 1$. The matrix elements of a can easily be worked out: the only non-zero matrix elements are $\langle n - 1|a|n\rangle = \sqrt{n}$. Of course you get the matrix elements of $a^\dagger$ just by transposing the matrix and taking the complex conjugate of the elements, so $\langle n|a^\dagger|n - 1\rangle = \sqrt{n}$ or $\langle n + 1|a^\dagger|n\rangle = \sqrt{n + 1}$

Finally we can represent the electric field as the displacement of our oscillator: $\vec{E} \propto q$. But $q \propto i(a^\dagger - a)$, so the electric field is proportional to $i(a^\dagger - a)/\sqrt{2}$. The output of the

square-law detector in a coherent system is then given by the operator

$$\mathcal{N} = -0.5(a^\dagger - a)^2 = 0.5[-a^\dagger a^\dagger + a^\dagger a + aa^\dagger - aa] \tag{97}$$

which is not the same as the photon number operator $n = a^\dagger a$. Figure 2 shows that even when the system is in a definite state there will still be a range of measured E^2's. The expected value of $\mathcal{N}$ in a thermal field with density matrix $\rho_{nn'} = \delta_{nn'}p_n$ is given by the trace of the density matrix times the operator. Since the density matrix is diagonal and the trace is just the sum of the diagonal elements, we only need to evaluate the diagonal elements of $\mathcal{N}$ giving

$$\overline{\mathcal{N}} = \text{Tr}(\mathcal{N}\rho) = \overline{n} + 0.5 \tag{98}$$

thus we estimate the photon number using $\tilde{n} = \overline{\mathcal{N}} - 0.5$. But we need the variance of $\mathcal{N}$ as well, so we look at

$$\begin{aligned}
\mathcal{N}^2 \ = 0.25 \ \ [&a^\dagger a^\dagger a^\dagger a^\dagger - a^\dagger a^\dagger a^\dagger a - a^\dagger a^\dagger aa^\dagger - \underline{a^\dagger a^\dagger aa} \\
&-a^\dagger aa^\dagger a^\dagger + \underline{a^\dagger aa^\dagger a} + \underline{a^\dagger aaa^\dagger} - a^\dagger aaa \\
&-aa^\dagger a^\dagger a^\dagger + \underline{aa^\dagger a^\dagger a} + \underline{aa^\dagger aa^\dagger} - aa^\dagger aa \\
&+\underline{aaa^\dagger a^\dagger} - aaa^\dagger a - aaaa^\dagger + aaaa]
\end{aligned} \tag{99}$$

Only the underlined terms above are diagonal and thus contribute to the expected value of $\mathcal{N}^2$, giving

$$\langle \mathcal{N}^2 \rangle = \text{Tr}(\mathcal{N}^2\rho) = Z^{-1}\sum_n e^{-nx}\left(\frac{3}{2}n^2 + \frac{3}{2}n + \frac{3}{4}\right) = \frac{3}{2}\overline{n}(1 + 2\overline{n}) + \frac{3}{2}\overline{n} + \frac{3}{4} \tag{100}$$

Then the variance is given by

$$\text{var}(\mathcal{N}) = \langle \mathcal{N}^2 \rangle - \overline{\mathcal{N}}^2 = 2\left(\overline{n} + \frac{1}{2}\right)^2 \tag{101}$$

But the Nyquist theorem states that 2 samples per cycle are necessary and sufficient to determine a band limited signal, and one can easily show for a Gaussian signal $E(t)$ with a power spectrum that is a boxcar of width B that samples of $E(t)^2$ separated by $1/(2B)$ seconds are independent. Therefore the effect of the E field not commuting with the thermal density matrix is to increase the variance of the photon number measurement from $\overline{n}(1+\overline{n})$ to $\overline{n}(1 + \overline{n}) + 0.25 = (\overline{n} + 0.5)^2$. Thus the noise after t seconds of integration is given by $\sigma(\overline{n}) = (\overline{n} + 0.5)/\sqrt{Bt}$. Since $T_A = h\nu\overline{n}/k$, this gives

$$\sigma(T_A) = \frac{T_A + h\nu/2k}{\sqrt{Bt}} \tag{102}$$

so the appropriate system temperature to use in the Dicke radiometer equation is

$$T_{sys} = T_A + \frac{h\nu}{2k} \tag{103}$$

The second term above evaluates to 0.024 K/GHz and is the *quantum limit* for coherent detection systems. It corresponds to one half photon per mode, which is the just the zero point energy of the harmonic oscillators making up the E&M field.

Practical systems now achieve 0.3 K/GHz (for the Cosmic Background Imager) using HEMT amplifiers, which corresponds to a photon number variance of $(\overline{n} + 7)^2$. Clearly coherent detection systems will be disfavored when $\overline{n} << 1$: in the Wien tail of the ambient blackbody spectrum, or for $\lambda < 48$ μm at 300 K. Hence most optical and IR detectors are incoherent, while most radio detectors are coherent.

4.4. Noise Equivalent Power

Practical incoherent detectors generate noise even when not being hit by photons. This noise is usually specified in terms of the Noise Equivalent Power (NEP), which is the noise in a 1 Hz post-detection bandwidth, or an 0.5 second integration time. Since the noise power in a single mode system is given by $\sigma(P) = kB\sigma(T_A)$, we can relate the system temperature of a coherent system to an equivalent NEP:

$$\text{NEP} = kT_{sys}\sqrt{2B} \tag{104}$$

For the state of the art HEMT with a 10 GHz bandwidth and 15 K amplifier noise at 30 GHz one gets an NEP of 3.5×10^{-17} W $(\text{Hz})^{-0.5}$ which has only been achievable with bolometers for the last few years. Clearly for smaller bandwidths B the effective NEP of a coherent system is lower.

5. Stimulated Emission

The transitions between different atomic levels are caused by a perturbation Hamiltonian $\Delta H = \vec{E} \cdot \vec{d}$ where $\vec{d}$ is the transition dipole moment and $\vec{E}$ is the electric field. So absorption and stimulated emission just correspond to time reversed version of the same interaction. But why is there *spontaneous* emission? If there is no radiation, why isn't $\Delta H = 0$ giving no transitions? The answer comes from our picture of the E&M field as a set of harmonic oscillators. We can represent the electric field as the displacement of our oscillator: $\vec{E} \propto q$. But $q \propto i(a^\dagger - a)$. The transition rate introduced by $\Delta H = \vec{E} \cdot \vec{d}$ is proportional to $|\Delta H|^2$, and the atomic physics part $\vec{d}$ does not depend on the number of photons in the mode. Thus the rate from a state with n photons to a state with n' photons is

$$\text{rate} \propto |\langle n'|(a^\dagger - a)|n\rangle|^2 = \begin{cases} n + 1 & \text{for } n' = n + 1; \\ n & \text{for } n' = n - 1. \end{cases} \tag{105}$$

Therefore the downward transitions of the atom proceed with a rate proportional to $n + 1$, where n is the initial number of photons in the mode. This gives an explanation for the

existence of spontaneous emission, and shows that the Einstein A_{ul} coefficient for spontaneous emission is equal to the Einstein B_{ul} coefficient for stimulated emission when the intensity is measured in units of photons per mode.

The Einstein B_{lu} coefficient for absorption differs from B_{ul} only by a factor of statistical weights: the rates are proportional to the number of "target states" for the transition, since the rates to each sublevel in the final state must be summed to get the final rate. Thus when intensities are measured in units of photons/mode,

$$
\begin{aligned}
A_{ul} &= B_{ul} \\
\frac{B_{ul}}{B_{lu}} &= \frac{g_l}{g_u}
\end{aligned}
\tag{106}
$$

5.1. Correction for Stimulated Emission

Note that for an intensity of 1 photon per mode, the number of absorptions per second per atom due to σ_{lu} is

$$
\int \sigma_{lu}(\nu) 8\pi (\nu/c)^2 d\nu = B_{lu} = \frac{g_u}{g_l} A_{ul}
\tag{107}
$$

so

$$
\int \sigma(\nu) d\nu = \frac{\lambda^2}{8\pi} \frac{g_u}{g_l} A_{ul}
\tag{108}
$$

Integrating R&L 1.74 gives

$$
\int \alpha_\nu d\nu = n_1 \frac{\lambda^2}{8\pi} \frac{g_2}{g_1} A_{21}
\tag{109}
$$

after substituting $\hat{B}_{12} = (g_2/g_1)\hat{B}_{21} = (g_2/g_1)(c^2/2h\nu^3)A_{21}$ which is consistent with the previous equation. [I use $\hat{B}$ for the Einstein B's scaled to cgs intensity units used by R&L.] Thus one can use any set of intensity units to define the Einstein B's but one has to be consistent. The Einstein A's always have units of s^{-1}.

Because stimulated emission is equivalent to *negative absorption*, a naive calculation of the optical depth using

$$
\tau = N_l \sigma_{lu},
\tag{110}
$$

with N_l being the column density in the lower level and σ_{lu} being the absorption cross-section from the lower to the upper level, will give the wrong answer. The corrected result is

$$
\tau = \sigma_{lu} \left(N_l - N_u \frac{B_{ul}}{B_{lu}} \right) = N_l \sigma_{lu} \left(1 - \frac{N_u/g_u}{N_l/g_l} \right).
\tag{111}
$$

This ratio of populations defines the *excitation* temperature T_{ex} for the transition, so

$$
\frac{N_u/g_u}{N_l/g_l} = \exp\left(\frac{-(E_u - E_l)}{kT_{ex}} \right)
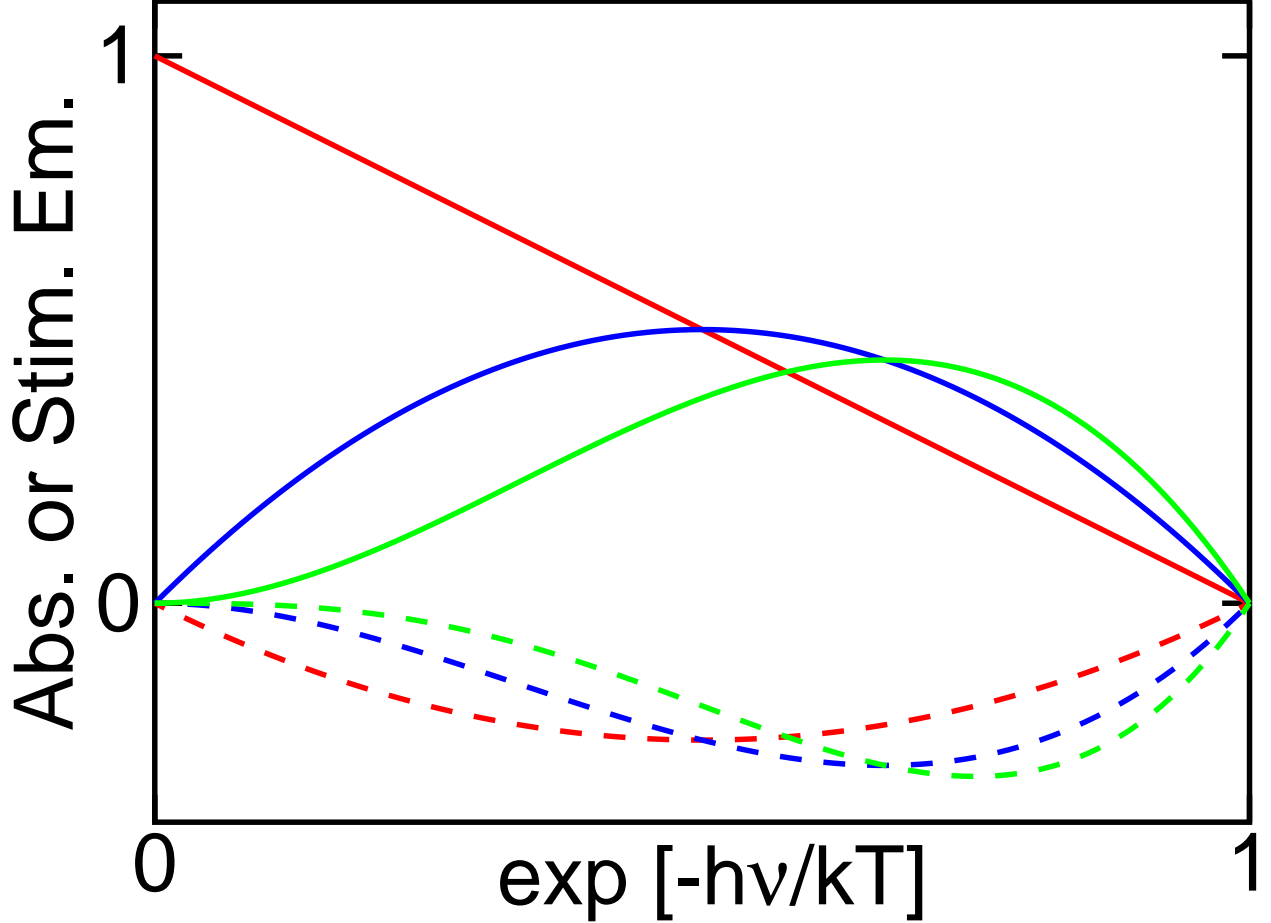\tag{112}
$$

Fig. 3.— Absorption and stimulated emission in the 0-1 transition of a harmonic oscillator [red], the 1-2 transition [blue], and the 2-3 transition [green], as a function of $\exp(-h\nu/kT)$.

and the correction for stimulated emission is a factor of

$$\left(1 - \exp\left(\frac{-h\nu}{kT_{ex}}\right)\right) \tag{113}$$

Figure 3 shows the absorption and stimulated emission (negative absorption) in different transitions of a harmonic oscillator.

An common example in astrophysics where stimulated emission is crucial is the 21 cm line of neutral hydrogen, with $\lambda = 21.10614$ cm, $\nu = 1.420406$ GHz, and $A_{ul} = 2.87E{-}15$ s^{-1}. Because $h\nu/k = 0.07$ K is so small compared to the actual temperatures even in the cold interstellar medium, the atoms will be distributed among states according to the statistical weight, with $g_l = 1$ and $g_u = 3$. Thus $\frac{3}{4}$ of the atoms will be in the upper state, and the total emission from a column density N_H of neutral hydrogen will be

$$\int\int j_\nu d\nu ds = \tfrac{3}{4} h\nu A_{ul} N_H/4\pi = 1.61 \times 10^{-33} N_H \text{ erg/sec/cm}^2/\text{sr} \tag{114}$$

For an optically thin column this gives the observed intensity, so $\int I_\nu d\nu = 1.61 \times 10^{-33} N_H$ erg/sec/cm^2/sr. A Rayleigh-Jeans brightness temperature of 1 K over a bandwidth of 1 km/sec gives $\int I_\nu d\nu = 2k(\nu/c)^2(1 \text{ km/sec}/c)\nu = 2.94 \times 10^{-15}$ erg/sec/cm^2/sr and is thus equivalent to a neutral hydrogen column density of $2.94 \times 10^{-15}/1.61 \times 10^{-33} = 1.822 \times 10^{18}$ atoms/cm^2.

Now the integrated absorption cross-section is

$$\int \sigma(\nu)d\nu = \frac{\lambda^2}{8\pi} \frac{g_u}{g_l} A_{ul} \tag{115}$$

and the integrated intensity without any stimulated emission corrections would be

$$\int I_\nu d\nu = N_H \int \sigma(\nu) B_\nu(T) d\nu = \tfrac{1}{4} N_H (4\pi)^{-1} \frac{g_u}{g_l} A_{ul} h\nu/(\exp(h\nu/kT) - 1) \tag{116}$$

this differs from the previous result by a factor of $kT/h\nu \approx 1 - \exp(-h\nu/kT)$.

5.2. No Stimulated Emission Correction for Dust

Since all atomic absorption calculations should include the correction for stimulated emission, why does it *not* apply to dust? The absorptions in dust grains are due to optically active lattice vibrations. Each vibrational mode can be treated as a harmonic oscillator, and the excitation level $n = 0$, 1, ... of the mode can be called the number of *phonons* in the mode, just as the energy level of a mode of the E&M field gives the the number of photons in the mode. The dipole transition matrix element $\vec{d}$ discussed before is due to the displacement of charges during the oscillation, and is thus proportional to the displacement q of the harmonic oscillator. Thus the matrix element for the transition from a state with n phonons to the next lower state with $n - 1$ phonons is proportional to $\sqrt{n}$, and thus the Einstein A which goes like the square of the matrix element (times a frequency factor which is constant because all the allowed transitions in a harmonic oscillator have the same frequency) varies like n. Thus the Einstein A's satisfy $A_{21} = 2A_{10}$, $A_{32} = 3A_{10}$, etc. Since the statistical weights are all unity we also have $B_{12} = B_{21} = 2B_{01}$, $B_{23} = B_{32} = 3B_{01}$. If we now compute the net absorptions from dust grains with column densities N_0, N_1, N_2 ... in the various levels we get

$$
\begin{aligned}
\text{net} &= N_0 B_{01} + N_1(B_{12} - B_{10}) + N_2(B_{23} - B_{21}) + \ldots \\
&= N_0 B_{01} + N_1(2B_{01} - B_{01}) + N_2(3B_{01} - 2B_{01}) + \ldots \\
&= N_{tot} B_{01} \tag{117}
\end{aligned}
$$

Thus the net absorptions in a dust grain are just proportional to the *total* column density (not the ground state column density) times the ground state absorption cross-section. This cancellation of the stimulated emission correction for dust is a consequence of the infinite ladder of states in a harmonic oscillator which all absorb at the same frequency. Figure 4 shows the total net absorption obtained by summing all the net absorptions up to a given transition for a harmonic oscillator.
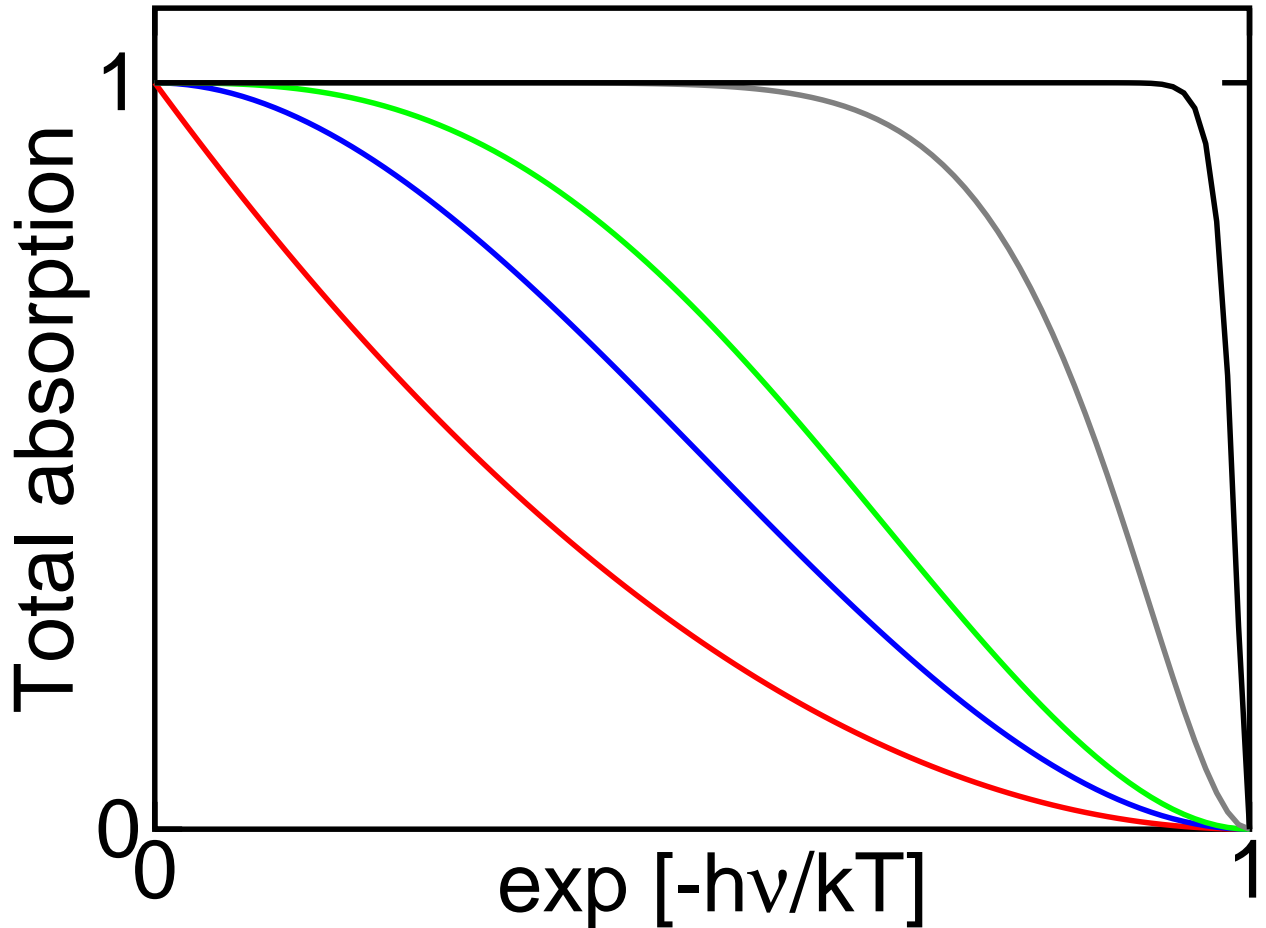
Fig. 4.— Absorption minus stimulated emission in the 0-1 transition of a harmonic oscillator [red], sum of the 0-1 and 1-2 transitions [blue], sum of the 0-1 to 2-3 transitions [green], sum of the 0-1 to 10-11 transitions [gray], and sum of the 0-1 to 100-101 transitions [green], as a function of $\exp(-h\nu/kT)$.

6. Solving Radiative Transfer Problems

The formal solution to the equation of radiative transfer

$$\frac{\partial I_\nu}{\kappa_\nu \rho ds} = -I_\nu + S_\nu \tag{118}$$

is trivial when the source functions is known:

$$I_\nu(\tau) = \int_0^\tau S(\tau') \exp(\tau' - \tau) d\tau' + \exp(-\tau) I_\nu(0) \tag{119}$$

with $d\tau = \kappa_\nu \rho ds$ taken along the direction the ray has come from, not the direction it is going toward. In R&L κ is taken to be in units of cm^2/gm, and the absorption coefficient α is taken to have unit cm^{-1} so $\alpha = \kappa\rho$.

27

But the source function in almost never specified except in problem sets. Instead one has various integral constraints like *radiative equilibrium,* which requires that radiation transport not lead to net sinks or sources of energy, so the energy density at each point in a model remains constant. This constraint is a non-linear function of the temperature, and it ties together the equations for light rays traveling in different directions. The simplest example of this is for a grey atmosphere. The power per unit volume absorbed by material is $4\pi\alpha J$, where J is the *bolometric mean intensity.* The power emitted is $4\pi\alpha S$, and radiative equilibrium then gives the equation $S = J$. The simplest geometry is a plane-parallel atmosphere or semi-infinite slab, and in this geometry we have

$$\mu\frac{\partial I}{\partial \tau} = I - S = I - J \tag{120}$$

where τ is the optical depth measured vertically into the slab, and μ is the cosine of the angle of the vertical. The ray coming straight out of the slab has $\mu = 1$ while a ray going straight down has $\mu = -1$.

This linear integro-differential equation can be profitably attacked using two different methods. These are the *moment expansion* method and the *discrete ordinate* method.

6.1. Moment Expansion

For the moment method, we multiply the equation by μ^n, and then average over μ. The result for $n = 0$ is

$$\frac{1}{2}\int_{-1}^{+1}\mu\frac{\partial I}{\partial \tau}d\mu = \frac{1}{2}\int_{-1}^{+1}(I - S)d\mu \tag{121}$$

If we define $H = \frac{1}{2}\int \mu I d\mu = \frac{1}{4\pi}F$ this gives

$$\frac{dH}{d\tau} = J - S = 0 \tag{122}$$

For the $n = 1$ case we define the *radiation pressure,* $K = \frac{1}{2}\int \mu^2 I d\mu$. The moment equation is

$$\frac{1}{2}\int_{-1}^{+1}\mu^2\frac{\partial I}{\partial \tau}d\mu = \frac{dK}{d\tau} = \frac{1}{2}\int_{-1}^{+1}\mu(I - S)d\mu = H \tag{123}$$

If we try the $n = 2$ case we find we need another moment $\frac{1}{2}\int \mu^3 I d\mu$ so the process leads to an infinite number of ordinary differential equations. To truncate this series we assume that $I(\mu) = a + b\mu$ which then gives $K = \frac{1}{3}J$. Thus we get the equations

$$\frac{dJ}{d\tau} = 3H$$
$$\frac{dH}{d\tau} = J - S = 0 \tag{124}$$

28

Note that if we don't assume $J = S$ from radiative equilibrium, we have a second order equation for J:

$$\frac{d^2 J}{d\tau^2} = 3(J - S) \tag{125}$$

which will be useful later. But if we assume radiative equilibrium, then $d^2 J/d\tau^2 = 0$ so $J = a + b\tau$. Now we need to set some boundary conditions to determine a and b. We want the flux to be $F = \sigma T_e^4$ so $H = \sigma T_e^4/4\pi$ and thus $b = dJ/d\tau = 3H = (3\sigma T_e^4/4\pi)$.

At the surface we want the downward intensity to be zero for all $\mu < 0$. If we assume the upward intensity is some constant value $I_\circ$ for all $\mu > 0$ then we find $J = I_\circ/2$ and $H = I_\circ/4$. Thus we get the boundary condition $J = 2H$ at the surface. Thus we find that

$$J = H(2 + 3\tau) = \frac{\sigma T_e^4}{\pi} \left(\frac{1}{2} + \frac{3}{4}\tau \right) \tag{126}$$

But $J = \sigma T^4/\pi$ so the final approximate solution for a grey plane-parallel atmosphere is

$$T^4 = \left(\frac{1}{2} + \frac{3}{4}\tau \right) T_e^4 \tag{127}$$

These approximations: truncating the moment series by assuming $K = J/3$ and the surface boundary condition $J = 2H = (2/3)dJ/d\tau$ are known as the Eddington approximation (see Eqn $226 \cdot 3$ in *The Internal Constitution of the Stars* by A. S. Eddington).

6.2. Discrete Ordinates

The second useful method for converting the radiative transfer equation into a set of ordinary differential equations that can be solved is the discrete ordinate method. Choose a set of μ's and solve for I at these discrete values of μ, and then use this small set of values to evaluate the integrals over μ that are needed. The weights w_n used to approximate integrals as

$$\int f(\mu)d\mu \approx \sum w_n f(\mu_n) \tag{128}$$

are also needed. With two μ's the natural choice for the *knots* (points where I is evaluated) are the knots for Gaussian quadrature with 2 points: $\mu_n = \pm 1/\sqrt{3}$ with weights $w_n = 1$. This gives a set of equations

$$\frac{1}{\sqrt{3}} \frac{dI^+}{d\tau} = I^+ - \frac{1}{2} \left(I^+ + I^- \right)$$
$$-\frac{1}{\sqrt{3}} \frac{dI^-}{d\tau} = I^- - \frac{1}{2} \left(I^+ + I^- \right) \tag{129}$$

We can transform these into

$$\frac{dI^+}{d\tau} = \frac{\sqrt{3}}{2}\left(I^+ - I^-\right)$$
$$\frac{dI^-}{d\tau} = \frac{\sqrt{3}}{2}\left(I^+ - I^-\right) \tag{130}$$

Subtracting these two equations shows us that $(I^+ - I^-)$ is a constant, and since

$$H = \frac{1}{2}\int \mu I d\mu = \frac{1}{2\sqrt{3}}\left(I^+ - I^-\right) \tag{131}$$

we see that $dI^+/d\tau = 3H$ and $dI^-/d\tau = 3H$, so $J = a + 3H\tau$ as before. The boundary condition at the surface is $I^- = 0$ giving $J = I^+/2$, $H = I^+/2\sqrt{3}$, so $J = \sqrt{3}H$ at $\tau = 0$. Hence $J = (\sqrt{3} + 3\tau)H$ and the 2-ordinate method (or two-stream method) gives

$$T^4 = \left(\frac{\sqrt{3}}{4} + \frac{3}{4}\tau\right)T_e^4 \tag{132}$$

Except for the small change in the constant term, this answer is identical to the Eddington approximation. Which is correct? Actually neither answer is correct, but both are reasonable approximations. Using 8 discrete ordinates Chandrasekhar (in *Radiative Transfer*, Dover) finds

$$T^4 = \left(0.5302 - 0.0629e^{-1.1032\tau} - 0.0271e^{-1.5918\tau} - 0.0071e^{-4.4581\tau} + \frac{3}{4}\tau\right)T_e^4 \tag{133}$$

The plot in Figure 5 shows the flux as function of optical depth for the Eddington, the two stream, and the eight stream approximations.

My own feeling is the the eight stream method is a lot of work for a rather small improvement.

6.3. Integral Operators

One can also use integral operators to represent radiative transfer, and then solving radiative transfer problems is equivalent to finding a particular eigenfunction of the operator. For example, the mean intensity in a grey semi-infinite slab can be found using

$$J(\tau) = 0.5\int_0^\infty E_1(|\tau' - \tau|)S(\tau')d\tau' \tag{134}$$

where $E_n(z) = \int_1^\infty [\exp(-zt)/t^n]dt$ is an exponential integral. We can show this by finding the contribution of an optical depth element at τ' to the intensity $I(\tau, \mu)$, which is $dI = \exp(|\tau' - \tau|/\mu)S(\tau')d\tau'/\mu$ which contributes $dJ = 0.5\exp(|\tau' - \tau|/\mu)S(\tau')d\tau'd\mu/\mu$ to the
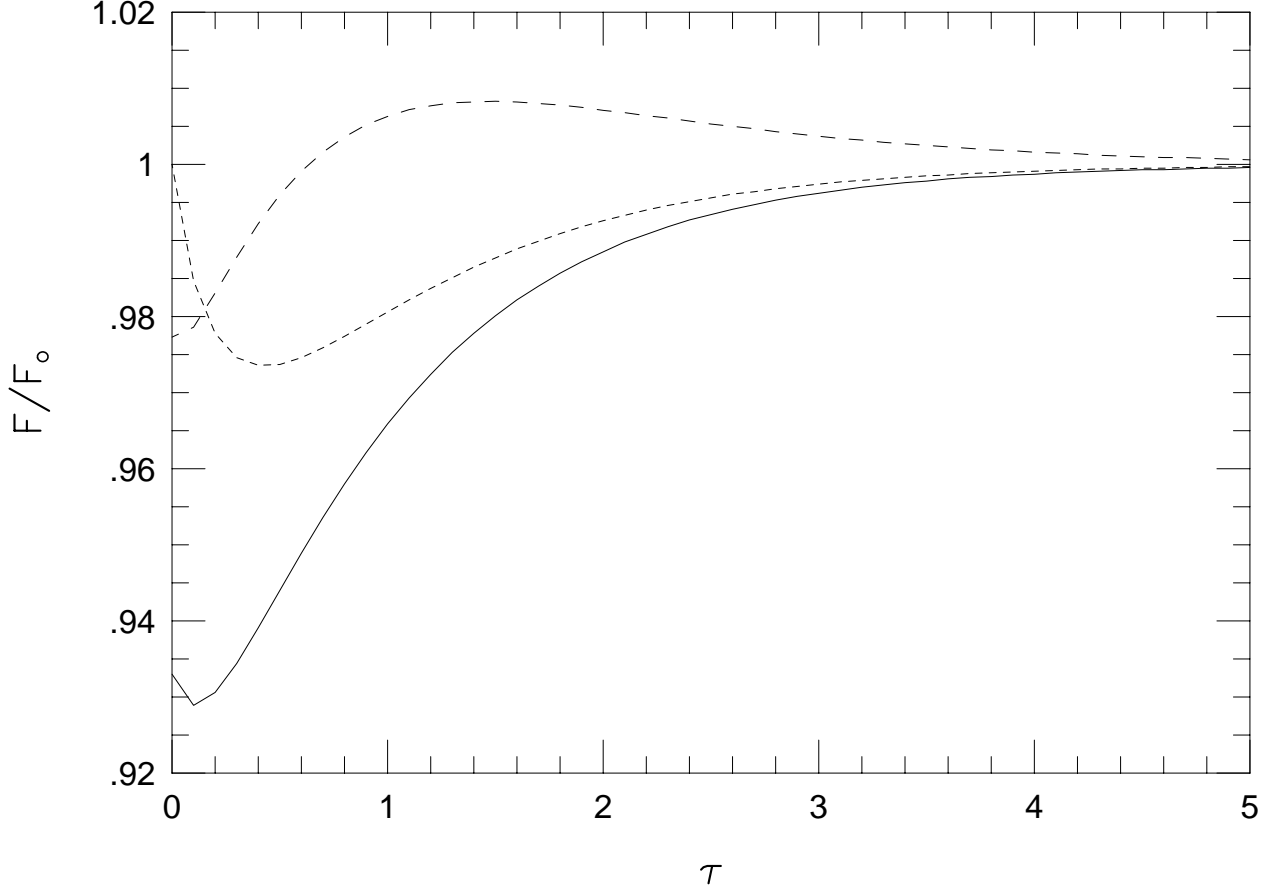
Fig. 5.— Ratio of the actual flux to the desired flux for three approximations: two-stream (solid), Eddington (short dash), and 8 stream (long dash).

mean intensity. Changing variables to $t = |\mu|^{-1}$ gives $dJ = 0.5E_1(|\tau' - \tau|)S(\tau')d\tau'$. Note that for $\tau' > \tau$, only upgoing rays with $\mu > 0$ contribute, while for $\tau' < \tau$, only downgoing rays with $\mu < 0$ contribute.

This relation can be written as a functional equation $\mathbf{J} = \Lambda\mathbf{S}$. Finding radiative equilibrium solutions then reduces to finding eigenfunctions of Λ with unit eigenvalue. There are an infinite number of these corresponding to different values of the flux. Generally if $\mathbf{S}$ is close to a radiative equilibrium solution, then the series $\Lambda\mathbf{S}$, $\Lambda^2\mathbf{S}$, ... will converge to a radiative equilibrium solution, a process known as Λ-iteration. Regions where J is greater than S will absorb more energy than they emit, so they will heat up, leading to a higher value of S. Λ-iteration is a simple computational implementation of this approach to equilibrium.

We can also write the flux as

$$H(\tau) = 0.5\left[\int_\tau^\infty E_2(\tau' - \tau)S(\tau')d\tau' - \int_0^\tau E_2(\tau - \tau')S(\tau')d\tau'\right] \qquad (135)$$

or $\mathbf{H} = \Phi\mathbf{S}$. If the optical depth range is discretized so that S and H are vectors and Φ is a
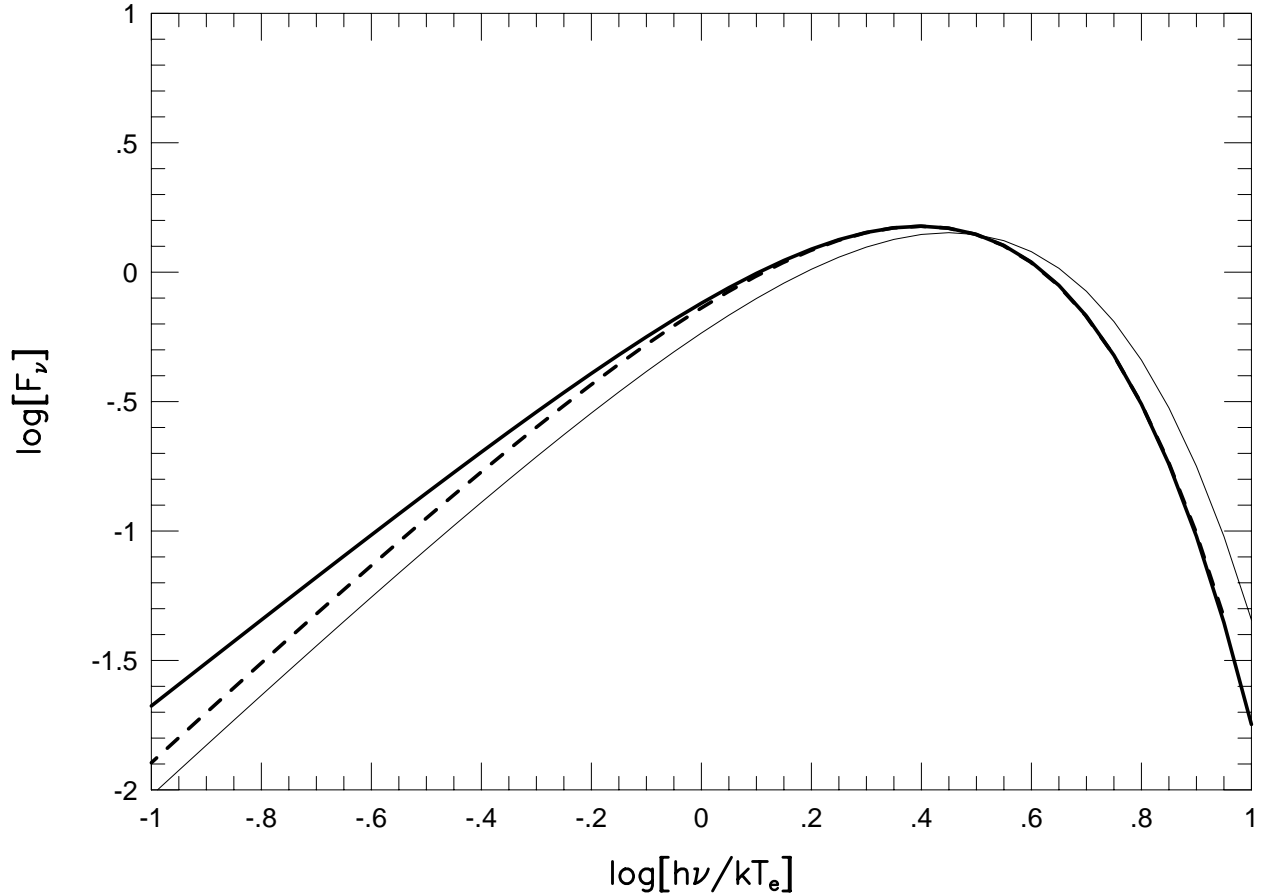
Fig. 6.— Plot of the surface flux for an atmosphere with $\kappa \propto \nu$ (heavy solid line) compared to a blackbody at T_e (light solid line) and a greybody $1.52B_\nu(0.887T_e)$ (dashed line).

matrix, then the source function S which gives a constant flux H is given by $S = \Phi^{-1}H$.

6.4.   Non-grey Problems

What if the opacity is not a constant? Then a reasonable first guess is to use

$$T^4 = \left(\frac{1}{2} + \frac{3}{4}\tau_R \right) T_e^4 \tag{136}$$

where τ_R is the Rosseland mean optical depth. The Rosseland mean absorption coefficient is a *harmonic mean* (so $1/\alpha$ is averaged instead of α) that is weighted by the derivative of the Planck function with respect to temperature:

$$\frac{1}{\alpha_R} = \frac{\int \alpha_\nu^{-1} \frac{\partial B}{\partial T} d\nu}{\int \frac{\partial B}{\partial T} d\nu} \tag{137}$$

The relationship between τ_R and the real depth in the slab has to be determined by

32

integrating the differential equation

$$d\tau_R = \kappa_R \rho dz = \alpha_R dz \tag{138}$$

For a simple example, let us consider a situation where $\kappa_\nu \rho = A\nu$. We know that $\partial B/\partial T \propto T^{-1}x^4 e^x/(e^x - 1)^2$ with $x = h\nu/kT$ so we need to find

$$\langle x \rangle_R = \left[\frac{\int x^{-1}x^4 e^x dx/(e^x - 1)^2}{\int x^4 e^x dx/(e^x - 1)^2} \right]^{-1} \tag{139}$$

These integrals can be done by parts, using $dq = e^x dx/(e^x - 1)^2 = d[-1/(e^x - 1)]$, so $\int x^n e^x dx/(e^x - 1)^2 = \int nx^{n-1}dx/(e^x - 1) = \Gamma(n+1)\zeta(n)$. Thus for $\alpha \propto \nu$ we get

$$\langle x \rangle_R = 4\zeta(4)/\zeta(3) = 3.60 \tag{140}$$

so $\kappa_R \rho = 3.6A(kT/h)$. Thus

$$d\tau_R = \kappa_R \rho dz = (3.6AkT_e/h) \left(\frac{1}{2} + \frac{3}{4}\tau_R \right)^{1/4} dz \tag{141}$$

or

$$\frac{d\tau_R}{\left(\frac{1}{2} + \frac{3}{4}\tau_R \right)^{1/4}} = (3.6AkT_e/h)dz \tag{142}$$

This has the solution

$$\frac{16}{9} \left[\left(\frac{1}{2} + \frac{3}{4}\tau_R \right)^{3/4} - \left(\frac{1}{2} \right)^{3/4} \right] = (3.6AkT_e/h)z \tag{143}$$

which can be rewritten as

$$T = T_e \left(2.03(AkT_e/h)z + \left(\frac{1}{2} \right)^{3/4} \right)^{1/3} \tag{144}$$

If we now want to evaluate the surface flux F_ν, we go back to the formal solution of the RT equation, since with $T(z)$ known we have specified the source function. If τ is the optical depth at ν along a ray with direction μ, the actual depth is

$$z = \frac{\tau\mu}{\kappa\rho} = \frac{\tau\mu}{A\nu} \tag{145}$$

Thus

$$T = T_e \left(2.03(kT_e/h\nu)\tau\mu + \left(\frac{1}{2} \right)^{3/4} \right)^{1/3} \tag{146}$$

and the flux is

$$F_\nu = 2\pi \int_0^1 \int_0^\infty e^{-\tau} B_\nu(T)d\tau\mu d\mu \tag{147}$$

33

The integral over μ can be eased by letting $y = \mu^2$, and the infinite integral over τ can be simplified by letting $w = \exp(-\tau)$. Then

$$F_\nu = \pi \int_0^1 \int_0^1 B_\nu(T) dw dy \tag{148}$$

A FORTRAN program to evaluate this is given below. Note that it takes $h = k = c = T_e = 1$, and drops the π in the definition of the flux:

```
C to compute flux from semi-infinite slab with tau \propto nu
C units: NU and T in units of T_e
C note that kappa_Ross = 3.6*T
REAL NU,NU3,MU
C13 = 1./3.
C1234 = 0.5**0.75
OPEN (UNIT=1,CARRIAGECONTROL='LIST',STATUS='NEW')
DO K=1,21
  NU = 0.1*10.**(0.1*(K-1))
  NU3 = NU**3
  SUM = 0.
  DO J=1,100
    Y = 0.01*(J-0.5)
    MU = SQRT(Y)
    DO I=1,100
      W = 0.01*(I-0.5)
      TAU = -ALOG(W)
      T = (2.025*MU*TAU/NU + C1234)**C13
      SUM = SUM + NU3/(EXP(NU/T)-1.)
    ENDDO
  ENDDO
  SUM = SUM/100**2
  B = NU3/(EXP(NU)-1.)
  WRITE (1,FMT='(F8.5,2F12.8)') NU,B,SUM
ENDDO
STOP
END
```

Figure 6 shows the surface flux computed by this program along with a blackbody at T_e and a cooler greybody with $T = 0.887 T_e$ and emissivity $\epsilon = 1.52$. Because the opacity is less at small frequencies, more radiation escapes at low frequencies, making the spectrum appear colder. But the total flux remains the same, so the emissivity needed to match the observed spectrum is greater than unity. If this were a brown dwarf with radius 0.1 $R_\odot$ and
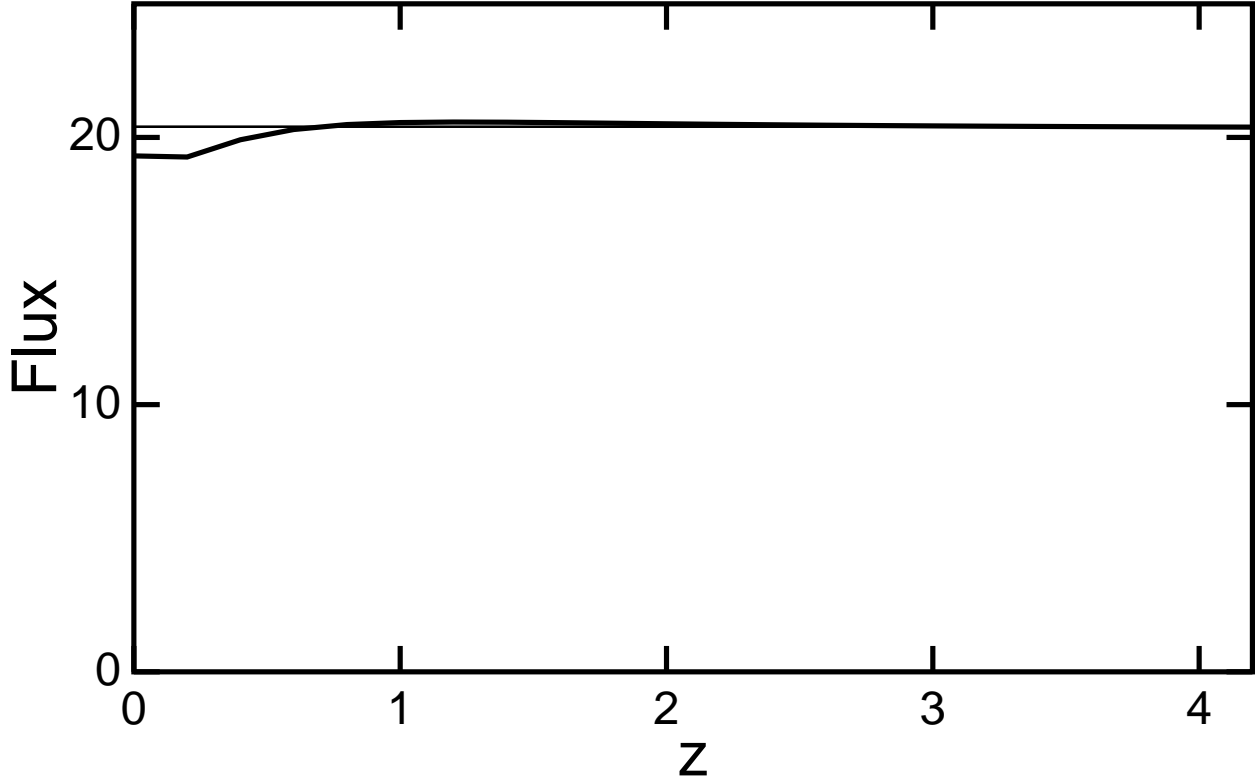
34

Fig. 7.— The bolometric flux versus depth in a non-grey semi-infinite slab with $\alpha_\nu = \nu$, $h = 1$, $k = 1$, $T_e = 1$, and $T = (0.5 + 0.75\tau_R)^{0.25}$.

$T_e = 1000$ K, a blackbody fit would give $T = 887$ K and $R = 0.123$ $R_\odot$. The flux versus depth is shown in Figure 7. The peak error in the flux is about 5% which is no worse than the two stream approximation to the grey problem.

7.   Non-LTE

In some circumstances the approximation that the source function is equal to the Planck function at the local temperature breaks down. One common example of this involves line formation in a *resonance* line. A resonance line is the lowest energy line connected to the ground state of a species. As a result, an ion excited to the upper level of the line will usually decay back down to the ground state by emitting another line photon. Thus the opacity in the line is dominated by scattering, while the continuum is dominated by absorption.
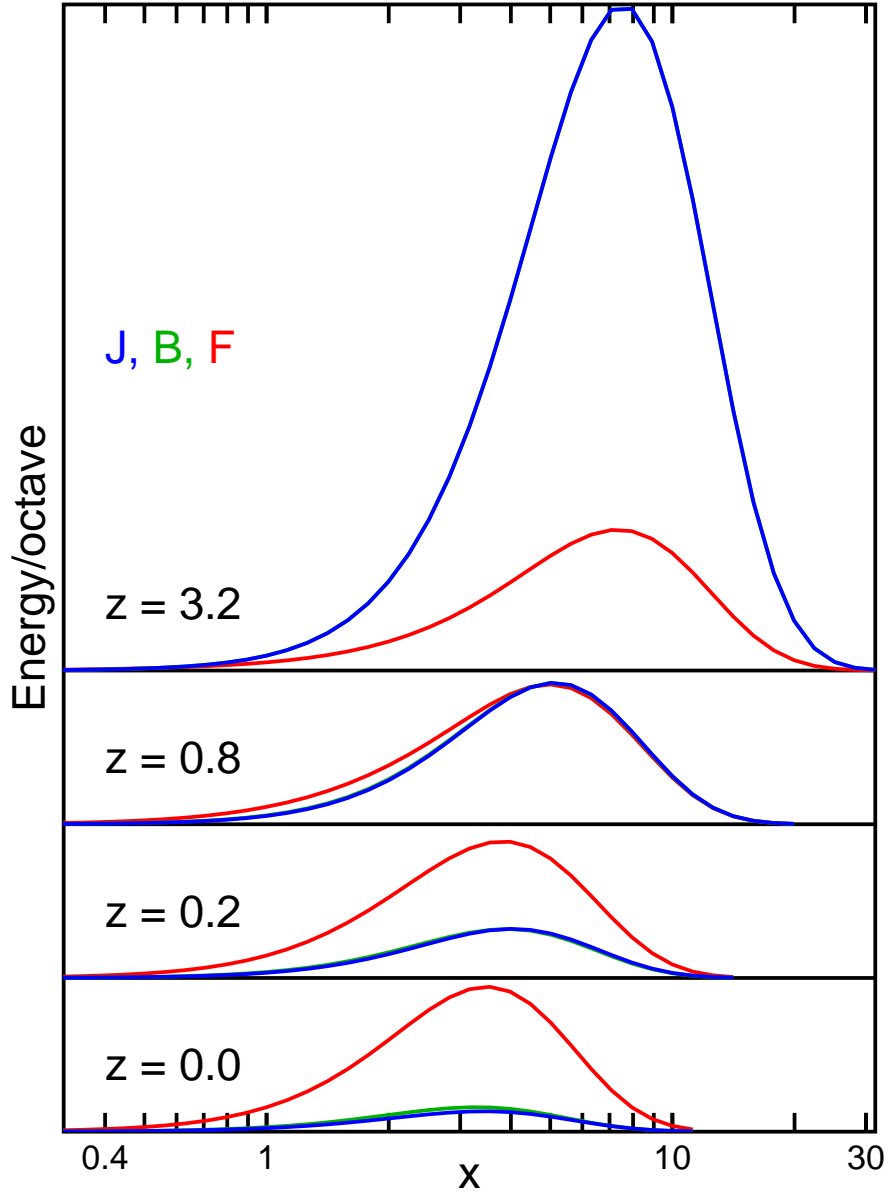
Fig. 8.— Plots of νJ_ν, νB_ν, and νF_ν versus ν for several depths in a non-grey semi-infinite slab with $\alpha_\nu = \nu$, $h = 1$, $k = 1$, $T_e = 1$, and $T = (0.5 + 0.75\tau_R)^{0.25}$. Note how the condition $J = B$ expected in radiative equilibrium is very quickly established, and how the isotropic part of the radiation field (J) becomes much greater than the anisotropic part (F) as the depth increases.

7.1. Two Level Atom

Consider a medium filled with idealized two level atoms, with two levels separated by $E = h\nu$. Let the collisional de-excitation rate be C, while the radiative de-excitation rate is the Einstein coefficient A_{ul}. Since there are only two levels, the subscripts on A are not needed and I will drop them. If electrons are responsible for the collisional de-excitations, then

$$C = n_e \langle \sigma v \rangle_{ul} \tag{149}$$

In thermal equilibrium the rate of collisional excitations $n_l n_e \langle \sigma v \rangle_{lu}$ has to equal the rate of collisional de-excitations $n_u n_e \langle \sigma v \rangle_{ul}$ by the principle of *detailed balance*. In thermal equilibrium

$$\frac{n_u}{n_l} = \frac{g_u}{g_l} \exp(-h\nu/kT) \tag{150}$$

so one gets

$$\frac{\langle \sigma v \rangle_{lu}}{\langle \sigma v \rangle_{ul}} = \frac{n_u}{n_l} = \frac{g_u}{g_l} \exp(-h\nu/kT) \tag{151}$$

In addition to collisional processes one also has spontaneous and stimulated radiative de-excitation, and radiative excitation. The total equation for the time derivative of n_u is

$$\frac{\partial n_u}{\partial t} = n_l n_e \langle \sigma v \rangle_{lu} - n_u n_e \langle \sigma v \rangle_{ul} - A_{ul} n_u - B_{ul} J n_u + B_{lu} J n_l \tag{152}$$

Using $B_{lu} = (g_u/g_l) B_{ul}$ and $B_{ul} = A_{ul}/J_1$ (where $J_1 = 2h\nu(\nu/c)^2$ is the intensity for 1 photon per mode), this becomes

$$\frac{\partial n_u}{\partial t} = n_l (g_u/g_l) e^{-h\nu/kT} C - n_u C + n_l (g_u/g_l) AJ/J_1 - n_u A - n_u AJ/J_1 \tag{153}$$

Assuming $dn_u/dt = 0$ gives *statistical equilibrium* and a population ratio

$$\frac{n_u/g_u}{n_l/g_l} = \frac{e^{-h\nu/kT} C + AJ/J_1}{C + A + AJ/J_1} \tag{154}$$

The source function is found from j/α. The emissivity is

$$j = \frac{n_u A h\nu}{4\pi \Delta\nu} \tag{155}$$

where the power in spontaneous emission has been spread into 4π steradians and a bandwidth $\Delta\nu$ which is just used to give the correct units and cancels out in the end. The absorption coefficient α comes from the absorption cross-section

$$\sigma_\nu = \frac{h\nu B_{lu}}{4\pi \Delta\nu} = \frac{(g_u/g_l) A h\nu}{4\pi J_1 \Delta\nu} \tag{156}$$

and with the correction for stimulated emission we get

$$\alpha_\nu = \sigma_\nu n_l \left(1 - \frac{n_u/g_u}{n_l/g_l}\right) \tag{157}$$
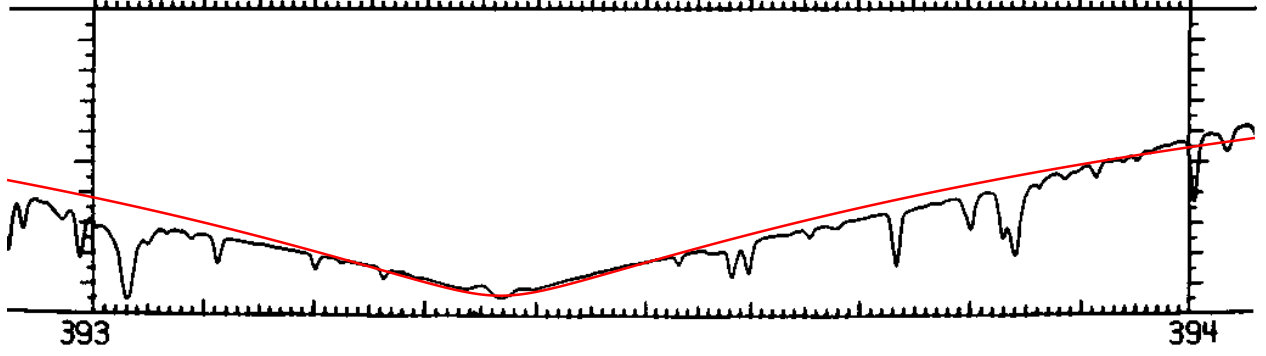
Fig. 9.— The residual intensity in the solar K line. Notice that the line profile is close to an absolute value function.

Taking the ratio gives the source function:

$$S = \frac{j}{\alpha} = J_1 \frac{n_u/g_u}{n_l/g_l} / \left(1 - \frac{n_u/g_u}{n_l/g_l}\right) = \frac{J_1}{\frac{n_l/g_l}{n_u/g_u} - 1} \tag{158}$$

If LTE holds then $(n_l/g_l)/(n_u/g_u) = e^{h\nu/kT}$ and

$$S = \frac{J_1}{e^{h\nu/kT} - 1} = B_\nu(T) \tag{159}$$

But LTE does not hold so we substitute the level ratio from statistical equilibrium giving

$$S = J_1 / \left(\frac{C + A(1 + J/J_1)}{e^{-x}C + AJ/J_1} - 1\right) = \frac{J_1 e^{-x}C + AJ}{(1 - e^{-x})C + A} \tag{160}$$

If we define

$$\epsilon = \frac{(1 - e^{-x})C}{(1 - e^{-x})C + A} \tag{161}$$

we get

$$S = \epsilon B_\nu + (1 - \epsilon)J_\nu \tag{162}$$

For cases where collisional rates are much less than the radiative rates (strong lines and low densities), ϵ can be very small.

7.2. Mixture of Scattering and Absorption

The source function in a pure scattering case is $S_\nu = J_\nu$, while in a pure absorption case it is B_ν. For a mixture of scattering and absorption we have

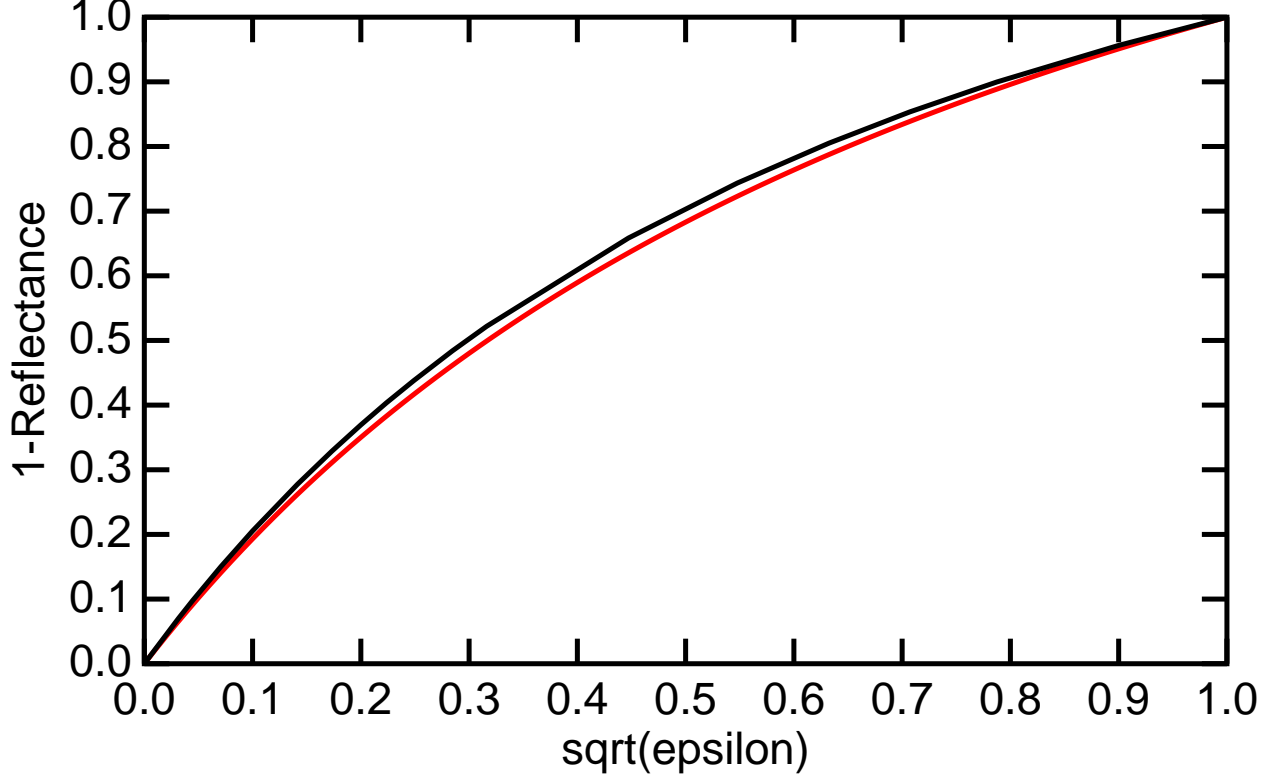$$S_\nu = \epsilon_\nu B_\nu + (1 - \epsilon_\nu)J_\nu \tag{163}$$

38

Fig. 10.— One minus the reflectivity *vs.* the fraction ϵ of absorptions in an isotropic scattering semi-infinite plane-parallel slab from the Monte Carlo code. The red curve is the prediction from Eqn 170.

where ϵ is the ratio of absorption to total extinction:

$$\epsilon_\nu = \frac{\alpha_\nu}{\alpha_\nu + \sigma_\nu} \tag{164}$$

This form for the source function is the same as that derived for the two level atom, but for a mixture of scattering and absorption the scatterers and absorbers can be different species instead of different de-excitation mechanisms.

If the line is strong, the optical depth in the line will be much larger than the optical depth in the continuum, so if the temperature follows $T^4 = T_e^4(0.5 + 0.75\tau_{cont})$, the Planck function will be effectively constant over the line formation region. If the line opacity were true absorption, then the residual intensity in the line would be $\approx B_\nu(0.5^{0.25}T_e)/B_\nu(T_e)$ which is fairly large. But many resonance lines have residual intensities which approach zero. This is caused by the line opacity being dominated by scattering.

Let us solve the equation of radiative transfer with constant B. The equation is

$$\frac{d^2 J}{d\tau^2} = 3(J - S) = 3(J - (\epsilon B + (1 - \epsilon)J)) = 3\epsilon(J - B) \tag{165}$$

39

This has the solution

$$J = B + A\exp(-\sqrt{3\epsilon}\tau) + C\exp(+\sqrt{3\epsilon}\tau) \tag{166}$$

Clearly the coefficient $C = 0$ to prevent the solution from blowing up at large depth. The coefficient A must be set to satisfy the boundary condition $J = 2H = (2/3)dJ/d\tau$. This gives

$$B + A = -(2/3)\sqrt{3\epsilon}A \tag{167}$$

so the solution is

$$A = -\frac{B}{1 + (2/3)\sqrt{3\epsilon}} \tag{168}$$

The outgoing flux is

$$H = \frac{1}{3}\frac{dJ}{d\tau} = \frac{B\sqrt{3\epsilon}}{3 + 2\sqrt{3\epsilon}} \tag{169}$$

Taking the value for $\epsilon = 1$ as the continuum in this model gives the residual intensity is

$$R = \frac{\sqrt{\epsilon}(3 + 2\sqrt{3})}{(3 + 2\sqrt{3\epsilon})} \tag{170}$$

which goes to zero as ϵ goes to zero. In fact, for a strong line with a Lorentz line profile, $\sigma_\nu = A\alpha_c/(1 + (\Delta\nu/w)^2)$, we have $\epsilon \propto (\Delta\nu)^2$, and the residual intensity is $R \propto \sqrt{(\Delta\nu)^2} \propto |\Delta\nu|$. Just this kind of triangular line profile is seen in the solar H and K lines. Figure 9 shows the actual solar K line along with the R given above using a line center σ_ν that is $A = 1250$ times larger than the continuum opacity, and a HWHM (Half Width at Half Maximum) for the scattering of $w = 0.046$ nm. This is a pretty good fit for such a simple model.

Why is the residual intensity of a scattering line so low? In the line, the atmosphere becomes a good reflector with a high albedo. An incident photon will scatter many times before being absorbed, and it has a high probability of escaping back into space. An object with a high albedo has a very low emissivity, so the the radiation from the star in the line is very low. To see how this works out with the equations, consider a cold scattering slab with an incident flux $F_\circ$ in the straight down direction. This flux gives a direct contribution to the mean intensity of

$$J_\circ = \frac{F_\circ e^{-\tau}}{4\pi} \tag{171}$$

There will also be a diffuse scattered radiation field J_s. The source function is

$$S = \epsilon B + (1 - \epsilon)J = (1 - \epsilon)J_s + (1 - \epsilon)\frac{F_\circ e^{-\tau}}{4\pi} \tag{172}$$

where the Planck function has been set to zero since the slab is cold. The Eddington approximation gives

$$\frac{d^2 J_s}{d\tau^2} = 3(J_s - S) = 3\epsilon J_s - 3(1 - \epsilon)\frac{F_\circ e^{-\tau}}{4\pi} \tag{173}$$

40

This has a solution of the form $J_s = C\exp(-\tau) + A\exp(-\sqrt{3\epsilon}\tau)$ and the coefficient C is given by

$$C = \frac{-3(1-\epsilon)}{1-3\epsilon}\frac{F_\circ}{4\pi} \tag{174}$$

The Eddington boundary condition $J = 2H = (2/3)dJ/d\tau$ gives $A+C = (2/3)[-C-\sqrt{3\epsilon}A]$, so

$$A = \frac{-(5/3)C}{1+(2/3)\sqrt{3\epsilon}} = \frac{5(1-\epsilon)}{(1-3\epsilon)[1+(2/3)\sqrt{3\epsilon}]}\frac{F_\circ}{4\pi} \tag{175}$$

and the surface mean intensity is

$$J(\tau=0) = A+C = \frac{(2-2\sqrt{3\epsilon})}{(1-3\epsilon)[1+(2/3)\sqrt{3\epsilon}]}\frac{F_\circ}{4\pi} = \frac{2F_\circ/4\pi}{(1+\sqrt{3\epsilon})[1+(2/3)\sqrt{3\epsilon}]} = 2H(\tau=0) \tag{176}$$

Since the reflected flux F is $4\pi H$, the reflectivity is

$$\frac{4\pi H}{F_\circ} = \frac{1}{(1+\sqrt{3\epsilon})[1+(2/3)\sqrt{3\epsilon}]} \tag{177}$$

and it goes to unity as $\epsilon \to 0$.

7.3. Monte Carlo Calculation

It is easy to do a Monte Carlo calculation of the reflection from a scattering atmosphere. The pseudocode would be:
Repeat many times for statistics:
Start a ray at $z=0$ with $\mu = \sqrt{U(0,1)}$
Step: compute path length $s = -\ln[U(0,1)]$; $z \leftarrow z+s\mu$
If $z < 0$ then REFLECTED
If $U(0,1) < \epsilon$ then ABSORBED
$\mu = 2U(0,1) - 1$; go to Step

$U(0,1)$ is a uniform random number generator for the interval 0..1. The ray starting procedure generates $p(\mu)d\mu = 2\mu d\mu$ using the square root.

An actual Monte Carlo code in FORTRAN follows:

```
C     MC-scat.f
      REAL MU
      REAL ABST(14)/.001,.002,.005,.01,.02,.035,.05,.07,.1,.15,.2,.35,.5,.7/
      EXTERNAL RAN
```

```
        CALL RANDOMIZE
        DO 60 IABS=1,14
          AB = ABST(IABS)
          NREF = 0
          DO 40 IMC=1,10000000
            Z = 0 ! start ray at surface with random down direction
            MU = SQRT(RAN(IR,JR))
20          U = RAN(IR,JR)          ! find new scattering position
            IF (U.LE.0.) GO TO 20
            S = -ALOG(U)
            Z = Z+S*MU
            IF (Z.LT.0.) GO TO 30            ! REFLECTED
            IF (RAN(IR,JR).LT.AB) GO TO 40  ! ABSORBED
              MU = 2*RAN(IR,JR)-1  ! pick new direction and loop
              NSCAT = NSCAT+1
              GO TO 20
30            NREF = NREF+1
40          CONTINUE
60          WRITE (*,FMT='(F5.3,F8.5,2H P)') AB,(1.E-7*NREF)
        STOP
        END
```

Figure 10 plots the output of this code compared to Eqn 170. The agreement is quite good.

8.   Atomic Structure

Chapter 10 of R&L

In this chapter the *semi-classical* theory of atomic transitions is worked out. The E&M field is treated classically while the atom is treated as a quantum system. Since a classical field has no zero-point motion, the semi-classical theory can not be used to compute spontaneous emission directly. Instead we calculate the rate of induced transitions which give the Einstein B coefficients, and then use the standard relationship between the Einstein B's and the Einstein A to get the spontaneous rate.

8.1. Electric Dipole Transitions

R&L take a roundabout way to reach the electric dipole approximation to the interaction Hamiltonian. But if we assume the wavelengths are much longer than the size of an atom, and that the electrons are moving much slower than c, it is clear that the energy shift caused by an incident traveling E&M wave is just

$$H_1 = \sum q\phi = \sum q\vec{r} \cdot \vec{E}(t) = \vec{d} \cdot \vec{E}(t) \tag{178}$$

The time dependent Schrödinger equation is

$$i\hbar\frac{\partial \Psi}{\partial t} = H\Psi \tag{179}$$

We work in the basis of stationary states of the atom in which the unperturbed Hamiltonian is diagonal. Each of these stationary states has a time dependence of $e^{-iE_n t/\hbar}$. If we start in state 1 and look for transitions to state 2, then we see that the probability amplitude, evaluated as a first-order perturbation, is

$$\langle 2|\Psi\rangle = \frac{1}{i\hbar}\int_0^T \langle 2|H_1|1\rangle dt = \frac{1}{i\hbar}\int_0^T \langle 2|\vec{d}|1\rangle \cdot \vec{E}(t)dt \tag{180}$$

Let $\hat{l}$ be the direction of the polarization of the electric field. Then the probability of a transition, which is the square of the complex magnitude of the amplitude, is

$$Prob(1 \rightarrow 2) = \frac{|\langle 2|\vec{d}|1\rangle \cdot \hat{l}|^2}{\hbar^2}\left|\int e^{i\omega t}E(t)dt\right|^2 \tag{181}$$

where the frequency is given by $\hbar\omega = E_2 - E_1$. The integral of the electric field times $e^{i\omega t}$ over time is the Fourier transform of the field:

$$\hat{E}(\omega) = \frac{1}{2\pi}\int e^{i\omega t}E(t)dt \tag{182}$$

We can also write this probability as

$$Prob(1 \rightarrow 2) = \int_0^T B_{12}J_\nu dt \tag{183}$$

where B_{12} is the Einstein B and J_ν is the mean intensity. The total fluence in the E&M pulse is

$$\int_0^T \int 4\pi J_\nu d\nu dt = \int_0^T S dt = \frac{c}{4\pi}\int_0^T E(t)^2 dt = c\int_0^\infty |\hat{E}(\omega)|^2 d\omega \tag{184}$$

where S is the Poyntyng vector. We see that

$$|\hat{E}(\omega)|^2 = \frac{2}{c}\int_0^T J_\nu dt \tag{185}$$

since $d\omega = 2\pi d\nu$ and thus

$$
\begin{aligned}
Prob(1 \to 2) &= \int_0^T B_{12} J_\nu dt \\
&= \frac{|\langle 2|\vec{d}|1\rangle \cdot \hat{l}|^2}{\hbar^2} \left| 2\pi \hat{E}(\omega) \right|^2 \\
&= \frac{8\pi^2}{\hbar^2 c} |\langle 2|\vec{d}|1\rangle \cdot \hat{l}|^2 \int_0^T J_\nu dt
\end{aligned}
\tag{186}
$$

Thus the Einstein B is

$$
B_{12} = \frac{8\pi^2}{\hbar^2 c} |\langle 2|\vec{d}|1\rangle \cdot \hat{l}|^2
\tag{187}
$$

Since the dipole moment and the polarization are usually randomly oriented, and the spherical average of $\cos^2 \theta$ is $1/3$, we get finally

$$
B_{12} = \frac{8\pi^2}{3\hbar^2 c} |\langle 2|\vec{d}|1\rangle|^2
\tag{188}
$$

which agrees with R&L Eqn(10.27). This is the correct value when the levels are not degenerate. If there are several degenerate levels then the formula becomes

$$
B_{12} = \frac{8\pi^2}{3\hbar^2 c} \frac{1}{g_1} \sum_{n=1}^{g_1} \sum_{m=1}^{g_2} |\langle 2_m|\vec{d}|1_n\rangle|^2
\tag{189}
$$

where we have averaged over the initial states and summed over final states. If we evaluate this for the reverse transition from $2 \to 1$, we have the same sum of degenerate levels but divide by g_2 instead of g_1, so we get

$$
g_1 B_{12} = g_2 B_{21}
\tag{190}
$$

So far it hasn't mattered which level has higher energy but for spontaneous emission it does. Assuming that level 2 has higher energy, we get

$$
A_{21} = 2h\nu \left(\frac{\nu}{c}\right)^2 B_{21} = \frac{64\pi^4 \nu^3}{3hc^3} \frac{1}{g_2} \sum_{n=1}^{g_1} \sum_{m=1}^{g_2} |\langle 2_m|\vec{d}|1_n\rangle|^2
\tag{191}
$$

8.2. Permanent Molecular Dipoles

For most atomic transitions the radial integral changes drastically between different angular momentum values, but for molecules with a permanent dipole moment μ the only things that change in the pure rotational series are the J and m of the rotational states. In this case there are some simple regularities. The overall transition rate between J, m and

44

J', m' is given by

$$B_{Jm,J'm'} = \frac{8\pi^2\mu^2}{3\hbar^2 c} \left[\left(\int Y_{Jm}^* \cos\theta \, Y_{J'm'} d\Omega \right)^2 \right.$$

$$\left. + \left(\int Y_{Jm}^* \sin\theta \cos\phi \, Y_{J'm'} d\Omega \right)^2 + \left(\int Y_{Jm}^* \sin\theta \sin\phi \, Y_{J'm'} d\Omega \right)^2 \right] \quad (192)$$

The quantity within the brackets is always a rational number with denominator $(2J+1)(2J'+1)$, and it vanishes unless $J' = J+1$ or $J' = J-1$ and $m' = m-1$, m or $m+1$. Table 1 shows the values of this transition strength for the individual subtransitions. The values have been multiplied by 105 to clear the denominators. Notice that the matrix is symmetric and that each row and column sums to unity (the $J = 3$ rows are incomplete because the transitions to $J = 4$ have not been included).

	J	0	1	1	1	2	2	2	2	2	3	3	3	3	3	3	3
	m	0	-1	0	1	-2	-1	0	1	2	-3	-2	-1	0	1	2	3
J'	m'																
0	0	0	35	35	35	0	0	0	0	0	0	0	0	0	0	0	0
1	-1	35	0	0	0	42	21	7	0	0	0	0	0	0	0	0	0
1	0	35	0	0	0	0	21	28	21	0	0	0	0	0	0	0	0
1	1	35	0	0	0	0	0	7	21	42	0	0	0	0	0	0	0
2	-2	0	42	0	0	0	0	0	0	0	45	15	3	0	0	0	0
2	-1	0	21	21	0	0	0	0	0	0	0	30	24	9	0	0	0
2	0	0	7	28	7	0	0	0	0	0	0	0	18	27	18	0	0
2	1	0	0	21	21	0	0	0	0	0	0	0	0	9	24	30	0
2	2	0	0	0	42	0	0	0	0	0	0	0	0	0	3	15	45
3	-3	0	0	0	0	45	0	0	0	0	0	0	0	0	0	0	0
3	-2	0	0	0	0	15	30	0	0	0	0	0	0	0	0	0	0
3	-1	0	0	0	0	3	24	18	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	9	27	9	0	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	18	24	3	0	0	0	0	0	0	0
3	2	0	0	0	0	0	0	0	30	15	0	0	0	0	0	0	0
3	3	0	0	0	0	0	0	0	0	45	0	0	0	0	0	0	0

Table 1: 105 times the value of $\langle Jm| \sin\theta \cos\phi |J'm'\rangle^2 + \langle Jm| \sin\theta \sin\phi |J'm'\rangle^2 + \langle Jm| \cos\theta |J'm'\rangle^2$.

If we do the summation over all the m's and m''s of the value in the [ ]'s in Eqn(192) we get $\max(J, J')$. When we evaluate the Einstein A we average over the initial states so we divide by $2J_u + 1$. This gives us

$$A_{J+1,J} = \frac{64\pi^4\nu^3\mu^2}{3hc^3} \frac{J+1}{2J+3} \quad (193)$$

If we write the radiated power instead we get

$$P_{rad} = h\nu A = \frac{64\pi^4 h\nu^4 \mu^2}{3hc^3} \frac{J+1}{2J+3} = \frac{4\omega^4 \mu^2}{3c^3} \frac{J+1}{2J+3} \tag{194}$$

In the classical limit this should agree with the Larmor formula:

$$P_{rad} = \frac{2}{3} \frac{q^2 a^2}{c^3} = \frac{2}{3} \frac{\ddot{\mu}^2}{c^3} = \frac{2}{3} \frac{\omega^4 \mu^2}{c^3} \tag{195}$$

The classical limit would be as $J \to \infty$ so $(J+1)/(2J+3) \to 1/2$ and our result does agree with the Larmor formula.

The energy of pure rotational states is given by $BJ(J+1)$ plus centrifugal distortion terms. Then the frequency of the $J+1 \to J$ transition is $2B(J+1)/h$. Thus the Einstein A's are given by

$$A_{J+1,J} = \frac{3A_{10}(J+1)^4}{2J+3} \tag{196}$$

This formula is pretty accurate for CO, and it says that A_{43} is 85.3 times higher than A_{10}. For this reason the excitation of the high-J lines in CO is generally considered to be an indication of high density which leads to high collision rates which can compete with the rapid radiative de-excitation.

The dipole moment of CO is 0.122 Debye, or 0.122×10^{-18} esu-cm. One electron charge at 1 Å radius would be a dipole of 4.8 Debye, so CO has a rather small dipole moment. This gives an Einstein $A_{10} = 8.8 \times 10^{-8}$ sec^{-1}. The cyanogen radical CN has a much larger dipole moment, $\mu = 1.45$ Debye, or 1.45×10^{-18} esu-cm. This gives an $A_{10} = 1.2 \times 10^{-5}$ sec^{-1}, for a radiative life-time less than a day. Thus the excitation temperature of the 1-0 transition in interstellar CN is mainly determined by the temperture of the CMB.

8.3. Harmonic Oscillator

If we consider a three dimensional harmonic oscillator with spring constant k, mass m and charge e, then we have $\omega = \sqrt{k/m}$. The raising and lowering operators

$$a_x^\dagger = (p_x/\sqrt{2m} + ix\sqrt{k/2})/\sqrt{h\nu}$$

and

$$a_x = (p_x/\sqrt{2m} - ix\sqrt{k/2})/\sqrt{h\nu}$$

(and similarly for y and z) allow us to find the matrix elements of the dipole moment. If we denote states by $n_x n_y n_z$ then there are three states $|100\rangle$, $|010\rangle$ and $|001\rangle$ with energy $h\nu$ above the ground state. Linear combinations of these states form a $J=1$ triplet. The x component of the dipole operator is $d_x = -ie\sqrt{h\nu/(2k)}(a_x^\dagger - a_x)$. This gives $\langle 000|d_x|100\rangle =$

$ie\sqrt{h\nu/(2k)}$ because $\langle n|a|n+1\rangle = \sqrt{n+1}$. The other matrix elements of d_x are zero: $\langle 000|d_x|010\rangle = \langle 000|d_x|001\rangle = 0$. The d_y matrix connects the $|010\rangle$ state to the ground state, and the d_z matrix connects the $|001\rangle$ state to the ground state. Thus

$$\sum_{n=1}^{g_1}\sum_{m=1}^{g_2}|\langle 2_m|\vec{d}|1_n\rangle|^2 = \frac{e^2h\nu}{2k}\times 3 \tag{197}$$

and the Einstein A is

$$A_{10} = \frac{64\pi^4\nu^3}{3hc^3}\frac{1}{g_1}\sum_{n=1}^{g_0}\sum_{m=1}^{g_1}|\langle 1_m|\vec{d}|0_n\rangle|^2 = \frac{64\pi^4\nu^3}{3hc^3}\frac{e^2h\nu}{2k} = \frac{2\omega^4e^2}{3kc^3} = \frac{2\omega^2e^2}{3mc^3} \tag{198}$$

where I have changed the 2_m notation to 1_m and 1_n to 0_n because the levels are the 0th and 1st levels. A classical oscillator with energy $E = kx^2_{max}/2$ has an average squared acceleration of $\langle a^2\rangle = \omega^4 x^2_{max}/2$ so the radiated power by the Larmor formula is $P = (2e^2\langle a^2\rangle/3c^3)$ and the rate of decay of the energy is

$$\frac{P}{E} = \frac{2e^2\omega^4}{3kc^3} = A_{10} \tag{199}$$

Note that the *oscillator strength* of this system, defined (see R&L Eq(10.29a)) via

$$f_{01} = \frac{h\nu mc}{4\pi^2e^2}B_{01} = \frac{8\pi^2h\nu mc}{4\pi^2e^23\hbar^2c}\frac{1}{g_0}\sum_{n=1}^{g_0}\sum_{m=1}^{g_1}|\langle 1_m|\vec{d}|0_n\rangle|^2 = \frac{8\pi^2h\nu mc}{4\pi^2e^23\hbar^2c}\frac{e^2h\nu}{2k}\times 3 = 1 \tag{200}$$

This is the origin of the term oscillator strength.

The integral of the absorption cross-section for the oscillator is given by

$$\int\sigma(\nu)d\nu = \frac{\lambda^2}{8\pi}\frac{g_u}{g_l}A_{10} = \frac{\lambda^2}{8\pi}\frac{g_u}{g_l}\frac{2e^2\omega^2}{3mc^3} = \frac{\pi e^2}{3mc}\frac{g_u}{g_l} \tag{201}$$

The higher levels have higher Einstein A's with $A_{n,n-1} = nA_{10}$. We discussed this for a one dimensional oscillator in §5.2 but we can also work it out for a three dimensional oscillator. There are six states that have $E = 2h\nu$: these are $|200\rangle$, $|110\rangle$, $|101\rangle$, $|020\rangle$, $|011\rangle$ and $|002\rangle$. The matrix element of x between any two states is non zero only if $n'_x = n_x \pm 1$ and $n'_y = n_y$ and $n'_z = n_z$. The squares of the non-zero elements of are given by the larger of n_x and n'_x. Then the individual sublevel line strengths are given by

	$|200\rangle$	$|110\rangle$	$|101\rangle$	$|020\rangle$	$|011\rangle$	$|002\rangle$
$|100\rangle$	2	1	1	0	0	0
$|010\rangle$	0	1	0	2	1	0
$|001\rangle$	0	0	1	0	1	2

and the sum over all sublevels is 12 for the 2-1 transition as opposed to 3 for the 1-0 transition. But the statistical weight for the $n = 2$ level is $g_2 = 6$ as opposed to $g_1 = 3$ so the Einstein A which is $\propto [\sum\sum|\langle 2_m|\vec{d}|1_n\rangle|^2]/g_2$ is exactly $A_{21} = (12/6)[A_{10}/(3/3)] = 2A_{10}$.

8.4. Line Broadening

Atomic and molecular lines have three kinds of broadening: the natural line width, collisional broadening, and Doppler broadening. The natural line width can be computed by considering the radiated field from an excited atom as a function of time. This will be given by

$$E(t) \propto \exp(-i\omega_\circ t)\exp(-At/2) \tag{202}$$

The $At/2$ term, when squared, gives an exponential decay in the power with a rate given by A. The Fourier transform of this wavetrain is

$$\hat{E}(\omega) \propto [i(\omega - \omega_\circ) - A/2]^{-1} \tag{203}$$

so the power goes like

$$|\hat{E}(\omega)|^2 \propto \frac{1}{\Delta\omega^2 + (A/2)^2} \tag{204}$$

and the FWHM in rad/sec is A and the FWHM in Hertz is $A/2\pi$.

In general a state can decay by many different lines, and the exponential decay of the E field will depend on the sum all the Einstein A's. Thus we should compute

$$\gamma_u = \sum_l A_{ul} \tag{205}$$

where the sum goes over all lower levels l. Thus the line profile goes like

$$\phi(\nu) \propto \frac{1}{\Delta\nu^2 + (\gamma_u/4\pi)^2} \tag{206}$$

.

Of course the lower level will also have a width given by γ_l. This will also broaden the lines. In order to allow for this broadening we need to convolve two Lorentzians. The Fourier transform of $1/(x^2 + a^2)$ is $\propto \exp(-|k|a)$, and the FT of $1/(x^2 + b^2)$ is $\propto \exp(-|k|b)$, so the product of the FTs is $\propto \exp(-|k|(a + b))$, which is the FT of the convolution and obviously just another Lorentzian with a width given by $a + b$. Hence the natural broadening of a line gives a Lorentzian profile with a FWHM in Hz of $(\gamma_u + \gamma_l)/2\pi$.

For collisional broadening one has a random process where the phase of the $\exp(-i\omega_\circ t)$ term is reset randomly with a mean time between resets of τ_c. One or more phase resets guarantee at zero expectation value between $E(t)^*$ and $E(t - \tau)$. The probability of no resets before τ is $\exp(-\tau/\tau_c)$ so the auto-correlation function is

$$A(\tau) \propto \exp(-\tau/\tau_c) \tag{207}$$

The power spectrum is the Fourier transform of the autocorrelation function, so this will give another Lorentzian. The net result is a Lorentzian profile with a FWHM in Hz of

$(\gamma_u + \gamma_l + 2/\tau_c)/2\pi$. For the Lyman-$\alpha$ line, the lifetime of the excited state is 1.6 ns, so $\gamma_u = 6 \times 10^8$ sec^{-1}. This leads to a FWHM of 10^8 Hz.

Doppler broadening due to thermal and turbulent motions will lead to a Gaussian broadening profile. The velocity distribution in a Maxwellian is

$$p(v_x) \propto \exp(-\tfrac{1}{2}mv^2/kT) \tag{208}$$

Thus the standard deviation of the radial velocity is $\sigma = \sqrt{kT/m}$. For hydrogen at 10^4 K this is 9 km/sec. The FWHM of Lyman-α caused by this dispersion is $\Delta\nu = 2.36\nu_\circ\sqrt{kT/mc^2} = 1.8 \times 10^{11}$ Hz. While this is much wider than the FWHM of the Lorentzian, the line profile will still be dominated by the Lorentzian in the wings. Even though the convolution of two Lorentzians with widths a and b is another Lorentzian with width $a+b$, and the convolution of two Gaussians with widths a and b is another Gaussian with width $\sqrt{a^2 + b^2}$, the convolution of a Gaussian and a Lorentzian is a new kind of function called a Voigt profile.

The Voigt function is given by

$$V(x; \sigma, \gamma) = \int G(x'; \sigma)L(x - x'; \gamma)dx' \tag{209}$$

with $G = \exp(-\tfrac{1}{2}(x/\sigma)^2)/(\sigma\sqrt{2\pi})$ and $L = \gamma/(\pi(x^2 + \gamma^2))$. The Gaussian has a FWHM $f_G = \sigma\sqrt{8\ln 2} \approx 2.36\sigma$, and the Lorentzian has a FWHM $f_L = 2\gamma$, and the FWHM of the Voigt is approximately $f_V = 0.5346f_L + \sqrt{0.2166f_L^2 + f_G^2}$. When $f_G >> f_L$, as is usually the case in astronomical lines, the FWHM is determined by the Doppler broadening. But the far wings will still be dominated by the Lorentzian, with $V(x) \approx L(x)$ for $|x| >> \sigma$.

For Lyman-α, with $A = 6 \times 10^8$ sec^{-1}, the integral $\int \sigma d\nu = 0.011$ cm^{2}Hz, and the Lorentzian HWHM is $\gamma = 5 \times 10^7$ Hz. Thus

$$\sigma = \frac{0.011 \times 5 \times 10^7}{\pi(\Delta\nu^2 + (5 \times 10^7)^2)} \tag{210}$$

This cross-section is 7×10^{-11} cm^2 at line center, which is huge, and it is $\sigma \approx 1.75 \times 10^5/\Delta\nu^2$ in the wings. For a column density of $N_H = 10^{20}$ cm^{-2}, optical depth unity will be reached at $\Delta\nu = 4.2 \times 10^{12}$ Hz, which is 2.1 Å off-center. The integral $\int(1 - \exp(-1/x^2))dx = 3.345$ so the equivalent width of the absorption line produced by $N_H = 10^{20}$ cm^{-2} will be 7 Å.

The *curve-of-growth* describes how the equivalent width of a line grows as a function of the column density of a line. This will depend on the line formation circumstances: an intervening interstellar cloud will have a different curve-of-growth than a stellar atmosphere. For the interstellar cloud, either absorption or scattering will remove a photon from the transmitted flux. In a stellar atmosphere, a scattered photon has a good chance of eventually escaping from the star. In the interstellar case, the equivalent width is given by

$$W_\nu = \int (1 - \exp(-\tau))d\tau \approx N_l \int \sigma(\nu)d\nu - \mathcal{O}(\tau^2) + \ldots \tag{211}$$

Now

$$\int \sigma(\nu)d\nu = \frac{\lambda^2}{8\pi}\frac{g_u}{g_l}A_{ul} = \frac{\lambda^2}{8\pi}\frac{2h\nu}{\lambda^2}B_{lu} = \frac{2h\nu}{8\pi}\frac{4\pi^2 e^2}{h\nu mc}f_{lu} = \frac{\pi e^2}{mc}f_{lu} \tag{212}$$

Now N_l is given by $N_l = Ng_l \exp(-E_l/kT)/Z$ so the equivalent width scales with Ngf for low τ.

Once the line becomes optically thick in the center, the curve-of-growth starts to saturate. The linear growth of W with Ngf turns over toward a $\sqrt{Ngf}$ form when the damping wings are optically thick. When the central optical depth is small for a Lorentzian line with HWHM $= \gamma$ the equivalent width is $\pi\gamma\tau_c$. When $\tau_c >> 1$, the damping wings reach $\tau = 1$ at $\Delta\nu = \gamma\sqrt{\tau_c}$ so the equivalent width is $3.345\gamma\sqrt{\tau_c}$ since $\int(1 - \exp(-1/x^2))dx = 3.345$. These two branches are equal when $\pi\tau_c = 3.345\sqrt{\tau_c}$ or $\tau_c = (3.345/\pi)^2 = 1.13$. When Doppler broadening is added in, the central optical depth is lower, and the linear part of the curve-of-growth stays on the same line but continues to a higher value before the line center gets optically thick. There is then a "flat" part of the curve-of-growth until the square root part is reached. The flat part is not completely flat, but only grows logarithmically with Ngf so it is nearly flat. The damping wings are not affected by the Doppler broadening, so the square root part of the curve is not affected.

When discussing line shapes, the width of the Doppler broadening Gaussian is usually specified by a b parameter, which is the half-width at $1/e$ in the Gaussian. Thus $b = \sqrt{2}\sigma$. This difference between b and σ is related to an alternate definition of the Voigt function,

$$H(a, u) = \frac{a}{\pi}\int \frac{\exp(-y^2)dy}{a^2 + (u - y)^2} \tag{213}$$

given in R&L Eqn(10.77). In terms of this function the line profile is

$$\phi(\nu) = \frac{H(a, u)}{\sqrt{\pi}\Delta\nu_D} \tag{214}$$

with $a = \Gamma/(4\pi\Delta\nu_D)$ and $u = (\nu - \nu_\circ)/\Delta\nu_D$.

9.   Diatomic Molecules

Chapter 11 in R&L.

The Born-Oppenheimer approximation calculates the electronic energy for various states with the nuclei in a molecule at a fixed separation r. Since the electrostatic potential is not spherically symmetric, the total angular momentum is not a good quantum number. But the potential is axially symmetric, so the component of the angular momentum along the internuclear separation axis is conserved. The orbital momentum along this axis is called Λ. If $\Lambda = 0, 1, 2, 3, \ldots$ the states are called Σ, Π, Δ, Φ, ...

For the nth electronic state, this gives an effective potential energy $V_n(r)$. For a bound state, there will be a minimum of this effective potential at some r_n. This will lead to a moment of inertia $I_n = \mu r_n^2$, where the reduced mass is $\mu = M_1 M_2/(M_1 + M_2)$. The rotational energy will be $E = \hbar^2 K(K+1)/2I$ where K is the rotational quantum number. This is usually written as $E_{rot} = BK(K+1)$ with $B = \hbar^2/2I$. The allowed lines with $\Delta K = 1$ have photon energies $h\nu_{K+1,K} = B[(K+1)(K+2) - K(K+1)] = 2B(K+1)$.

There will be a ladder of vibrational states that are close to equally spaced as in a harmonic oscillator. The spring constant in the nth electronic state is $k_n = \partial^2 V_n(r)/\partial r^2$ evaluated at r_n. The vibrational spacing is then $\Delta E_v = \hbar\sqrt{k_n/\mu}$. The same spring constant controls the centrifugal distortion of the rotational energies. The kinetic energy due trnasverse velocities is $L^2/(2\mu r^2)$ which leads to a centrifugal pseudo-force $L^2/(\mu r^3)$. This then gives the stretch of the internuclear separation,

$$k_n \Delta r = \frac{L^2}{\mu r^3} \tag{215}$$

This gives a change in the moment of inertia,

$$\frac{\Delta I}{I} = 2\frac{\Delta r}{r} = \frac{2L^2}{k_n \mu r^4} = \frac{2\mu \hbar^2 K(K+1)}{k_n I^2} \tag{216}$$

In addition, the stretching of the bond gives an energy change of

$$\Delta E = \frac{1}{2}k_n \Delta r^2 = \frac{L^4}{2k_n \mu^2 r^6} \tag{217}$$

The rotational energy is

$$E_{rot} = \frac{L^2}{2I} = \frac{L^2}{2I(1+\Delta I/I)} = \frac{L^2}{2I} - \frac{L^2}{2I}\frac{2L^2}{k_n \mu r^4} = \frac{L^2}{2I} - \frac{L^4}{k_n \mu^2 r^6} \tag{218}$$

The stretching energy plus the rotational energy becomes

$$\begin{aligned} E(L) &= \frac{L^2}{2I} - \frac{L^4}{2k_n \mu^2 r^6} \\ &= \frac{\hbar^2 K(K+1)}{2I} - \frac{\hbar^4 K^2(K+1)^2}{2k_n \mu^2 r^6} \\ &= BK(K+1) - DK^2(K+1)^2 + \dots \end{aligned} \tag{219}$$

where $B = \hbar^2/2I$ and the correction term for rotational distortion D is

$$D = \frac{\hbar^4}{2k_n \mu^2 r^6} = \frac{\hbar^4 \mu}{2k_n(\mu r^2)^3} = 4\left(\frac{\hbar^2}{2\mu r^2}\right)^3\left(\frac{\mu}{\hbar^2 k_n}\right) = \frac{4B^3}{(\hbar\omega_v)^2} \tag{220}$$

For CO, with $B = 1.9225$ cm^{-1} and $\hbar\omega_v = 2169.756$ cm^{-1}, one estimates $D = 0.00000604$ cm^{-1} compared to the measured 0.00000612 cm^{-1}. For H_2, with $B = 60.853$ and $\hbar\omega_v =$

4401.121, both in cm^{-1}, the estimated D is 0.0465 cm^{-1} compared to the measured 0.0471 cm^{-1}. The agreement of Eqn(219) with the empirical data is quite good, but note that Eqn(11.29) in R&L gives a distortion correction (the $K^2(K+1)^2$ term) that is a factor two too large because they did not include the energy due to stretching the bond.

The frequencies of the $K + 1 \rightarrow K$ lines are given by

$$
\begin{aligned}
h\nu_{K+1,K} &= B(K+1)(K+2) - BK(K+1) - D(K+1)^2(K+2)^2 + DK^2(K+1)^2 \\
&= 2B(K+1) - 4D(K+1)^3 \qquad (221)
\end{aligned}
$$

The selection rules for pure rotational lines are

1. $d \neq 0$

2. $\Delta K = \pm 1$

9.1.   Vibrational-Rotational Bands

If the vibrational level changes during a transition, the selection rules are:

1. $d(d)/dr|_{r=r_\circ} \neq 0$

2. $\Delta v = \pm 1$

3. $\Delta K = \pm 1$ for $\Lambda = 0$
   $\Delta K = 0, \pm 1$ for $\Lambda \neq 0$

The energy of the lines are

$$
h\nu = \hbar\omega_{v'v} + B_{v'}K'(K'+1) - B_v K(K+1) \qquad (222)
$$

where $\hbar\omega_{v'v}$ is the difference in energy between the v' and v vibrational levels. If $K' = K - 1$, one has the "R" branch. This has frequencies

$$
h\nu = \hbar\omega_{v'v} + B_{v'}(K^2 - K) - B_v(K^2 + K) = \hbar\omega_{v'v} - (B_{v'} + B_v)K + (B_{v'} - B_v)K^2 \quad (223)
$$

If $K' = K$, one has the "Q" branch. This has frequencies

$$
h\nu = \hbar\omega_{v'v} + (B_{v'} - B_v)K(K+1) \qquad (224)
$$

If $K' = K + 1$, one has the "P" branch. This has frequencies

$$
h\nu = \hbar\omega_{v'v} + B_{v'}(K^2 + 3K + 2) - B_v(K^2 + K) = \hbar\omega_{v'v} + (B_{v'} + B_v)(K+1) + (B_{v'} - B_v)(K+1)^2 \tag{225}
$$

Because of the K^2 terms, the line frequencies can become stationary at some value of K in the P or R branch. This causes a pile up of lines at the stationary frequency, creating a *band head*. The band head will occur at $2(K+1) = -((B_{v'} + B_v)/(B_{v'} - B_v))$ in the P branch if this is positive, or at $2K = ((B_{v'} + B_v)/(B_{v'} - B_v))$ in the R branch if this is positive. One of these quantities will always be positive, so band heads are a universal phenomenon.

Usually in vibrational transitions, the mean radius and hence the rotational constant B do not change much, so the band head may occur at such a high value of K that it is not very prominent.

9.2. Electronic Transitions

If there is a non-zero dipole moment matrix element between two electronic states, there will be a series of vibrational rotational bands associated with transitions between the two electronic states. In the case of electronic transitions, the mean separation of the nuclei is usually quite different between the upper and lower states, so the B values are quite different, and the band heads are very obvious.

10. Ionization Loss

Consider a particle of charge Ze moving with Lorentz factor γ through a medium with free electrons. Let the particle have an impact parameter b relative to an electron. We will assume the motion of the electron is negligible during the passage of the ion, so the force can be calculated without following the motion of the electron. The force gives the acceleration which we integrate to get the final velocity imparted to the electron. The field seen at the initial position of the electron is

$$\vec{E}(t) = \frac{-\gamma Ze(vt, b, 0)}{(b^2 + (\gamma vt)^2)^{3/2}} \tag{226}$$

Note that the ion is at $y = b$ so the the electron is at $y = -b$ relative to the ion, and this gives the minus sign above. Thus the acceleration of the electron is

$$\vec{a}(t) = \frac{\gamma Ze^2}{m_e} \frac{(vt, b, 0)}{(b^2 + (\gamma vt)^2)^{3/2}} \tag{227}$$

the integrated velocity change is all in the y direction:

$$\Delta v_y = \frac{\gamma Ze^2}{m_e} \int \frac{b\,dt}{(b^2 + (\gamma vt)^2)^{3/2}} = \frac{Ze^2}{m_e bv} \int \frac{dx}{(1 + x^2)^{3/2}} \tag{228}$$

with $x = \gamma vt/b$. Setting $x = \tan\theta$ gives $dx = \sec^2\theta d\theta = (1+x^2)d\theta$ so

$$\int \frac{dx}{(1+x^2)^{3/2}} = \int_{-\pi/2}^{\pi/2} \cos\theta d\theta = 2 \tag{229}$$

Thus the velocity change of the electron is

$$\Delta\vec{v} = \frac{2Ze^2\hat{j}}{m_e bv} \tag{230}$$

Then the kinetic energy of the electron after the passage of the ion is

$$\mathrm{KE} = \frac{1}{2}m_e v^2 = \frac{2Z^2 e^4}{m_e b^2 v^2} \tag{231}$$

This energy is lost by the ion. The total energy loss is obtained by integrating over the number of collisions with impact parameter b that occur while the ion travels a distance dx. This number of collisions is $2\pi b\, db\, dx\, n_e$, giving

$$-\frac{dE}{dx} = \frac{4\pi Z^2 e^4 n_e}{m_e v^2} \int \frac{db}{b} \tag{232}$$

This logarithmically divergent integral is very common is problems involving the Coulomb interaction. We need to identify the minimum and maximum impact parameters that are relevant. For a solid material the electrons can only be treated as free if the effective frequency of the electric field pulse of the ion is greater than the ionization threshold of the material. The effective frequency is $\omega \approx \gamma v/b$. Setting $\hbar\omega > I$ gives

$$b_{max} = \frac{\hbar\gamma v}{I} \tag{233}$$

where I is the effective ionization energy for the material. The minimum impact parameter is set by the larger of two limits: the classical limit where $\Delta v = v$ which gives

$$b_{min}^C = \frac{2Ze^2}{m_e v^2} \tag{234}$$

or the quantum limit where the orbital angular momentum of the electron in the ion's rest frame, $\gamma m_e vb$, is $\hbar$:

$$b_{min}^Q = \frac{\hbar}{\gamma m_e v} \tag{235}$$

We need to use the larger of these limits. If we set them equal we get

$$\frac{2Ze^2}{mv^2} = \frac{\hbar}{\gamma mv}$$
$$\frac{v}{c} = 2\gamma Z\frac{e^2}{\hbar c} = \frac{2\gamma Z}{137} \tag{236}$$

Usually the velocities are greater than $c/137$ so the quantum limit is used. Then we get

$$-\frac{dE}{dx} = \frac{4\pi Z^2 e^4 n_e}{m_e v^2} \ln\left(\frac{\gamma^2 m_e v^2}{I}\right) \tag{237}$$

Typically $m_e v^2$ is a few hundred keV and I is a few tens of electron volts so the logarithm is about $10 + 2\ln\gamma$. For practical purposes dE/dx is usually converted into energy loss per mass column density using $n_e = Z_m \rho / A_m m_H$ where Z_m and A_m are the material atomic number and atomic weight, and ρ is the material density. This gives

$$-\frac{dE}{\rho dx} = \frac{4\pi Z^2 e^4}{m_H m_e v^2} \frac{Z_m}{A_m} \ln\left(\frac{\gamma^2 m_e v^2}{I}\right) \tag{238}$$

For almost all materials $Z_m/A_m \approx 0.5$, so we can write

$$
\begin{aligned}
-\frac{dE}{\rho dx} &\approx \frac{2\pi Z^2 e^4}{m_H m_e c^2} \beta^{-2} \frac{2 Z_m}{A_m}\left(10 + 2\ln\gamma\right) \\
&= \frac{3}{4} \frac{8\pi e^4}{3 m_e^2 c^4} \frac{m_e c^2}{m_H} \frac{2 Z_m}{A_m} \frac{Z^2}{\beta^2}\left(10 + 2\ln\gamma\right) \\
&\approx \frac{7.5 \sigma_T m_e c^2}{m_H} \frac{2 Z_m}{A_m} \frac{Z^2}{\beta^2}\left(1 + 0.2\ln\gamma\right) \\
&= 1.5 \frac{2 Z_m}{A_m} \frac{Z^2}{\beta^2}\left(1 + 0.2\ln\gamma\right) \text{ MeV/(gm/cm}^2) \tag{239}
\end{aligned}
$$

The minimum value of $(1 + 0.2\ln\gamma)/\beta^2$ is 1.36 at $\beta = 0.96$, so the minimum ionization loss is about $2Z^2$ MeV/(gm/cm^2).

11. Bremsstrahlung

When the electron is moving instead of the ion, then the electron loses energy by radiation. This radiation is called *bremsstrahlung* which is German for braking radiation. It is also called free-free emission because the electron goes from one unbound orbit around the ion to another unbound orbit. But since the Larmor formula only applies in the electron's rest frame, we will analyze exactly the same collision that we considered for ionization loss. Thus

$$\vec{a}(t) = \frac{\gamma Z e^2}{m_e} \frac{(vt, b, 0)}{(b^2 + (\gamma vt)^2)^{3/2}} \tag{240}$$

and the power radiated is $P = (2/3)e^2 a^2 / c^3$. Since the collision has an effective duration of $b/\gamma v$, the total energy radiated is approximately

$$W = \int P dt \approx \frac{b}{\gamma v} \frac{2\gamma^2 Z^2 e^6}{3 m_e^2 b^4 c^3} \approx \frac{2\gamma Z^2 e^6}{3 m_e^2 b^3 c^3 v} \tag{241}$$
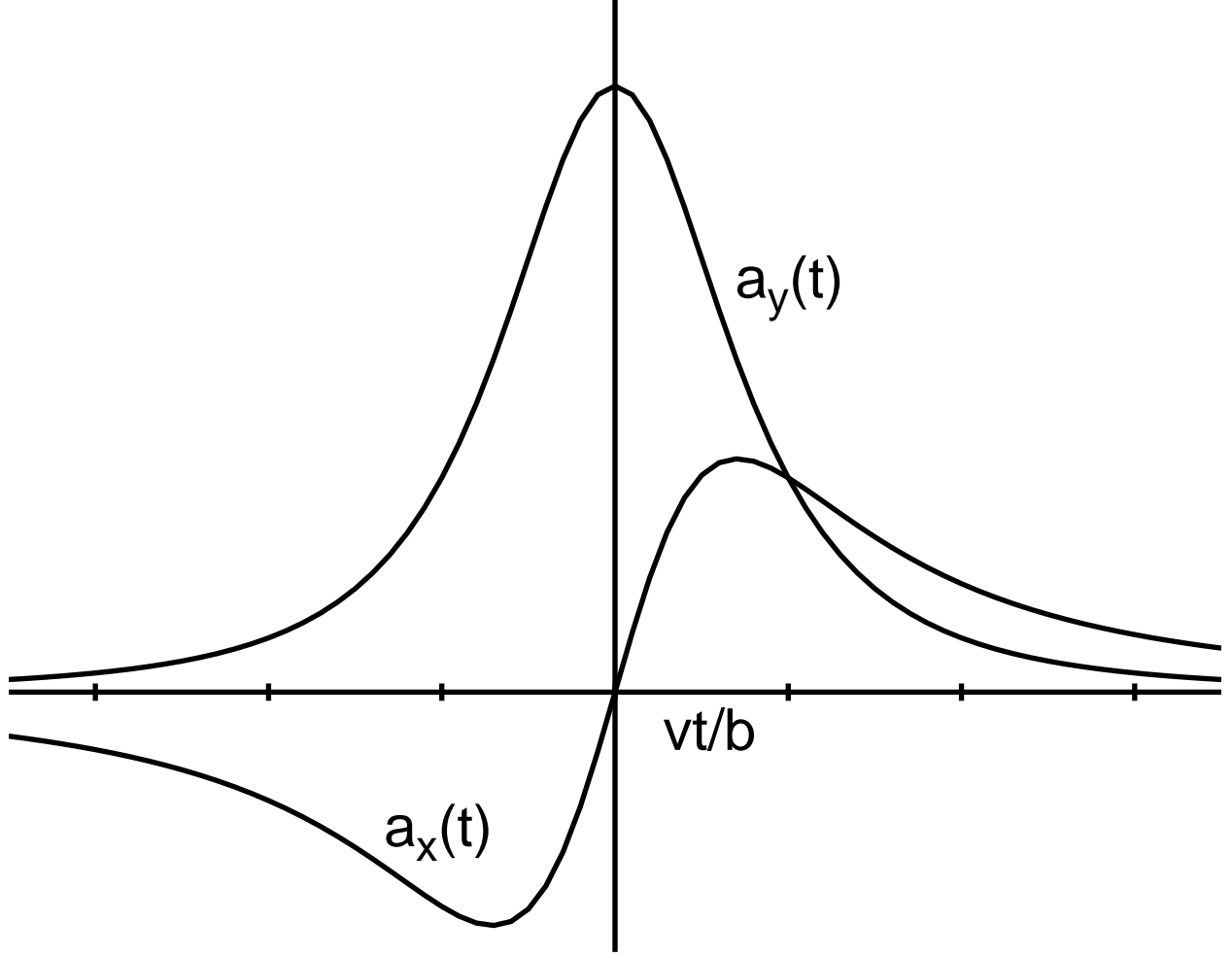
Fig. 11.— x and y axis acceleration vs time.

If we integrate this over the distribution of impact parameters, we find that the integral goes like $\int b^{-2}db$ and is dominated by close encounters. Since we have not treated the orbits in a self-consistent way, and since the closest encounters need a quantum treatment, we cannot accurately evaluate the total radiated power.

We can do much better with the power per unit frequency, however. By Parseval's Theorem

$$\int_{-\infty}^{\infty} a(t)^2 dt = 2\pi \int_{-\infty}^{\infty} |\hat{a}(\omega)|^2 d\omega = 4\pi \int_0^{\infty} |\hat{a}(\omega)|^2 d\omega \qquad (242)$$

where $\hat{a}(\omega) = (2\pi)^{-1} \int e^{i\omega t} a(t) dt$. Since the total power can be written several ways:

$$W = \frac{2e^2}{3c^3} \int a(t)^2 dt = \frac{8\pi e^2}{3c^3} \int_0^{\infty} |\hat{a}(\omega)|^2 d\omega = \int_0^{\infty} \frac{dW}{d\omega} d\omega \qquad (243)$$

we can identify the power per unit frequency as

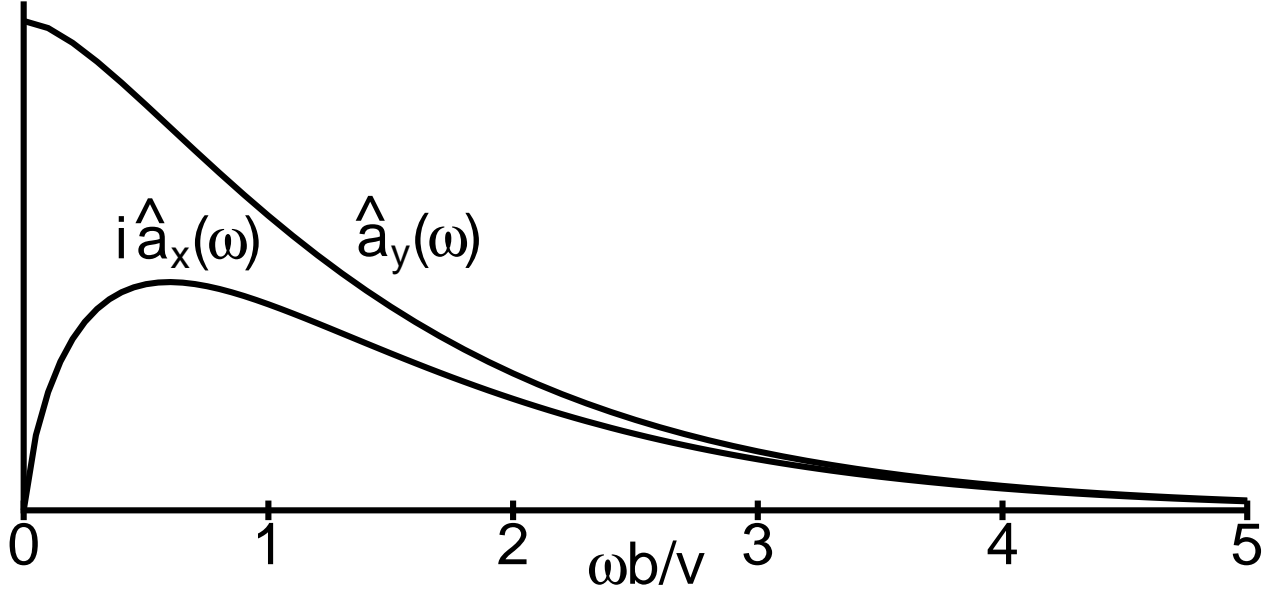$$\frac{dW}{d\omega} = \frac{8\pi e^2}{3c^3} |\hat{a}(\omega)|^2 \qquad (244)$$

56

Fig. 12.— Fourier transforms of the x and y axis accelerations.

Thus we only need to find the Fourier transforms of the acceleration components:

$$\hat{a}_y(\omega) = \frac{1}{2\pi}\frac{\gamma Ze^2}{m_e b^2}\int \frac{\exp(i\omega t)dt}{(1+(\gamma vt/b)^2)^{3/2}}$$

$$\hat{a}_x(\omega) = \frac{1}{2\pi}\frac{Ze^2}{m_e b^2}\int \frac{(\gamma vt/b)\exp(i\omega t)dt}{(1+(\gamma vt/b)^2)^{3/2}}$$

$$= \frac{v}{ib}\frac{d\hat{a}_y(\omega)}{d\omega} \tag{245}$$

We can easily see that the singularity closest to the real t axis in the upper half of the complex t plane is at $t = ib/\gamma v$. Since this is the start of a branch cut, we cannot use contour integration to easily evaluate the integral, but we can predict that the large ω behaviour of the integrals will go like $\exp(-|\omega|b/\gamma v)$. The low ω limit can be found as well: for the x-component $\hat{a}_x(0) = 0$ because the function being integrated is odd, while for the y-component we get

$$\hat{a}_y(\omega) = \frac{1}{2\pi}\frac{Ze^2}{m_e bv}\int \frac{dx}{(1+x^2)^{3/2}} = \frac{Ze^2}{\pi m_e bv} \tag{246}$$

By looking in the NBS Mathematical Functions book we find an integral representation of a modified Bessel function of the second kind, giving

$$2kK_1(k) = \int \frac{e^{ikx}dx}{(1+x^2)^{3/2}} \tag{247}$$

Since $K_1(x) \approx x^{-1}$ for small x, and $K_1(x) \approx \sqrt{\pi/2x}e^{-x}$ for large x, we find that our simple determinations of the limiting behaviour of $\hat{a}_y(\omega)$ are correct.

Now we can evaluate $dW/d\omega$ at low frequencies, which is produced only by a_y. This gives

$$\frac{dW}{d\omega} = \frac{8\pi e^2}{3c^3} \left| \frac{Ze^2}{\pi m_e bv} \right|^2 = \frac{8Z^2 e^6}{3\pi m_e^2 b^2 c^3 v^2} \tag{248}$$

The energy radiated per Hertz is given by

$$\frac{dW}{d\nu} = 2\pi \frac{dW}{d\omega} = \frac{16Z^2 e^6}{3m_e^2 b^2 c^3 v^2} \tag{249}$$

Now we have to integrate this result over the number of collisions per second. In the lab frame, with the ions at rest, the collisions go as $2\pi\, n_i\, b\, db\, vdt$. In the electron's frame, there is an extra factor of γ. This factor of γ can be viewed as coming from the Lorentz contraction of the x-sepeartion of the ions or it can be viewed as coming from time dilation: 1 second of the electron's clock is γ seconds in the lab frame. This gives the energy per unit time per unit frequency in the electron's rest frame as

$$\frac{dW}{d\nu dt} = \frac{32\pi\gamma Z^2 e^6 n_i}{3m_e^2 c^3 v} \int \frac{db}{b} \tag{250}$$

We need to define the limits of the logarithmically divergent integral over b. The outer limit is set not by the ionization potential of the material but rather by the frequency being observed: we need the argument of the $\exp()$ function in the power, $2|\omega|b/\gamma v$, to be < 1 or

$$b < b_{max} = \frac{\gamma v}{2|\omega|} \tag{251}$$

The quantum inner limit is

$$b_{min}^Q = \frac{\hbar}{\gamma m_e v} \tag{252}$$

so

$$\int \frac{db}{b} = \ln\left(\frac{b_{max}}{b_{min}}\right) = \ln\left(\frac{\gamma^2 m_e v^2}{2\hbar\omega}\right) = \ln\Lambda \tag{253}$$

Note that in the non-relativistic limit this automatically gives a cutoff when $\hbar\omega = m_e v^2/2$ which is very reasonable since an electron cannot radiate a photon with more energy than it starts with.

11.1. Thermal Bremsstrahlung

Thermal bremsstrahlung occurs from non-relativistic plasmas where the electrons have a Maxwellian velocity distribution. We need to average the radiation as a function of velocity over this distribution. First let us write the radiated power per unit volume:

$$\frac{dW}{d\omega dV dt} = \frac{16e^6 n_e n_i Z^2}{3m_e^2 c^3 v} \ln\left(\frac{b_{max}}{b_{min}}\right) \tag{254}$$

58

The Gaunt factor g is defined as

$$g_{ff}(v, \omega) = \frac{\sqrt{3}}{\pi} \ln \left(\frac{b_{max}}{b_{min}} \right) \tag{255}$$

so we get

$$\frac{dW}{d\omega dV dt} = \frac{16\pi e^6}{3\sqrt{3}m_e^2 c^3} \frac{n_e n_i Z^2}{v} g_{ff}(v, \omega) \tag{256}$$

Now we need the normalized integral of $1/v$ over the Maxwellian for velocities such that the kinetic energy of the electron is greater than $\hbar\omega$. The Maxwellian is $v^2 \exp(-mv^2/(2kT))dv$ so we want

$$\int_{v(min)}^\infty v^{-1} v^2 \exp \left(\frac{-mv^2}{2kT} \right) dv = \frac{kT}{m_e} \int_{mv(min)^2/2kT}^\infty e^{-x} dx = \frac{kT}{m_e} \exp \left(\frac{-h\nu}{kT} \right) \tag{257}$$

with $x = mv^2/(2kT)$. The normalization we need is given by

$$\int_0^\infty v^2 \exp \left(\frac{-mv^2}{2kT} \right) dv = \sqrt{\frac{2kT}{m}} \frac{kT}{m} \int_0^\infty \sqrt{x} e^{-x} dx \tag{258}$$

This integral

$$\int_0^\infty \sqrt{x} e^{-x} dx = \Gamma(1.5) = (0.5)! = \frac{\sqrt{\pi}}{2} \tag{259}$$

Therefore the normalized integral of $1/v$ over the Maxwellian is

$$\frac{\int_{v(min)}^\infty v e^{-x} dv}{\int_0^\infty v^2 e^{-x} dv} = \sqrt{\frac{2m}{\pi kT}} \exp \left(\frac{-h\nu}{kT} \right) \tag{260}$$

Thus we get

$$\frac{dW}{d\omega dV dt} = \frac{16\pi e^6}{3\sqrt{3}m_e^2 c^3} n_e n_i Z^2 \sqrt{\frac{2m}{\pi kT}} \exp \left(\frac{-h\nu}{kT} \right) \overline{g_{ff}} \tag{261}$$

or in terms of ν:

$$\frac{dW}{d\nu dV dt} = \frac{32\sqrt{2}\pi^{3/2} e^6}{3\sqrt{3}m_e^{3/2} k^{1/2} c^3} n_e n_i Z^2 T^{-1/2} \exp \left(\frac{-h\nu}{kT} \right) \overline{g_{ff}} = 4\pi j_\nu(ff) \tag{262}$$

A practical formula is

$$4\pi j_\nu(ff) = (6.8 \times 10^{-38} \text{ erg/sec/cm}^3/\text{Hz}) n_e n_i Z^2 T^{-1/2} \exp \left(\frac{-h\nu}{kT} \right) \overline{g_{ff}} \tag{263}$$

where T is in Kelvin. Note that the Gaunt factor $\overline{g_{ff}}(T, \nu)$ is the average of the Gaunt factor $g_{ff}(v, \nu)$:

$$\overline{g_{ff}}(T, \nu) = \frac{\int_{v(min}^\infty v e^{-x} g_{ff}(v, \nu) dv}{\int_{v(min}^\infty v e^{-x} dv} \tag{264}$$

For $10^{-4} < h\nu/kT <$ a few, the range of the Gaunt factor is $1 < \overline{g_{ff}}(T, \nu) < 5$. It shows a slow logarithmic increase toward low frequencies due to the ω^{-1} dependence of b_{max}/b_{min}, and this causes the freqency dependence of $j_\nu(ff)$ to be close to $\nu^{-0.1}$ instead of ν^0.

11.2. Total Free-free Cooling Power

The total power radiated by thermal bremsstrahlung is given by

$$
\frac{dW}{dt\,dV} = \int_0^\infty \frac{dW}{d\nu dV dt}d\nu = \frac{kT}{h}\frac{32\sqrt{2}\pi^{3/2}e^6}{3\sqrt{3}m_e^{3/2}k^{1/2}c^3}n_e n_i Z^2 T^{-1/2}\overline{g_B}
$$
$$
= (1.4\times10^{-27}\ \text{erg/sec/cm}^3)n_e n_i Z^2 T^{1/2}\overline{g_B} \tag{265}
$$

where T is in Levin. The bolometric Gaunt factor is in the range $1.1 < \overline{g_B} < 1.5$.

11.3. Free-free Absorption

Kirchoff's Law gives us

$$
j_\nu(ff) = \alpha_\nu(ff)B_\nu(T) = \frac{1}{4\pi}\frac{dW}{d\nu dV dt} \tag{266}
$$

Therefore

$$
\alpha_\nu(ff) = \frac{4e^6}{3mhc}\sqrt{\frac{2\pi}{3km}}T^{-1/2}n_e n_i Z^2 \nu^{-3}(1 - e^{-h\nu/kT})\overline{g_{ff}} \tag{267}
$$

Thus

$$
\alpha_\nu(ff) \propto n_e n_i Z^2 \overline{g_{ff}}
\begin{cases}
\nu^{-2}T^{-3/2} & \text{for } h\nu << kT \\
\nu^{-3}T^{-1/2} & \text{for } h\nu >> kT
\end{cases} \tag{268}
$$

The Rosseland mean opacity is needed when deep inside a star. This is roughly the opacity at $h\nu = 4kT$ for a smoothly varying opacity. This means that $\exp(h\nu/kT)$ is large, so

$$
\alpha_R(ff) \propto T^{-1/2}(kT/h)^{-3}n_e n_i Z^2 \propto n^2 T^{-7/2} \tag{269}
$$

In practical units this is

$$
\alpha_R(ff) = (1.7\times10^{-25}\ \text{cm}^{-1})\,T^{-7/2}Z^2 n_e n_i \overline{g_R} \tag{270}
$$

If air with $n = 3\times10^{19}$ cm^{-3} were singly ionized at $T = 10^4$ K, this opacity would be $\alpha_R(ff) = 1.7$ cm^{-1}: so you would not be able to see the end of your nose!

When the opacity is given in terms of κ in cm^2/gm then since $\kappa\rho = \alpha \propto \rho^2$, we have

$$
\kappa \propto \rho/T^{7/2} \tag{271}
$$

This is known as Kramer's law for opacity.

60

11.4. Relativistic Bremsstrahlung

When $\gamma > 1$ we have to transform $dW/d\nu dt$ from the electron's rest frame in which we have evaluated it into the lab frame. The radiation in the rest frame has zero net 3-momentum, so the total power is the same in all frames. But the maximum frequency in the lab frame is $\nu'_{max} = \gamma \nu_{max}$ in the rest frame. Thus the power per unit frequency in the lab frame is reduced by a factor of γ^{-1}, but there was a factor of γ^{+1} in the collision rate due to the Lorentz contraction of the ion distribution. Thus the power radiated per electron per unit frequency in the lab frame is

$$\left(\frac{dW}{d\omega}\right)' = \frac{16Z^2 e^6 n_i}{3m_e^2 vc^3} \ln\left(\frac{b_{max}}{b_{min}}\right) \tag{272}$$

As we did in the non-relativistic case, we set $b_{max} = \gamma v/2\omega$. We set b_{min} so that the logarithm goes to zero when $\hbar\omega = \gamma m_e c^2$. This gives $b_{min} = \hbar/m_e c$, the reduced Compton wavelength of the electron. We get the total energy loss by integrating $dW/d\omega$ from 0 to $E/\hbar$, giving

$$-\frac{dE}{dt} = \frac{16Z^2 e^6 n_i E}{3m_e^2 c^4 \hbar} \left\langle \ln\left(\frac{b_{max}}{b_{min}}\right) \right\rangle \tag{273}$$

It is more common to express this energy loss per unit mass column density instead of per unit time. Then $n_i = \rho/Am_H$ and $dx = cdt$, so

$$
\begin{aligned}
-\frac{dE}{\rho dx} &= \frac{16}{3}\left(\frac{Z^2}{A}\right)\frac{e^6}{m_H m_e^2 \hbar c^5} E \left\langle \ln\left(\frac{b_{max}}{b_{min}}\right)\right\rangle \\
&= \left[\frac{2}{\pi}\frac{e^2}{\hbar c}\frac{\sigma_T}{m_H}\left(\frac{Z^2}{A}\right)\left\langle \ln\left(\frac{b_{max}}{b_{min}}\right)\right\rangle\right] E = E/\xi_\circ \tag{274}
\end{aligned}
$$

The quantity $\xi_\circ$ is known as the radiation length. It is 36.5 gm/cm^2 for air, 58 gm/cm^2 for hydrogen, and 5.8 gm/cm^2 for lead. The short radiation length for lead means that a given mass column density of lead is more effective at stopping relativistic electrons (and γ-rays with $E > 1$ MeV) than the same mass column density of other materials. A practical formula for the radiation length is

$$\xi_\circ = \frac{716A \text{ gm/cm}^2}{Z(Z+1.3)\left[\ln\left(\frac{183}{Z^{1/3}}\right) + \frac{1}{8}\right]} \tag{275}$$

11.5. Pair Production

The relativistic bremsstrahlung reaction $e^- + Z \to Z + e^- + \gamma$ and the pair production reaction $\gamma + Z \to Z + e^- + e^+$ are identical under the crossing symmetry where a particle is replaced by its time-reversed anti-particle. A photon is its own anti-particle. This means that the ratio of pair production to relativistic bremsstrahlung is the same as the ratio of the
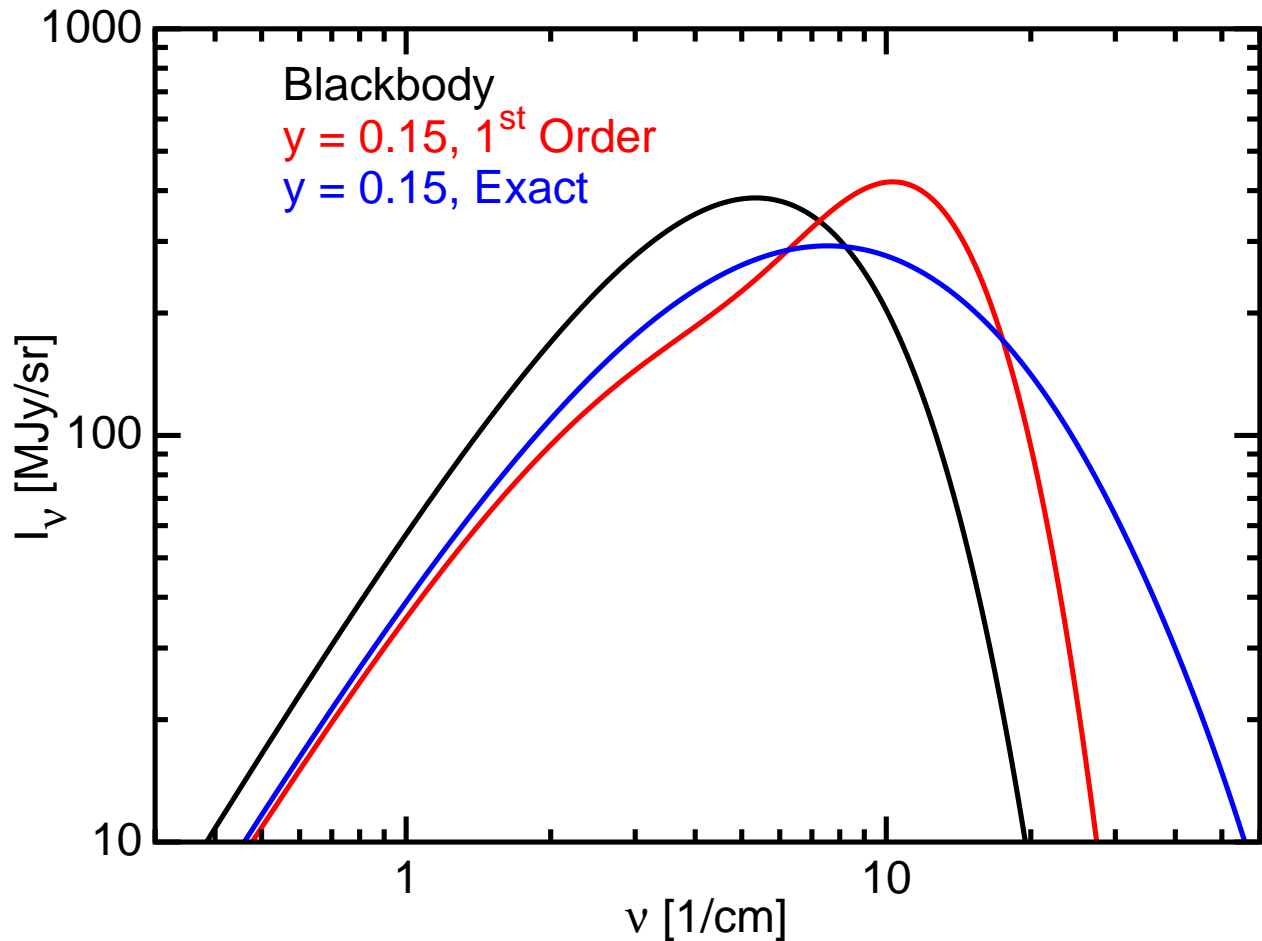
Fig. 13.— An input blackbody, and an S-Z distortion for $y = 0.15$ computed incorrectly using the 1st order result outside of its range of validity (in red), and the correct result using the convolution with a Gaussian in s in blue.

phase space available for the outgoing particles. As long as the γ-ray energy is $>> 1$ MeV, there are two relativistic particles in each final state, so the rates are the same. This means that the radiation length is also the mean distance for pair production by high energy γ-rays.

12. Sunyaev-Zel'dovich Effect

The Sunyaev-Zel'dovich effect is the distortion of an input spectrum by scattering off moving electrons. The CMB spectrum is the usual input spectrum, and the electrons are usually taken to be non-relativistic, giving restrictions

$$hv << kT << m_e c^2 \tag{276}$$

These lead to substantial simplifications. When a photon collides with a stationary electron, its wavelength gets longer due to the original flavor Compton shift, increasing by an amount

of order the Compton wavelength, $\lambda_c = h/mc = 2.4 \times 10^{-10}$ cm. If the photon wavelengths are much longer than this, the wavelength shifts are all due to Doppler shifts, and just multiply the input frequency by a factor that is independent of the input wavelength. If kT is much less than $m_e c^2$, then non-relativistic Maxwellian velocity distributions and Doppler shifts can be used. In a typical application of the SZ effect, CMB photons with energies $h\nu \approx 1$ meV are scattered by an electron population with $kT \approx 10$ keV, so $h\nu$ is 5×10^8 times smaller than $m_e c^2$, but a typical electron has an energy only 25 times smaller than $m_e c^2$. Thus the $kT << m_e c^2$ restriction is more likely to be violated.

Consider an electron moving with $\beta = v/c$ through an isotropic distribution of photons with frequency ν.

A photon approaching with angles $\pi - \theta$ to the electron's velocity has 4-momentum

$$p^\mu = \frac{h\nu}{c}(1; -\cos\theta, -\sin\theta, 0) \tag{277}$$

if we make the x axis the electron's velocity vector and put the photon direction in the xy plane. Then the photon momentum in the electron's rest frame is

$$p^{\mu\prime} = \frac{h\nu}{c}(\gamma(1 + \beta\mu); -\gamma(\beta + \mu), -\sin\theta, 0) \tag{278}$$

Now let the photon scatter off the electron, leaving with angle θ' to the x axis. For now we ignore the true Compton shift, so the energy of the outgoing photon is $h\nu'' = h\nu' = (\gamma(1 + \beta\mu))h\nu$. Note that $\theta' = 0$ corresponds to a photon traveling in the same direction as the electron, while $\theta = 0$ corresponds to a photon traveling in the opposite direction as the electron. The 4 momentum after scattering is

$$p^{\mu\prime\prime} = \frac{h\nu''}{c}(1; \cos\theta', \sin\theta', 0) \tag{279}$$

Now let $\mu' = \cos\theta'$. Then the 4 momentum after scattering back in the lab frame is

$$p^{\mu\prime\prime\prime} = \frac{h\nu''}{c}(\gamma(1 + \beta\mu'); \gamma(\beta + \mu'), \sin\theta', 0) \tag{280}$$

We finally get the ratio of the lab frame energy after scattering to the lab frame energy before scattering:

$$\frac{\nu'''}{\nu} = \gamma^2(1 + \beta\mu)(1 + \beta\mu') \tag{281}$$

We want to find the average value of this ratio, which will tell us the total energy transfer from the electron to the photons, and also the typical frequency of the output photons. In order to find the average, we need to know the probability distributions for the scattering angles μ and μ'.

A first guess would be that $p(\mu)$ should be a uniform distribution since this corresponds to an isotropic initial photon field because $d\Omega = 2\pi d\mu$. But head-on collisions ($\mu = 1$)

will happen more frequently than stern chase collisions ($\mu = -1$). In fact, the rate of photon collisions has exactly the same Doppler shift factor as the photon frequency, so $p(\mu) \propto \gamma(1 + \beta\mu)$. Normalizing this in the range $-1 < \mu < 1$ gives the probability density function (pdf):

$$p(\mu) = \frac{1 + \beta\mu}{2} \tag{282}$$

Note the function you integrate to find probabilities in a given distribution is called the probability density function and should not be confused with the distribution. Rybicki & Lightman approximate the distribution of μ' with a uniform distribution, corresponding to isotropic scattering. This gives a pdf

$$p'_{R\&L}(\mu') = \frac{1}{2} \tag{283}$$

With this approximation we can easily find the average ratio:

$$
\begin{aligned}
\left\langle \frac{\nu'''}{\nu} \right\rangle &= \int \int \gamma^2 (1 + \beta\mu)(1 + \beta\mu') p(\mu) p'(\mu') d\mu d\mu' \\
&= \gamma^2 \int \frac{(1 + \beta\mu)^2}{2} d\mu \int \frac{1}{2}(1 + \beta\mu') d\mu' \\
&= \gamma^2 \left(1 + \frac{\beta^2}{3}\right)
\end{aligned}
\tag{284}
$$

The ultrarelativistic limit is $(4/3)\gamma^2$.

However, electrons do not scatter isotropically. If the actual scattering angle is ψ, the scattering rate is proportional to $(1 + \cos^2\psi)$. We can write $\cos\psi$ as

$$\cos\psi = -\left(\mu\mu' + \sqrt{1 - \mu^2}\sqrt{1 - \mu'^2}\cos\phi\right) \tag{285}$$

where ϕ is the dihedral angle between the plane defined by the x axis and the incoming photon and the plane defined by the x axis and the outgoing photon. The rate of scattering into μ' will be proportional to the average of $(1 + \cos^2\psi)$ over all ϕ's. Thus the true pdf $p'(\mu')$ for Rayleigh scattering electrons is given by

$$
\begin{aligned}
p'(\mu'; \mu) &= \frac{3}{8} \int_0^{2\pi} (1 + \cos^2\psi) \frac{d\phi}{2\pi} \\
&= \frac{3}{8} \int_0^{2\pi} \left[1 + \mu^2\mu'^2 + 2\mu\mu'\sqrt{1 - \mu^2}\sqrt{1 - \mu'^2}\cos\phi + (1 - \mu^2)(1 - \mu'^2)\cos^2\phi\right] \frac{d\phi}{2\pi} \\
&= \frac{3}{8} \left[1 + \mu^2\mu'^2 + \frac{1}{2}(1 - \mu^2)(1 - \mu'^2)\right] \\
&= \frac{3}{8} \left[\frac{3}{2}(1 + \mu^2\mu'^2) - \frac{1}{2}(\mu^2 + \mu'^2)\right] \\
&= \frac{3}{8} \left[\mu^2 \left(\frac{3}{2}\mu'^2 - \frac{1}{2}\right) + \frac{3}{2} - \frac{1}{2}\mu'^2\right] \\
&= \left[\frac{1}{4} P_2(\mu) P_2(\mu') + \frac{1}{2}\right]
\end{aligned}
\tag{286}
$$

where $P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}$ is the 2nd order Legendre polynomial. The $(3/8)$ factor normalizes the pdf:

$$\int p'(\mu';\mu)d\mu' = \frac{3}{8}\left[3 + \mu^2 - \mu^2 - \frac{1}{3}\right] = 1 \tag{287}$$

for any value of μ. Note that the distribution of μ' does depend on μ, but is this dependence is not the probability density function for μ. By writing μ following a semicolon, we mean that μ is a parameter that modifies the μ' distribution. The joint pdf for μ and μ' is given by

$$p(\mu, \mu') = p(\mu)p'(\mu';\mu) \tag{288}$$

But the pre-scattering angle we need is not really μ, which is measured in the lab frame, but rather the pre-scattering angle in the electron's rest frame, whose cosine is given by $(\beta + \mu)/(1 + \beta\mu)$. We can now find the average ratio:

$$
\begin{aligned}
\left\langle \frac{\nu'''}{\nu} \right\rangle &= \int\int \gamma^2(1 + \beta\mu)(1 + \beta\mu')p(\mu)p'(\mu';(\beta + \mu)/(1 + \beta\mu))d\mu d\mu' \\
&= \gamma^2 \int \frac{(1 + \beta\mu)^2}{2}d\mu \\
&\quad \times \int \frac{3}{8}\left[\frac{3}{2}\left(1 + \left[\frac{\beta + \mu}{1 + \beta\mu}\right]^2 \mu'^2\right) - \frac{1}{2}\left(\left[\frac{\beta + \mu}{1 + \beta\mu}\right]^2 + \mu'^2\right)\right](1 + \beta\mu')d\mu' \\
&= \gamma^2\left(1 + \frac{\beta^2}{3}\right) \tag{289}
\end{aligned}
$$

To get the last line of this result, note that $p'(\mu')$ is always an even function of μ', so $\int \beta\mu'p'(\mu')d\mu' = 0$ and the integral $\int(1 + \beta\mu')p'(\mu')d\mu' = 1$ Thus the average is exactly the same as the result calculated in the isotropic scattering approximation used by Rybicki & Lightman.

12.1. Variance

The mean shift is quadratic in β, while the typical Doppler shift is $\pm\beta$. The square of $\pm\beta$ will interact with the second derivative of the input spectrum to produce a contribution quadratic in β, and thus of the same order as the mean shift. Therefore we need to know the variance of the frequency shift as well as the mean shift to compute the effects of electron scattering to the leading order in β. The maximum possible output:input ratio is $\gamma^2(1+\beta)^2$, but since $\gamma^2 = (1 - \beta^2)^{-1} = (1 + \beta)^{-1}(1 - \beta)^{-1}$, the maximum energy can be written as $(1 + \beta)/(1 - \beta)$, which is the square of the relativistic Doppler shift formula. The minimum possible output to input ratio is $\gamma^2(1 - \beta)^2$ which can be written as $(1 - \beta)/(1 + \beta)$. Note that this is just one over the maximum ratio.

Let $R = \nu'''/\nu$ be the output:input ratio. Then we can find the pdf for R using

$$
\begin{aligned}
p_R(R) &= \int\int \delta(R - \gamma^2(1+\beta\mu)(1+\beta\mu'))p(\mu)p'(\mu'; (\beta+\mu)/(1+\beta\mu))d\mu d\mu' \\
&= \int\int \delta(R - \gamma^2(1+\beta\mu)(1+\beta\mu'))\frac{1+\beta\mu}{2}\left[\frac{1}{4}P_2(\mu')P_2\left(\frac{\beta+\mu}{1+\beta\mu}\right) + \frac{1}{2}\right]d\mu d\mu'
\end{aligned}
\tag{290}
$$

In the isotropic approximation this is easy to evaluate: do the integral over μ' first, so the delta function integrates to $[\gamma^2\beta(1+\beta\mu)]^{-1}$. This cancels the μ dependence of $p(\mu)$, so the result for $p_R(R)$ is proportional to the amount of the μ range that is compatible with a given R. This is all of it for $R = 1$, and from $\mu = \beta^{-1}[(1-\beta)R - 1]$ to 1 for $R > 1$. Thus $p_R(R)$ declines as a linear function of R from a maximum at $R = 1$ to zero at $R = R_{max} = (1+\beta)/(1-\beta)$. For $R < 1$ the range in μ is from -1 to $\mu = \beta^{-1}[(1+\beta)R - 1]$, so again $p_R(R)$ declines linearly from the maximum at $R = 1$ to zero at $R = R_{min} = (1-\beta)/(1+\beta)$.

Finally we can calculate the mean squared output:input ratio which will give the variance. In the isotropic approximation used by R&L, this quantity is

$$
\begin{aligned}
\left\langle \left(\frac{\nu'''}{\nu}\right)^2 \right\rangle &= \int\int \gamma^4(1+\beta\mu)^2(1+\beta\mu')^2 p(\mu)p'(\mu'; (\beta+\mu)/(1+\beta\mu))d\mu d\mu' \\
&= \gamma^4\int\frac{(1+\beta\mu)^3}{2}\int\frac{(1+\beta\mu')^2}{2}d\mu'd\mu \\
&= \gamma^4\int\frac{(1+\beta\mu)^3}{2}\left[1 + \frac{\beta^2}{3}\right]d\mu \\
&= \gamma^4\left[1 + \beta^2\right]\left[1 + \frac{\beta^2}{3}\right] \\
&= \gamma^4\left[1 + \frac{4}{3}\beta^2 + \frac{1}{3}\beta^4\right]
\end{aligned}
\tag{291}
$$

This differs from the exact result only in the β^4 term which is sub-dominant for $kT << m_e c^2$.

12.2. Effect on an Input Blackbody

The Planck function is a fairly broad input. In terms of a logarithmic frequency variable, $s = \ln(h\nu/kT_\circ)$, with $T_\circ$ being the temperature of the input blackbody, the number of photons per unit of s is $\propto B_\nu \propto e^{3s}/(\exp(e^s) - 1)$. The peak of B_ν is when $x = h\nu/kT = 2.82$ or $s = 1.04$. The second derivative of the Planck function at the peak is given by $d^2\ln(B_\nu)/d^2s = -2.46$. This is $-1/\sigma^2$ where σ is the standard deviation of a Gaussian that osculates the Planck function in the s-$\ln(I_\nu)$ plane at the peak, so the effective "standard deviation" of the Planck function in terms of $\ln\nu$ is $\sigma = 0.64$. Another way to estimate an

effective standard deviation is to use $\int G(x)dx = \sqrt{2\pi}\sigma G_{pk}$ for a Gaussian $G(x)$. This gives

$$\sigma \approx \frac{\int B_s ds}{\sqrt{2\pi}B_{pk}} = \frac{2.40411}{2.5066 \times 1.42144} = 0.67474 \tag{292}$$

This is quite wide, and it can hide a number of details of the scatterings in the S-Z process. In particular, only the things that add like the mean shift and the variance of the shift are important in lowest order.

In terms of $\Delta s = \ln(R) \approx (R-1) - \frac{1}{2}(R-1)^2 + \ldots$ the mean shift and the variance of the shift are given to lowest order in β by

$$\langle \Delta s \rangle = \langle R-1 \rangle - \langle \frac{1}{2}(R^2 - 2R + 1) \rangle + \ldots \approx \frac{4}{3}\beta^2 - \frac{1}{2}\left(\frac{10}{3} - \frac{8}{3}\right)\beta^2 = \beta^2 + \ldots \tag{293}$$

and

$$\text{var}(\Delta s) = \langle (R-1)^2 \rangle = \frac{2}{3}\beta^2 + \ldots \tag{294}$$

Now the standard deviation of v for any one component of $\vec{v}$ is $\sqrt{kT/m_e}$, so the mean of the squared magnitude of β is $\langle \beta^2 \rangle = 3kT/m_e c^2$. The mean number of scattering events is given by τ, and the overall spectral shift and variance will both scale like $y = (kT/m_e c^2)\tau$ with the mean shift being $3y$ and the variance of the shift being $2y$. Hence the distorted spectrum will be a convolution of a Gaussian having mean $3y$ and $\sigma = \sqrt{2y}$ with the input spectrum:

$$
\begin{aligned}
I_s(y) &= \int I_{s+s'}(y=0)p_s(-s')ds' \\
&= \int I_{s+s'}(y=0)\frac{\exp(-(s'+3y)^2/4y)}{\sqrt{4\pi y}}ds'
\end{aligned}
\tag{295}
$$

Here I_s is I_ν at $s = \ln(\nu/\nu_\circ)$ where $\nu_\circ$ can be chosen for convenience if the input spectrum is a blackbody using $h\nu_\circ/kT_\circ = 1$. The first above line can be used for $kT \sim m_e c^2$ if the correct form is used for $p_s(s)$, while the second line used a Gaussian approximation to $p_s(s)$ that is good for $kT << m_e c^2$. If we expand $I_{s+s'} \approx I_s + s'(\partial I/\partial s) + \frac{1}{2}s'^2(\partial^2 I/\partial s^2) + \ldots$ then we get a partial differential equation that approximates the integral equation above:

$$\frac{\partial I}{\partial y} = -3\frac{\partial I}{\partial s} + \frac{\partial^2 I}{\partial s^2} \tag{296}$$

Note that this vanishes for either a constant I_ν or for a constant number of photons/mode, $I_\nu \propto \nu^3 \propto e^{3s}$. If I_ν follows the Rayleigh-Jeans law with $I_\nu \propto e^{2s}$, then $\partial I/\partial y = -2I$.

We can transform this into an equation in terms of n, the number of photons per mode, by letting $I = e^{3s}n$, which gives

$$\frac{\partial I}{\partial s} = 3I + e^{3s}\frac{\partial n}{\partial s}$$

$$\frac{\partial^2 I}{\partial s^2} = 3\frac{\partial I}{\partial s} + 3e^{3s}\frac{\partial n}{\partial s} + e^{3s}\frac{\partial^2 n}{\partial s^2}$$

$$e^{3s}\frac{\partial n}{\partial y} = 3e^{3s}\frac{\partial n}{\partial s} + e^{3s}\frac{\partial^2 n}{\partial s^2}$$

$$\frac{\partial n}{\partial y} = 3\frac{\partial n}{\partial s} + \frac{\partial^2 n}{\partial s^2} \tag{297}$$

Finally we can write this in terms of $x = e^s$ as

$$
\begin{aligned}
\frac{\partial n}{\partial y} &= \frac{1}{x^2}\frac{\partial}{\partial x}\left(x^4\frac{\partial n}{\partial x}\right) \\
&= \frac{1}{x^3}\frac{\partial}{\partial s}\left(x^3\frac{\partial n}{\partial s}\right) \\
&= \frac{1}{x^3}\left(3x^3\frac{\partial n}{\partial s} + x^3\frac{\partial^2 n}{\partial s^2}\right) \\
&= 3\frac{\partial n}{\partial s} + \frac{\partial^2 n}{\partial s^2}
\end{aligned} \tag{298}
$$

This is thus one part of the Kompane'ets (1957, Sov. Phys. JETP, 4, 730) equation,

$$\frac{\partial n}{\partial y} = \frac{1}{x^2}\frac{\partial}{\partial x}\left[x^4\left(n + n^2 + \frac{\partial n}{\partial x}\right)\right] \tag{299}$$

Since $x = h\nu/kT_e$, when $h\nu << kT_e$ both x and dx are very small, so the $\partial n/\partial x$ term is much bigger than the $n + n^2$ terms. Without the $n + n^2$ terms the equation is homogeneous in x so there is no tendency to drive the photons into equilibrium with the electron temperature. This is an adequate approximation when $h\nu << kT_e$.

Note that the energy density of the photon field, $U \propto \int nx^3 dx$, grows in a simple way with y:

$$
\begin{aligned}
\frac{\partial}{\partial y}\int nx^3 dx &= \int\frac{\partial n}{\partial y}x^3 dx = \int x\frac{\partial}{\partial x}\left(x^4\frac{\partial n}{\partial x}\right)dx \\
&= -\int\left(x^4\frac{\partial n}{\partial x}\right)dx = \int 4x^3 n\, dx
\end{aligned} \tag{300}
$$

where the last line above involves integrating by parts twice. We see that $\partial U/\partial y = 4U$.

There is a simple solution to the Kompane'ets equation with non-zero $\partial n/\partial y$ which gives the Sunyaev-Zel'dovich (1969, Ap. & Sp. Sci., 4, 301) or y spectral distortion. For a blackbody $n = 1/(e^x - 1)$, and neglecting the $n + n^2$ terms in the Kompane'ets equation we get

$$
\begin{aligned}
\frac{\partial n}{\partial y} &= x^{-2}\frac{\partial}{\partial x}x^4\frac{\partial n}{\partial x} \\
&= -x^{-2}\frac{\partial}{\partial x}\frac{x^4 e^x}{(e^x - 1)^2}
\end{aligned}
$$

68

$$= x^{-2} \frac{2x^4 e^{2x} - (e^x - 1)(4x^3 e^x + x^4 e^x)}{(e^x - 1)^3}$$

$$= \frac{x^2 e^{2x} - 4x e^{2x} + 4x e^x + x^2 e^x}{(e^x - 1)^3}$$

$$= e^x \frac{x^2(e^x + 1) - 4x(e^x - 1)}{(e^x - 1)^3}$$

$$= \frac{x e^x}{(e^x - 1)^2} \left(x \frac{e^x + 1}{e^x - 1} - 4 \right) \tag{301}$$

Thus the change in intensity due to y is (to first order)

$$\frac{\partial I_\nu}{\partial y} = \frac{2h\nu^3}{c^2} \frac{x e^x}{(e^x - 1)^2} \left(x \frac{e^x + 1}{e^x - 1} - 4 \right) \quad \text{with} \quad x = h\nu/kT_\circ \tag{302}$$

The derivative of the Planck function with respect to T is given by

$$T \frac{\partial B_\nu(T)}{\partial T} = \frac{2h\nu^3}{c^2} \frac{x e^x}{(e^x - 1)^2} \tag{303}$$

so $\partial I_\nu/\partial y$ can be described as a changed Planck brightness temperature:

$$T_\nu = T_\circ \left[1 + y \left(x \frac{e^x + 1}{e^x - 1} - 4 \right) + \dots \right]. \tag{304}$$

These formulae all assume the initial spectrum is a blackbody, and that $y << 1$, but there are many figures in the literature that are inaccurate because they used these first-order formulae for $y \approx 0.15$ and the higher order terms are already quite important. Figure 4 in Sunyaev & Zel'dovich (1980, ARAA, 18, 537) and Figure 1 in Carlstrom, Holder & Reese (2002, ARAA, 40, 643) are examples of these inaccurate figures. Figure 13 shows the inaccurate result in red and the correct result in blue.

12.3. High Temperature Corrections

If the temperature violates the $kT << m_e c^2$ restriction one can still use the first line of Eqn(295) with the correct $p_s(s)$ but the Gaussian approximation to $p_s(s)$ becomes inaccurate. When a single scattering can cause a shift equal to the "standard deviation" of the blackbody, $\sigma_s \approx 0.64$, one expects to see deviations from the Gaussian approximation. This happens for $(1 + \beta)/(1 - \beta) = e^{0.64}$, so $\beta = (e^{0.64} - 1)/(e^{0.64} + 1) = 0.31$, or for a kinetic energy of 26 keV. The median electron energy is higher than this value for $kT_e = 22$ keV. This is not that much higher than the temperatures of X-ray emitting clusters of galaxies, so one can expect corrections at the 10% level to be common. Since the corrections are produced by scatterings that produce the biggest shifts, and these are enhanced by the effects of non-isotropic scattering, we need to work with the full $p'(\mu'; \mu)$ instead of the R&L isotropic approximation.
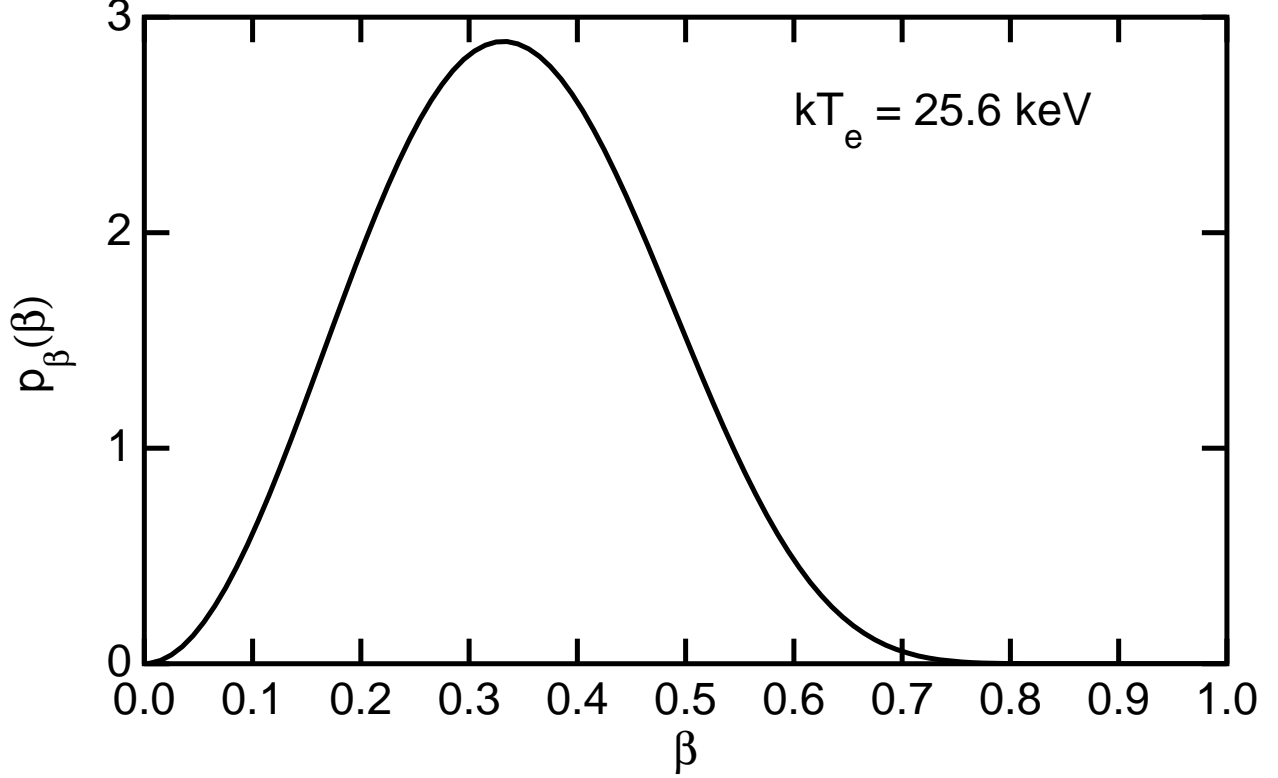
Fig. 14.— Probability distribution for electron speed β when $kT_e = 0.05m_ec^2 = 25.6$ keV. The median $\beta = 0.334$.

Note that one does not need to modify the Thomson scattering cross-section unless $h\nu$ in the rest frame is comparable to m_ec^2. Since the CMB photon energy is about 3 meV this is not necessary. Even for a very high electron energy with $\beta = 0.9$ the photon energy in the electron rest frame is less than 20 meV which is $<<<< m_cc^2$ so the Thomson scattering cross-section applies.

To find $p_s(s)$ one can modify Eqn(290):

$$p_s(s; \beta) = \int \int \delta(s - \ln[\gamma^2(1 + \beta\mu)(1 + \beta\mu')])p(\mu)p'(\mu'; \frac{\beta + \mu}{1 + \beta\mu})d\mu d\mu' \qquad (305)$$

The integral over μ' can be performed (see Wright 1979, ApJ, 232, 348) giving

$$p_s(s; \beta) = \int \left(\frac{1 + \beta\mu'}{\beta} \right) \left(\frac{1 + \beta\mu}{2} \right) \left[\frac{1}{2} + \frac{1}{4}P_2(\mu')P_2 \left(\frac{\beta + \mu}{1 + \beta\mu} \right) \right] d\mu \qquad (306)$$

with

$$\mu' = \beta^{-1} \left[\frac{e^s}{\gamma^2(1 + \beta\mu)} - 1 \right]$$

and with the range in μ running from $\mu_{min} = \beta^{-1}[(1 - \beta)e^s - 1]$ to 1 for $s > 0$. For $s < 0$, one can use the fact that $p_s(-s) = e^{-3s}p_s(s)$.
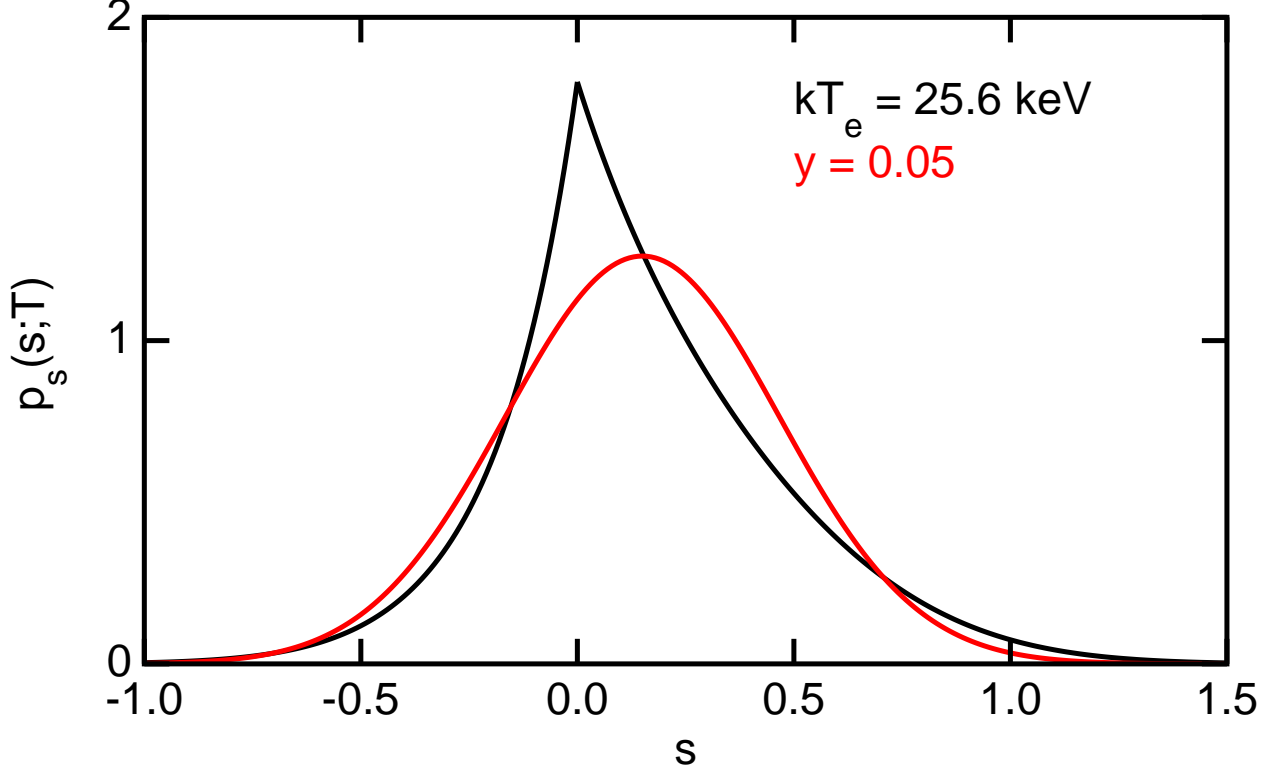
Fig. 15.— The probability density function for the frequency shift s in a single scattering by an electron gas with $kT_e = 0.05 m_e c^2 = 25.6$ keV. Also shown in red is the Gaussian approximation for $y = 0.05$.

Once $p_s(s; \beta)$ is found, the integral over the relativistic Maxwellian is needed. The momentum of an electron is $p = m\gamma\beta c$. Phase space goes like $p^2 dp$, and $dp = mc\gamma^3 d\beta$, so the probability of an electron having a given beta is

$$p_\beta(\beta; T) d\beta \propto \gamma^5 \beta^2 \exp\left(\frac{-(\gamma - 1)m_e c^2}{kT_e}\right) d\beta \tag{307}$$

The final single scattering probability density function for s is then

$$p_s(s; T) = \int p_s(s; \beta) p_\beta(\beta; T) d\beta \tag{308}$$

Figure 15 shows this function for $kT_e/m_e c^2 = 0.05$ along with the Gaussian approximation for $y = 0.05$. While these functions are similar in centroid and width, the differences are easily seen.

But the single scattering probability density function is not the final answer, since even with $\tau = 1$ there will be some photons that are never scattered, and some photons that are scattered twice. The probability density function for unscattered photons is $\delta(s)$, while the pdf for twice-scattered photons is $p_s(s; T) \star p_s(s; T)$, where $\star$ denotes a convolution. The
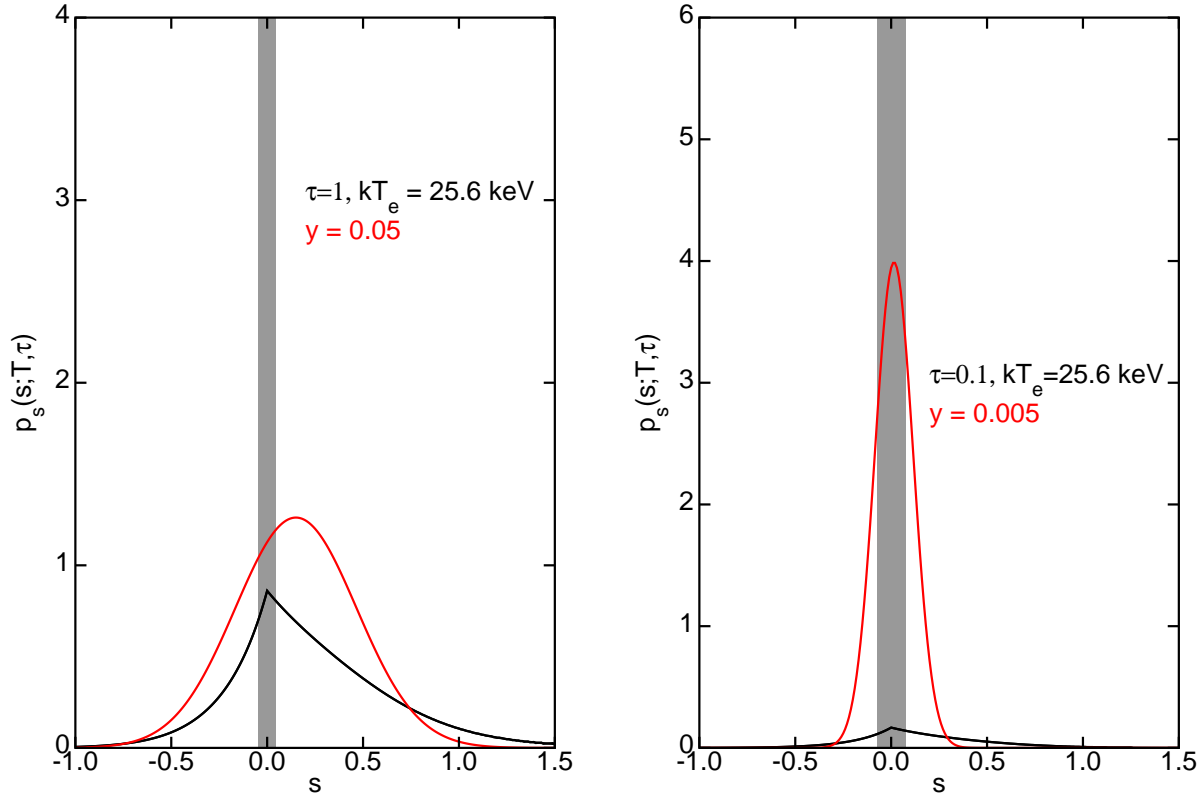
Fig. 16.— Left: The frequency shift probability density function for $\tau = 1$ and $kT_e = 0.05 m_e c^2 = 25.6$ keV compared to the Gaussian approximation for $y = 0.05$. The grey bar has the correct area for the delta function representing unscattered photons. Right: the same for $\tau = 0.1$ and $y = 0.005$.
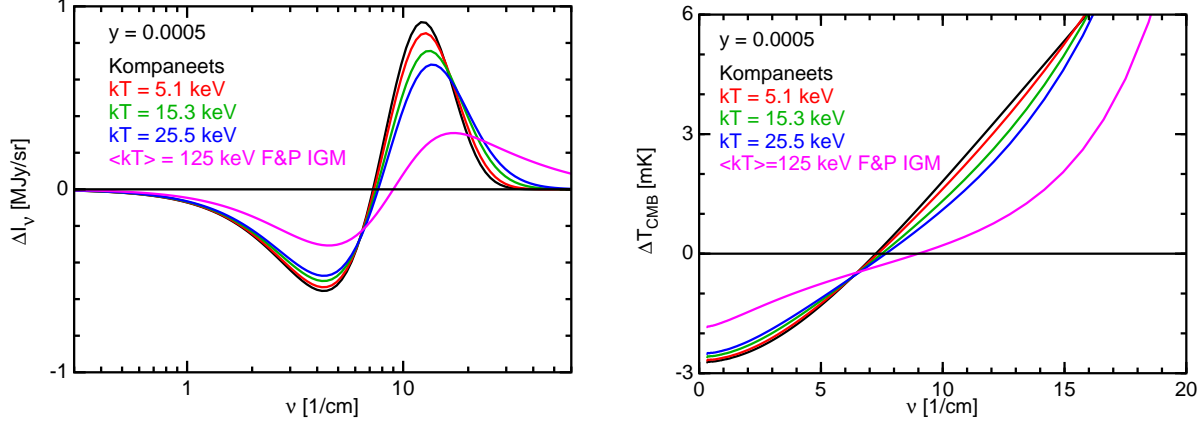
Fig. 17.— Left: Intensity change caused by $y = 0.0005$ for several different fixed temperatures and the hot intergalactic medium proposed by Field & Perrenod (1977, ApJ, 215, 717). The crossover frequency clearly goes up as the temperature goes up. Right: the intensity changes converted into ΔT_{CMB}.

resulting pdf for s after an optical depth of τ is given by

$$p_s(s; T, \tau) = e^{-\tau} \left[\delta(s) + \frac{\tau^1}{1!} p_s(s; T) + \frac{\tau^2}{2!} p_s(s; T) \bigstar p_s(s; T) + \dots \right] \quad (309)$$

Figure 16 shows this density function compared to the Gaussian approximation with the same $y = 0.05$. There are clear differences between the densities.

One can compute the intensity change produced by a small τ by convolving the input Planck function with $p_s(s; T)$ and then subtracting the input Planck function:

$$\frac{kT_e}{m_e c^2} \frac{\partial I_\nu}{\partial y} = \frac{\partial I_\nu}{\partial \tau} = p_s(s; T) \bigstar B_{\nu e^s}(T_\circ) - B_\nu(T_\circ) \quad (310)$$

Figure 17 shows the distortion produced for $y = 0.0005$ at several temperatures using this first order equation. The frequency at which the effect goes from negative to positive depends on the temperature as

$$\frac{h\nu_c}{kT_\circ} = x_c \approx 3.83 + 4.345 \frac{kT_e}{m_e c^2} \quad (311)$$

The slope of the T vs. ν curve through the cross-over also changes significantly with $kT_e/m_e c^2$. This dependence can be fit fairly well as

$$\frac{\partial^2 \ln(T)}{\partial x \partial y}\bigg|_{x=x_c} = \left(1.149 + 9.5 \frac{kT_e}{m_e c^2} \right)^{-1} \quad (312)$$

as shown in Figure 18. Finally the sensitivity of observations at 30.3 GHz as done by the Combined ARray for Millimeter-wave Astronomy (CARMA) goes down slightly as T_e goes
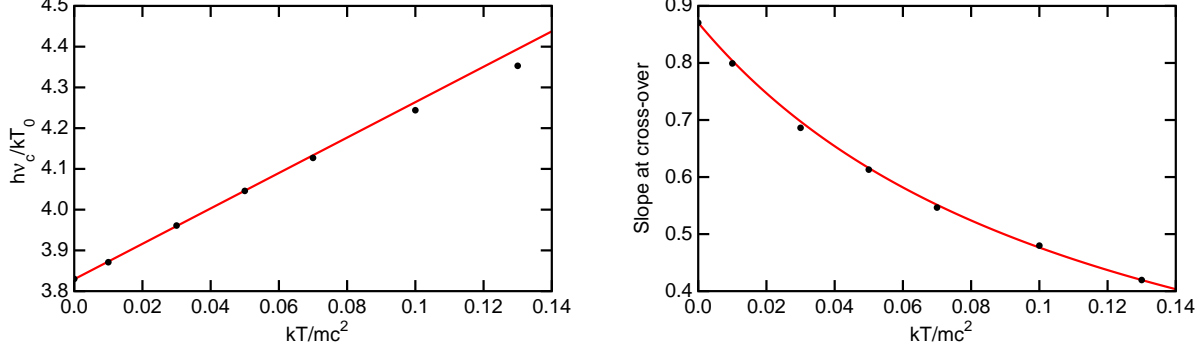
73

Fig. 18.— Left: Crossover frequency *vs.* electron temperature using the the exact $p_s(s; T, \tau)$ formulation. The point at zero temperature is the crossover for the usual S-Z formula. The line shown in red is a fit $x_c = 3.83 + 4.345(kT_e/m_ec^2)$. Right: the slope, $\partial^2 \ln(T)/\partial x \partial y$, at the crossover frequency *vs.* T_e. The red curve is $\partial^2 \ln(T)/\partial x \partial y = 1/(1.149 + 9.5(kT_e/m_ec^2))$.

up, being 8% down from the standard low frequency result of $\Delta T/T = -2y$ at $kT_e = 15$ keV, as shown in Figure 19. An approximation to the plotted results is

$$\left.\frac{\Delta T}{T_\circ}\right|_{\nu=30.3\,\mathrm{GHz}} = \frac{-y}{0.512 + 0.998(kT_e/m_cc^2)} \tag{313}$$

12.4. Kinematic S-Z Effect

The electrons in a cluster of galaxies will also distort the spectrum of the CMB if there is net radial velocity of the cluster relative to the CMB rest frame. The change in temperature for the scattered photons is given by $\Delta T = -(v_{rad}/c)T_\circ$ since a positive radial velocity corresponds to a redshift. The change after passing through a cluster with optical depth τ is $\Delta T/T_\circ = -\tau v_{rad}/c$. This shift raises or lowers the crossover frequency, but since the temperature distortion near the crossover depends on three parameters, τ, kT_e/m_ec^2, & v_{rad}/c, and measuring the crossover frequency and slope only provides two independent data values, one will definitely need other information like an X-ray measurement of T_e to be able to find v_{rad}.

Ignoring all the changes to the cross-over frequency and slope with T_e, one still needs to know T_e in order to convert the measured y into an optical depth. Assume one has found the cross-over frequency x_c, so

$$\frac{\Delta T}{T_\circ} \approx -\frac{\tau v_{rad}}{c} + 0.87(x_c - 3.83)\tau\frac{kT_e}{m_ec^2} = 0 \tag{314}$$

Then

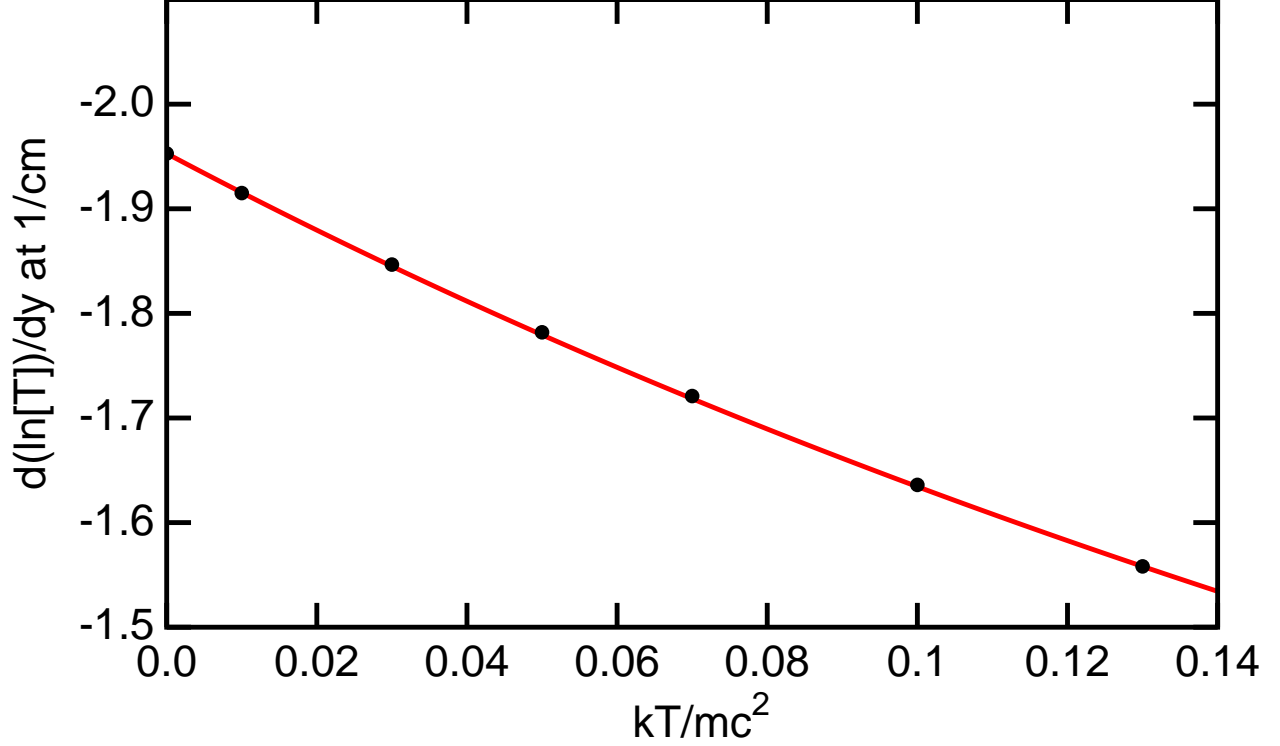$$\frac{v_{rad}}{c} = 0.87(x_c - 3.83)\frac{kT_e}{m_ec^2} \tag{315}$$

74

Fig. 19.— S-Z signal at $\nu = 1.01$/cm per unit y as a function kT_e/m_ec^2. The red curve shows a simple fit to the points, $d\ln T/dy = -(0.512 + 0.998kT_e/m_ec^2)^{-1}$.

When we include the changes in slope and cross-over due to high temperature effects, the equation changes to

$$\frac{v_{rad}}{c} = \frac{x_c - (3.83 + 4.345kT_e/m_ec^2)}{1.149 + 9.5kT_e/m_ec^2}\frac{kT_e}{m_ec^2} \tag{316}$$

If $x_c = 3.83$ and $kT_e = 10.2$ keV, then the difference between these formulae is $\Delta v = 400$ km/sec. Clearly the high temperature corrections to the S-Z effect are important when looking for the kinematic S-Z effect.

12.5. The Full Kompane'ets Equation

See §7.6 in R&L.

The formulae we have just developed all assumed that $h\nu << kT_e$. This assumption is very well satisfied when dealing with the CMB distorted by X-ray emitting gas, but equation we have been solving:

$$\frac{\partial n}{\partial y} = \frac{1}{x^2}\frac{\partial}{\partial x}\left(x^4\frac{\partial n}{\partial x}\right) \tag{317}$$

cannot possibly be correct in general because it is homogeneous in x and linear in n. For

any solution $f(x, y)$ to Eqn(317), the function $Af(Bx, y)$ is also a solution. This makes our convolution analysis possible, but does not have the property that a blackbody at T_e will not be affected by electron scattering. Since twice a blackbody is not a blackbody, the complete equation must be non-linear. It also must have terms which make a blackbody a stationary point, so something must be added to $\partial n/\partial x$ to cancel this term for a blackbody at the electron temperature, $n = 1/(e^x - 1)$. This form gives

$$\frac{\partial n}{\partial x} = \frac{d}{dx}\left(\frac{1}{e^x - 1}\right) = \frac{-e^x}{(e^x - 1)^2} = -(n + n^2) \tag{318}$$

so the completed equation has to be

$$\frac{\partial n}{\partial y} = \frac{1}{x^2}\frac{\partial}{\partial x}\left(x^4\left[n + n^2 + \frac{\partial n}{\partial x}\right]\right) \tag{319}$$

which is the Kompane'ets equation.

Since the Kompane'ets equation is describing electron scattering, which preserves the number of photons, one finds that the y derivative of the photon density $N = 8\pi(kT_e)^3/(hc)^3 \int nx^2 dx$ vanishes:

$$\begin{aligned}\frac{\partial N}{\partial y} &\propto \int x^2 \frac{\partial n}{\partial y} dx \\ &= \int \frac{\partial}{\partial x}\left[x^4\left(n + n^2 + \frac{\partial n}{\partial x}\right)\right] dx \\ &= \left[x^4\left(n + n^2 + \frac{\partial n}{\partial x}\right)\right]\Big|_0^\infty \\ &= 0 \end{aligned} \tag{320}$$

since the x^4 factor vanishes at zero and n and dn/dx both vanish at infinity.

The stationary solutions of the equation $\partial n/\partial y = 0$ are the photon distributions in thermal equilibrium with the electrons. Since photons are conserved, the photon number density does not have to agree with the photon number density in a blackbody at the electron temperature. Thus a more general Bose-Einstein thermal distribution is allowed: $n = 1/(\exp(x + \mu) - 1)$. This gives

$$\frac{\partial n}{\partial x} = \frac{-\exp(x + \mu)}{(\exp(x + \mu) - 1)^2} = -(n + n^2) \tag{321}$$

Therefore $\partial n/\partial y = 0$ for all μ.

Note that without the n^2 term we would find that $n + dn/dx = 0$ implies $n = \exp(-x)$ which gives a Wien distribution instead of a Planck distribution. This is also what we saw for thermal equilibrium when ignoring stimulated emission. Thus the derivation of the Kompane'ets equation will involve stimulated Compton scattering.

12.6.   Derivation of the Kompane'ets Equation

If we ignore relativistic effects, the Doppler shift in scattering through an angle θ is easily described by considering the electron velocity component along the bisector of the scattering angle. Call this component x, so the shift is

$$\Delta\nu = \nu\frac{v_x}{c}2\cos(\theta/2) \tag{322}$$

The mean squared value of this is

$$\mathrm{var}\left(\frac{\Delta\nu}{\nu}\right) = 4\langle\cos^2(\theta/2)\rangle\frac{\langle v_x^2\rangle}{c^2} = 4\times\frac{1}{2}\times\frac{kT_e}{m_ec^2} = 2kT_e/m_ec^2 \tag{323}$$

This agrees with $\mathrm{var}(s) = 2y$.

The mean shift corresponds to an energy transfer from the electron to the photons. This is due to photon drag. In the electrons rest frame the intensity is anisotropic with

$$I(\theta) = I_\circ\left(1 + \frac{4v}{c}\cos\theta\right) \tag{324}$$

The scattered photons have zero net momentum so the total momentum transfer is

$$\langle p_x\rangle = \frac{1}{4\pi}\int\left(\frac{h\nu}{c}\right)\left(1 + \frac{4v}{c}\cos\theta\right)\cos\theta d\Omega = \frac{4}{3}\frac{v_x}{c}\left(\frac{h\nu}{c}\right) \tag{325}$$

per scattering. The work done is

$$W = p_xv_x + p_yv_y + p_zv_z = \frac{4}{3}\left(\frac{h\nu}{c}\right)\frac{\vec{v}^2}{c} = 4h\nu\frac{kT_e}{m_ec^2} \tag{326}$$

In addition to the photon drag, there is also energy transfer from the photons to the electrons due to the Compton shift. Scattering through angle θ gives a momentum transfer of

$$\Delta p = \frac{h\nu}{c}2\cos(\theta/2) \tag{327}$$

This gives a net energy transfer of

$$\left\langle\frac{\Delta p^2}{2m}\right\rangle = \frac{(h\nu)^2}{m_ec^2} \tag{328}$$

Thus the total energy change per scattering is

$$\langle h\Delta\nu\rangle = h\nu\frac{4kT - h\nu}{m_ec^2} \tag{329}$$

77

The overall rate equation is

$$\frac{\partial n(\nu)}{\partial t} = \left(\frac{\text{scatterings}}{\text{second}}\right) \times \int \left[P(\nu_1 \to \nu)n(\nu_1)(1 + n(\nu)) - P(\nu \to \nu_1)n(\nu)(1 + n(\nu_1))\right] d\nu_1 \tag{330}$$

Now consider the electron distribution in phase space, $f_e(p)$. Consider $p_1 + \nu_1 \leftrightarrow p + \nu$. Given $\vec{p}$, ν, ν_1 and the scattering angle Ω we know $\vec{p_1}$. Then

$$\frac{\partial n(\nu)}{\partial t} = n_e c \int d^3 p \frac{d\sigma}{d\Omega} \left[f_e(p_1)n(\nu_1)(1 + n(\nu)) - f_e(p)n(\nu)(1 + n(\nu_1))\right] d\nu_1 \tag{331}$$

Now we expand in powers of $\nu_1 - \nu = \Delta$. We now that $f_e(p_1) = f_e(p) \exp(h\Delta/kT_e)$ since if ν_1 has higher energy then p_1 has lower energy.

Note that this assumes the electrons are not degenerate. Do not use the Kompane'ets equation in a white dwarf.

Use $x = E/kT_e$ as the energy variable, and write $h\nu/kT_e = x$

Now expand $n(\nu_1)$ in powers of Δ giving

$$\begin{aligned}\frac{\partial n(\nu)}{\partial t} &= n_e c \int d^3 p \frac{d\sigma}{d\Omega} f_e(p) \left[(1 + \Delta + \tfrac{1}{2}\Delta^2)(n + n'\Delta + \tfrac{1}{2}n''\Delta^2)(1 + n)\right. \\ &\quad - \left. n(1 + n + n'\Delta + \tfrac{1}{2}n''\Delta^2)\right] d\Delta\end{aligned} \tag{332}$$

The 0th order in Δ cancels.

The 1st order in Δ gives

$$n(1 + n) + n'(1 + n) - nn'.$$

The 2nd order in Δ gives

$$\tfrac{1}{2}n(1 + n) + n'(1 + n) + \tfrac{1}{2}n''(1 + n) - \tfrac{1}{2}n''n.$$

Thus

$$\frac{\partial n}{\partial t} = n_e c \sigma_T \left[(n' + n + n^2)\langle\Delta\rangle + (\tfrac{1}{2}n'' + n'(1 + n) + \tfrac{1}{2}n(1 + n))\langle\Delta^2\rangle\right] \tag{333}$$

But $\langle\Delta\rangle = x(4 - x)(kT_e/m_e c^2)$ and $\langle\Delta^2\rangle = 2x^2(kT_e/m_e c^2)$. Therefore

$$\begin{aligned}\frac{\partial n}{\partial t} &= n_e c \sigma_T \frac{kT_e}{m_e c^2} \left[x^2(n'' + 2n' + 2nn' + n + n^2) + x(4 - x)(n' + n + n^2)\right] \\ &= n_e c \sigma_T \frac{kT_e}{m_e c^2} \left[x^2(n'' + 2n' + 2nn' + n + n^2 - n' - n - n^2) + 4x(n' + n + n^2)\right] \\ &= n_e c \sigma_T \frac{kT_e}{m_e c^2} \left[4x(n' + n + n^2) + x^2(n'' + n' + 2nn')\right]\end{aligned} \tag{334}$$

Now $dy = n_e c\sigma_T (kT_e/m_e c^2) dt$ and

$$\frac{\partial}{\partial x}\left(x^4[n' + n + n^2]\right) = 4x^3[n' + n + n^2] + x^4[n'' + n' + 2nn']$$ (335)

so the equation becomes

$$\frac{\partial n}{\partial y} = \frac{1}{x^2}\frac{\partial}{\partial x}\left(x^4\left[\frac{\partial n}{\partial x} + n + n^2\right]\right)$$ (336)

which is the Kompane'ets equation.

The Kompane'ets equation can easily handle situations where $h\nu \approx kT_e$ and it includes the Compton shift, but its derivation used the small shift approximations requiring that $kT_e <<< m_e c^2$. I do not know of any easy way to handle situations where both $h\nu \approx kT_e$ and $kT_e \approx m_e c^2$. Monte Carlo simulations are simple to program except for the stimulated scattering effect though.

13. Synchrotron Radiation

13.1. Power

Consider an electron moving in a magnetic field. The plasma in space is a good electrical conductor so we expect zero electric field in the lab frame. Let the B field be aligned along the z-axis. The 4 momentum of the electron will be

$$p^\mu = mc(\gamma; \gamma\beta \sin\phi \cos\theta, \gamma\beta \sin\phi \sin\theta, \gamma\beta \cos\phi)$$ (337)

where ϕ is the *pitch angle* which remains constant, while θ is the angle around the electron's circular orbit and varies like $d\theta/dt = \Omega$ with

$$\Omega = \frac{eB}{\gamma mc}$$ (338)

For non-relativistic electrons the frequency is the *electron cyclotron frequency*: $\omega_c = eB/mc$, or $\nu_c = (2.8 \text{ MHz/gauss})B$.

If we do a Lorentz boost along the z-axis into the *guiding center* frame, the electron is moving perpendicular to the magnetic field:

$$\begin{aligned} p^{\mu\prime} &= mc(\gamma'; \gamma'\beta' \cos\theta, \gamma'\beta' \sin\theta, 0) \\ B' &= B \end{aligned}$$ (339)

Note that while θ is unchanged in this transformation, t does change, so the new $d\theta/dt' = \Omega' = eB/\gamma'mc$.

Finally do a boost along the x-axis into a frame moving at the same speed as the electron when $\theta = 0$. This is an inertial frame, known as the *momentarily comoving inertial frame* or MCIF, while the electron's rest frame is an accelerating frame. Usually when we talk about the electron's "rest frame" we mean the MCIF. In this frame we have

$$
\begin{aligned}
p^{\mu'''} &= mc(1; 0) \\
B'' &= \gamma' B \\
E'' &= \gamma' \beta' B = \gamma \beta \sin \phi \, B
\end{aligned}
\tag{340}
$$

In this frame only the electric field causes accelerations, so

$$
a'' = \frac{\gamma \beta \sin \phi \, eB}{m}
\tag{341}
$$

and the Larmor formula gives the radiated power as

$$
P'' = \frac{2}{3} \frac{e^2 (a'')^2}{c^3} = \frac{2}{3} \frac{e^4}{m^2 c^3} \gamma^2 \beta^2 \sin^2 \phi \, B^2
\tag{342}
$$

But this power is radiated in a dipole pattern in the MCIF, so the radiated power has no net 3 momentum. Thus the radiated energy-momentum is parallel to the electron's 4 velocity, and therefore the power dE_{rad}/dt, which is the ratio of the time components of two parallel 4 vectors, will be the same in all frames. Thus the lab power equals the MCIF power: $P = P''$.

If the pitch angles are isotropically distributed, then the average value of $\sin^2 \phi$ is $\langle \sin^2 \phi \rangle = 2/3$, giving an average synchrotron power per electron of

$$
P = \frac{4}{9} \frac{e^4}{m^2 c^3} \gamma^2 \beta^2 \, B^2 = \frac{4}{3} \gamma^2 \sigma_T c u_B
\tag{343}
$$

with the magnetic field energy density being $u_B = B^2/8\pi$ and the Thomson scattering cross-section being $\sigma_T = 8\pi e^4/3m^2 c^4$.

13.2. Spectrum

The electron's motion is periodic with frequency Ω, but in the lab frame the received radiated pulse is far from sinusoidal. The dipole pattern radiated by the electron has nulls along the direction of acceleration, which are along the y''-axis in our example. When transformed into the guiding center frame, momentum of a photon along the y'' axis becomes:

$$
\begin{aligned}
p^{\mu'''} &= (h\nu''/c)(1; 0, 1, 0) \\
&\rightarrow \\
p^{\mu'} &= (h\nu''/c)(\gamma'; \gamma'\beta', 1, 0)
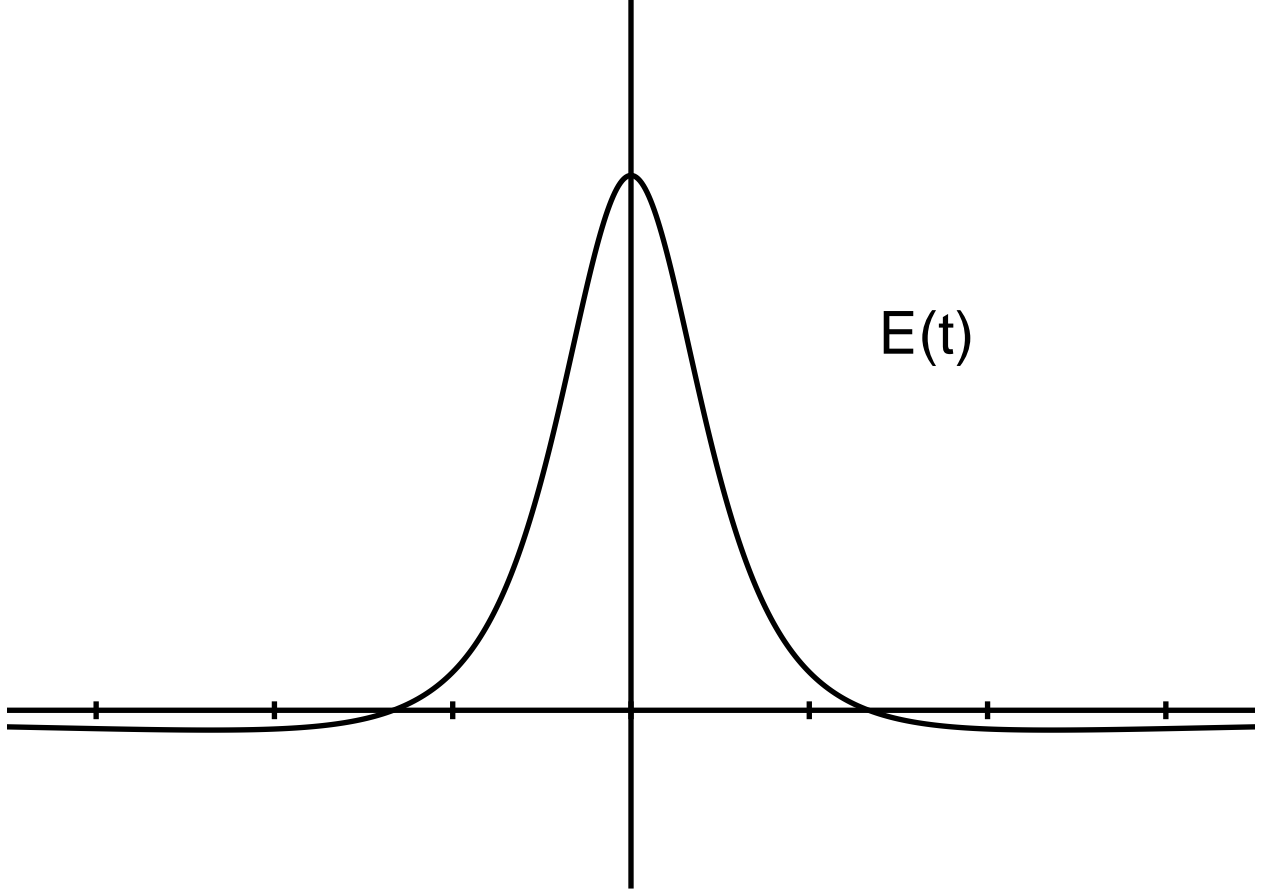\end{aligned}
\tag{344}
$$

Fig. 20.— The pulse shape by an observer on the plane of the electron's synchrotron orbit. Ticks every $mc/(2\gamma^2 eB)$ on the t-axis.

Thus the peak to null angle is $\sin^{-1}(1/\gamma')$. For $\gamma >> 1$ we can use small angle approximations. Therefore the electron only turns through an angle $\Delta\theta \approx 1/\gamma'$ during the pulse. Turning through this angle takes $\Delta t' = \Delta\theta/\Omega' = mc/eB$. But the pulse as received by an observer in front of the electron is seen in a time given by

$$\Delta t_{rec} = (1 - \beta')\Delta t' \approx \frac{mc}{2(\gamma')^2 eB} \tag{345}$$

Thus a typical frequency for synchrotron radiation will be

$$\omega' = \frac{1}{\Delta t_{rec}} = \frac{2(\gamma')^2 eB}{mc} \tag{346}$$

In the lab frame instead of the guiding center frame, $\gamma' = \gamma \sin\phi$ but there is a Doppler shift factor $1/\sin\phi$, so this becomes

$$\omega_{crit} \propto \gamma^2 \frac{eB}{mc} \sin\phi \tag{347}$$

We can calculate the exact shape of the pulse seen by an observer situated on the plane of the electron's orbit in the guiding center frame. For simplicity I will drop the primes

81

on γ and β in this frame. The acceleration in the MCIF is $a' = \gamma\beta eB$ so the radiated electric field in the MCIF is $E'_{rad} \propto \gamma\beta eB \cos\theta'$. The 4 momentum of a photon in the MCIF, $(h\nu'/c)(1; \cos\theta', \sin\theta', 0)$ becomes $(h\nu'/c)(\gamma(1+\beta\cos\theta'); \gamma(\cos\theta'+\beta), \sin\theta', 0)$ in the guiding center frame. Thus the Doppler shift factor is

$$\frac{\nu}{\nu'} = \gamma(1 + \beta\cos\theta') \tag{348}$$

and the angle in the guiding center frame is given by

$$\sin\theta = \frac{p_y}{|p|} = \frac{\sin\theta'}{\gamma(1 + \beta\cos\theta')} \tag{349}$$

Now the radiated field transforms like the Doppler shift factor squared to keep the number of photons per mode invariant. Hence

$$E_{rad} \propto eB\gamma^3\beta(1 + \beta\cos\theta')^2 \cos\theta' \tag{350}$$

Now we need to calculate the received time, $t_{rec} = t - x/c$. We know that $x = R\sin\Omega t = R\sin\theta$. Let's let

$$y = \frac{\sin\theta'}{1 + \beta\cos\theta'} \tag{351}$$

Then

$$
\begin{aligned}
t_{rec} &= \frac{\theta}{\Omega} - \frac{R}{c}\sin\theta \\
&= \frac{1}{\Omega}\sin^{-1}\left(\frac{y}{\gamma}\right) - \frac{R}{c}\frac{y}{\gamma} \\
&= \frac{\gamma mc}{eB}\left(\frac{y}{\gamma} + \frac{1}{6}\left(\frac{y}{\gamma}\right)^3 + \ldots\right) - \frac{\gamma\beta mc}{eB}\frac{y}{\gamma} \\
&= \frac{mc}{eB}\left[(1-\beta)y + \frac{y^3}{6\gamma^2} + \ldots\right] \\
&= \frac{mc}{2\gamma^2 eB}\left[y + \frac{y^3}{3}\right]
\end{aligned}
\tag{352}
$$

We have already assumed $\beta \approx 1$ in $1 - \beta = 1/(2\gamma^2)$, so continuing to use $\beta \approx 1$ gives us

$$y^2 = \frac{\sin^2\theta'}{(1+\cos\theta')^2} = \frac{1-\cos^2\theta'}{(1+\cos\theta')^2} = \frac{1-\cos\theta'}{1+\cos\theta'} \tag{353}$$

Hence

$$\cos\theta' = \frac{1-y^2}{1+y^2} \tag{354}$$

and

$$1 + \cos\theta' = \frac{2}{1+y^2} \tag{355}$$

82

Thus we get

$$E_{rad}(t_{rec}) \propto eB\gamma^3 \frac{1-y^2}{(1+y^2)^3} \tag{356}$$

at

$$t_{rec} = \frac{mc}{2\gamma^2 eB}\left[y + \frac{y^3}{3}\right] \tag{357}$$

We can of course find the spectrum by Fourier transforming this pulse, giving

$$\hat{E}(\omega) \propto \int \frac{1-y^2}{(1+y^2)^3}e^{i\omega t}dt \tag{358}$$

Inverting the equation for y as a function of t introduces branch cuts but we can still find the dominant high frequency behaviour by locating the pole closest to the real t-axis in the upper half of the complex t plane. This is at $y = +i$ which gives

$$t_{pole} = \frac{mc}{2\gamma^2 eB}\left(i - \frac{i}{3}\right) = \frac{imc}{3\gamma^2 eB} \tag{359}$$

The high frequency cutoff in $\hat{E}(\omega)$ thus goes like

$$\begin{aligned}
\hat{E}(\omega) &\propto \exp\left[-\frac{mc}{3\gamma^2 eB}\omega\right] \\
|\hat{E}(\omega)|^2 &\propto \exp\left[-\frac{2mc}{3\gamma^2 eB}\omega\right] \\
&= \exp\left[-\frac{\omega}{\omega_{crit}}\right]
\end{aligned} \tag{360}$$

with a critical frequency that is just $1.5\gamma^2$ times the cyclotron frequency:

$$\begin{aligned}
\omega_{crit} &= \frac{3}{2}\frac{eB}{mc}\gamma^2 \\
\nu_{crit} &= \frac{\omega_{crit}}{2\pi} = (4.2\text{ MHz/gauss})B
\end{aligned} \tag{361}$$

The integral for $\hat{E}(\omega)$ can be done exactly by changing the variable of integration to y with $dt = (1+y^2)dy$, giving

$$\hat{E}(\omega) \propto \int \frac{1-y^2}{(1+y^2)^2}\exp\left[\frac{i\omega mc}{2\gamma^2 eB}\left(y + \frac{y^3}{3}\right)\right]dy \tag{362}$$

Integrate by parts ($\int p\,dq = pq - \int q\,dp$) using $q = y/(1+y^2)$ so $dq = (1-y^2)dy/(1+y^2)^2$, and $p = \exp[\ldots]$ so $dp = (i\omega mc/(2\gamma^2 eB))(1+y^2)\exp[\ldots]$, giving

$$\hat{E}(\omega) \propto \omega \int y\exp\left[\frac{i\omega mc}{2\gamma^2 eB}\left(y + \frac{y^3}{3}\right)\right]dy \propto \omega K_{2/3}(\omega/(3\gamma^2 eB/mc)) \tag{363}$$

83

which has the exponential high frequency cutoff we found earlier and behaves like $\omega^{1/3}$ for low frequency. This would give a power going like $\omega^{2/3}$ but the total power radiated, integrated over all angles instead of just for an observer in the plane of the orbit, is given by

$$
\begin{aligned}
P(\omega) &\propto F\left(\frac{\omega}{\omega_{crit}}\right) \\
F(x) &= x\int_x^\infty K_{5/3}(y)dy
\end{aligned}
\tag{364}
$$

which goes likes $\omega^{1/3}$ for low frequencies.

As discussed earlier, the critical frequency varies with pitch angle as $\sin\phi$. For a tangled magnetic field and an isotropic pitch angle distribution, we observe a power $\propto \sin^2\phi$ with critical frequency $1.5\sin\phi\,\gamma^2 eB/mc$. The power weighted mean of $\sin\phi$ is

$$
\frac{\int \sin^3\phi\,d\Omega}{\int \sin^2\phi\,d\Omega} = \frac{2\pi\int \sin^4\phi\,d\phi}{8\pi/3} = \frac{3}{4}\int_0^\pi [(1-\cos(2\phi))/2]^2 d\phi = \frac{9\pi}{32}
\tag{365}
$$

Thus the power weighted mean critical frequency is

$$
\frac{3}{2}\times\frac{9\pi}{32}\times\frac{eB}{mc}\gamma^2 = 1.325\frac{eB}{mc}\gamma^2
\tag{366}
$$

I recommend remembering this factor 1.325 as $4/3$.

13.3. Power Law Distributions

In most astrophysical sources the relativistic electrons producing synchrotron radiation have γ's that cover a wide range: usually a few decades or more. Because of this, the synchrotron radiation spans several decades, but the exact spectral function $F(x)$ is not so wide. It is therefore appropriate to ignore the width of $F(x)$ and to approximate it by a delta function! We will assume the following very simple form for the power emitted by mono-energetic electrons with random pitch angles in a tangled magnetic field:

$$
P(\nu) = \frac{4}{3}\gamma^2 c\sigma_T u_B \delta\left(\nu - \frac{4}{3}\frac{\gamma^2 eB}{2\pi mc}\right)
\tag{367}
$$

With this approximation it is easy to compute the synchrotron power generated by a population $N(\gamma)$ of electrons per unit γ. With $2\pi\nu_\circ = (4eB)/(3mc)$ we get

$$
L_\nu = \int \frac{4}{3}\gamma^2 c\sigma_T u_B \delta(\nu - \nu_\circ\gamma^2)N(\gamma)d\gamma = \left.\frac{2\gamma N(\gamma)c\sigma_T u_B}{3\nu_\circ}\right|_{\gamma=\sqrt{\nu/\nu_\circ}}
\tag{368}
$$

An easier way to remember this is

$$
\nu L_\nu = \left(\frac{\gamma N(\gamma)}{2}\right)\left(\frac{4}{3}\gamma^2 c\sigma_T u_B\right)
\tag{369}
$$

84

which is (power/octave) = (electrons/half-octave) times (power/electron). Only a half-octave in the electron distribution is used because the quadratic dependence of ν on γ means that a half-octave in γ covers a full octave in ν.

For a power law $N(\gamma) \propto \gamma^{-\beta}$, we get a power law spectrum. From

$$L_\nu \propto \gamma N(\gamma) \propto \sqrt{\nu/\nu_\circ} N\left(\sqrt{\nu/\nu_\circ}\right) \propto \nu^{(1-\beta)/2} \tag{370}$$

we see that the power index in $L_\nu \propto \nu^{-\alpha}$ is

$$\alpha = \frac{\beta - 1}{2} \tag{371}$$

Thus the power law index $\beta = 2.5$ seen in the cosmic rays translates into $\alpha = 0.75$ which is a typical spectral index for a non-thermal radio source.

13.4. Energy in Relativistic Particles

The relation between synchrotron emission and the number of electrons with γ's close to $\sqrt{\nu/\nu_\circ}$ allows us to compute the total energy in relativistic electrons in a radio source from its synchrotron spectrum. We can write

$$E_{part} = \int N(\gamma)\gamma mc^2 d\gamma \tag{372}$$

and

$$\gamma N(\gamma) = \frac{3\nu_\circ L_\nu}{2c\sigma_T u_B} \tag{373}$$

But $d\gamma = \gamma d\nu/(2\nu)$ so

$$E_{part} = \int \frac{3\nu_\circ \gamma mc^2 L_\nu}{4c\sigma_T u_B} \frac{d\nu}{\nu} = \int \frac{3\sqrt{\nu_\circ \nu} mc^2 L_\nu}{4c\sigma_T u_B} \frac{d\nu}{\nu} \tag{374}$$

which becomes

$$E_{part} = \frac{\sqrt{24\pi emc}}{\sigma_T} B^{-3/2} \int \sqrt{\nu} L_\nu d\ln\nu = 1.5 \times 10^{12} B^{-3/2} \int \sqrt{\nu} L_\nu d\ln\nu \tag{375}$$

in cgs units. Thus if we know the magnetic field, we can compute the energy in relativistic electrons needed to produce the observed synchrotron radiation.

In the local ISM the cosmic ray energy density approximately equals the magnetic field energy density. Since the cosmic rays are only held to the galaxy by the magnetic field, if the magnetic field energy density was less than the cosmic ray energy density the cosmic rays would rapidly escape. And apparently if the cosmic ray energy density is less than

the magnetic field energy density Alfven turbulence accelerates cosmic rays until the energy densities are comparable. If we assume the energy density in relativistic electrons equals the magnetic field energy density, we derive the *equipartition* magnetic field strength for a radio source. As seen above, the particle energy is $E_{part} = AB^{-3/2}$ where is A depends on observed quantities and constants. The magnetic field energy is $E_{mag} = B^2V/8\pi$ where V is the volume of the source. Setting these equal gives

$$B_{eq} = \left(\frac{8\pi A}{V}\right)^{2/7} \tag{376}$$

We can also find the B that minimizes the total energy needed to produce the observed synchrotron emission,

$$B_{ME} = \left(\frac{6\pi A}{V}\right)^{2/7} \tag{377}$$

The difference between these is an insignificant factor of $0.75^{2/7} = 0.92$.

A more significant problem is the protons. In the local cosmic rays, we see about $dN/dE = 1 \times E^{-2.5}$ protons/cm^2/sec/sr/GeV and $0.01 \times E^{-2.5}$ electrons/cm^2/sec/sr/GeV. So it might be reasonable to assume that E_{part} should be multiplied by 100 to allow for relativistic protons which we cannot see because they don't radiate synchrotron emission. This would increase B_{eq} by a factor of $100^{2/7} = 3.7$. But if we compute $N(\gamma)$ at $\gamma = 2000$ for the protons we find 6.1×10^{-9} while for electrons we find $N(\gamma = 2000) = 4.8 \times 10^{-6}$. Thus while electrons of a given energy are 100 times less common that protons of the same energy, electrons of a given γ are 800 times more numerous than protons of the same γ. As a result, we don't know whether to apply the proton correction to B_{eq}.

13.5. Synchrotron Self-Absorption

A population of relativistic electrons produces a synchrotron emissivity of

$$4\pi j_\nu = \left. \frac{2\gamma n(\gamma) c\sigma_T u_B}{3\nu_\circ} \right|_{\gamma=\sqrt{\nu/\nu_\circ}} \tag{378}$$

where $n(\gamma)$ is the number density of electrons per unit γ. If there were no corresponding absorption process, then a line of sight with a very long path length through a synchrotron source would have an intensity that grew without limit as the path length increased. But there is a synchrotron absorption process, and we can calculate its magnitude using Kirchoff's law: $B_\nu(T_{ex}) = j_\nu/\alpha_\nu$. But what is the excitation temperature T_{ex} for the transitions involved in synchrotron emission? What are the quantum levels?

The orbits of electrons around the magnetic field lines are actually quantized with angular momenta in units of $\hbar$. The angular momentum is pR and $R = pc/eB$ so the quantum

number is $n = p^2c/\hbar eB = \gamma^2\beta^2m^2c^3/\hbar eB$. We can write this as $n = \gamma^2\beta^2 B_{crit}/B$ where $B_{crit} = 5 \times 10^{13}$ gauss is the magnetic field that makes the cyclotron frequency equivalent to the electron rest mass. For a typical $\gamma = 10^4$ and $B = 10^{-4}$ gauss, $n = 5 \times 10^{21}$, so a purely classical description of synchrotron emission is appropriate. The typical transition in synchrotron emission is not $\Delta n = 1$, but rather $\Delta n \approx \gamma^3$, so both n and Δn can be taken to be continuous variables. This gives

$$kT_{ex} = \frac{\Delta E}{\ln(n_1/g_1) - \ln(n_2/g_2)} \quad \Rightarrow \quad kT_{ex} = -\left(\frac{\partial \ln(n/g)}{\partial E}\right)^{-1} \tag{379}$$

Now we need to know the statistical weight g for states in a unit range of γ. This is given by phase space volume in units of h^3, $g = g_s d^3p\, d^3x/h^3$, where $g_s = 2$ is the statistical weight for spins. The volume factor dx^3 cancels out, and for the ultrarelativistic limit, $\gamma >> 1$, we have $d^3p \propto \gamma^2 d\gamma$. Therefore, when $n(\gamma) \propto \gamma^{-\beta}$, the excitation temperature is

$$kT_{ex} = -\left(\frac{\partial \ln(\gamma^{-(\beta+2)})}{mc^2\partial\gamma}\right)^{-1} = \frac{\gamma mc^2}{\beta + 2} \tag{380}$$

Since $\gamma = \sqrt{\nu/\nu_\circ}$, and since kT_{ex} is always much greater than the energy of the photons seen from synchrotron sources, when a source is optically thick to synchrotron emission its surface brightness varies like $I_\nu \propto \nu^{2.5}$.

Thus observing the surface brightness of an optically thick sychrotron source gives the γ factor, and γ plus ν together give the magnetic field. Since the surface brightness is very high, with brightness temperatures $>> mc^2/k = 6 \times 10^9$ K, optically thick synchrotron sources always have very small angular sizes and can only be resolved by Very Long Baseline Interferometry (VLBI). A source with brightness temperature of 10^{12} K at 5 GHz with an angular size of $0.001''$ has a flux of $F_\nu = 2kT(\nu/c)^2\Omega \approx 20$ Jy.

14. Inverse Compton Scattering

Consider an electron moving relativistically through an isotropic distribution of photons with energy $h\nu$. A photon approaching with angles $\pi - \theta$ to the electron's velocity ($\theta = 0$ for a head-on collision) has 4-momentum

$$p^\mu = \frac{h\nu}{c}(1; -\cos\theta, -\sin\theta, 0) \tag{381}$$

if we make the x axis the electron's velocity vector and put the photon direction in the xy plane. Let $\mu = \cos\theta$. Then the photon momentum in the electron's rest frame is

$$p^{\mu\prime} = \frac{h\nu}{c}(\gamma(1 + \beta\mu); -\gamma(\beta + \mu), -\sin\theta, 0) \tag{382}$$

Now let the photon scatter off the electron, leaving with angle θ' to the x axis. For now we ignore the true Compton shift, so the energy of the outgoing photon is $h\nu'' = h\nu' = (\gamma(1 + \beta\mu))h\nu$. Note that $\theta' = 0$ corresponds to a photon traveling in the same direction as the electron, while $\theta = 0$ corresponds to a photon traveling in the opposite direction as the electron. The 4 momentum after scattering is

$$p^{\mu'''} = \frac{h\nu''}{c}(1; \cos\theta', \sin\theta', 0) \tag{383}$$

Now let $\mu' = \cos\theta'$. Then the 4 momentum after scattering back in the lab frame is

$$p^{\mu''''} = \frac{h\nu''}{c}(\gamma(1 + \beta\mu'); \gamma(\beta + \mu'), \sin\theta', 0) \tag{384}$$

We finally get the ratio of the lab frame energy after scattering to the lab frame energy before scattering:

$$\frac{\nu'''}{\nu} = \gamma^2(1 + \beta\mu)(1 + \beta\mu') \tag{385}$$

We want to find the average value of this ratio, which will tell us the total energy transfer from the electron to the photons, and also the typical frequency of the output photons. In order to find the average, we need to know the probability distributions for the scattering angles μ and μ'.

A first guess would be that $p(\mu)$ should be a uniform distribution since this corresponds to an isotropic initial photon field because $d\Omega = 2\pi d\mu$. But head-on collisions ($\mu = 1$) will happen more frequently than stern chase collisions ($\mu = -1$). In fact, the rate of photon collisions has exactly the same Doppler shift factor as the photon frequency, so $p(\mu) \propto \gamma(1 + \beta\mu)$. Normalizing this in the range $-1 < \mu < 1$ gives the probability density function (pdf):

$$p(\mu) = \frac{1 + \beta\mu}{2} \tag{386}$$

Note the function you integrate to find probabilities in a given distribution is called the probability density function and should not be confused with the distribution. Rybicki & Lightman approximate the distribution of μ' with a uniform distribution, corresponding to isotropic scattering. This gives a pdf

$$p'_{R\&L}(\mu') = \frac{1}{2} \tag{387}$$

With this approximation we can easily find the average ratio:

$$\begin{aligned}
\left\langle \frac{\nu'''}{\nu} \right\rangle &= \int\int \gamma^2(1 + \beta\mu)(1 + \beta\mu')p(\mu)p'(\mu')d\mu d\mu' \\
&= \gamma^2 \int \frac{(1 + \beta\mu)^2}{2}d\mu \int \frac{1}{2}(1 + \beta\mu')d\mu' \\
&= \gamma^2\left(1 + \frac{\beta^2}{3}\right) \tag{388}
\end{aligned}$$

The ultrarelativistic limit is $(4/3)\gamma^2$.

However, electrons do not scatter isotropically. If the actual scattering angle is ψ, the scattering rate is proportional to $(1 + \cos^2 \psi)$. We can write $\cos \psi$ as

$$\cos \psi = -\left(\mu\mu' + \sqrt{1 - \mu^2}\sqrt{1 - \mu'^2}\cos\phi\right) \tag{389}$$

where ϕ is the dihedral angle between the plane defined by the x axis and the incoming photon and the plane defined by the x axis and the outgoing photon. The rate of scattering into μ' will be proportional to the average of $(1 + \cos^2 \psi)$ over all ϕ's. Thus the true pdf $p'(\mu')$ for Rayleigh scattering electrons is given by

$$\begin{aligned}
p'(\mu';\mu) &= \frac{3}{8}\int_0^{2\pi}(1 + \cos^2\psi)\frac{d\phi}{2\pi} \\
&= \frac{3}{8}\int_0^{2\pi}\left[1 + \mu^2\mu'^2 + 2\mu\mu'\sqrt{1 - \mu^2}\sqrt{1 - \mu'^2}\cos\phi + (1 - \mu^2)(1 - \mu'^2)\cos^2\phi\right]\frac{d\phi}{2\pi} \\
&= \frac{3}{8}\left[1 + \mu^2\mu'^2 + \frac{1}{2}(1 - \mu^2)(1 - \mu'^2)\right] \\
&= \frac{3}{8}\left[\frac{3}{2}(1 + \mu^2\mu'^2) - \frac{1}{2}(\mu^2 + \mu'^2)\right] \tag{390}
\end{aligned}$$

The $(3/8)$ factor normalizes the pdf:

$$\int p'(\mu';\mu)d\mu' = \frac{3}{8}\left[3 + \mu^2 - \mu^2 - \frac{1}{3}\right] = 1 \tag{391}$$

for any value of μ. Note that the distribution of μ' does depend on μ, but is this dependence is not the probability density function for μ. By writing μ following a semicolon, we mean that μ is a parameter that modifies the μ' distribution. The joint pdf for μ and μ' is given by

$$p(\mu, \mu') = p(\mu)p'(\mu';\mu) \tag{392}$$

But the pre-scattering angle we need is not really μ, which is measured in the lab frame, but rather the pre-scattering angle in the electron's rest frame, whose cosine is given by $(\beta + \mu)/(1 + \beta\mu)$. We can now find the average ratio:

$$\begin{aligned}
\left\langle\frac{\nu'''}{\nu}\right\rangle &= \int\int\gamma^2(1 + \beta\mu)(1 + \beta\mu')p(\mu)p'(\mu';(\beta + \mu)/(1 + \beta\mu))d\mu d\mu' \\
&= \gamma^2\int\frac{(1 + \beta\mu)^2}{2}d\mu \\
&\quad\times\int\frac{3}{8}\left[\frac{3}{2}\left(1 + \left[\frac{\beta + \mu}{1 + \beta\mu}\right]^2\mu'^2\right) - \frac{1}{2}\left(\left[\frac{\beta + \mu}{1 + \beta\mu}\right]^2 + \mu'^2\right)\right](1 + \beta\mu')d\mu' \\
&= \gamma^2\left(1 + \frac{\beta^2}{3}\right) \tag{393}
\end{aligned}$$

To get the last line of this result, note that $p'(\mu')$ is always an even function of μ', so $\int \beta\mu' p'(\mu')d\mu' = 0$ and the integral $\int (1 + \beta\mu')p'(\mu')d\mu' = 1$ Thus the average is exactly the same as the result calculated in the isotropic scattering approximation used by Rybicki & Lightman.

14.1. Power Law Sources

Using the mean energy increase factor, we find that an ultrarelativistic electron loses energy by the inverse Compton process at a rate

$$P_{IC} = \frac{4}{3}\gamma^2 c\,\sigma_T\,u_{ph} \tag{394}$$

where u_{ph} is the energy density of the photons the electron is colliding with. The typical output frequency is

$$\nu_{IC} = \frac{4}{3}\gamma^2\nu_{in} \tag{395}$$

where ν_{in} is the typical frequency of the photons the electron is colliding with. Thus a distribution of relativistic electrons $N(\gamma)$ produces an inverse Compton source with luminosity

$$\nu L_\nu = \left(\frac{\gamma N(\gamma)}{2}\frac{4}{3}\gamma^2 c\sigma_T u_{ph} \right)\Bigg|_{\gamma=\sqrt{\nu/\nu_1}} \tag{396}$$

with $\nu_1 = (4/3)\nu_{in}$. The relationship between $N(\gamma)$ and the spectrum is exactly same for inverse Compton sources and synchrotron sources. The the shapes of the spectra on a log-log plot will be identical, but the scaling in luminosity and in frequency will be different.

But what are the typical sources of input photons? Three are important: the cosmic background radiation, starlight, and a source's own synchrotron radiation for compact radio sources. The CMBR is universal, so we will concentrate on it. The energy density is

$$u_{CMBR} = aT_\circ^4 = \frac{4\sigma}{c}T_\circ^4 = 4 \times 10^{-13} \text{ erg/cm}^3 \tag{397}$$

This is equivalent to the magnetic field energy density from a field of $B = 3.3\ \mu$G. The typical frequency is $\approx 3kT_\circ/h = 1.7 \times 10^{11}$ Hz. Thus if both the IC off the CMBR and synchrotron emission from a source are detected, the magnetic field can be determined from both the ratio of νL_ν's and from the ratio of frequencies. The relations to use are:

$$\frac{(\nu L_\nu)_{IC}}{(\nu L_\nu)_S} = \left(\frac{3.3\ \mu\text{G}}{B} \right)^2 \tag{398}$$

and

$$\frac{\nu_{IC}}{\nu_S} = \frac{6 \times 10^4 \text{ G}}{B} \tag{399}$$

Thus for a typical radio source the IC luminosity is much less than the synchrotron luminosity but it peaks at a much higher frequency. Typically the IC flux is in X-rays or γ-rays. An upper limit on the X-ray or γ-ray flux from a radio source gives a lower limit on the magnetic field.

When a radio source is very compact the synchrotron flux it emits can have an energy density larger than the CMBR. This energy density of synchrotron photons can then be scattered by relativistic electrons, producing X-rays by what is called the synchrotron-self-Compton process. The fraction of synchrotron photons scattered is about τ, where τ is the optical depth to electron scattering from the relativistic electrons. The energy of the scattered photons is boosted by a factor of $(4/3)\gamma^2$. If the product $(4/3)\gamma^2\tau$ is greater than one, then the IC photons will have a larger energy density than the synchrotron photons, and one will get second-order IC emission. This process continues until the energy in the electron's rest frame is comparable to mc^2. The IC photons produced in this step have energy γmc^2 – which is all of the electron's energy. This cascading of the IC process puts out energy in the γ-ray region of the spectrum, instead of the radio region. Since some quasars are compact radio sources, the IC cascade has been called "the Compton catastrophe" in quasar models. But since quasars are also γ-ray sources, it isn't all bad.

14.2. Details

We already have the most important results: the mean energy increase factor. But the distribution of output energies is also good to know. The maximum possible output:input ratio is $\gamma^2(1+\beta)^2$, but since $\gamma^2 = (1-\beta^2)^{-1} = (1+\beta)^{-1}(1-\beta)^{-1}$, the maximum energy can be written as $(1+\beta)/(1-\beta)$, which is the square of the relativistic Doppler shift formula. The minimum possible output to input ratio is $\gamma^2(1-\beta)^2$ which can be written as $(1-\beta)/(1+\beta)$. Note that this is just one over the maximum ratio.

Let $R = \nu'''/\nu$ be the output:input ratio. Then we can find the pdf for R using

$$p_R(R) = \int \int \delta(R - \gamma^2(1 + \beta\mu)(1 + \beta\mu'))p(\mu)p'(\mu';\mu)d\mu d\mu' \qquad (400)$$

In the isotropic approximation this is easy to evaluate: do the integral over μ' first, so the delta function integrates to $[\gamma^2\beta(1 + \beta\mu)]^{-1}$. This cancels the μ dependence of $p(\mu)$, so the result for $p_R(R)$ is proportional to the amount of the μ range that is compatible with a given R. This is all of it for $R = 1$, and from $\mu = \beta^{-1}[(1 - \beta)R - 1]$ to 1 for $R > 1$. Thus $p_R(R)$ declines as a linear function of R from a maximum at $R = 1$ to zero at $R = R_{max} = (1 + \beta)/(1 - \beta)$. For $R < 1$ the range in μ is from -1 to $\mu = \beta^{-1}[(1 + \beta)R - 1]$, so again $p_R(R)$ declines linearly from the maximum at $R = 1$ to zero at $R = R_{min} = (1 - \beta)/(1 + \beta)$.

Finally we should calculate the mean squared output:input ratio. In the isotropic ap-

proximation used by R&L, this quantity is

$$
\begin{aligned}
\left\langle \left(\frac{\nu'''}{\nu}\right)^2 \right\rangle &= \int\int \gamma^4 (1+\beta\mu)^2 (1+\beta\mu')^2 p(\mu) p'(\mu';\mu) d\mu d\mu' \\
&= \gamma^4 \int \frac{(1+\beta\mu)^3}{2} \int \frac{(1+\beta\mu')^2}{2} d\mu' d\mu \\
&= \gamma^4 \int \frac{(1+\beta\mu)^3}{2} \left[1 + \frac{\beta^2}{3}\right] d\mu \\
&= \gamma^4 \left[1+\beta^2\right] \left[1 + \frac{\beta^2}{3}\right] \\
&= \gamma^4 \left[1 + \frac{4}{3}\beta^2 + \frac{1}{3}\beta^4\right] \tag{401}
\end{aligned}
$$

This differs from the exact result only in the β^4 term.

15. Plasma Effects

Chapter 8 of R&L

15.1. Dispersion

Assume that the electrons in the plasma are moving with $\beta << 1$. An electromagnetic wave has $|\vec{E}| \approx |\vec{B}|$ (exactly equal in a vacuum of course) which means that the force due to the wave's magnetic field can be neglected. However, there can be a static magnetic field in the plasma, $\vec{B}$. For a wave with time dependence $\propto e^{-i\omega t}$, we get

$$
-im\omega\vec{v} = -e(\vec{E} + c^{-1}\vec{v} \times \vec{B}) \tag{402}
$$

We need to calculate the dielectric constant of the medium, ϵ, in order to calculate the phase velocity of the wave, $v_p = c/\sqrt{\epsilon\mu}$, since $\mu = 1$.

Now the dielectric constant is

$$
\epsilon = 1 + 4\pi\chi \tag{403}
$$

where the permittivity χ is the ratio of the dipole density to the applied electric field. The dipole density is $-en_e x = -en_e \vec{v}/(-i\omega)$. Since both the electric field and the dipole density are vectors, χ is a tensor. In order to simplify the analysis, take the magnetic field to be along the z-axis. Then an electric field along the z-axis generates velocities along the z-axis, which give zero cross product with B, so the z behavior separates out, giving:

$$
v_z = \frac{-ieE_z}{m\omega} \tag{404}
$$

Since $x = v/(-i\omega)$ and the dipole density is $-n_e e x$, one has

$$\chi_{zz} = \frac{-e^2 n_e}{m\omega^2} \tag{405}$$

$$\epsilon = 1 - \frac{4\pi e^2 n_e}{m\omega^2} = 1 - \left(\frac{\omega_p}{\omega}\right)^2 \tag{406}$$

where the *plasma frequency* is $\omega_p = \sqrt{4\pi e^2 n_e/m}$.

The index of refraction is $n = \sqrt{\epsilon}$. Because of variations in the electron density, rays are bent as they travel through the interstellar medium. This leads to interstellar scintillation for point-like radio sources. The equation for ray paths in refraction is

$$\frac{d(n\hat{k})}{ds} = \vec{\nabla} n \tag{407}$$

following Eqn (8.18) in R&L. Since n is less than one, convex concentrations of electrons yield diverging lenses.

The intensity I is not then conserved along a ray, but rather I/n^2.

Phase *vs.* group velocity

Note that for $\omega < \omega_p$, the dielectric constant is negative, which gives an *imaginary* phase velocity. Thus the wave vector $k = \omega/v_p$ is also imaginary, so the spatial propagation $\propto e^{ikx}$ gives a decaying exponential instead of a propagating wave. Thus waves below the plasma frequency are absorbed. The relationship between k and ω, known as the *dispersion relation*, is

$$c^2 k^2 = (\omega^2 - \omega_p^2) \tag{408}$$

For $\omega > \omega_p$ the dielectric constant is positive but less than unity. This implies a phase speed that is faster than the speed of light. How can this be if *no thing* can travel faster than the speed of light? The wave crests in a pulse don't correspond to signals, but the peak of the amplitude of a pulse composed of several frequencies does correspond to a photon or signal. Consider a propagating pulse with complex amplitude $A(\omega)$. Then the electric field will be

$$E(x,t) = \int A(\omega) \exp(ikx - i\omega t) d\omega \tag{409}$$

The amplitude of this pulse will be maximized at the x where the phase of the integrand is constant near the peak of $|A|$. Let $A = |A|e^{i\phi}$ and assume that $d\phi/d\omega = 0$ at the peak of $|A|$. This corresponds to the pulse peaking at $x = 0$ when $t = 0$. At a later time $t > 0$, the location of the peak of the pulse is found by setting

$$d[\phi + kx - \omega t]/d\omega = 0 = d\phi/d\omega + x dk/d\omega - t = 0 \tag{410}$$

Thus the peak is at $x = (d\omega/dk)t$ and the *group velocity* which describes the motion of the peak is

$$v_g = \frac{d\omega}{dk} \tag{411}$$

The value of the group velocity can be found using

$$v_p v_g = \frac{\omega d\omega}{k dk} = \frac{d(\omega^2)}{d(k^2)} = c^2 \tag{412}$$

since the dispersion relation gives a linear relation between k^2 and ω^2 with a slope of c^2.

Thus the group velocity of the wave with an electric field parallel to the static magnetic field is

$$v_g = c\sqrt{1 - \left(\frac{\omega_p}{\omega}\right)^2} \tag{413}$$

If a broadband pulse is emitted simultaneously at all frequencies, then it will be received first at high frequencies, then later and later at lower and lower frequencies. This turns the pulse into a falling chirp instead of a click. The time it takes for a given frequency to arrive is given by

$$t_\nu = \int \frac{ds}{v_g} \approx \int (1 + \frac{\nu_p^2}{2\nu^2})\frac{ds}{c} = \int \frac{ds}{c} + \frac{e^2}{2\pi mc\nu^2} \int n_e ds \tag{414}$$

With the dispersion measure DM being $\int n_e ds$ in units of cm^{-3}pc, the delay for frequency ν is

$$\Delta t_\nu = 0.415 \, DM \left(\frac{100 \text{ MHz}}{\nu}\right)^2 \text{ sec} \tag{415}$$

For pulsars the difference in delay times between two different frequencies is easily measured, giving a determination of DM. For a few pulsars distances can be estimated, and the typical ratio of DM to distance is the mean electron density in the ISM, $\langle n_e \rangle = 0.03$ cm^{-3}. For other pulsars distances can be estimated from the dispersion measure: distance $\approx DM/\langle n_e \rangle$.

15.2. Rotation Measure

The components of the velocity perpendicular to the magnetic field get coupled by the cross product, so the equations for v_x and v_y are:

$$\begin{aligned}
-im\omega v_x &= -e(E_x + c^{-1}v_y B_z) \\
-im\omega v_y &= -e(E_y - c^{-1}v_x B_z)
\end{aligned} \tag{416}$$

If we consider $v_+ = v_x + iv_y$ we get an equation

$$v_x + iv_y = -i(e/m\omega)(E_x + iE_y + c^{-1}B_z(v_y - iv_x)) \tag{417}$$

94

or

$$v_+ = -i\frac{e}{m\omega}E_+ - \frac{eB_z}{mc\,\omega}v_+ \tag{418}$$

with E_+ being $E_x + iE_y$, a circularly polarized wave. Thus

$$v_+ = -i\frac{e}{m(\omega + \omega_B)}E_+ \tag{419}$$

where ω_B is the cyclotron frequency eB/mc. For the opposite sense for circular polarization, it is clear that

$$v_- = -i\frac{e}{m(\omega - \omega_B)}E_- \tag{420}$$

The displacement x is still $x = v/(-i\omega)$, so the dielectric constants for the two circular polarizations are

$$\epsilon_{R,L} = 1 - \frac{\omega_p^2}{\omega(\omega \pm \omega_B)} \tag{421}$$

Over a path, the accumulated phase is

$$\phi_{R,L} \approx \int \left(\omega - \frac{\omega_p^2}{2(\omega \pm \omega_B)}\right)\frac{ds}{c} \tag{422}$$

Thus the phase difference between the right and left circularly polarized waves is

$$\phi_R - \phi_L = \Delta\phi \approx \int \frac{\omega_p^2\omega_B}{\omega^2}\frac{ds}{c} \tag{423}$$

The rotation of a plane-polarized wave is $\Delta\theta = \Delta\phi/2$, so

$$\Delta\theta = \int \frac{\omega_p^2\omega_B}{2\omega^2}\frac{ds}{c} \tag{424}$$

which for a constant B is just $\Delta\theta = \omega_B\Delta t_\nu$. The average magnetic field along the line of sight, weighted by electron density, is found using

$$\langle B_{||}\rangle = \frac{\int n_e\vec{B}\cdot\vec{ds}}{\int n_e ds} = \frac{mc\Delta\theta}{e\Delta t_\nu} \tag{425}$$

In general we don't know the original angle of polarization (or the time of emission), but we can measure the change in polarization angle between two frequencies and compare it to the change in arrival time between the same two frequencies for pulsars, giving a well determined $\langle B_{||}\rangle$. In the ISM this field is 2.5 μG. In typical synchrotron sources the emission is polarized so $\int n_e\vec{B}\cdot\vec{ds}$ can be found, but without pulses one needs to assume an n_e to get a magnetic field estimate.
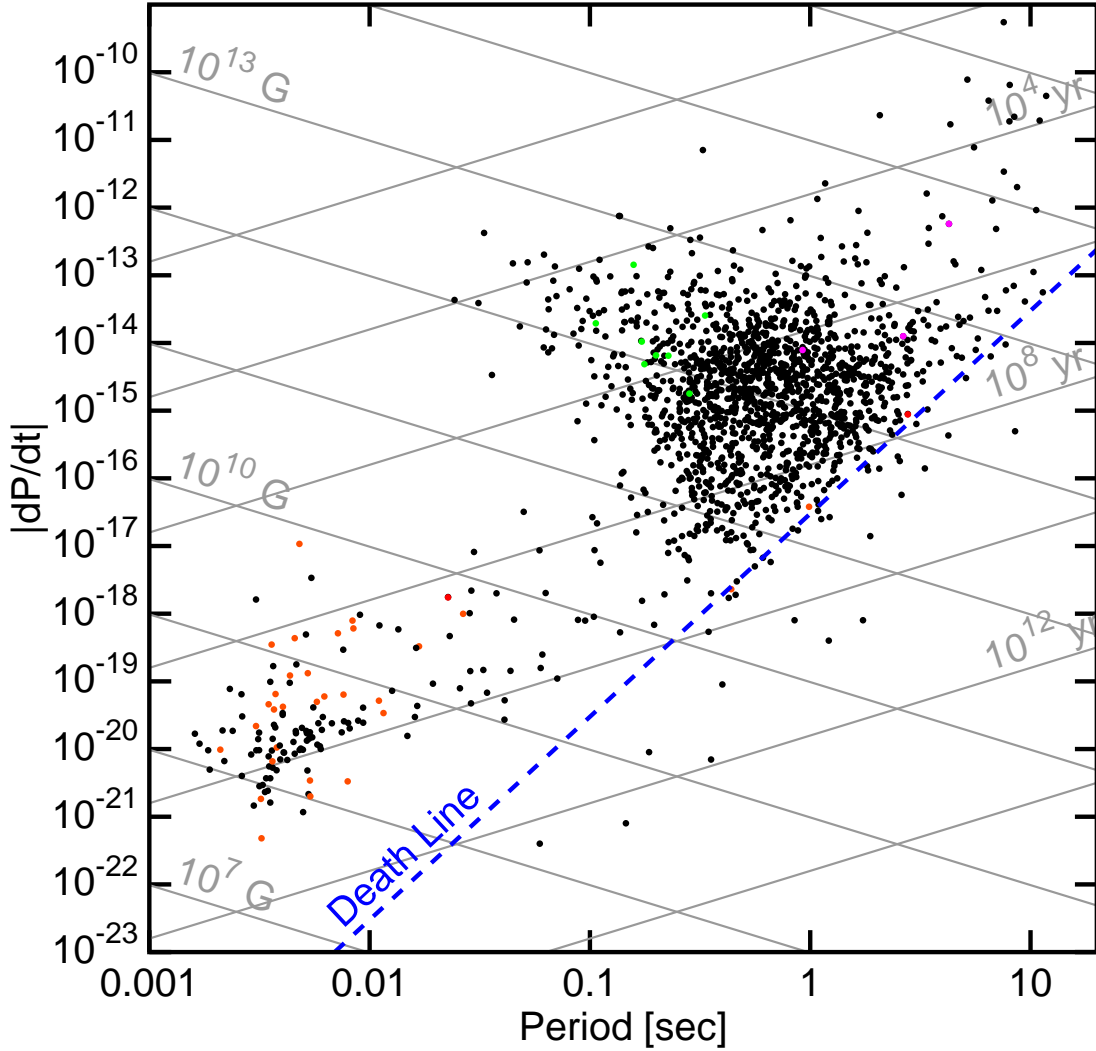
Fig. 21.— The distribution of 1743 known pulsars in the P-$\dot{P}$ plane. The orange dots have negative $\dot{P}$, probably due to gravitational acceleration in the galaxy or a globular cluster. The magenta dots are recently discovered occasional pulsars that only emit a pulse several times per day, even though their rotation periods are a few seconds (astro-ph/0511587). The green dots are γ-ray only pulsars from Pletsch *et al.*(arXiv:1111.0523). The red dots are the components of the double pulsar, J0737-3039. The blue dashed line shows the empirical "death line" for radio pulsars, with $\dot{P} \propto P^3$. The few points in the far upper right are the so-called "magnetars" or anomalous X-ray pulsars. The main clump of points are the "young" pulsars. The points in the lower left are the "recycled" pulsars that were spun up by accretion in a binary system.

16.   Pulsar Emission

One of the most interesting astrophysical E&M emissions are pulses from pulsars. This is also one of the least well understood processes, but a few basic ideas can be easily stated. Pulsars are rapidly rotating neutron stars with very large magnetic fields. The energy that drives the emission is coming from the rotational kinetic energy for radio pulsars. Therefore the pulse rate is slowing down. One can define a spindown or characteristic age using the pulse period P and the spindown rate $\dot{P}$. A characteristic time is $P/\dot{P}$. The moment of inertia of a neutron star is about 10^{45} gm cm^2, and the rotational kinetic energy is $E = 0.5\,I\omega^2 \approx 10^{48}(140$ ms$/P)^2$ ergs. For a rapidly spinning pulsar like the Crab pulsar this kinetic energy is larger that the total light output of a supernova, 10^{49} ergs. The spindown luminosity is

$$L = -I\omega\dot{\omega} = 4\pi^2 I\dot{P}/P^3. \tag{426}$$

This luminosity is usually assumed to be magnetic dipole radiation. An electric dipole with a charge $-q$ at the origin and a charge q at $z << R$ has a dipole moment of $\mu_e = qz$ and a field at the pole $z = R$ of $E_p = 2qz/R^3$. Therefore $\mu_e = 0.5\,E_pR^3$. Thus the magnetic dipole moment of a neutron star with polar surface field of B_p and radius R is $\mu_m = 0.5\,B_p\,R^3$. The power radiated by a spinning electric dipole is $P = (2\mu_e^2\omega^4/3c^3)$ and the power radiated by a spinning magnetic dipole is $P = (2\mu_m^2\omega^4/3c^3)$ due to the nice symmetry between electric and magnetic quantities in cgs units. Setting this equal to the spindown luminosity gives

$$\frac{B_p^2\,R^6\,\omega^4}{6c^3} = -I\omega\dot{\omega} \tag{427}$$

so an estimate for the surface magnetic field is given by

$$B_p = \left(\frac{6c^3 I}{4\pi^2 R^6}\right)^{0.5}\sqrt{P\dot{P}} = 2\times 10^{12}\text{ gauss}\left(\frac{I}{10^{45}}\right)^{0.5}\left(\frac{R}{10\text{ km}}\right)^{-3}\left(\frac{P}{1\text{ sec}}\frac{\dot{P}}{10^{-15}}\right)^{0.5} \tag{428}$$

The equatorial magnetic field is half this, so $B_{eq} = 3.2\times 10^{19}\sqrt{P\dot{P}}$ G. Note that this assumes an oblique rotator where there is an angle $\theta = \pi/2$ between the dipole axis and the spin axis. For general θ the luminosity goes like $\sin^2\theta$. Because $\dot{\omega} \propto -\omega^3$ if B is constant, we find that $\Delta\omega^{-2} \propto \Delta t$ or $P \propto t^{1/2}$. Therefore the characteristic age of a pulsar is taken to be

$$\tau = \frac{P}{2\dot{P}}. \tag{429}$$

For the Crab nebula pulsar the characteristic age is 1240 years while the actual age is 960 years. The Crab nebula pulsar emits about $10^{30.5}$ erg/sec in the radio and 10^{35} erg/sec at IR through γ-ray wavelengths, but the spin down luminosity is $10^{38.4}$ erg/sec which is much larger. Clearly 99.9% of the spindown luminosity must go into accelerating relativistic particles which power the Crab nebula itself.

The "death line" is the observed boundary of the pulsar distribution, as shown in Figure 21. It has been described as $B/P^2 = 0.17 \times 10^{12}$ Gauss/sec^2 which would give $\dot{P} \propto P^3$

(Faucher-Giguere & Kaspi, astro-ph/0512585, heareafter F-G&K). This corresponds to a spin-down luminosity of $-\dot{E} = 10^{30.1}$ erg/sec. F-G&K find that the radio luminosity of a pulsar is on average proportional to the square root of the spin-down luminosity.

F-G&K find that pulsars are born with a broad range of periods, $P_{birth} = 300 \pm 150$ milliseconds.

Pulsar positions can be measured very accurately by timing. The motion of the Earth produces a $\pm 500 \cos \beta$ second variation of pulse arrival times with a one year period whose phase gives the ecliptic longitude. β is the ecliptic latitude. Since pulse arrival times can be measured to μsec accuracy the angular position can be found to milli-arcsecond accuracy. Proper motions and occasionally parallaxes can also be measured. The studies of proper motions show that pulsars are typically born with a "kick" velocity broadly distributed around a mean of 380 km/sec.

The interstellar plasma causes a dispersion of pulses into chirps with low frequencies arriving later, and this dispersion measure can be used to estimate distances to pulsars.

Since pulsar pulses are quite narrow in rotational longitude, it is expected that the beam is also narrow in latitude. Then only a small fraction of pulsars will be observable from the Earth. Using a model which gives a 10% average for this beaming fraction, F-G&K get a pulsar birth rate of 2.8/century in the Milky Way.

Pulsar radio emission mechanisms are not well understood. But it is clear that coherent emission by a large number of charges moving together must be involved. For a flux of 1 Jy at a frequency of 100 MHz from a source of radius 10 km at a distance of 1 kpc the brightness temperature is

$$T_e = \frac{I_\nu}{2k(\nu/c)^2} = \frac{10^{-23} \times 90000 \times (3.08 \times 10^{21})^2}{2 \times \pi \times (10^6)^2 \times 1.38 \times 10^{-16}} = 10^{28} \text{ K} \tag{430}$$

which is too high to come from single electron processes. In order to produce a large number of charges it is thought that γ-ray photons produced by curvature radiation must have enough energy to allow pair production by colliding with the virtual photons of the pulsar magnetic field. The cyclotron frequency in a Teragauss field is 11 keV so 25 MeV γ-rays are needed. Radiating 25 MeV photons by synchrotron radiation in a field with a cyclotron frequency of 11 keV requires electrons with $\gamma = 50$. An electron with this energy has a synchrotron lifetime of 10^{-17} sec in a magnetic field of 10^{12} gauss so the pitch angles will rapidly decay to zero. Then the radiation is caused by curvature radiation, where the radius is determined by the curvature of the field line instead of pc/eB.

The field lines in a dipole field are in fact circles, but even for more general shapes we only need to look at the local radius of curvature. If we calculate an effective magnetic field that would give the same radius as the curvature of the field line we get $\gamma\beta mc^2 = eB_{eff}R$. Then

the critical energy for curvature radiation can be found using the formulae for synchrotron radiation in this field. We should not average over pitch angles in this calculation. This gives

$$\omega_c = 1.5\gamma^2 \frac{eB_{eff}}{mc} = 1.5\gamma^3 \beta \frac{c}{R} \tag{431}$$

For this to reach 25 MeV requires that $\gamma = 10^6$. We can calculate the power radiated just from the motion of the electron around a circle, with an acceleration $a = v^2/R$ which is increased by a factor of γ^2 in the MCIF. The power per electron radiated by curvature radiation is then

$$P = \frac{2e^2a^2}{3c^3} = \frac{2\gamma^4\beta^4e^2c}{3R^2} \tag{432}$$

For $\gamma = 10^6$ this is 4600 erg/sec for $R = 10$ km. Thus an electric field of $E = P/ec$ is needed to keep the electrons going fast enough to emit 25 MeV γ-rays. Converting the power into eV/sec and then dividing by c gives an electric field of $E = 10^5$ Volt/cm. In cgs units this is 300 statvolts/cm which is 3 billion times smaller than the magnetic field.

16.1. γ-ray Pulsars

Many radio pulsars have been seen in pulsed γ-rays, but with the Fermi Gamma-ray Space Telescope the γ-ray observations are sensitive enough to find pulsars that cannot be seen in radio observations. Pletsch *et al.*(arXiv:1111.0523) found 9 γ-ray pulsars but only one could be seen in radio emission.

17.   Magnetic Dipole Emission

As discussed for pulsars, a time-varying magnetic dipole produces the same radiated power as a time-varying delectric dipole because in cgs units the magnetic field has the units as the electric filed, and a magnetic dipole has the same units as an electric dipole. But magnetic dipole transitions are considered to be forbidden transitions. Why is this so?

Consider an electron orbiting around with an orbit radius of $a_\circ = 0.53$ Å, with velocity v. If there is an applied magnetic field B in the plane of the orbit, the electron feels a force $F = e(v/c)B\cos\theta$ with a lever arm $a_\circ\cos\theta$, where θ is the angle around the orbit, so the torque obtained by averaging around the orbit is $T = 0.5e(v/c)a_\circ B$ and the magnetic dipole is $\mu_{orb} = 0.5e(v/c)a_\circ$. Typically $(v/c) = \alpha = 1/137$ so the magnetic dipole is more than 100 times smaller than a typical electric dipole of $\mathcal{O}(ea_\circ)$.

Note that $\hbar^2/(2m_e a_\circ^2) = 0.5e^2/a_\circ$ giving $a_\circ = \hbar^2/(m_e e^2)$, $v_t = \hbar/(m_e a_\circ) = e^2/\hbar = \alpha c$. Thus $\mu_{orb} = 0.5e(e^2/\hbar c)(\hbar^2/(m_e e^2)) = 0.5e\hbar/m_e c = \mu_B$, the Bohr magneton.

Thus magnetic dipole transitions typically have Einstein A values that are α^2 or $\mathcal{O}(10^4)$ times smaller than electric dipole transitions.

18.   Molecular Hydrogen Emission

The most common molecule in the Universe is H_2, but because it is homo-nuclear it has no permanent dipole moment. As a result, the vibrational-rotational emission from H_2 is "forbidden", and occurs through the weak channel of electric quadrupole radiation.

For a typical quadrupole moment of $Q = ea_\circ^2$ compared to a typical dipole moment of $\mu = ea_\circ$, the electric quadrupole interaction energy which goes like $\vec{\nabla}\vec{E}Q$ is $ka_\circ$ times smaller than the dipole interaction energy $\mu \cdot \vec{E}$. Since the wavenumer k is about $10^5/$cm for electronic transitions and $10^3/$cm for infrared transitions, the electric quadrupole lines will be 10^6 to 10^{10} times weaker than electric dipole lines. If an electric dipole line at 82,300 cm^{-1} has an Einstein $A = 6 \times 10^8$ sec^{-1} like Lyman-α, at 354 cm$^-$ one would expect $A \approx 48$ sec^{-1} since $A \propto \nu^3$ for dipole transitions. The actual Einstein A for the 2-0 rotational line of H_2 is 10^{12} times smaller than this.

Consider two charges q at positions given by $z = \pm[z_\circ + \Delta z \cos \omega t]$, along with a stationary charge $-2q$ at $z = 0$. Each charge produces a radiated field given by

$$E_{rad} = \frac{-qa(t - r/c)\sin\theta}{c^2 r}\hat{t} \tag{433}$$

where θ is the angle to the z-axis, a is the acceleration $-\omega^2\Delta z$, $\hat{t}$ is a unit vector perpendicular to $\vec{r}$ in the plane defined by $\vec{a}$ and $\vec{r}$. The two accelerating charges are accelerating in opposite directions so the radiated field will cancel out except for the differences in r, θ and the phase produced by the displacements $\pm z_\circ$.

We want to evaluate the radiated power in the limit that $\lambda >> z_\circ$. Also the oscillatory motion $\Delta z << z_\circ$. And finally $r >> z_\circ$.

Difference in θ are $\mathcal{O}(z_\circ/r)$, so the field difference is $\mathcal{O}(z_\circ/r^2)$ and does not lead to a radiated field which has to go like $\mathcal{O}(1/r)$. The same happens for differences in the $1/r$ factor. The only difference that leads to a radiated field is the r in the retarded time. The resultant field is

$$E_{rad} = \frac{(2qz_\circ\Delta z)\omega^3}{c^3 r}\sin\theta\cos\theta\sin(\omega(t - r/c)) \tag{434}$$

We obtain the radiated power by integrating $cE_{rad}^2/4\pi$ over the area of a sphere at radius r. The average of $\sin^2\theta\cos^2\theta$ over 4π is $\int_0^1(1 - \mu^2)\mu^2 d\mu = 1/3 - 1/5 = 2/15$. The time average of $\sin^2\omega t$ is $(1/2)$. This gives

$$P = \frac{4q^2 z_\circ^2 \Delta z^2 \omega^6}{15c^5} \tag{435}$$

The *quadrupole moment tensor* is

$$Q_{\alpha\beta} = \sum_i q_i(3x_\alpha x_\beta - r^2\delta_{\alpha\beta}) \tag{436}$$

which for this model is diagonal with $Q_{zz} = 4qz_\circ^2$ and $Q_{xx} = Q_{yy} = -2qz_\circ^2$. The third time derivative of this is

$$\frac{d^3Q}{dt^3} = (-4qz_\circ\Delta z, -4qz_\circ\Delta z, 8qz_\circ\Delta z)\sin(\omega(t - r/c)) \tag{437}$$

The summed squared amplitude of the third time derivative of Q is

$$\sum_{\alpha\beta}\left|\frac{d^3Q_{\alpha\beta}}{dt^3}\right|^2 = 96q^2z_\circ^2\Delta z^2\omega^6\sin^2(\omega(t - r/c)) \tag{438}$$

So in terms of the time averaged squared third derivative,

$$P = \frac{1}{180c^5}\sum_{\alpha\beta}\left\langle\left|\frac{d^3Q_{\alpha\beta}}{dt^3}\right|^2\right\rangle \tag{439}$$

The Einstein A's for the $2-0$, $3-1$, $4-2$, $5-3$ & $6-4$ lines of H_2 are $A = 2.9\times10^{-11}$, 4.7×10^{-10}, 2.7×10^{-9}, 9.8×10^{-9} & 2.6×10^{-8} sec^{-1}. The quadrupole moment is $\langle 0, J|Q|0, J+2\rangle \approx 0.98ea_\circ^2 = 1.32\times10^{-26}$ esu-cm^2 where $a_\circ$ is the Bohr radius 5.29×10^{-9} cm. This is the component along the inter-nuclear axis. Note that the electrons spend more time between the nuclei, leading to a net positive charge at larger radius. Call this component Q and represent the molecule as two charges of $+e$ at position $x = \pm\sqrt{Q/4e} \approx \pm 0.5a_\circ$. Then as the molecule rotates at frequency ω, the quadrupole tensor is $Q_{xx} = Q[0.75\cos(2\omega t) + 0.25]$, $Q_{yy} = Q[-0.75\cos(2\omega t) + 0.25]$, $Q_{zz} = -0.5Q$, and $Q_{xy} = Q_{yx} = Q[0.75\sin(2\omega t)]$. The power radiated at wavenumber $\tilde\nu = 2\omega/2\pi c$ is

$$P = \frac{64\omega^6Q^2}{180c^5}\left(\frac{1}{2}\right)\left(4\times\frac{9}{16}\right) = \frac{64\pi^6\tilde\nu^6cQ^2}{180}\times\frac{9}{8} = \frac{2\pi^6\tilde\nu^6cQ^2}{5} \tag{440}$$

The Einstein A is $P/h\nu$ so this classical calculation gives

$$A \approx \frac{2\pi^6\tilde\nu^6cQ^2}{5h\nu} = \frac{2\pi^6\tilde\nu^5Q^2}{5h} \tag{441}$$

The quantum Einstein A is given by

$$A(v'J'v''J'') = 1.4258\times10^4(E_{v'J'} - E_{v''J''})^5|\langle v'J'|Q|v''J''\rangle|^2f(J'J'')\text{ sec}^{-1} \tag{442}$$

where E is in atomic units of $e^2/a_\circ = 27.2$ eV, and Q is in units of $ea_\circ^2 = 1.344\times10^{-26}$ esu-cm^2 (Jean Turner, Kirby-Docken & Dalgarno, 1977, ApJS, 35, 281). The $f(J'J'')$ factor are the Clebsch-Gordon coefficents, given by:

$$f(J'J'') = \begin{cases} \frac{3(J'+1)(J'+2)}{2(2J'+1)(2J'+3)} & \text{for } J'' = J' + 2 \\ \frac{J'(J'+1)}{(2J'-1)(2J'+3)} & \text{for } J'' = J' \\ \frac{3J'(J'-1)}{2(2J'-1)(2J'+1)} & \text{for } J'' = J' - 2 \end{cases} \tag{443}$$
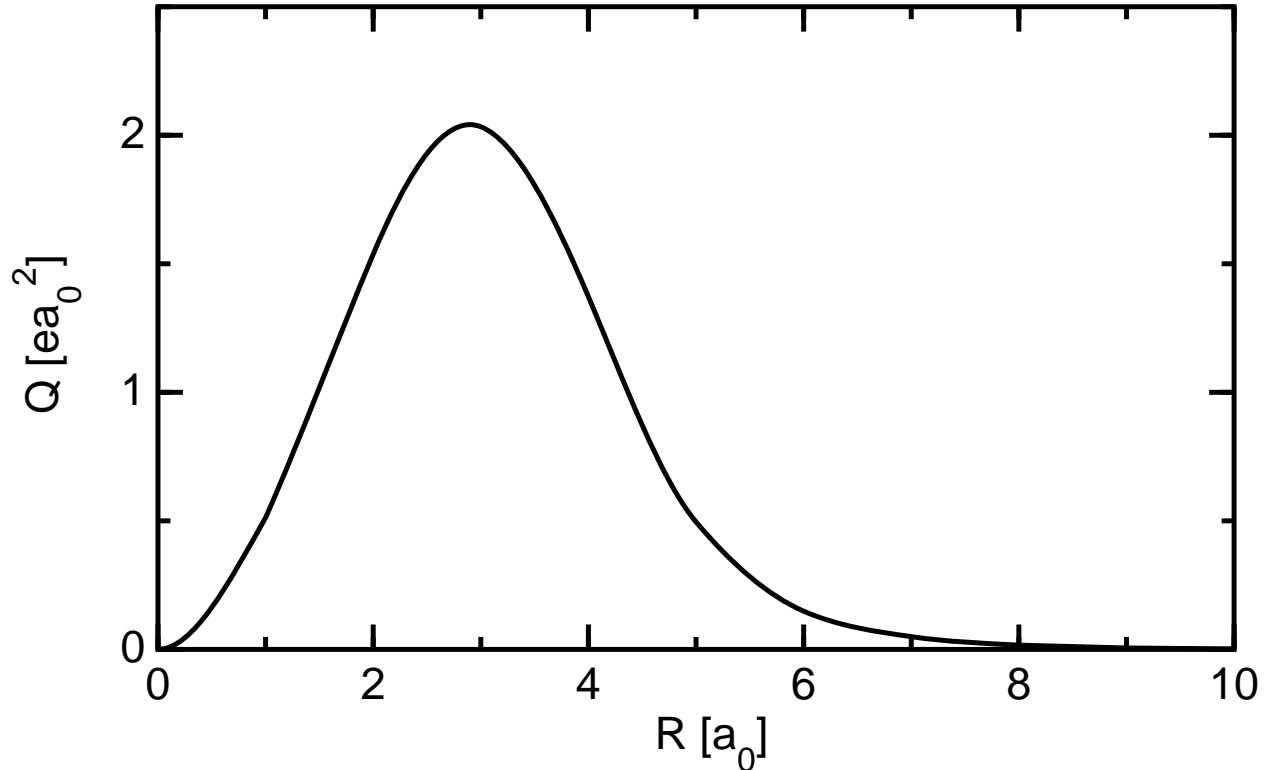
101

Fig. 22.— The electric quadrupole moment of the hydrogen molecule in its ground electronic state as a function of the internuclear separation. Atomic units in terms of $a_\circ = 0.539$ Å for distances and $ea_\circ^2$ for the quadrupole are used. Based on Figure 1 of Dalgarno, Allison & Browne (1969, JAtS, 26, 946).

For $J'' = 0$ and $J' = 2$, $f = 1/5$. The 28 μm line of H_2 has $E = 0.00162$ atomic units so the $A_{20} = 2.9 \times 10^{-11}$ sec^{-1} implies $\langle 0, 0|Q|0, 2\rangle = 0.96$ $ea_\circ^2$. The quadrupole moment is taken from Dalgarno, Allison & Browne (1969, J.Atm.Sci., 26, 946) and is defined as $Q(R) = R^2 - 2\langle\phi|3z^2 - r^2|\phi\rangle$. The R^2 term is $3z^2 - r^2$ for two protons each at $z = \pm R/2$.

The formula in Eqn(441) gives $A = 3.16 \times 10^{-8}$ for the 6-4 line for $Q = 1.00$ $ea_\circ^2$ which is appropriate for this transition because the $Q(R)$ curve for H_2 reaches a maximum at R about 1.5 times the ground state separation. The actual A_{64} is 82% of this so the formulae work well. The last line of Eqn(443) gives $f = 0.315$ which is 84% of the asymptotic $J \to \infty$ limit of 3/8.

18.1. Excitation Mechanisms

Molecular hydrogen has a very small moment-of-inertia, so the rotational B constant is large. In addition, the lowest allowed transition is the 2-0 quadrupole line, which has an
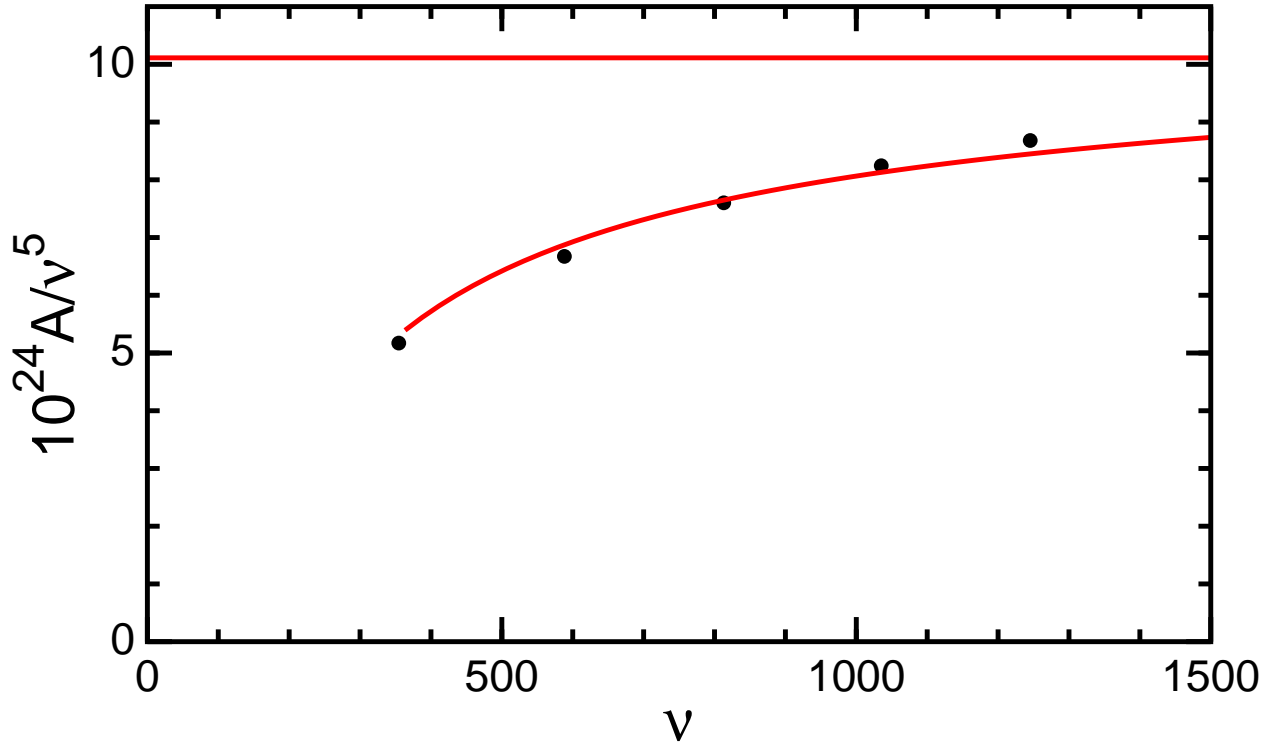
Fig. 23.— Einstein A values, normalized by $\tilde{\nu}^5$, for molecular hydrogen. The black dots are the 2-0, 3-1, 4-2, 5-3 and 6-4 lines, while the red horizontal line is the classical prediction. The red curve shows the quantum prediction for a constant value of Q which is nearly correct.

excitation energy equivalent to 512 K. Thus most molecular clouds in the ISM do not radiate much H_2 emission at all.

The protons in H_2 have spin $\frac{1}{2}$, so they can combine to make a total nuclear spin of 0 or 1. In the $K = 0$ state the rotational wavefunction is symmetric between the two nuclei so the protons have to have spins in different directions, leading to a total nucelar spin of 0. This also applies to all even K. For odd K, the rotational wavefunction is anti-symmetric so the total nuclear spin has to be 1. The interaction between the nuclear magnetic moments of the protons is very tiny, so the radiative transition between $K = 1$ and $K = 0$ by nuclear magnetic dipole radiation has a lifetime much longer than the age of the Universe. As a result, H_2 is divided into two species: para-H_2 with even K, and ortho-H_2 with odd K. The staitstical weight of even K states are $2K+1$, while the odd K states have weights $3(2K+1)$. At high temperatures one gets an ortho:para ratio of 3:1.

In the laboratory when preparing solid hydrogen, used for cooling WISE among other things, it is important to make sure that all the hydrogen is converted to the para form. In the lab this is done using a catalyst made of powdered iron which interacts well with the nuclear magnetic moments. But even though iron filings are not present in the interstellar
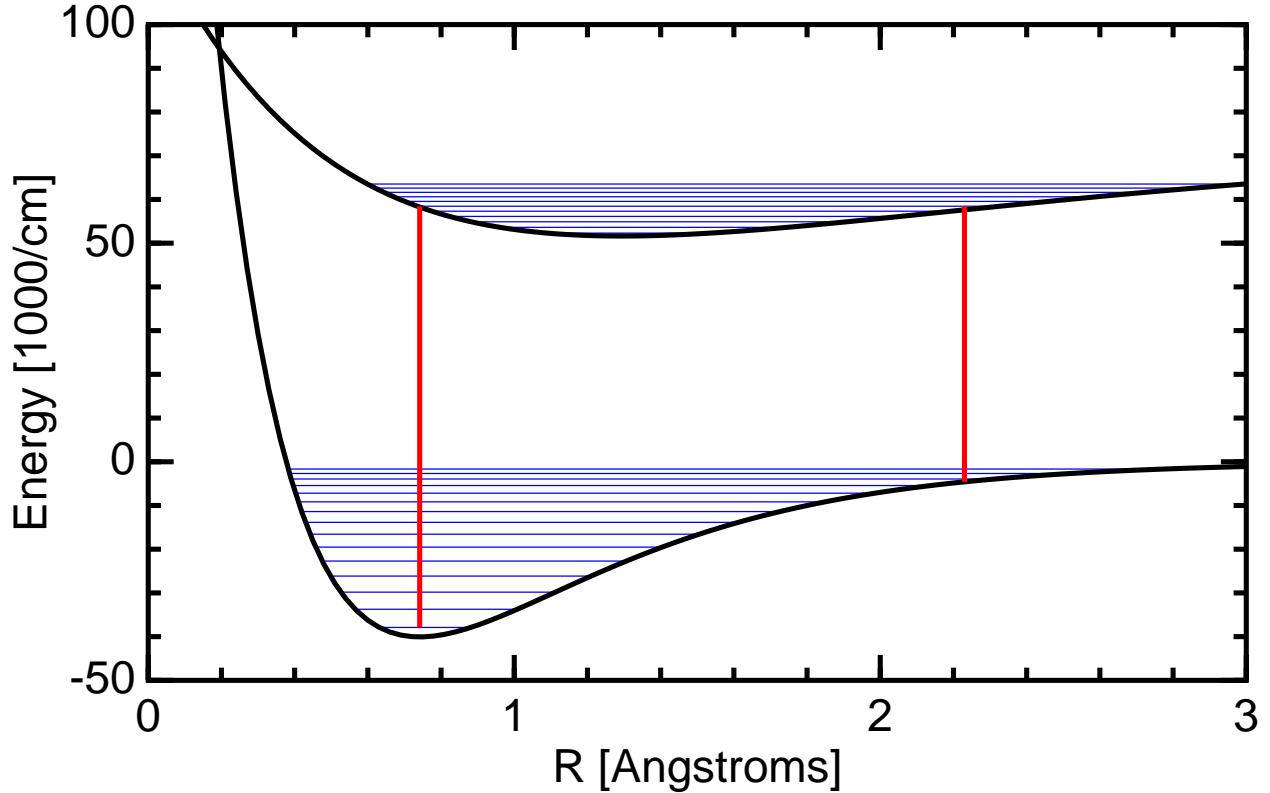
Fig. 24.— Morse potentials fitted to the $X^1\Sigma_g^+$ ground electronic state of H_2 and the $B^1\Sigma_u^+$ upper state of the Lyman bands. The vibrational levels in the Morse potential are shown as thin blue lines. The red line on the left shows the ultraviolet absorption at about 104 nm from the ground state to the $v = 5$ state in the upper levels that gives the largest overlap integral. About half the time the decay to the ground state occurs via the right red line into the $v \approx 12$ level in the ground state.

medium, the ortho:para ratio of H_2 in the ISM is usually close to the equilibrium value. The reaction which enables this equilibrium is the proton exchange process: $H_2 + p \rightarrow p + H_2$. In a cold cloud, this will happen much more often when the final state is the lowest energy $K = 0$ para state, leading to an ortho-para ratio is thermal equilibrium with the kinetic temperature of the medium.

Even though thermal excitation of H_2 requires high temperatures, rotational-vibrational lines from highly excited H_2 molecules are seen from photo-dissociation regions (PDRs), where UV light from hot young stars is absorbed by a molecular cloud. The mechanism which produces the excited H_2 is shown in Figure 24. A UV photon excites the molecule into an excited vibrational level in an excited electronic state. The vibrational frequency is about 3×10^{13} Hz, so the molecule makes a large number of the vibrations in the 10 ns or so that it takes to decay electronically. The most likely decay is then where the vibrating molecule spends most of its time, which is out at about 2.2 Å separation, so when it decays

to the electronic ground state it is most likely to decay into the $v = 11$ or 12 state. A cascade of vibrational transitions then occurs, each one adding or subtracting 2 units of angular momentum from the rotational state. By the time the $v = 0$ state is reached, a large fraction of the molecules are in states as high at $K = 10$ or more, even though the $K = 10$ level is about 10^4 K above the ground state.

Note that a direct dissociation, by an electronic transition into the vibrational continuum of $B^1\Sigma$ requires a photon with energy greater than 13.6 eV which are very rare in the ISM. It only takes 5 eV to excite H_2 into the vibrational continuum of the ground state, but this transition is forbidden since it is a quadrupole transition. The most common dissociation channel parallels the fluorescence process but the electronic decay goes into the vibrational continuum of the ground state. 11% of the decays lead to dissociations.

19. Gravitational Radiation

Gravitational waves also come from time varying mass quadrupoles. The formula for radiation from a binary star is

$$-\frac{dE}{dt} = \frac{32G^4 M_1^2 M_2^2 (M_1 + M_2)}{5c^5 a^5} \frac{1 + \frac{73}{24}e^2 - \frac{37}{96}e^4}{(1 - e^2)^{1/2}} \tag{444}$$

Taking $e = 0$ for simplicity this can written in terms of the quadrupole moment $Q^{ij} = \sum M(3x^i x^j - \delta^{ij} r^2)$. This is 3 times a traceless version of the moment of inertia tensor. For convenience put the stars along the x^1 axis, with M_1 at $x_1 = aM_2/(M_1 + M_2)$ and M_2 at $x_2 = -aM_1/(M_1 + M_2)$, then

$$\begin{aligned} Q^{11} &= 2\left(a^2(M_1 M_2^2 + M_2 M_1^2)/(M_1 + M_2)^2\right) = 2\left(a^2 M_1 M_2/(M_1 + M_2)\right) \\ Q^{22} &= Q^{33} = -\tfrac{1}{2}Q^{11} \end{aligned} \tag{445}$$

and the non-diagonal terms are zero. When the binary is rotated by $45°$ we get $Q^{12} = \frac{3}{2}(a^2 M_1 M_2/(M_1 + M_2))$. The square of the third time derivative of the quadrupole moment gives the radiated power. As the binary rotates (in the $x^1 - x^2$ plane Q^{11}, $Q^{12} = Q^{21}$ and Q^{22} vary with frequency twice the binary angular frequency given by $\omega^2 = G(M_1 + M_2)/a^3$ and amplitude $\frac{3}{2}(a^2 M_1 M_2/(M_1 + M_2))$. Thus

$$\left\langle \sum |d^3 Q^{ij}/dt^3|^2 \right\rangle = (4 \times \tfrac{1}{2} \times 64\omega^6 \times \tfrac{9}{4})\left(\frac{a^2 M_1 M_2}{M_1 + M_2}\right)^2 = 9 \times 32\omega^6 \left(\frac{a^2 M_1 M_2}{M_1 + M_2}\right)^2 \tag{446}$$

where $\langle \rangle$ is a time average over the period. Therefore

$$-\frac{dE}{dt} = \frac{G}{45c^5}\left\langle \sum \left|\frac{d^3 Q^{ij}}{dt^3}\right|^2 \right\rangle. \tag{447}$$

105

This is the gravitational wave equivalent of the Larmor formula for electric dipole radiation. Identifying Gm^2 with q^2 this power is 4 times larger than the electromagnetic analog in Eqn(439). The actual frequency of the radiated waves is 2ω since the quadrupole goes through two cycles per orbit. Gravitational radiation causes the period of the orbit to decrease and if this effect can be detected it will allow one to find M_1 and M_2.

Note that $dE/dt \propto a^{-5} \propto E^5$, so $E^{-4} \propto t$ where t is the time before merger. Thus a system with binary period $P_b \propto |E|^{3/2}$ and binary period derivative $\dot{P}_b$ will merge in a time given by $\tau = |E|/4|\dot{E}| = P_b^{2/3}/4(2/3)P_b^{-1/3}\dot{P}_b = \frac{3}{8}P_b/|\dot{P}_b|$.

The flux carried by gravitational waves can be estimated by considering the force on a mass which is $F = \frac{1}{2}mc^2\vec{\nabla}h_{00}$. Here h is the difference between the metric and the Minkowski metric, $h_{\mu\nu} = g_{\mu\nu} - \eta_{\mu\nu}$. If we make this equivalent to $F = qE$ with the charge given by $q = \sqrt{G}m$, we get an equivalence for E as $\frac{1}{2}c^2\vec{\nabla}h/\sqrt{G}$. Thus the flux given by $cE^2/4\pi$ becomes $c^5k^2h^2/16\pi G$. Note that since the Planck mass is $M_{pl} = \sqrt{\hbar c/G} = 22\ \mu$g and the Planck time is $t_{pl} = \hbar/M_{pl}c^2 = 5.3 \times 10^{-43}$ sec, the Planck luminosity is $L_{pl} = M_{pl}c^2/t_{pl} = c^5/G = 3.6 \times 10^{59}$ erg/sec. This gives

$$\text{Flux} = \frac{c^3}{16\pi G}\omega^2(h_+^2 + h_\times^2) \tag{448}$$

which is the correct formula for the gravitational wave flux. This flux was first measured by advanced LIGO on 14 Sep 2015.

The flux formula basically says that for $h = 1$ the flux is the Planck luminosity in an area of about a square wavelength, since $\sqrt{16\pi} \approx 2\pi$. Since the Planck luminosity is so large, any reasonable flux corresponds to a very small strain amplitude h.

Note that the actual definition of h_+ and $h_\times$ used above are for a metric with a different "gauge" than the Newtonian gauge used for the estimate above. The actual metric is

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1+h_+ & h_\times & 0 \\ 0 & h_\times & -1-h_+ & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{449}$$

for a wave traveling in the z direction with both h_+ and $h_\times$ varying like $\exp(ikz-i\omega t)$. In this metric stationary objects like the end mirrors of LIGO do not move, but the speed of light is modulated by h. In the x-direction, $c^2dt^2 - (1-h_+)dx^2 = 0$ for light, so $dx/dt = c/\sqrt{1-h_+}$. In the y-direction $dy/dt = c/\sqrt{1+h_+}$. Thus in an "L" shaped interferometer like LIGO, with arm lengths L, the path length difference from the wave is hL. The $h_\times$ term affects an "L" turned by $45°$.

The real accelerations produced by gravitational wave are tidal effects with a time varying gradient in the acceleration of gravity. This is not like an electric field, so the analogy used above is only an approximate guide.

If we consider an example with two neutron stars orbiting with a period of 60 seconds, each having a mass of 1.4 $M_\odot$, then the separation is $a = [G(M_1+M_2)P^2/(4\pi^2)]^{1/3} = 3.2\times10^9$ cm and the orbital velocity of each NS is 1700 km/sec. The energy is $\frac{1}{2}GM_1M_2/a = 1.6\times10^{50}$ ergs. The gravitational wave luminosity is 5.2×10^{39} erg/sec. The time before merger is $\tau = 1.6 \times 10^{50}/(4 \times 5.2 \times 10^{39}) = 7.7 \times 10^9$ seconds or 240 years. At a distance of 10 kpc, the flux will be 4.3×10^{-7} erg/cm^2/sec. The gravitational waves have a period of 30 seconds so $\omega = 0.2$ and $h = 1.1 \times 10^{-21}$. LISA would have been able to detect this signal easily, but LISA is being reorganized.

If we put the neutron stars only 32 km apart, or 1000 times closer than before, the velocity is now 50,000 km/sec and the period is 2 msec. The luminosity is 10^{15} times higher, the time before merger is 10^{12} times shorter or 8 msec, the frequency of the waves is 1000 Hz, and $h = 1.2 \times 10^{-18}$. aLIGO could detect this signal from a NS-NS merger out to 10 Mpc. The best sensitivity of advanced LIGO, which started in 2015, is a strain $h \approx 5 \times 10^{-23}$ at 100 Hz. Since the arm lengths of LIGO are 4 km, this corresponds to a path length difference of 2×10^{-17} cm, which is about 10^4 times smaller than the diameter of a proton. This strain corresponds to a flux of 8×10^{-3} erg/cm^2/sec.

20. Polarization

Consider an electromagnetic wave traveling in the z direction with a transverse electric field given

$$\vec{E} = \Re\left([\hat{x}E_1 + \hat{y}E_2]e^{-i\omega t}\right). \tag{450}$$

Resolve the complex amplitudes E_1 and E_2 into magnitudes and phases, $E_j = |E_j|e^{i\phi_j}$.

Then the x-component is a sinusoid with amplitude $|E_1|$ and the y-component is a sinusoid with amplitude $|E_2|$. The total power in the wave is given by

$$S = \frac{c}{4\pi}\left\langle|\vec{E}|^2\right\rangle = \frac{c}{8\pi}(|E_1|^2 + |E_2|^2). \tag{451}$$

This can be related to the parameter I in the set of *Stokes parameters:*

$$
\begin{aligned}
I &= |E_1|^2 + |E_2|^2 \\
Q &= |E_1|^2 - |E_2|^2 \\
U &= 2|E_1||E_2|\cos(\phi_1 - \phi_2) \\
V &= 2|E_1||E_2|\sin(\phi_1 - \phi_2).
\end{aligned} \tag{452}
$$

The parameter I is the total intensity, Q is the difference in the x-polarized intensity and the y-polarized intensity, U is the linearly polarized intensity along an axis at 45° to the xy-axes, and V is the circular polarization which has the x and y components of the field 90° out of phase.

Note that for this single sinusoidal wave, there is an identity:

$$I^2 = Q^2 + U^2 + V^2 \tag{453}$$

so a single sinusoid is always 100% polarized in some combination of linear and circular polarization.

The more general case involves many sinusoids. We need to come up with a more general way of measuring the Stokes parameters. In addition, the definitions above are far from how one would actually measure the polarization. To measure polarization one uses a polarizer turned to various angles to isolate a given axis. Let this angle be θ with $\theta = 0$ for the x-axis and $\theta = 90°$ for the y-axis. Then if $P(\theta)$ is the power measured at angle θ, one gets

$$
\begin{aligned}
I &= P(0) + P(90°) = P(45°) + P(135°) \\
Q &= P(0) - P(90°) \\
U &= P(45°) - P(135°).
\end{aligned}
\tag{454}
$$

But what about the circular polarization V? This requires correlating the x field delayed by 0.25 cycles with the y-field. The delay can be accomplished using a *quarter-wave plate* which delays one axis relative to the other by $\lambda/4$. The correlation can be accomplished using a polarizer at $\theta = 45°$. Thus

$$V = P(45°) - P(135°) \text{ after passing through a } \lambda/4\text{-plate} \tag{455}$$

A quarter-wave plate is made using an appropriate thickness of a bi-refringent crystal. Bi-refringence is having different indices of refraction for light polarized in different directions. Calcite is an example of a bi-refringent crystal, with ordinary and extraordinary refractive indices of 1.658 and 1.486 at $\lambda = 590$ nm, and light entering a calcite block at an angle will split into two beams. Light entering calcite at $45°$ to the ordinary axis will resolve into two beams with wavelengths λ_{vac}/n_o and λ_{vac}/n_e, so in the thickness L there will be a phase shift of $(n_o - n_e)L/\lambda_{vac}$ waves between the two beams. A calcite layer 860 nm thick will act as a quarter-wave plate for 590 nm light. Since a quarter-wave plate is a fairly narrow-band device, the Stokes parameters apply to quasi-monochromatic light.

Given the definition of the Stokes parameters in terms of powers observed through polarizers at various angles, it is clear that the Stokes parameters are additive when several different signals are combined, as long as there are no correlations between the phases of the different signals. When the phases are uncorrelated, powers add, so IQU&V add as well.

While for a single sinusoid $I^2 = Q^2 + U^2 + V^2$, when several signals are combined one gets

$$I^2 \geq Q^2 + U^2 + V^2 \tag{456}$$

because it is possible for the QU&V terms to partially cancel since they can be positive or negative, while I is always positive. The polarization fraction is given by

$$\Pi = \frac{\sqrt{Q^2 + U^2 + V^2}}{I} \tag{457}$$

which is usually given as a percentage between 0 and 100%.

The power observed through a polarizer at arbitrary orientation is given by

$$P(\theta) = \frac{I + Q\cos(2\theta) + U\sin(2\theta)}{2} \tag{458}$$

Many materials become birefringent when placed under stress. The strain induced by the stress makes the lattice spacing different along different axes, causing the indices of refraction to be different for different polarizations. When a block of such material is placed between crossed polarizers, the polarization change induced by the stress causes light to be transmitted revealing the pattern of the strain.

In addition to the quarter-wave plate which converts linear to circular polarization and vice versa, a half-wave plate is also an interesting device. If one starts with linearly polarized light at angle θ, then $E_x = E\cos\theta$ and $E_y = E\sin\theta$. A half-wave plate will change the relative phase between the x and y fields by π radians, so one has $E_x = E\cos\theta$ and $E_y = -E\sin\theta$ at the output, a wave polarized at angle $-\theta$.

In the frame of reference of the input polarization, the half-wave plate has its axes at angle $-\theta$ and the output wave has polarization angle -2θ. This means that a rotating half-wave plate will rotate the output plane of polarization at twice its rotation rate. This provides a way to measure linear polarization: a linear polarizer will normally have different gains in its two channels making the raw difference a biased estimate of Q or U. But by rotating a half-wave plate in front the linear polarizer, the outputs are modulated at 4 times the rotation rate and modulated signals will not be biased.

Of course the gold standard for eliminating biases is to rotate the entire telescope and instrument and see that the measured polarization stays fixed on the sky instead of turning with the telescope.

A half-wave plate gives a very different kind of rotation than the plasma effect we looked at earlier. The plasma effect is due to *circular dichroism*, where different senses of circular polarization have different indices of refraction. A plate made of circular dichroic material that introduced a half-wave of phase shift would always rotate the plane of polarization by 90°, no matter how it is oriented relative to the incoming polarization. But a half-wave plate is made of birefringent material which has linear dichroism, and a half-wave plate does not produce any rotation for an input polarization lined up with its principal axes.

Polarizing devices can be used to make some very useful instruments. The FIRAS spectrometer on COBE, which measured the spectrum of the sky over a factor of 100 in frequency, used a polarizing Michelson interferometer. The wavelength range was 1 cm to 100 μm, so the polarizers were made using wire grids having a 25 μm spacing and wires 12.5 μm in diameter. These wires were wound over a support frame on a lathe, and then were soldered to the frame. This gave two layers of wires, so one was cut away, leaving a free standing wire grid polarizer.

FIRAS used input polarizers to make a polarized beam, then used another polarizer tilted at 45° to make a 50:50 beamsplitter that worked over a 100:1 range in frequency. Radiation that reflected off the beamsplitter, went to the end mirror, returned to the beamsplitter, and then needed to be transmitted. This requires a 90° change in the orientation of the polarization. This could not be achieved with with a half-wave plate or a circular dichroic plate because of the very wide range of wavelengths. But a clever arrangement of mirrors could do the job. When horizontally polarized radiation hits a mirror, the reflected waves are horizontally polarized. For vertically polarized waves, the reflected waves are vertically polarized. So a simple mirror will not work. But a two mirror arrangement called a dihedral will work. This has two mirrors joined at a 90° angle along a vertical axis. A wave polarized at 45° that hits the dihedral on the right side will hit the top right side of the dihedral, bounce over to the top left side of the dihedral, and come back polarized at 135°. This is not a rotation, but rather a flip of the polarization angle around the axis of the dihedral.

So in FIRAS, radiation polarized at 45° is transmitted at the beamsplitter, goes down arm 1, is reflected and flipped in polarization by the dihedral, comes back to the beamsplitter and is now reflected. Meanwhile radiation polarized at −45° is reflected at the beamsplitter, goes down arm 2, is reflected and flipped in polarization by the dihedral, comes back to the beamsplitter and is now transmitted. All the radiation is recombined but with a variable phase shift of $2(L_1 - L_2)/\lambda$ waves, where L_1 and L_2 are the lengths of the two arms, so the FIRAS interferometer is basically a variable wave plate. When the phase shift is an integral number of waves the linearly polarized input is unchanged at the output. When the phase shift is a quarter-wave plus an integer the vertically polarized input becomes circularly polarized at the output. When the phase shift is a half-wave plus an integer the linearly polarized input is rotated 90° at the output. So the radiation that gets through an output polarizer is 100%, 50%, 0%, 50% & 100% as the path length difference is changed from 0 to $\lambda/4$, $\lambda/2$, $3\lambda/4$ & λ.

Radar astronomers usually transmit circular polarization to avoid problems with the rotation measure of the Earth's ionosphere. When circularly polarized radiation is reflected by a mirror, the rotation of the electric field is still the same but the direction of propagation is reversed. As a result, the sense of the circular polarization is reversed, so left circular polarization becomes right circular polarization and vice versa. But if the radiation is reflected by needles, the reflected power will be linearly polarized, corresponding to an equal mix of
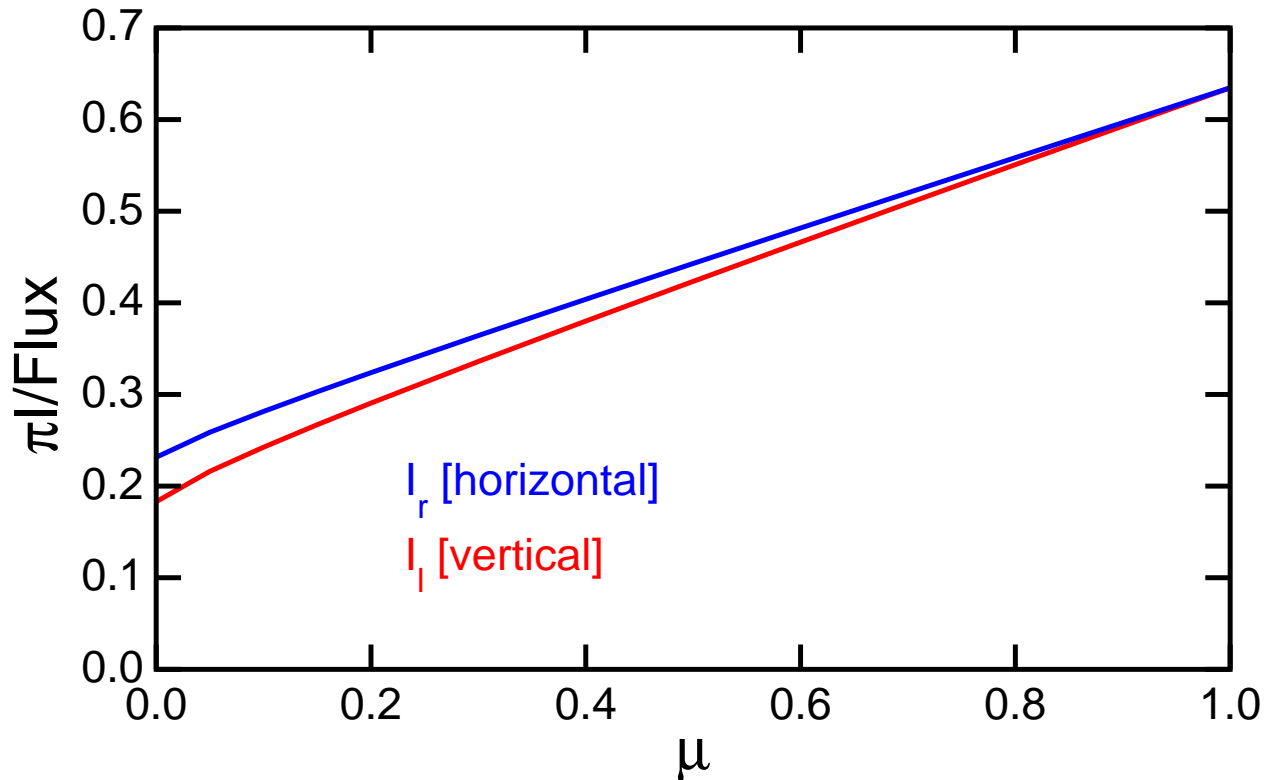
Fig. 25.— The limb darkening curves in orthogonal polarizations for an electron scattering plane-parallel atmosphere in radiative equilibirum, from Chandrasekhar (1960, *Radiate Transfer*, Dover).

left and right circular polarization. Thus the ratio of circular polarizations in the received echo is another way to study the nature of the reflecting body. This is called the SC/OC ratio, for the "same circular" to "opposite circular" ratio. Note that a mirror or specular reflection gives SC/OC << 1, randomly oriented needles ("chaff") will give SC/OC = 1, and a dihedral mirror will give SC/OC > 1.

Note that radar astronomy is limited by the $1/R^4$ dependence of its signal to noise ratio. One get an inverse square law for the outgoing pulse and another inverse square law for the received pulse.

21. Polarized Radiative Transfer

Because the Stokes parameters are additive, it is mathematically reasonable to take derivatives of the $(IQUV)$ vector and to operate on this vector with a matrix. Let **I** represent the vector of $IQUV$. Scattering usually produces a polarization, while true absorption and emission does not. The source function has to be a four-dimensional vector in the Stokes

parameter space. There will be a phase matrix $\mathbf{P}(\hat{n}, \hat{n}')$ describing how the input polarization is transformed into the output polarization. Thus

$$\frac{d\mathbf{I}(\hat{n})}{ds} = n \int \frac{d\sigma}{d\Omega} \mathbf{P}(\hat{n}, \hat{n}') \mathbf{I}(\hat{n}') d\Omega(\hat{n}') + \alpha B - (\sigma + \alpha) \mathbf{I}(\hat{n}) \qquad (459)$$

where α is the true absorption per unit length, B is the Planck function, and σ is the scattering per unit length but $d\sigma/d\Omega$ is the differential cross-section.

In a plane-parallel atmosphere, it seems easiest to work out the phase matrix $\mathbf{P}$ for a polarization basis I_l and I_r, where I_l is polarization in the plane containing the vertical and the line-of-sight, while I_r is perpendicular to that plane.

The net result for electron scattering is an equation

$$\mu \frac{\partial}{\partial \tau} \begin{pmatrix} I_l(\mu) \\ I_r(\mu) \end{pmatrix} = \begin{pmatrix} I_l(\mu) \\ I_r(\mu) \end{pmatrix}$$

$$- \tfrac{3}{8} \int \begin{pmatrix} 2(1-\mu^2)(1-\mu'^2) + \mu^2 \mu'^2 & \mu^2 \\ \mu'^2 & 1 \end{pmatrix} \begin{pmatrix} I_l(\mu') \\ I_r(\mu') \end{pmatrix} d\mu' \qquad (460)$$

taken from Chandrasekhar (1960, *Radiative Transfer*, Dover)

Chandrasekhar also solves these equation for the case of a constant flux F coming from below and no incident radiation. This would be a star in radiative equilibrium with an electron scattering atmosphere. Figure 25 show the limb darkening curves in the two polarizations. The fractional polarization is zero at $\mu = 0$ as it must be by symmetry, and rises to 11.7% at the limb. But the radiation is primarily moving outward at the surface, the light scattered toward the horizon is primarily horizontally polarized.

22. Albedo and Asteroids

When dealing with planets the *albedo* or reflectivity is often needed. Two different kinds of albedo are defined.

The geometric albedo (p_V for the V band) is the ratio of the reflected light at phase angle $\alpha = 0$ to the light reflected by a white Lambertian disk of the same size. A Lambertian disk has the same surface brightness at all angles, but because the observed surface area goes like $\cos \alpha$ the amount of reflected light goes down as α increases. Because a Lambertian disk scatters into several steradians, it is possible to have an object that reflects more light straight back to the source giving a $p_V > 1$ A cat's eye is an example of this, or a corner cube. Bicycle front and rear reflectors do it, as does Scotchlite tape.

The Bond albedo A is the ratio of total reflected light to the total incident light. To get the total reflected light we need to integrate over phase angles. If the phase function $\phi(\alpha)$ is

normalized to unity at $\alpha = 0$ then for a Lambertian disk the integral

$$\int \phi(\alpha)d\Omega = 2\pi \int_0^1 \mu d\mu = \pi \tag{461}$$

The phase integral q is defined as

$$q = \frac{1}{\pi} \int \phi(\alpha)d\Omega \tag{462}$$

so it is normalized to unity for a Lambertian disk. The Bond albedo is defined as

$$A = pq \tag{463}$$

For a white Lambertian sphere the geometric albedo is given by

$$p = \frac{\int_0^{\pi/2} \cos\theta d(\sin^2\theta)}{\int_0^{\pi/2} d(\sin^2\theta)} = \frac{\int_0^{\pi/2} 2\cos^2\theta\sin\theta d\theta}{1} = \int_0^1 2\mu^2 d\mu = \frac{2}{3} \tag{464}$$

Note that the projected radius on the sky goes like $\sin\theta$ where θ is the angle from the sub-solar point, so $d(\sin^2\theta)$ is $d(\text{Area})$. Since this is still a white object that reflects all the incident light, the phase integral for a Lambertian sphere is $q = 3/2$ so the Bond albedo is still $A = 1$. For the white disk, $\phi(90°) = 0$, but for the white sphere

$$\phi(90°) = \frac{2\int_0^1 \sqrt{1-x^2}\,xdx}{(2/3)\pi} = \frac{\int_0^1 \sqrt{1-y}\,dy}{(2/3)\pi} = \frac{1}{\pi} \tag{465}$$

The larger $\phi(90°)$ explains why $q = 1.5 > 1$ for the white sphere.

The absolute magnitude of an object that is seen by reflected sunlight is defined as the magnitude it would have if it were 1 AU from the Sun, and 1 AU from the observer, and at a phase angle $\alpha = 0$. This is an unlikely situation since it implies the observer is located at the Sun, but it gives a well-defined geometry. This absolute magnitude is called H, which should not be confused with the 1.6 μm Johnson magnitude H.

The Minor Planet Center uses a particular parameterized model for the phase function to convert observed magnitudes into absolute magnitudes. This phase function is parameterized by a "slope parameter" G:

$$\begin{aligned}
\phi(\alpha) &= (1-G)\phi_1(\alpha) + G\phi_2(\alpha) \\
\phi_1(\alpha) &= \exp[-3.33\tan(\alpha/2)^{0.63}] \\
\phi_2(\alpha) &= \exp[-1.87\tan(\alpha/2)^{1.22}]
\end{aligned} \tag{466}$$

$$\tag{467}$$

Since ϕ_2 declines less steeply with α than ϕ_1, a large G means a shallower slope for the phase function. The default value for the phase function is $G = 0.15$. The value for the phase

integral is $q = 0.29 + 0.684G$ and the value of the phase function at $90°$ is $\phi(\alpha = 90°) = 0.0358 + 0.1185G$. For the default slope parameter $G = 0.15$, the phase integral is $q = 0.384$ and $\phi(\alpha = 90°) = 0.0535$. Note that the white Lambertian sphere has a much shallower phase function (and hence a larger phase integral) than that given by any reasonable value of G in the range 0 to 1.

With these definitions the absolute magnitude H is related to the observed magnitude V by

$$H = V - 5\log_{10}(r\Delta) + 2.5\log_{10}[\phi(\alpha)] \tag{468}$$

where r is the distance from the Sun, and Δ is the distance from the observer. The relationship between H and the diameter D of the asteroid is given by

$$D = (1329 \text{ km})10^{-0.2H}/\sqrt{p_V}. \tag{469}$$

For example, the threshold for potentially hazardous asteroids, $H < 22$, corresponds to $D > 0.14$ km for $p_V = (1329/0.14)^2 10^{-0.4H} = 0.143$.

22.1. Thermal Models

Asteroids are generally too small to resolve with a telescope unless the telescope is the HST and the asteroid is one of the largest. Thus to determine the diameter one must either assume an albedo or obtain some other data about the asteroid. Observations of the thermal infrared radiation emitted by the asteroid are very useful for albedo determinations because the fraction $(1 - A)$ of the sunlight that is not scattered is absorbed and reradiated in the thermal infrared. If one could observe an asteroid from all directions and compute its "luminosity" in scattered light and infrared reradiation, L_{opt} and L_{IR}, then the Bond albedo is determined from

$$\frac{A}{1 - A} = \frac{L_{opt}}{L_{IR}}. \tag{470}$$

The diameter would also be well-determined, since

$$L_{opt} + L_{IR} = \pi(D/2)^2 \frac{L_\odot}{4\pi r^2}. \tag{471}$$

Of course one usually only observes the asteroid from one angle, and one needs to know the optical and infrared phase functions to extrapolate from one phase angle to the integral over all phase angles. Knowing the infrared phase function involves modelling the temperature distribution on the surface of the asteroid, and several different models are commonly used.

Clearly the simplest possible thermal model is an isothermal asteroid with the same temperature at all points on the surface. This has a constant infrared phase function, $\phi = 1$

for all angles, and thus a phase integral of 4. If the surface of the asteroid radiates like a greybody with emissivity ϵ then the temperature is given by

$$\epsilon \sigma_{SB} T^4 [4\pi (D/2)^2] = \frac{(1-A)\pi (D/2)^2 L_\odot}{4\pi r^2}. \tag{472}$$

For $\epsilon = 1$ and $A = 0$, this is $T = 279[r/(1\ \text{AU})]^{-1/2}$ K. The isothermal model is reasonable for very small asteroids, rapidly spinning asteroids, and for planets with atmospheres which have a large thermal inertia like the Earth, which explains why 279 K is such a pleasant temperature. A more detailed model of the Earth would include the albedo, $A \approx 0.24$, which makes the temperature colder, and also the warming due to the greenhouse effect in the atmosphere, again giving a pleasant surface temperature.

For asteroids larger than tens of meters, conduction is inefficient at transferring heat, so there is a large temperature gradient around the surface. This is often modelled using the Near Earth Asteroid Thermal Model (NEATM, Harris, A. W. 1998, Icarus, 131, 291-301), which assumes that the surface temperature at an angle θ away from the sub-solar point is given by $T(\theta) = T_{ss} \max(0, \cos \theta)^{1/4}$ where the sub-solar temperature is given by

$$\epsilon \eta \sigma_{SB} T_{ss}^4 = \frac{(1-A)L_\odot}{4\pi r^2}. \tag{473}$$

The "beaming parameter" η is a fudge factor that can be adjusted to make the model spectrum match the observed spectrum. The model spectrum for a given phase angle is obtained by integrating $\epsilon B_\nu (T(\theta))$ over the observed disk of the asteroid. Note that the integrated, or bolometric, infrared phase curve of the NEATM is exactly the same as the phase function of the perfectly diffusing Lambertian sphere. Thus the phase integral is 1.5. For wavelengths shorter than the peak of the IR spectrum the phase curve will be steeper and the phase integral smaller, while the reverse is true for wavelengths longer than the peak.

The NEATM with $\eta \neq 1$ does not satisfy radiative equilibrium. Thus the derived value of η has to vary with viewing angle. When viewing the morning side of a rotating asteroid, the temperature of the visible parts of the asteroid will be cooler than equilibrium giving a high η, while in the afternoon the temperature will be warmer leading to a low η. These variations are due to physical effects that are not included in the NEATM, such as heat conducted into the surface during the day that is radiated later at night. Thermophysical models do include such effects.

Another commonly used simple thermal model is the Fast Rotating Model, or FRM. In this model the Sun is over the equator of a rapidly spinning asteroid. Then the heat input depends on the latitude θ and local hour angle of the Sun as

$$F(\text{LHA}) = \cos \theta \max[0, \cos(\text{LHA})] \frac{(1-A)L_\odot}{4\pi r^2} \tag{474}$$

which averaged over LHA gives

$$\langle F \rangle = \frac{\cos\theta}{\pi} \frac{(1-A)L_\odot}{4\pi r^2} \tag{475}$$

This leads to a temperature on the equator that corresponds to the sub-solar temperature of the NEATM with $\eta = \pi$. But the phase function for the FRM is a constant when the observer is over the equator. This is the usual assumption. When the observer is over the pole at $\alpha = 90°$ the phase function is smaller. The phase integral integrated over solid angle is high so the FRM is in radiative equilibrium. The day side is cooler, but the night side is just as hot as the day side.

All of these simple models assume a spherical asteroid. The infrared flux is found by integrating over the observed disk of the asteroid. For the NEATM at $0°$ phase angle the flux is

$$F_\nu = \frac{(D/2)^2}{\Delta^2} \epsilon \int_0^1 B_\nu(T_{ss}[1-x^2]^{1/8}) 2\pi x dx \tag{476}$$

where x is the normalized radius on the observed disk, which is $\sin\theta$ if θ is the angle from the sub-solar point.

For the NEATM at $90°$ phase angle, the flux is

$$F_\nu = \frac{(D/2)^2}{\Delta^2} \epsilon \int_0^1 B_\nu(T_{ss}x^{1/4}) 4\sqrt{1-x^2} dx \tag{477}$$

where now $x = \cos\theta$ is the normalized x coordinate on the disk.

For the FRM seen equator-on, the flux for any phase angle is

$$F_\nu = \frac{(D/2)^2}{\Delta^2} \epsilon \int_0^1 B_\nu(T_{ss}[1-y^2]^{1/8}) 4\sqrt{1-y^2} dy \tag{478}$$

where y is the sine of the latitude.

Finally for the FRM seen pole-on, at $90°$ phase angle with a particular clock angle, the flux is

$$F_\nu = \frac{(D/2)^2}{\Delta^2} \epsilon \int_0^1 B_\nu(T_{ss}x^{1/4}) 2\pi x dx \tag{479}$$

where x is now the cosine of the latitude which is the normalized radius on the disk.