Astronomy 282 Lecture Notes, Winter 2012
©Edward L. Wright, 2012

High energy astrophysics is an exciting field where exotic processes and space-based observations lead to the discovery of new classes of objects and phenomena. This year marks the centennial of the first high-energy astrophysics observations.

1. History

Demo of Geiger counter and electroscope

Radioactivity: 1896 by Becquerel

X-rays: 1895 by Roentgen

Ionization of air decreased with altitude, 1900-1910

1912: Victor Hess took electroscopes to 5300 meters in a balloon found 4x more ionization at altitude - Nobel 1936

"cosmic ray" coined by Millikan, thought they were caused by $\gamma \rightarrow e$ by Compton scattering in the atmosphere

1928: Geiger-Mueller counter

1930: coincidence circuit - Bruno Rossi

East-West effect. Positive particles orbit in a retrograde sense around the Earth, moving East to West, but the lowest energy and hence most common cosmic rays are moving West-East when close to the surface.

1934: air showers

1954: first air shower array in Harvard MA, Rossi's MIT group

Auger project is a descendant

1 rad is 100 erg/gm. 1 Gray is 1 J/kg = 100 rad.

1 rem is a rad multiplied by the relative biological effectiveness (RBE)
100 rem is 1 Seivert

CRs are 0.4 mSv/yr out of typical total of 3 mSv/yr. Goes up to 1 mSv/yr in places

like Leadville CO.

Air crews get 2.2 mSv/yr extra, so rate is 10 mSv/yr

Conductivity of air $= n_i \mu$

μ = mobility, perhaps 1.5 cm^2/V/sec

ionization rates typically 10 /cc/sec

$n_i = 100 - 200$ /cc

conductivity 2-100 fS/m, S = Siemens = mho

decay times near 1 hour

2. 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

$$\vec{E}(t) = \frac{-\gamma Ze(vt, b, 0)}{(b^2 + (\gamma vt)^2)^{3/2}} \tag{1}$$

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

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

the integrated velocity change is all in the y direction:

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

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

$$\int \frac{dx}{(1 + x^2)^{3/2}} = \int_{-\pi/2}^{\pi/2} \cos\theta d\theta = 2 \tag{4}$$

Thus the velocity change of the electron is

$$\Delta\vec{v} = \frac{2Ze^2\hat{j}}{m_e bv} \tag{5}$$

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

$$\mathrm{KE} = \frac{1}{2} m_e v^2 = \frac{2Z^2 e^4}{m_e b^2 v^2} \tag{6}$$

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

$$-\frac{dE}{dx} = \frac{4\pi Z^2 e^4 n_e}{m_e v^2} \int \frac{db}{b} \tag{7}$$

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

$$b_{max} = \frac{\hbar \gamma v}{I} \tag{8}$$

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

$$b^C_{min} = \frac{2Ze^2}{m_e v^2} \tag{9}$$

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

$$b^Q_{min} = \frac{\hbar}{\gamma m_e v} \tag{10}$$

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

$$\begin{aligned} \frac{2Ze^2}{mv^2} &= \frac{\hbar}{\gamma m v} \\ \frac{v}{c} &= 2\gamma Z \frac{e^2}{\hbar c} = \frac{2\gamma Z}{137} \end{aligned} \tag{11}$$

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

$$-\frac{dE}{dx} = \frac{4\pi Z^2 e^4 n_e}{m_e v^2} \ln\left(\frac{\gamma^2 m_e v^2}{I}\right) \tag{12}$$

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

$$-\frac{dE}{\rho dx} = \frac{4\pi Z^2 e^4}{m_H m_e v^2} \frac{Z_m}{A_m} \ln\left(\frac{\gamma^2 m_e v^2}{I}\right) \tag{13}$$

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

$$\begin{aligned} -\frac{dE}{\rho dx} &\approx \frac{2\pi Z^2 e^4}{m_H m_e c^2}\beta^{-2}\frac{2Z_m}{A_m}\Big(10 + 2\ln\gamma\Big) \\ &= \frac{3}{4}\frac{8\pi e^4}{3m_e^2 c^4}\frac{m_e c^2}{m_H}\frac{2Z_m}{A_m}\frac{Z^2}{\beta^2}\Big(10 + 2\ln\gamma\Big) \\ &\approx \frac{7.5\sigma_T m_e c^2}{m_H}\frac{2Z_m}{A_m}\frac{Z^2}{\beta^2}\Big(1 + 0.2\ln\gamma\Big) \\ &= 1.5\frac{2Z_m}{A_m}\frac{Z^2}{\beta^2}\Big(1 + 0.2\ln\gamma\Big)\ \mathrm{MeV/(gm/cm^2)} \end{aligned} \tag{14}$$

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).

3. Ionization detectors

3.1. Proportional or "digital"?

Once charges are created they drift in an electric field toward electrodes. In solid state detectors the charge is usually just collected without multiplication.

In a gas one can get multiplication. The velocity gain between collisions is $dv = (qE/m)/(n\sigma v_{th})$. By having a high E field or by having a modest density n one can achieve a situation where a particle with dv has a chance to ionize new molecules. This gives a gain so one charge created becomes many charges collected. A way to achieve this is to have a thin wire electrode with a high positve potential, say +900 Volts. As the electrons approach the wire the E field goes up, and multiplication occurs. This is the way a PROPORTIONAL COUNTER works.

A proportional counter gives pulses with variable amplitudes that are proportional to the deposited energy.

If the E field gets really high, the ions can reach a high enough speed to ionize new molecules. Since the ions are drifting upstream, this leads to a positive feedback and a breakdown. The Geiger counter works in this mode. The pulse height is determined by the charge stored in the capacitor formed by the wire electrode and the surrounding cylinder. All the pulses are the same size.

Since minimum ionizing particles all lose the same amount of energy, preserving the pulse height is not important, but getting a convenient way to measure the track is.

One can make multi-wire counters with an array of anode wires. Which wire gets hit

gives one coordinate, while the perpendicular coordinate can be gotten by:
1) charge division between the ends of the anode wire
2) using a perpendicular array of cathode wires.
And the third coordinate can be obtained in a drift chamber, where the anode and cathode wire planes are separated by a large "drift" volume with uniform electric field. The time difference beween the ion collection on the cathode and the electron collection on the anode gives the z coordinate.

The first tracking ionization detectors were cloud chambers where the ions serve as condensation nuclei in a supersaturated vapor. A denser version is the bubble chamber filled with a superheated liquid.

Once can have multi-wire proportional counters or one can drive the voltage higher so the counter breaks down when hit. This is a SPARK CHAMBER.

3.2. conversion to light

One can also convert ionization loss into light Materials that do this efficiently are called scintillators. Sodium iodide doped with thallium [NaI(Tl)] is a good example. It produces about 38 photons/keV.
It is also fairly dense, 3.67 gm/cc, so 1 cm of a minimum ionizing track makes 300,000 photons.
$LaBr_3$(Ce), lanthanum bromide doped with cerium, is more efficent, faster, and denser; but NaI is the workhorse.

There are plastic scintillators like polyethylene napthalate too. And borosilicate glass works as a scintillator.

But other materials also work: air is very cheap and converts some energy into light. A minimum ionizing particle loses 2 MeV/gm/cm^2 and produces about 40 photons per meter in STP air in the 310-400 nm band: near-UV to blue. (Fakimoto etal., icrc951-1047.pdf)
A meter of air is about 0.1 gm/cm^2, so 200 keV of energy lost makes 40 photons, or about 5 keV per photon. Since the energy of the photons are only 4 eV, only 0.1% of the energy lost in air is emitted as light. But air is very cheap so the Auger array can use 30 billion tonnes of air an a detector. This is 3000 km^2. The 10^{20} eV CRs Auger is looking for have a flux of 1 /km^2/century.

3.3. Statistics

Fano factor = σ^2/mean = 1 for Poisson process

For ionization detectors generate N carriers with $< I >$ energy per carrier Thus the amount of ionization gives the deposited energy $E_{dep} = N < I >$ with errors $\sigma(E_{dep}) =< I > \sigma(N)$ so the resolution is $\sigma(E_{dep})/E_{dep} = \sqrt{FF/[E_{dep}/ < I >]}$ where FF is the Fano factor.

Many ionization detectors have small Fano factors: Si = 0.115, Ge = 0.13, CZT = 0.089(5), Ar = 0.20. This happens when there is only one way to create a carrier

For proportional counters and avalanche photodiodes, the randomness of the first multiplication event usually increases the Fano factor. Typically the distrbution of gains follows an exponential distribution so $p(q) = \exp(-q/Q)/Q$ where Q is the mean charge per carrier. This has mean Q and variance Q^2. Thus if 100 ± 5 carriers are made, FF=0.25, and each one is multiplied to give 100 ± 100 electrons, the final charge collected is 1E4 electrons with noise $\pm 10\%$ from the multiplication and $\pm 5\%$ from the generation, giving a total noise of $\pm 11\%$.

Photomultipliers do not have this noise source since the dynodes define the number of multiplications precisely.

3.4. μ-calorimeters

conversion to HEAT

If none of the charge carriers escape from the detector, all of the inoization energy loss becomes heat. Detectors that measure the heat of a pulse are called bolometers, or calorimeters. There is a thermodynamic limit to the accuracy of measuring the energy of a detector attached to a heat bath.

The definition of entropy is $S = k \ln(\Omega)$ where $\Omega(E)$ is the number of states. The law $dS = dQ/T$ defines the temperature in terms of $\Omega(E)$, since $dS = k(d\ln(\Omega)/dE)dE = dE/T$ hence $1/(kT) = (d\ln(\Omega)/dE)$
Adding energy to the detector typically raises its temperature so
$d[1/(kT)]/dE = (d^2\ln(\Omega)/dE^2) = -1/(kT^2)(dT/dE)$
Now the heat capacity C of the detector is dE/dT, and the linear slope $d\ln(\Omega)/dE$ will be canceled by the opposite slope from the heat bath, so one is left with $\ln(\Omega) \approx const - 0.5(E - E_{eq})^2/(kCT^2) + \ldots$.
Thus the energy content of the the detector will have a Gaussian distribution with a sigma

of $\sigma = \sqrt{kCT^2}$.

If we evaluate this for a gram of water at 300 K we get C = 1 gm-cal/K = 4.2 Joule/K = 4.2E7 erg/K k = 1.4E-16 erg/K sigma = 0.023 erg = 14 GeV. This is very lousy resolution. So we need a smaller heat capacity and temperature. At low temperatures the Debye form of the heat capacity applies. This gives an energy density just like the energy density of a blackbody but with the speed of light replaced by the speed of sound and with 3 polarizations instead of 2. For photons $u = 2 * 4\pi * (kT/hc)^3 * kT * \Gamma(4)\zeta(4)$. For silicon $c_p = 8$ km/sec, $c_s = 5$ km/sec So $(c^3/2)(1/c_p^3 + 2/c_s^3) = 2.5 \times 10^{13}$, and for a blackbody $a = 7.6 \times 10^{-15}$, so $u = 0.22T^4$ erg/cc. Obviously $C = 4 \times 0.22T^3V$ so a low temperature is needed. With T = 0.1 K and a volume of 10^{-3} cc (a 1 mm cube) we get $C = 10^{-6}$ erg/K The thermal noise is then $\sigma(E) = \sqrt{kCT^2} = 1 \times 10^{-12}$ erg = 0.6 eV
micro-calorimeters have achieved dE's of a few eV even for 6 keV X-rays giving very good resolution.

4. 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

$$\vec{a}(t) = \frac{\gamma Z e^2}{m_e} \frac{(vt, b, 0)}{(b^2 + (\gamma v t)^2)^{3/2}} \tag{15}$$

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

$$W = \int P dt \approx \frac{b}{\gamma v} \frac{2\gamma^2 Z^2 e^6}{3m_e^2 b^4 c^3} \approx \frac{2\gamma Z^2 e^6}{3m_e^2 b^3 c^3 v} \tag{16}$$

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

$$\int_{-\infty}^{\infty} a(t)^2 dt = 2\pi \int_{-\infty}^{\infty} |\hat{a}(\omega)|^2 d\omega = 4\pi \int_0^{\infty} |\hat{a}(\omega)|^2 d\omega \tag{17}$$

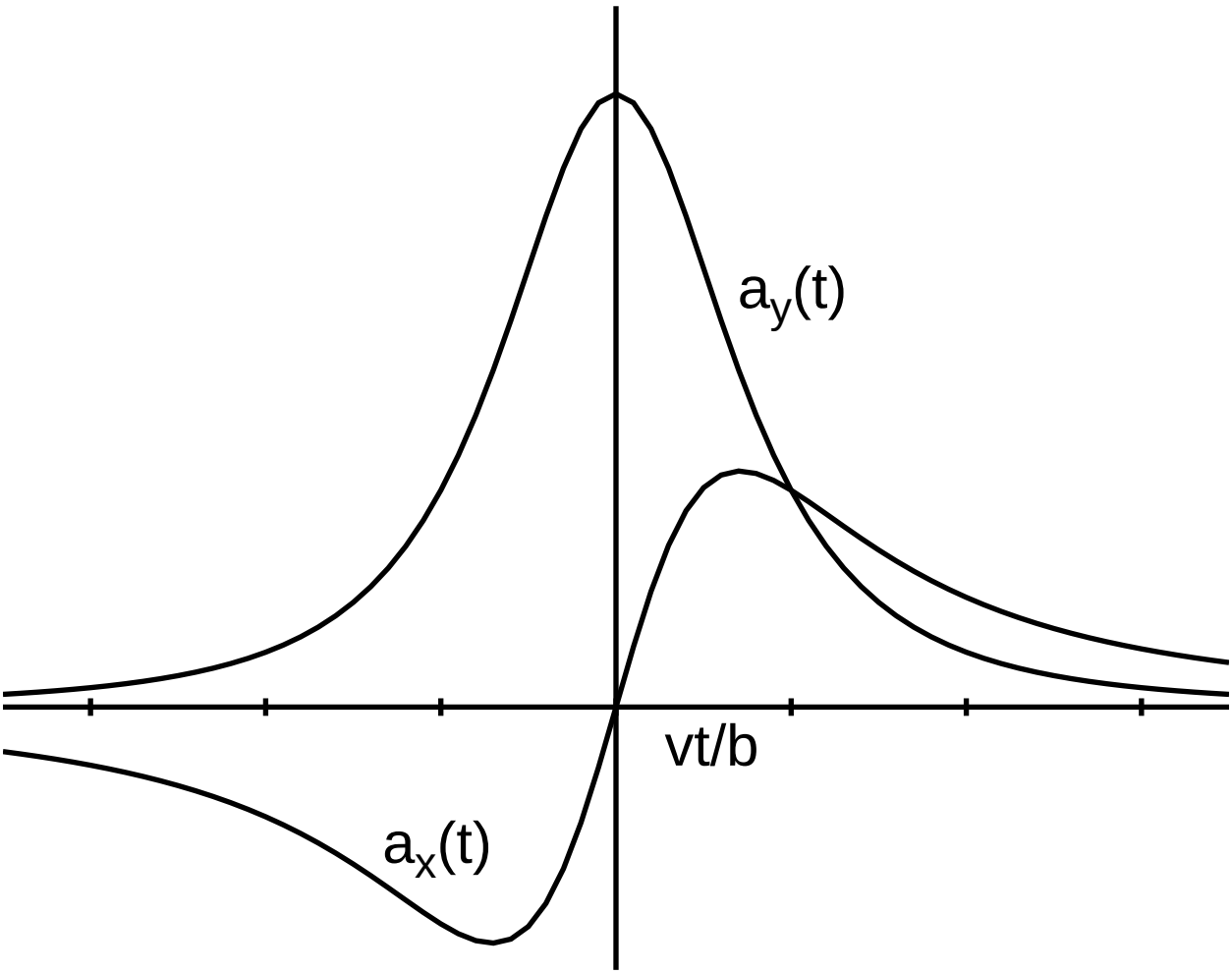

Fig. 1.— x and y axis acceleration vs time.

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

$$W = \frac{2e^2}{3c^3} \int a(t)^2 dt = \frac{8\pi e^2}{3c^3} \int_0^\infty |\hat{a}(\omega)|^2 d\omega = \int_0^\infty \frac{dW}{d\omega} d\omega \tag{18}$$

we can identify the power per unit frequency as

$$\frac{dW}{d\omega} = \frac{8\pi e^2}{3c^3} |\hat{a}(\omega)|^2 \tag{19}$$

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

$$\begin{aligned} \hat{a}_y(\omega) &= \frac{1}{2\pi} \frac{\gamma Z e^2}{m_e b^2} \int \frac{\exp(i\omega t) dt}{(1 + (\gamma v t/b)^2)^{3/2}} \\ \hat{a}_x(\omega) &= \frac{1}{2\pi} \frac{Z e^2}{m_e b^2} \int \frac{(\gamma v t/b) \exp(i\omega t) dt}{(1 + (\gamma v t/b)^2)^{3/2}} \\ &= \frac{v}{ib} \frac{d\hat{a}_y(\omega)}{d\omega} \end{aligned} \tag{20}$$

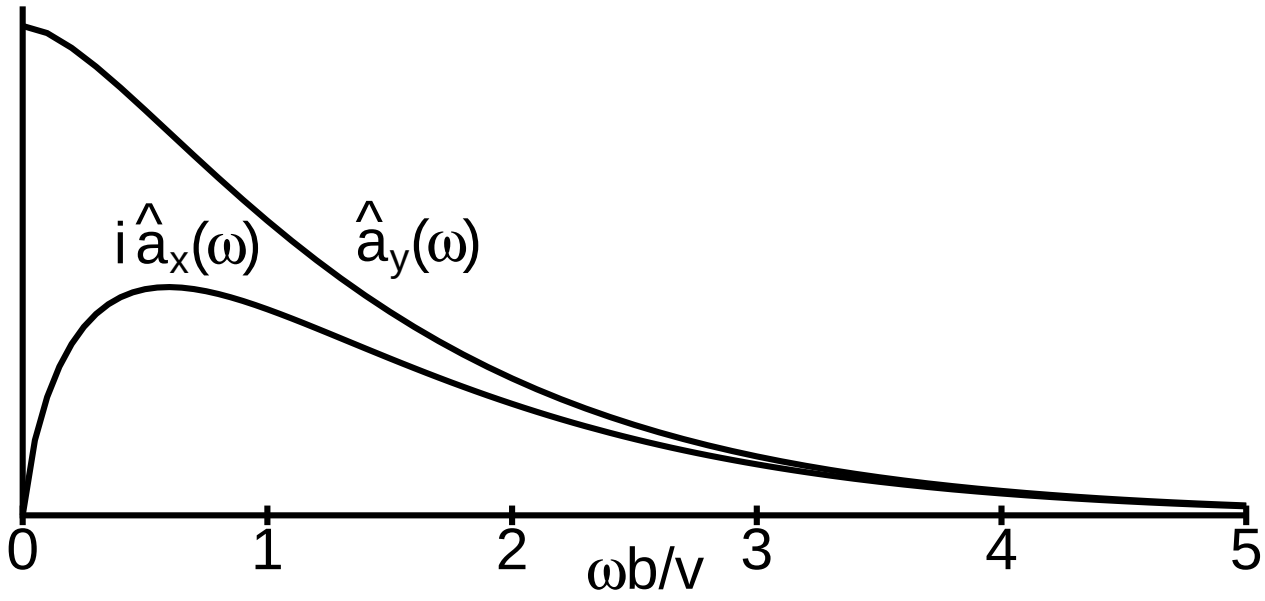

Fig. 2.— Fourier transforms of the x and y axis accelerations.

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

$$\hat{a}_y(\omega) = \frac{1}{2\pi}\frac{Ze^2}{m_e bv}\int \frac{dx}{(1+x^2)^{3/2}} = \frac{Ze^2}{\pi m_e bv} \tag{21}$$

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

$$2kK_1(k) = \int \frac{e^{ikx}dx}{(1+x^2)^{3/2}} \tag{22}$$

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

$$\frac{dW}{d\omega} = \frac{8\pi e^2}{3c^3}\left|\frac{Ze^2}{\pi m_e bv}\right|^2 = \frac{8Z^2e^6}{3\pi m_e^2 b^2 c^3 v^2} \tag{23}$$

The energy radiated per Hertz is given by

$$\frac{dW}{d\nu} = 2\pi\frac{dW}{d\omega} = \frac{16Z^2e^6}{3m_e^2 b^2 c^3 v^2} \tag{24}$$

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 $2\pi\, n_i\, b\, db\, v dt$. 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

$$\frac{dW}{d\nu dt} = \frac{32\pi\gamma Z^2 e^6 n_i}{3m_e^2 c^3 v} \int \frac{db}{b} \tag{25}$$

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

$$b < b_{max} = \frac{\gamma v}{2|\omega|} \tag{26}$$

The quantum inner limit is

$$b_{min}^Q = \frac{\hbar}{\gamma m_e v} \tag{27}$$

so

$$\int \frac{db}{b} = \ln\left(\frac{b_{max}}{b_{min}}\right) = \ln\left(\frac{\gamma^2 m_e v^2}{2\hbar\omega}\right) = \ln\Lambda \tag{28}$$

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.

4.1. 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:

$$\frac{dW}{d\omega dV dt} = \frac{16e^6 n_e n_i Z^2}{3m_e^2 c^3 v} \ln\left(\frac{b_{max}}{b_{min}}\right) \tag{29}$$

The Gaunt factor g is defined as

$$g_{ff}(v,\omega) = \frac{\sqrt{3}}{\pi} \ln\left(\frac{b_{max}}{b_{min}}\right) \tag{30}$$

so we get

$$\frac{dW}{d\omega dV dt} = \frac{16\pi e^6}{3\sqrt{3}m_e^2 c^3} \frac{n_e n_i Z^2}{v} g_{ff}(v,\omega) \tag{31}$$

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

$$\int_{v(min)}^{\infty} v^{-1} v^2 \exp\left(\frac{-mv^2}{2kT}\right) dv = \frac{kT}{m_e} \int_{mv(min)^2/2kT}^{\infty} e^{-x} dx = \frac{kT}{m_e} \exp\left(\frac{-h\nu}{kT}\right) \tag{32}$$

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

$$\int_0^{\infty} v^2 \exp\left(\frac{-mv^2}{2kT}\right) dv = \sqrt{\frac{2kT}{m}} \frac{kT}{m} \int_0^{\infty} \sqrt{x} e^{-x} dx \tag{33}$$

This integral

$$\int_0^{\infty} \sqrt{x} e^{-x} dx = \Gamma(1.5) = (0.5)! = \frac{\sqrt{\pi}}{2} \tag{34}$$

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

$$\frac{\int_{v(min)}^{\infty} v e^{-x} dv}{\int_0^{\infty} v^2 e^{-x} dv} = \sqrt{\frac{2m}{\pi kT}} \exp\left(\frac{-h\nu}{kT}\right) \tag{35}$$

Thus we get

$$\frac{dW}{d\omega dV dt} = \frac{16\pi e^6}{3\sqrt{3} m_e^2 c^3} n_e n_i Z^2 \sqrt{\frac{2m}{\pi kT}} \exp\left(\frac{-h\nu}{kT}\right) \overline{g_{ff}} \tag{36}$$

or in terms of ν:

$$\frac{dW}{d\nu dV dt} = \frac{32\sqrt{2}\pi^{3/2} e^6}{3\sqrt{3} m_e^{3/2} k^{1/2} c^3} n_e n_i Z^2 T^{-1/2} \exp\left(\frac{-h\nu}{kT}\right) \overline{g_{ff}} = 4\pi j_\nu(ff) \tag{37}$$

A practical formula is

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

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

$$\overline{g_{ff}}(T, \nu) = \frac{\int_{v(min}^{\infty} v e^{-x} g_{ff}(v, \nu) dv}{\int_{v(min}^{\infty} v e^{-x} dv} \tag{39}$$

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 freqeuncy dependence of $j_\nu(ff)$ to be close to $\nu^{-0.1}$ instead of ν^0.

4.2. Total Free-free Cooling Power

The total power radiated by thermal bremsstrahlung is given by

$$\begin{aligned}\frac{dW}{dt\,dV} &= \int_0^\infty \frac{dW}{d\nu dV dt} d\nu = \frac{kT}{h}\frac{32\sqrt{2}\pi^{3/2}e^6}{3\sqrt{3}m_e^{3/2}k^{1/2}c^3} n_e n_i Z^2 T^{-1/2}\overline{g_B} \\ &= (1.4\times10^{-27}\ \mathrm{erg/sec/cm^3}) n_e n_i Z^2 T^{1/2}\overline{g_B} \end{aligned} \tag{40}$$

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

4.3. 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

$$\left(\frac{dW}{d\omega}\right)' = \frac{16Z^2e^6n_i}{3m_e^2vc^3}\ln\left(\frac{b_{max}}{b_{min}}\right) \tag{41}$$

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

$$-\frac{dE}{dt} = \frac{16Z^2e^6n_iE}{3m_e^2c^4\hbar}\left\langle\ln\left(\frac{b_{max}}{b_{min}}\right)\right\rangle \tag{42}$$

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

$$\begin{aligned}-\frac{dE}{\rho dx} &= \frac{16}{3}\left(\frac{Z^2}{A}\right)\frac{e^6}{m_H m_e^2\hbar c^5}E\left\langle\ln\left(\frac{b_{max}}{b_{min}}\right)\right\rangle \\ &= \left[\frac{2}{\pi}\frac{e^2}{\hbar c}\frac{\sigma_T}{m_H}\left(\frac{Z^2}{A}\right)\left\langle\ln\left(\frac{b_{max}}{b_{min}}\right)\right\rangle\right]E = E/\xi_\circ \end{aligned} \tag{43}$$

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

$$\xi_\circ = \frac{716A\ \mathrm{gm/cm^2}}{Z(Z+1.3)\left[\ln\left(\frac{183}{Z^{1/3}}\right)+\frac{1}{8}\right]} \tag{44}$$

4.4. Pair Production

The relativistic bremsstrahlung reaction $e^- + Z \rightarrow Z + e^- + \gamma$ and the pair production reaction $\gamma + Z \rightarrow Z + e^- + e^+$ 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 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.

5. Cherenkov Radiation

If a particle moving along the path

$$x = vt \tag{45}$$

emits light which travels outward at speed c/n_r, where n_r is the index of refraction of the medium, then an event in spacetime at position $(ct; x, y)$ receives light emitted at times given by the equation

$$\begin{aligned} (x - vt_{em})^2 + y^2 &= [(c/n_r)(t - t_{em})]^2 \\ (v^2 - (c/n_r)^2)t_{em}^2 + (2(ct/n_r)(c/n_r) - 2xv)t_{em} + (y^2 - (ct/n_r)^2) &= 0 \end{aligned} \tag{46}$$

The discriminant of this quadratic equation is:

$$D = 4(ct/n_r)^2(c/n_r)^2 - 8xv(ct/n_r)(c/n_r) + 4x^2v^2 - 4(y^2 - (ct/n_r)^2)(v^2 - (c/n_r)^2) \tag{47}$$

$$D/v^2 = 4(ct/n_r)^2 - 8x\left(\frac{c}{n_r v}\right)(ct/n_r) + (4x^2 - 4y^2(1 - \left(\frac{c}{n_r v}\right)^2)) \tag{48}$$

If this discriminant is positive, then there are two different t_{em}'s which put the emitting particle on the past light cone of the event. As the received time t gets smaller, the discriminant gets smaller, until $D = 0$. For this particular past light cone, the world line of the emitting particle is tangent to the light cone. This gives a very large contribution to the potential. Finally, for even smaller received times t the discriminant is negative and the world line of the emitting particle does not intersect the light cone at all. When the derivative of the potential is evaluated at the received time where $D = 0$, a large derivative is found that corresponds to a radiated field falling off as $1/r$ unlike static fields which fall as $1/r^2$. This radiated field is generated even though the particle does not accelerate. The received time t when $D = 0$ is

$$\frac{ct}{n_r} = x\left(\frac{c}{n_r v}\right) + \sqrt{(x^2 + y^2)(1 - \left(\frac{c}{n_r v}\right)^2)} \tag{49}$$

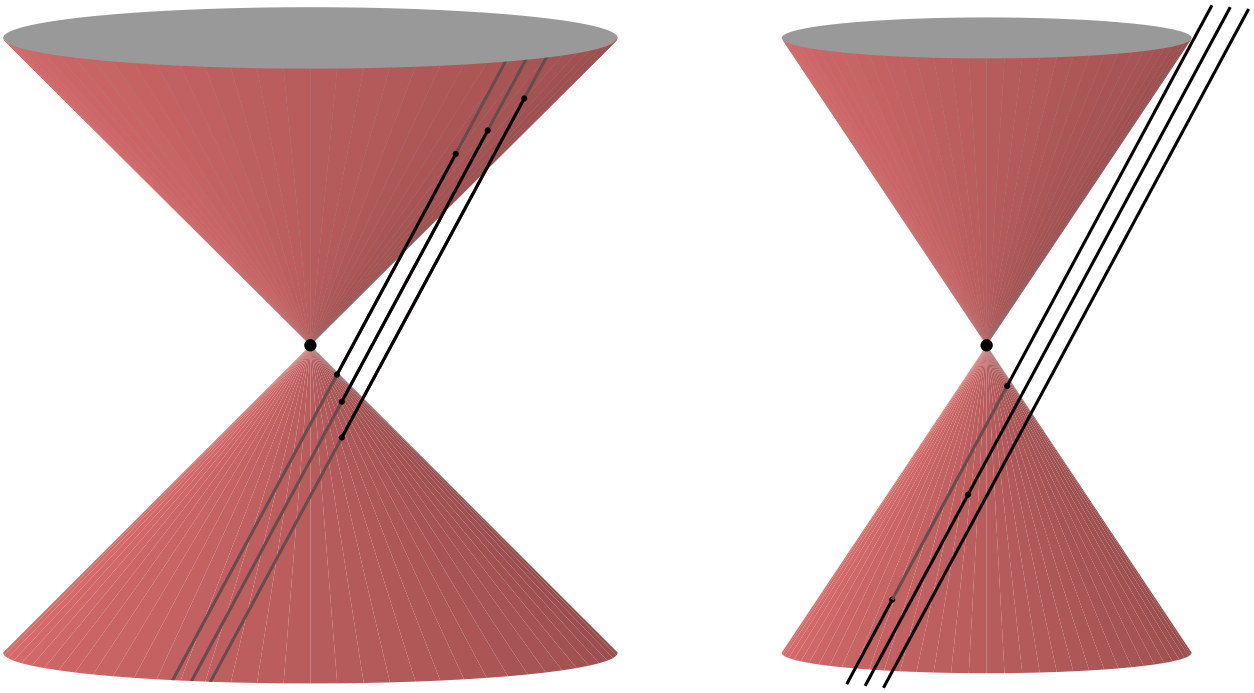

Fig. 3.— In a vacuum the worldline of a particle will intersect the past light cone once and the future light cone once, as shown on the left. In a dielectric medium light travels at c/n as shown on the right with $n = 1.5$. This allows the worldlines of particles with $\beta > 1/n$ to intersect the past light cone twice, be tangent to the light cone, or miss it entirely as shown by the three worldlines with $\beta = 0.8$. The cusp discontinuity in the potential produced when the worldline is tangent to the light cone produces a $1/r$ radiated field when differentiated to give $F_{\mu\nu}$.

There is also a solution with a minus sign in front of the square root but this corresponds to the world line being tangent to the future light cone. Note that if $n_r v < c$ then the above equation has no real solutions, so Cherenkov radiation only occurs when $n_r v > c$.

The magnitude of Cherenkov radiation can be worked out in detail: see §14.9 of Jackson. For a non-dispersive medium the energy radiated per unit frequency per unit path length is given by

$$-\frac{dE_\nu}{ds} = \frac{4\pi^2 e^2 \nu}{c^2}\left[1 - \frac{1}{\beta^2 n_r^2}\right] \tag{50}$$

Note that the integral of this over frequency diverges, which just shows that no material can be non-dispersive into the ultraviolet. Also note that this is a fairly blue light source: the $B - V$ color of a Cherenkov radiator is -0.13.

The amount of Cherenkov radiation emitted by a $\beta = 1$ particle going through water with $n_r = 1.33$ for frequencies less than 10^{15} Hz is 1.3 keV per cm, which is much less than the minimum ionizing energy loss of about 2 MeV/cm. Since the photon energy is about

4 eV for 300 nm wavelength, this corresponds to about 4-6 keV per photon. For air the energy loss per gm/cm^2 is about the same, with n a 1000 times closer to 1 and a density 1000 times smaller. Thus the energy loss per Cherenkov photon in air is about the same as the atmospheric flouresence, but the Cherenkov light is concentrated in a much smaller solid angle, so if one is in the cone the light is brighter. So particle detectors that depend on Cherenkov radiation must deal with a very inefficient light source. Detectors made of materials which convert some of the ionization loss into light are known as *scintillators*, and these produce much more light, but it is hard to beat the price of water when you are building a 50,000 tonne detector like Super-Kamiokande.

6. Photoelectric absorption

Peak absorption of ISM is at 13.6 eV and reaches $\approx 10^{-17}$ cm$^2/n_H$. This peak gives 5×10^6 cm^2/gm.

σ goes down as $E^{-7/2}$ for a given element for $E >> I$, but for the ISM with its mixture of elements $\sigma/n_H \propto E^{-8/3}$. Dust extinction at V gives $\sigma/n_H \approx 5 \times 10^{-22}$ cm^2 or 30 cm^2/gm. X-ray absorption matches this for two energies straddling the oxygen K edge at about 600 eV.

7. Compton Scattering

The Thomson scattering cross-section is

$$\sigma_T = \frac{8\pi e^4}{3m^2c^4} = (2/3) \times 10^{-24} \text{ cm}^2 \tag{51}$$

which corresponds to 0.4 cm^2/gm.

8. Cosmic Ray Astrophysics

See Chapter 15 in Longair.

The overall spectrum of cosmic rays is given by Longair Eq(15.1) as

$$I_N \approx 1.8 \times 10^4 E^{-2.7} \text{ nucleons/m}^2\text{/sr/sec/Gev} \tag{52}$$

with E being the energy per nucleon in GeV. This fit reaches a flux of $EdN/dE = 3/\text{m}^2/\text{yr}$ at an energy of 5×10^{15} eV. Since it is not practical to fly a detector as large as a square

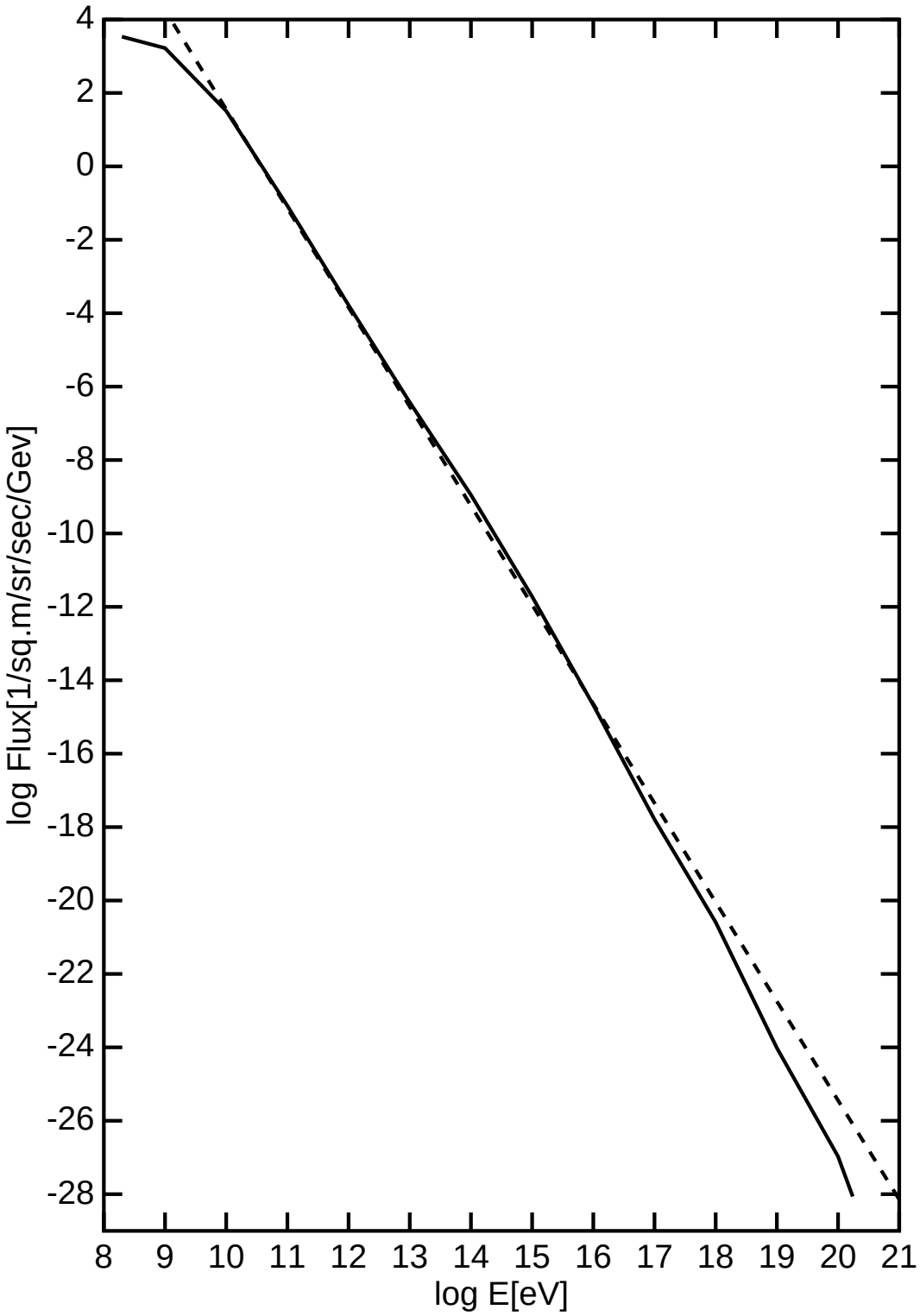

Fig. 4.— Flux of cosmic rays as a function of total energy [solid line] compared to a power-law fit with slope -2.7.

meter for a year at an altitude high enough to directly measure the composition, Figure 4 shows the measured flux of cosmic rays as a function of total energy compared to this fit. It looks like a good fit, but note that the y-axis of this log-log plot covers 32 orders of magnitude in flux. The measured flux is slightly less steep than the power-law fit for $E < 10^{15}$ eV, the "knee" in the spectrum, then becomes steeper until $E \approx 10^{18}$ eV, the "ankle" in the spectrum. Cosmic rays with energy greater than 5×10^{19} eV are above the Greisen-Kuzmin-Zatsepin (or GZK) cutoff, and have to come from sources closer than ~ 100 Mpc, and have such high rigidities that their arrival direction will reveal the location of the source. At the GZK cutoff, the fit gives $EdN/dE = 10^{-14}$ m^{-2} sr^{-1} s^{-1}, or one particle per square km per 3 years. Measurements of these highest energy cosmic rays always look at the many secondary particles produced in the *extensive air showers* they produce.

The atmosphere is quite thick so observations of the primary cosmic rays have to be made above all of the Earth's atmosphere in a rocket or satellite, or above most of the Earth's atmosphere in balloons floating above all but about 3 gm/cm^2 of the air. This is above 99.7% of the atmosphere and requires altitudes greater than 40 km. But by directly detecting the primary cosmic ray, very good diagnostics can be done.

The main tool for measuring the composition of cosmic rays is a combination of two ionization detectors. The first is a thin layer that measures the ionization loss rate dE/dx. The second is a thick detector that measures the total energy of the cosmic ray E. Thick detectors to measure the total energy are called calorimeters. There are three parameters to measure for a cosmic ray: charge, mass and energy. So with only two measurements, the solution can be ambiguous. But for non-relativistic cosmic rays, which can be measured near the Earth close to the magnetic poles where the geomagnetic cutoff is removed, or in interplanetary space, there is a fortunate coincidence, since the ionization loss

$$-\frac{dE}{\rho dx} = 1.5\frac{2Z_m}{A_m}\frac{Z^2}{\beta^2}\Big(1 + 0.2\ln\gamma\Big)\ \text{MeV/(gm/cm}^2\text{)} \tag{53}$$

which goes like β^{-2}, while the kinetic energy is $(\gamma - 1)Am_pc^2 \approx 0.5Am_p\beta^2c^2$, so product EdE/dx is proportional to Z^2A which uniquely identifies both the change Z and the mass A of the primary cosmic ray.

For high energies this coincidence gets distorted, but the particles also have a greater range, so if the calorimeter is divided into separately instrumented layers one can measure dE/dx vs x right up to where the particle stops, and the multiple samples of dE/dx give a unique determination of Z and A.

8.1. Abundances

The measured abundances are given in Table 15.1 and Figure 15.4 of Longair. The abundances of cosmic rays broadly follow the cosmic abundances with two main exceptions:

1. The heavier elements are somewhat over abundant at constant E.
2. The very rare elements like lithium, beryllium and boron are very overabundant - by factors up to 10^8. Flourine is also overabundant, as is scandium.

The overabundant rare elements all occur when the cosmic abundance is very low compared to the abundance of slightly heavier elements. This situation arises when the main route of nucleosynthesis skips over is given Z, as in the case of the tripe-α reaction making carbon straight from helium, and thus skipping over lithium, beryllium and boron.

But the large abundance of carbon then provides a "beam" of carbon cosmic ray primaries that can interact with interstellar medium protons via

$$^{12}C + p \rightarrow {}^{9}Be + 3p + n \tag{54}$$

This is a *spallation* reaction. Based on the cross-sections for spallation reactions and the observed abundances of cosmic rays, Hagen *et al.*. (1977, ApJ, 212, 262) found that mean amount of material traversed by the cosmic rays was $\lambda_e = 4$ gm/cm^2. This was derived assuming an exponential distribution of path lengths, $p(\lambda)d\lambda = \exp(-\lambda/\lambda_e)d\lambda/\lambda_e$. Simpson *et al.* (1983, ARNPS,33, 323) finds a path length of 5 gm/cm^2

8.2. Age

Of particular interest is the abundance of ^{10}Be, which is 11% of all beryllium. This isotope is radioactive with a half-life of 1.5×10^6 years. The spallation reaction cross-sections would make 19% of the beryllium in the ^{10}Be isotope, so some of the ^{10}Be has decayed, giving a measurement of the mean age of the cosmic rays. The beryllium isotopic ratio was measured in cosmic rays with $\beta = 2/3$. The mean age, or the leakage rate of a "leaky box" model, was found to be $\tau_e = 5^{+6}_{-3} \times 10^6$ yr. When combined with the 4 gm/cm^2, this gives a mean density of the medium through which these low energy cosmic rays travel of $n_H = 0.7^{+1.0}_{-0.4}$ /cc. Simpson *et al.*(1983) derived $\tau = 10 \times 10^6$ yr, while Yanasak *et al.*(2001, ApJ, 563, 768) derived $\tau = 15 \times 10^6$ yr. These longer ages lead to a lower mean density, $n_H = 0.34 \pm 0.04$ /cc. This desnity is still quite large, so this shows that low energy cosmic rays are mainly trapped within the disk of the Milky Way.

The mean field in the interstellar medium is about 3.3 μG, and for a proton with $\beta = 2/3$ the momentum is $pc = 834$ MeV, so the radius of the particle's orbit given by $R = pc/qB$ is

$834 \times 10^6/(300$ V/gauss-cm $\times 3.3 \times 10^{-6}) = 8.5 \times 10^{11}$ cm. Thus the cyclotron orbit is very small, and trapping in the disk ISM is to be expected.

8.3. Electrons

The electron spectrum in terms of dN/dE is 0.01 of the proton spectrum for energies up to 5 GeV. It then steepens to

$$dN/dE = 700E^{-3.3} \text{ /m}^2\text{/sec/sr} \tag{55}$$

for high energies according to Eq 15.2 of Longair. Figure 15.3 of Longair shows the electron cosmic rays are 300 times less than the protons at 100 GeV. At 1000 GeV they should be 1200 times less than the protons.

The steepening of the electron spectrum is caused by energy losses due to synchrotron emission in the galactic magnetic field.

8.4. Positrons

The electron to positron ratio is 0.2 at 1 GeV, 0.1 and 2 Gev, and 0.05 at 5-20 GeV according to Longair. But recently the PAMELA experiment has claimed to see a rise in the positron to electron ratio, going from 0.05 at 5 GeV to 0.12 at 100 GeV (Adriani *et al.*, 2010, arXiv:1001.3522). This claim is controversial and I am suspicious because the claimed rise in e^+/e^- tracks the well-known rise in p/e^-, so the positron to proton ratio is not changing in the PAMELA positron excess. This makes me worry about mis-identifying protons as positrons.

8.5. Total Energy Density

The total energy density is given by

$$u = 4\pi \int E^2 \frac{dN}{dE} \frac{d\ln E}{v(E)}. \tag{56}$$

The integrand peaks at about 10 GeV where $E^2 dN/dE = 3 \times 10^3$ GeV/m^2/sr/sec. The velocity is c so the integrand is 0.14 eV/cc. At 1 GeV the integrand is 0.1 eV/cc, while at high E it goes down like $E^{-0.7}$. The whole integral gives about 1 eV/cc for $E > 1$ GeV. But the low energy cosmic rays that dominate the integral are hard to measure due to solar modulation, so Neronov *et al.*(arXiv:1112.5541) used gamma ray observations of

local Giant Molecular Clouds to estimate that there is a break at 9 Gev and a total energy density of 0.9 ± 0.3 eV/cc. The energy density of the magnetic field is 0.26 eV/cc for 3.3 μG. The similarity of these two energy densities is not a coincidence. The energy density of starlight and the energy density of the CMB are also about 0.25 eV/cc but this similarity is coincidental.

8.6. Anisotropy

The cosmic rays are very isotropic at low energies, with dipole variations less than 0.1% at $E = 3 \times 10^{14}$ eV. Since a motion of the Solar System through an isotropic population would produce a dipole anisotropy through the Compton-Getting effect, one can rule drifts larger than 50 km/sec between the frame in which the cosmic rays are perfectly isotropic and the solar system. Thus the low energy cosmic rays have to move with the local ISM, and cannot, for example, be isotropic in a non-rotating halo of the Milky Way. But at the highest energy there appears to be some anisotropy, with $\mathcal{Q}(10\%)$ anisotropy at 10^{19} eV.

9. Air showers

9.1. Hadronic Interactions

The force between two quarks is described by Quantum ChromoDynamics (QCD). The force carriers in QCD are gluons, which are charged, unlike photons. As a result, the field lines of the color force do not spread out uniformly over 4π steradians. Instead, they attract each other leading to nearly one dimensional flux tubes. Hence for large distances, QCD interactions lead to a potential that is linear in separation, as seen in Figure 5. A typical slope is 1.2 GeV/Fermi. This is corresponds to the tension in a string. In normal units it is $T = (1.2 \times 10^9/6 \times 10^{11})/10^{-13} = 2 \times 10^{10}$ ergs/cm or the weight of 20 tonnes. This is a very heavy weight to be carried by a sub-atomic particle.

In a high energy collision between two nucleons, usually only 2 quarks get heavily involved. If the nucleons have Lorentz factor γ in the center-of-mass frame, then the 2 quarks will have a maximum Lorentz factor of γ. Now $\gamma = \cosh\psi$, where ψ is the rapidity. If the QCD string stretches for a time period $\tau = \hbar/m_\pi c^2$, then the string will follow the path in space-time $(ct; x) = (c\tau\cosh\psi; c\tau\sinh\psi)$ which is a set of events a proper time τ from the collision at $(0; 0)$. Figure 6 shows a diagram of this path. The length of the string is given by $ds^2 = (c\tau d\psi)^2(\cosh^2\psi - \sinh^2\psi)$ so $ds = c\tau d\psi$. The total length of this path is $2c\tau\psi$. A piece of string has energy $T\Delta\psi$ and if this is greater than the pion rest mass a lower energy state can be created by breaking the string by creating a new quark-antiquark

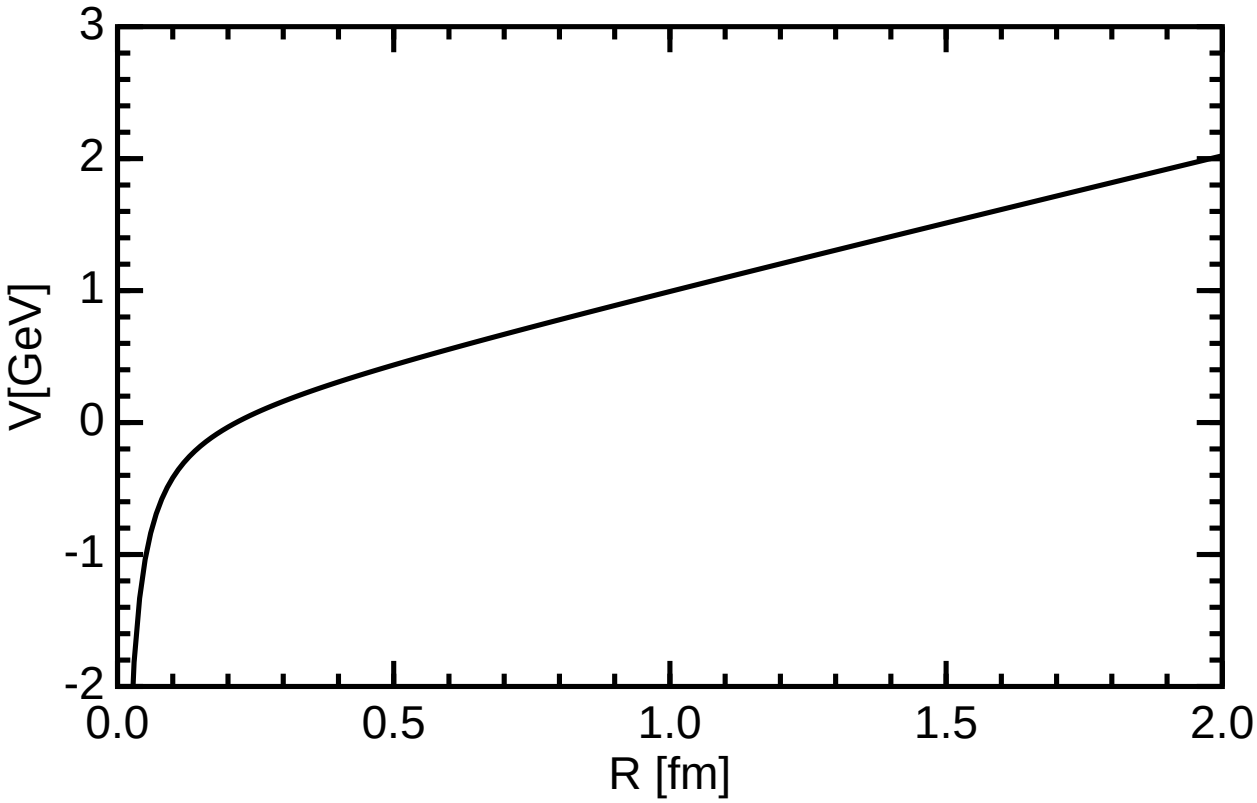

Fig. 5.— Sketch of the potential between two quarks as a function of distance. For short distances it goes like $1/R$, but for large distances it goes like R.

pair. So the collision makes a number of pions $n_\pi \approx \mathcal{O}(2c\tau\psi T/m_\pi c^2) = \mathcal{O}(2\hbar\psi T/m_\pi^2 c^3)$.

Now in the CoM frame the magnitude of the 4-momentum is $-4(m_p\gamma c)^2$ while in the lab frame it is $(m_p c)^2[-(1+\gamma')^2 + (\gamma'\beta')^2]$ which is $-2(m_p c)^2[1+\gamma']$. Thus $\gamma \approx \sqrt{\gamma'/2} \approx e^\psi/2$ and $\psi \approx \ln\gamma'/2$. Finally $n_\pi \approx \hbar \ln\gamma' T/m_\pi^2 c^3$. For $\gamma' = 2.7\times10^7$, equivalent to the LHC at 3.5 TeV, one gets an order of magnitude estimate of $n_\pi \approx 200$. The range in $\psi \approx 17$. The observed number of π^- is about one per unit of ψ so the total number of pions is about 50, showing that our order of magnitude estimate was off by a factor of 4. But the multiplicity going like $\ln E$ is a good fit to the data, as is a uniform distribution in rapidity.

The pions produced in the collision decay into leptons and photons, with the π^0's decaying in 1.8×10^{-16} sec into 2 photons, while the $\pi^\pm$ decay in 2.6×10^{-8} sec, usually into a muon and a muon neutrino.

In the lab frame, the leading nucleon still has a reasonable fraction of its initial energy.

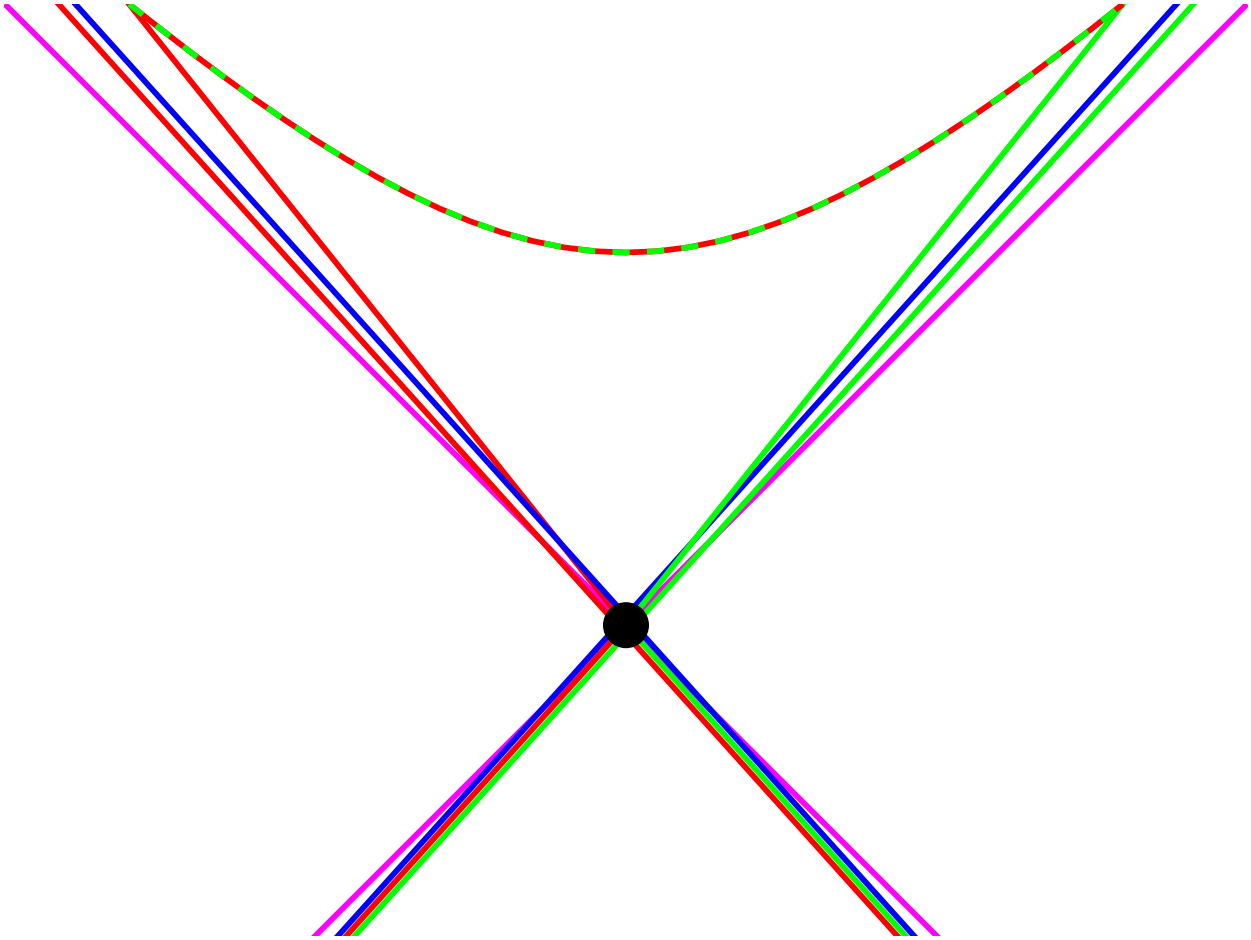

Fig. 6.— Cartoon space-time diagram showing a hard collision between two baryons in the center-of-mass frame. Magenta lines show the light cones. Two quarks get knocked out and a string of "glue" stretches between them on a hyperbola of constant proper time since the collision.

9.2. Total Cross-section

The total cross-section for pp collisions is 70 mb, where 1 barn (b) is 10^{-24} cm^2. This corresponds to a circle of radius 1.5 fm. 1 fermi (fm) is 10^{-13} cm or 1 femtometer (fm). Since a nucleus of atomic weight A has a radius of $1.2A^{1/3}$ fm, we seen that a nucleon is close to opaque to nucleons. Taking the radius of a nitrogen nucleus as 3 fm using the formula, and using a cross-section of πR^2, one gets 280 mb, which gives an absorption length in air of 80 gm/cm^2.

The pions produced in hadronic collisions can also collide with air nuclei. But since mesons have only two quarks instead of the three in a baryon, their interaction length is about 120 gm/cm^2. This amount of air is in 1 km at the surface of the Earth, which is 131 times longer than c times the $\pi^\pm$ lifetime. So at the surface, charged pions usually decay

unless their energy is greater than $131 m_\pi c^2 = 18$ GeV, while at a depth of 80 gm/cm^2 one needs 12.5 km of air to give 120 gm/cm^2 and they will usually decay unless their energy is > 220 GeV.

10. X-ray Astronomy History

1. 1948 - NRL rocket experiment detects the Sun

2. Predictions from Giacconi, Clark & Rossi, 1960:

Source	$E > 600$ eV γ/cm^2/sec
(a) Sun	10^6
(b) Sirius if $L_X \approx L_{opt}$	0.25
(c) Moon	$0 - 1.6 \times 10^3$

3. Rocket flight in 1962 to detect Moon. Detected source at azimuth $195 \pm 15°$ when Moon was at 165°. (But the figure shows 210°, which is actually much closer to the true position.) Flux from this source, called Sco X-1, was 28 γ/cm^2/sec. (Giacconi *et al.* (1962, PRL, 9, 439).

   The Moon was not detected. This experiment has 3 Geiger counters. One had 7 mg/cm^2 of mica for a window, the others had 1.4 mg/cm^2 of mica. The FWHM of the detectors was about 120°. The rocket spun around the vertical, with the detector faces 55° from the spin axis. The flight reached peak altitude at 00:02 MST above White Sands on 18 June 1962. The latitude and longitude are 32.5646°, 106.3591° W. The GMT was 07:02 18 June 1962, so the Greenwich Sidereal Time was 00:45:48 and the Local Sidereal Time was 17:40:22.

   The quoted source position was $(\alpha,\ \delta) = ((17.7 \pm 0.5)^h, (-42 \pm 5)°)$.

4. Bowyer *et al.* (1964, AJ, 69, 135) flew a rocket with a large area proportional counter on 29 Apr 1963. The detector area was 86 cm^2. A honeycomb collimator in front of the detector restricted the field of view to 8° FWHM. A strong sources was seen at $(\alpha,\ \delta) = (16^h 10^m, -18°)\ \pm 1°$.

   Note that Giacconi's position was neither precise nor accurate, being 25° off and also 5σ off. But Nobel Prize winning work.

5. Gursky *et al.* (1966, ApJ, 146, 310) used a modulation collimator to get possible positions of $(16^h 17^m 7^s,\ -15°30'54'')$ and $(16^h 17^m 19^s,\ -15°35'20'')$, both $(\pm 4^s,\ \pm 30'')$. A modulation collimator has two wire grids spaced apart, giving a periodic series of sensitivity peaks. Gursky *et al.* had two MCs, one with a spacing of 302″ and the other

with a spacing of 317″. Uncertainty in deciding which fringe was being observed led to the two possible positions.

6. Sandage *et al.* (1966, ApJ, 146, 316) used these positions to optically identify Sco X-1 with a UV-bright star having colors V=12.6(2), B-V=0.30(5), U-B=-0.8(1) at $(\alpha,\ \delta) = (16^h17^m4.3^s,\ -15°31'13'')$.

7. Gottlieb, Wright & Liller (1974, IAU Circular 2704) give a 3.74 day period based on more than a 1000 plates in the Harvard plate stacks covering 1889-1953.

8. Gottlieb, Wright & Liller (1975, ApJL, 195, L33) correct to a 0.787313 day period. $1/0.787313 - 1/3.74001 = 1.00276$ cycles/day or 1 cycle per sidereal day. All of the magnitudes were taken at about the same Greenwich sidereal time so this kind of *aliasing* is not uncommon. Full amplitude is 0.22 magnitudes.

11. X-ray Satellites

See `http://imagine.gsfc.nasa.gov/docs/sats_n_data/xray_missions.html` for full list, highlights below:

1. Vela satellites: 3-750 keV
   (a) Launched 1969 - secret project to look for nuclear tests in space
   (b) 3-12 keV All Sky Monitor, 6° FWHM, 26 cm^2
   (c) 6 γ-ray detectors, 150-750 keV, $\sim$ 60 cc of CsI
   (d) 4 units: Vela 5A & B, Vela 6A & B
   (e) Circular orbit, 118 Mm from Earth, period 112 hours
   (f) Discovered γ-ray bursts, declassified in 1973
   (g) Grindlay, Wright & McCrosky (1974, ApJL, 192, L113) showed the optical flux from GRB720514 was fainter than 6^{th} magnitude using Prairie Network meteor camera films

2. Uhuru
   (a) launched into 540 km, 3° inclination, 96 minute period orbit on 12 Dec 1970. Operated until 1973.
   (b) Spin period 12 minutes
   (c) 2-20 keV
   (d) Sky survey with sensitivity of 0.5 mCrab

(e) Two sets of proportional counters, each with 840 cm^2 area

(f) Angular resolution 0.5° and 5° FWHM

(g) Made catalog of 339 celestial X-ray sources.

(h) 1 UFU or Uhuru Flux Unit is 1.7×10^{-11} erg/cm^2/sec in 2-11 keV for the Crab spectrum or one counts per second in the Uhuru detectors. 1 μJy is 10^{-29} erg/cm^2/sec/Hz is about 1 UFU.

(i) Last catalog was the 4^{th}, leading to "4U" source names.

3. Astronomische Nederlandse Satelliet (ANS)

(a) launched 30 Aug 1974 into a nearly polar, Sun-synchronous orbit
500 km circular orbit desired, 280x1150 km achieved so high background

(b) operate until 14 June 1977

(c) 3 instruments:

i. UV instrument, 150-330 nm

ii. Soft X-ray instrument

iii. hard X-ray instrument, 1-30 keV, 40 cm^2, FOV = $10' \times 3°$

4. SAS-3 (Small Astronomy Satellite)

(a) 0.1-60 keV

(b) Launched 7 May 1975 into equatorial circular 95 minute orbit and spin period

(c) Modulation collimators, 2-11 keV, 4.5′ FWHM in 12° FOV, on the satellite spin axis.

(d) Slat and tube collimators
Slat collimators give a single triangular shaped beam in one dimension but a long sensitive area perpendicular to that axis. Three long FOVs, two with $1 \times 32°$ FWHM tilted like \ and /, and a $0.5 \times 32°$ FWHM like |.

(e) Soft X-ray experiment, 0.15-1 keV, 2.9° FOV

5. HEAO 1

(a) Launched 12 Aug 1977, operated until 9 Jan 1979

(b) Mass 3000 kg, 4 instruments

i. A1: 0.25-25 keV, 10^4 cm^2, FWHMs $1 \times 0.5°$ to $1 \times 4°$ in 7 different modules

ii. A2: cosmic X-ray background experiment, 2-60 keV with 3 modules, FOV $3 \times 1.5°$, $3 \times 3°$ & $3 \times 6°$

iii. A3: Modulation collimator, 0.9-13.3 keV, 700 cm^2

iv. A4: hard X-ray & γ-ray

A. 15-200 keV, 200 cm^2, $1.7 \times 20^\circ$ FWHM

B. 80-2000 keV, 180 cm^2, 17° FWHM

C. 120 keV to 10 MeV, 100 cm^2, 37° FWHM

6. HEAO 2 - *Einstein*

   (a) 12 Nov 1978 to April 1981

   (b) 0.2-20 keV

   (c) Monitor Proportional Counter, 667 cm^2, 1.5-20 keV

   (d) First imaging X-ray telescope in space
   Wolter Type I grazing incidence telescope, 0.1-4 keV, with 4 instruments

      i. Imaging Proportional Counter, 0.4-4 keV, 100 cm^2 effective area, FOV $75'$, FWHM $1'$

      ii. High Resolution Imager, 0.15-3 keV, 5-20 cm^2 eff. area, FOV $25'$, FWHM $2''$

      iii. Solid State Spectrometer, 200 cm^2, $E/\Delta E$ of 3-25, FOV $6'$

      iv. Focal Plane Crystal Spectrometer, 0.42-2.6 keV, 0.1-1 cm^2, $E/\Delta E$ of 50-100 for $E < 0.4$ keV, 100-1000 at $E > 0.4$ keV

      v. Also Objective Grating Spectrometer used with HRI, $E/\Delta E$ of 50

   Giacconi & Rossi (1960, JGR, 65, 773) first proposed using a focusing telescope working at grazing incidence for X-rays, but their first suggestion was a steep paraboloid. The plasma frequency is $2\pi\nu_p = \sqrt{4\pi e^2 n_e/m_e}$ which gives $\nu_p = 8 \times 10^{15}$ Hz for aluminum assuming all the electrons are free. This is about 30 eV photon energy. Above the plasma frequency the index of refraction is less than 1, $n = \sqrt{1-(\nu_p/\nu)^2}$, so *total external reflection* is possible. This requires $\sin i > n$. In term of the grazing angle $g = 90^\circ - i$, this gives $\sin g < \nu_p/\nu$. For 1 keV photons and an aluminum mirror, the grazing angle has to be less than 2°. Materials with the highest electron density are better, and these are heavier metals like gold. Two mirrors are needed to get a good image with grazing incidence. The Abbé sine condition states that the sine of the ray angle at the image has to be proportional to sine of the angle at the object. For astronomy, the object is at infinity, so this reduces to requiring that the intersections between the rays at the focus and the incoming rays fall on a sphere centered on the focus. With a steep paraboloid followed by a steep hyperboloid, this can be approximated. See `http://imagine.gsfc.nasa.gov/docs/science/how_l2/xtelescopes_systems.html`

7. ROSAT: The Roentgen Satellite

   (a) 1 June 1990 to 12 Feb 1999

   (b) ROSAT All Sky Survey, 1.5×10^5 sources

   (c) Wide Field Camera: EUV 62-206 eV

(d) Focussing X-ray telescope with 2 instruments:

i. Position Sensitive Proportional Counter (PSPC), 2° FOV, $E/\Delta E \approx 0.4/\sqrt{E \text{ in keV}}$, eff. area 240 cm^2

ii. High Resolution Imager, 38′ FOV, eff. area 80 cm^2, 2″ FWHM

8. ASCA: The Advanced Satellite for Cosmology and Astrophysics

(a) 20 Feb 1993 to 15 June 2000

(b) 0.4-10 keV

(c) First use of CCDs in an X-ray astronomy

(d) Foil mirror X-ray telescopes, 120 nested Au-coated Al foil surfaces
1300 cm^2 eff. area at 1 keV, spatial resolution 3′, FOV 24′ at 1 keV. 2 instruments:

i. Gas Imaging Spectrometer, 0.8-12 keV, 50 cm^2 eff. area

ii. Solid State Imaging Spectrometer, 0.4-12 keV, FOV 22′, energy resolution 2% at 6 keV, eff. area 105 cm^2

9. Rossi X-ray Timing Explorer (RXTE)

(a) 30 Dec 1995 - 4 Jan 2012

(b) 2-250 keV

(c) large area proportional counters, high time resolution

(d) Discovered kiloHertz Quasi Periodic Oscillations (QPOs)

(e) 6500 cm^2, 1 μsec time resolution

10. Chandra (née AXAF)

(a) 23 July 1999 - present

(b) orbit 64 hour period highly elliptical

(c) Wolter Type I telescope (2 concentric rings) with 800 cm^2 area at 0.25 keV and 400 cm^2 at 5 keV, 4 instruments:

i. ACIS: AXAF Charge-Coupled Imaging Spectrometer
$E/\Delta E = 20 - 50$ at 1-6 keV.
FOV is 16′

ii. High Resolution Camera, 30′ FOV, 0.5″ resolution

iii. High Energy Transmission Grating, $E/\Delta E$ = 60-1000, 0.5-10 keV

iv. Low Energy Transmission Grating, $E/\Delta E$ = 30-2000, 0.08-6 keV

11. XMM - Newton

(a) 10 Dec 1999 - present

(b) foil mirror telescope with very large collecting area

(c) 1500 cm^2 at 1 keV, 6″ resolution

(d) EPIC - European Photon Imaging Camera, 33′ FOV, $E/\Delta E$ = 20-50

(e) Reflection Grating Spectrograph

12. Suzaku (née Astro-E2)

(a) 10 July 2005 - present

(b) 0.2-600 keV

(c) First μ-calorimeter in space

(d) μ-calorimeter failed due to cryogenics mistake on 8 Aug 2005

(e) Angular resolution 2′

(f) CCD imaging spectrometer, 18′ FOV, 340 cm^2 eff. area

(g) ΔE = 130 eV at 6 keV

(h) Hard X-ray Detector not focused, FWHM = 34′ E < 100 keV, 4.5° E > 100 keV

12. Synchrotron Radiation

12.1. 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

$$p^\mu = mc(\gamma; \gamma\beta \sin\phi \cos\theta, \gamma\beta \sin\phi \sin\theta, \gamma\beta \cos\phi) \tag{57}$$

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

$$\Omega = \frac{eB}{\gamma mc} \tag{58}$$

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:

$$\begin{aligned} p^{\mu'} &= mc(\gamma'; \gamma'\beta' \cos\theta, \gamma'\beta' \sin\theta, 0) \\ B' &= B \end{aligned} \tag{59}$$

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

$$\begin{aligned} p^{\mu''} &= mc(1;0) \\ B'' &= \gamma' B \\ E'' &= \gamma'\beta' B = \gamma\beta \sin\phi\, B \end{aligned} \tag{60}$$

In this frame only the electric field causes accelerations, so

$$a'' = \frac{\gamma\beta \sin\phi\, eB}{m} \tag{61}$$

and the Larmor formula gives the radiated power as

$$P'' = \frac{2}{3}\frac{e^2(a'')^2}{c^3} = \frac{2}{3}\frac{e^4}{m^2c^3}\gamma^2\beta^2 \sin^2\phi\, B^2 \tag{62}$$

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

$$P = \frac{4}{9}\frac{e^4}{m^2c^3}\gamma^2\beta^2\, B^2 = \frac{4}{3}\gamma^2\sigma_T c u_B \tag{63}$$

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

12.2. 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:

$$\begin{aligned} p^{\mu''} &= (h\nu''/c)(1;0,1,0) \\ &\rightarrow \\ p^{\mu'} &= (h\nu''/c)(\gamma';\gamma'\beta',1,0) \end{aligned} \tag{64}$$

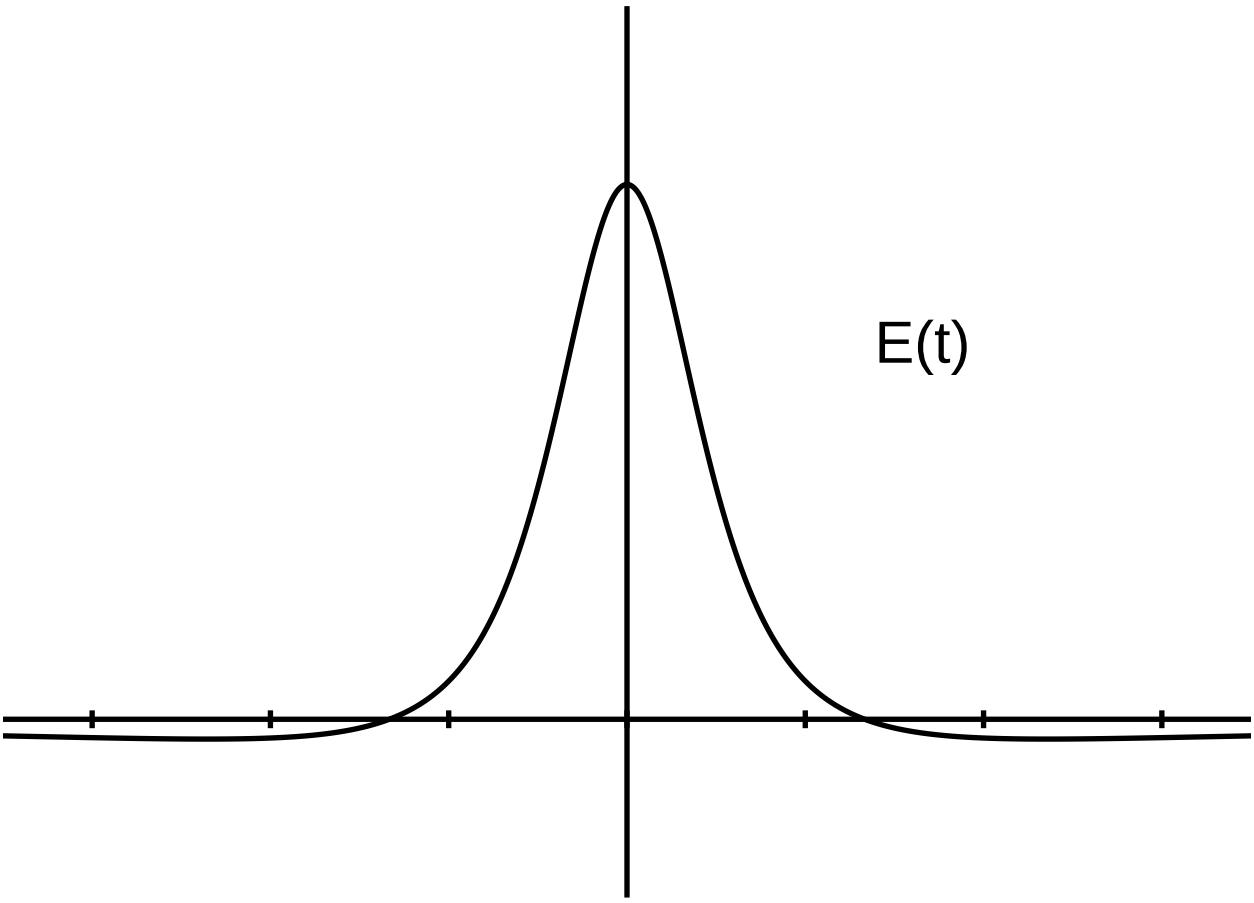

Fig. 7.— The pulse shape by an observer on the plane of the electron's synchrotron orbit. Ticks every $mc/(2\gamma^2 eB)$ on the t-axis.

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

$$\Delta t_{rec} = (1 - \beta')\Delta t' \approx \frac{mc}{2(\gamma')^2 eB} \tag{65}$$

Thus a typical frequency for synchrotron radiation will be

$$\omega' = \frac{1}{\Delta t_{rec}} = \frac{2(\gamma')^2 eB}{mc} \tag{66}$$

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

$$\omega_{crit} \propto \gamma^2 \frac{eB}{mc} \sin\phi \tag{67}$$

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

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

$$\frac{\nu}{\nu'} = \gamma(1+\beta\cos\theta') \tag{68}$$

and the angle in the guiding center frame is given by

$$\sin\theta = \frac{p_y}{|p|} = \frac{\sin\theta'}{\gamma(1+\beta\cos\theta')} \tag{69}$$

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

$$E_{rad} \propto eB\gamma^3\beta(1+\beta\cos\theta')^2\cos\theta' \tag{70}$$

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

$$y = \frac{\sin\theta'}{1+\beta\cos\theta'} \tag{71}$$

Then

$$\begin{aligned} t_{rec} &= \frac{\theta}{\Omega} - \frac{R}{c}\sin\theta \\ &= \frac{1}{\Omega}\sin^{-1}\left(\frac{y}{\gamma}\right) - \frac{R}{c}\frac{y}{\gamma} \\ &= \frac{\gamma mc}{eB}\left(\frac{y}{\gamma} + \frac{1}{6}\left(\frac{y}{\gamma}\right)^3 + \ldots\right) - \frac{\gamma\beta mc}{eB}\frac{y}{\gamma} \\ &= \frac{mc}{eB}\left[(1-\beta)y + \frac{y^3}{6\gamma^2} + \ldots\right] \\ &= \frac{mc}{2\gamma^2 eB}\left[y + \frac{y^3}{3}\right] \end{aligned} \tag{72}$$

We have already assumed $\beta \approx 1$ in $1-\beta = 1/(2\gamma^2)$, so continuing to use $\beta \approx 1$ gives us

$$y^2 = \frac{\sin^2\theta'}{(1+\cos\theta')^2} = \frac{1-\cos^2\theta'}{(1+\cos\theta')^2} = \frac{1-\cos\theta'}{1+\cos\theta'} \tag{73}$$

Hence

$$\cos\theta' = \frac{1-y^2}{1+y^2} \tag{74}$$

and

$$1+\cos\theta' = \frac{2}{1+y^2} \tag{75}$$

Thus we get

$$E_{rad}(t_{rec}) \propto eB\gamma^3 \frac{1-y^2}{(1+y^2)^3} \tag{76}$$

at

$$t_{rec} = \frac{mc}{2\gamma^2 eB}\left[y + \frac{y^3}{3}\right] \tag{77}$$

We can of course find the spectrum by Fourier transforming this pulse, giving

$$\hat{E}(\omega) \propto \int \frac{1-y^2}{(1+y^2)^3} e^{i\omega t} dt \tag{78}$$

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

$$t_{pole} = \frac{mc}{2\gamma^2 eB}\left(i - \frac{i}{3}\right) = \frac{imc}{3\gamma^2 eB} \tag{79}$$

The high frequency cutoff in $\hat{E}(\omega)$ thus goes like

$$\begin{aligned} \hat{E}(\omega) &\propto \exp\left[-\frac{mc}{3\gamma^2 eB}\omega\right] \\ |\hat{E}(\omega)|^2 &\propto \exp\left[-\frac{2mc}{3\gamma^2 eB}\omega\right] \\ &= \exp\left[-\frac{\omega}{\omega_{crit}}\right] \end{aligned} \tag{80}$$

with a critical frequency that is just $1.5\gamma^2$ times the cyclotron frequency:

$$\begin{aligned} \omega_{crit} &= \frac{3}{2}\frac{eB}{mc}\gamma^2 \\ \nu_{crit} &= \frac{\omega_{crit}}{2\pi} = (4.2\ \mathrm{MHz/gauss})B \end{aligned} \tag{81}$$

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

$$\hat{E}(\omega) \propto \int \frac{1-y^2}{(1+y^2)^2} \exp\left[\frac{i\omega mc}{2\gamma^2 eB}\left(y + \frac{y^3}{3}\right)\right] dy \tag{82}$$

Integrate by parts ($\int pdq = pq - \int qdp$) 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

$$\hat{E}(\omega) \propto \omega \int y \exp\left[\frac{i\omega mc}{2\gamma^2 eB}\left(y + \frac{y^3}{3}\right)\right] dy \propto \omega K_{2/3}(\omega/(3\gamma^2 eB/mc)) \tag{83}$$

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

$$\begin{aligned} P(\omega) &\propto F\left(\frac{\omega}{\omega_{crit}}\right) \\ F(x) &= x\int_x^\infty K_{5/3}(y)dy \end{aligned} \tag{84}$$

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

$$\frac{\int \sin^3\phi\, d\Omega}{\int \sin^2\phi\, d\Omega} = \frac{2\pi\int \sin^4\phi\, d\phi}{8\pi/3} = \frac{3}{4}\int_0^\pi [(1-\cos(2\phi))/2]^2 d\phi = \frac{9\pi}{32} \tag{85}$$

Thus the power weighted mean critical frequency is

$$\frac{3}{2}\times\frac{9\pi}{32}\times\frac{eB}{mc}\gamma^2 = 1.325\frac{eB}{mc}\gamma^2 \tag{86}$$

I recommend remembering this factor 1.325 as 4/3.

12.3. 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:

$$P(\nu) = \frac{4}{3}\gamma^2 c\sigma_T u_B \delta\left(\nu - \frac{4}{3}\,\frac{\gamma^2 eB}{2\pi mc}\right) \tag{87}$$

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

$$L_\nu = \int \frac{4}{3}\gamma^2 c\sigma_T u_B \delta(\nu-\nu_\circ\gamma^2)N(\gamma)d\gamma = \left.\frac{2\gamma N(\gamma)c\sigma_T u_B}{3\nu_\circ}\right|_{\gamma=\sqrt{\nu/\nu_\circ}} \tag{88}$$

An easier way to remember this is

$$\nu L_\nu = \left(\frac{\gamma N(\gamma)}{2}\right)\left(\frac{4}{3}\gamma^2 c\sigma_T u_B\right) \tag{89}$$

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

$$L_\nu \propto \gamma N(\gamma) \propto \sqrt{\nu/\nu_\circ} N\left(\sqrt{\nu/\nu_\circ}\right) \propto \nu^{(1-\beta)/2} \tag{90}$$

we see that the power index in $L_\nu \propto \nu^{-\alpha}$ is

$$\alpha = \frac{\beta - 1}{2} \tag{91}$$

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.

12.4. 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

$$E_{part} = \int N(\gamma)\gamma mc^2 d\gamma \tag{92}$$

and

$$\gamma N(\gamma) = \frac{3\nu_\circ L_\nu}{2c\sigma_T\, u_B} \tag{93}$$

But $d\gamma = \gamma d\nu/(2\nu)$ so

$$E_{part} = \int \frac{3\nu_\circ \gamma mc^2 L_\nu}{4c\sigma_T\, u_B} \frac{d\nu}{\nu} = \int \frac{3\sqrt{\nu_\circ \nu} mc^2 L_\nu}{4c\sigma_T\, u_B} \frac{d\nu}{\nu} \tag{94}$$

which becomes

$$E_{part} = \frac{\sqrt{24\pi emc}}{\sigma_T} B^{-3/2} \int \sqrt{\nu} L_\nu d\ln\nu = 1.5 \times 10^{12} B^{-3/2} \int \sqrt{\nu} L_\nu d\ln\nu \tag{95}$$

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

$$B_{eq} = \left(\frac{8\pi A}{V}\right)^{2/7} \tag{96}$$

We can also find the B that minimizes the total energy needed to produce the observed synchrotron emission,

$$B_{ME} = \left(\frac{6\pi A}{V}\right)^{2/7} \tag{97}$$

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}.

12.5. Synchrotron Self-Absorption

A population of relativistic electrons produces a synchrotron emissivity of

$$4\pi j_\nu = \left.\frac{2\gamma n(\gamma) c\sigma_T u_B}{3\nu_\circ}\right|_{\gamma=\sqrt{\nu/\nu_\circ}} \tag{98}$$

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 $n = p^2c/\hbar eB = \gamma^2\beta^2 m^2c^3/\hbar eB$. 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

$$kT_{ex} = \frac{\Delta E}{\ln(n_1/g_1) - \ln(n_2/g_2)} \quad \Rightarrow \quad kT_{ex} = -\left(\frac{\partial \ln(n/g)}{\partial E}\right)^{-1} \tag{99}$$

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

$$kT_{ex} = -\left(\frac{\partial \ln(\gamma^{-(\beta+2)})}{mc^2 \partial\gamma}\right)^{-1} = \frac{\gamma mc^2}{\beta + 2} \tag{100}$$

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.

13. Cosmic Ray Acceleration

Cosmic rays are definitely accelerated in supernova remnants which have strong shock waves. γ-rays produced by the $p + p \rightarrow p + p + \pi^\circ$ reaction have been seen coming from SNRs by the HESS telescope. The mechanism for the acceleration involves the cosmic rays, isotropic in the ISM rest frame, going through the shock front. If the shock is traveling at speed U, and is strongly supersonic, then the gas behind the shock is 4 times denser than the preshock gas and thus traveling at $0.75U$. After passing through the shock front, the cosmic rays re-isotropize but now in the frame moving at $0.75U$ relative to the ISM. This results in an energy gain that is effectively a Doppler blue shift. Now the cosmic rays pass through the shock in the other direction, once again experiencing an energy gain and re-isotropization. Every time the cosmic rays cycle across the shock and back they gain a factor β in energy. But there is also a loss of cosmic rays, since some diffuse away from the shock and never

cross it again. If the probability of remaining at the shock is P in one cycle, and the starting energy is $E_\circ$, then it takes $n = \ln(E/E_\circ)/\ln\beta$ cycles to reach energy E and only a fraction P^n of the cosmic rays remain. Thus

$$\frac{N(>E)}{N_\circ} = P^{\ln(E/E_\circ)/\ln\beta} = \exp(\ln P \ln(E/E_\circ)/\ln\beta) = \left(\frac{E}{E_\circ}\right)^{\ln P/\ln\beta}. \qquad (101)$$

Thus this mechanism naturally produces a power low distribution of cosmic ray energies, as observed at the Solar System, and deduced from the power law spectrum seen in synchrotron sources.

In order to compute the index of the power law we have to evaluate the energy gain β and escape probability $1 - P$. The energy gain in one passage across the shock for particle approaching at angle θ such that $\mu = \cos\theta$ is a factor of $1 + 0.75\mu U/c$. Now the rate of particles approaching the shock goes like $\mu d\mu$ which integrates to 0.5 for particles with $\mu > 0$, while the energy gain integrates to $0.25U/c$. Thus the net gain factor is $1 + 0.5U/c$. In the full cycle across the shock and back, the gain factor is $\beta = 1 + U/c$.

To compute the escape probability, note that for cosmic ray density n, all traveling at c isotropically, the rate of crossing the shock front is given by $0.5nc\int_0^1 \mu d\mu = nc/4$. But the cosmic rays downstream of the shock are isotropic in the frame of the postshock gas which is moving away from the shock at speed $U/4$. Thus cosmic rays are advected away from the interaction region at a rate of $nU/4$. Since $nc/4$ cross the shock and only $nU/4$ are lost, the escape probability is U/c and $P = 1 - U/c$. Thus $\ln P/\ln\beta = -1$, and the cosmic ray spectrum is $N(>E) \propto E^{-1}$, corresponding to $dN/dE \propto E^{-2}$. This spectrum is generally flatter than the observed spectral slopes, but the radio spectral index for this spectrum is $F_\nu \propto \nu^{-0.5}$ which is the boundary between flat and steep spectrum radio sources. For $dN/dE \propto E^{-2}$ the integral for the total energy in relativistic particles is logarithmically divergent, so for steeper spectra like $dN/dE \propto E^{-2.7}$ the energy is dominated by low energy particles and for steeper spectra like $dN/dE \propto E^{-1.5}$ it is dominated by high energy particles.

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

$$p^\mu = \frac{h\nu}{c}(1; -\cos\theta, -\sin\theta, 0) \qquad (102)$$

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

$$p^{\mu'} = \frac{h\nu}{c}(\gamma(1 + \beta\mu); -\gamma(\beta + \mu), -\sin\theta, 0) \qquad (103)$$

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

$$p^{\mu''} = \frac{h\nu''}{c}(1; \cos\theta', \sin\theta', 0) \tag{104}$$

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

$$p^{\mu'''} = \frac{h\nu''}{c}(\gamma(1+\beta\mu'); \gamma(\beta+\mu'), \sin\theta', 0) \tag{105}$$

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

$$\frac{\nu'''}{\nu} = \gamma^2(1+\beta\mu)(1+\beta\mu') \tag{106}$$

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):

$$p(\mu) = \frac{1+\beta\mu}{2} \tag{107}$$

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

$$p'_{R\&L}(\mu') = \frac{1}{2} \tag{108}$$

With this approximation we can easily find the average ratio:

$$\begin{aligned}
\left\langle \frac{\nu'''}{\nu} \right\rangle &= \int\int \gamma^2(1+\beta\mu)(1+\beta\mu')p(\mu)p'(\mu')d\mu d\mu' \\
&= \gamma^2 \int \frac{(1+\beta\mu)^2}{2} d\mu \int \frac{1}{2}(1+\beta\mu')d\mu' \\
&= \gamma^2\left(1+\frac{\beta^2}{3}\right)
\end{aligned} \tag{109}$$

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

$$\cos\psi = -\left(\mu\mu' + \sqrt{1-\mu^2}\sqrt{1-\mu'^2}\cos\phi\right) \tag{110}$$

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

$$\begin{aligned} p'(\mu';\mu) &= \frac{3}{8}\int_0^{2\pi}(1+\cos^2\psi)\frac{d\phi}{2\pi} \\ &= \frac{3}{8}\int_0^{2\pi}\left[1+\mu^2\mu'^2+2\mu\mu'\sqrt{1-\mu^2}\sqrt{1-\mu'^2}\cos\phi+(1-\mu^2)(1-\mu'^2)\cos^2\phi\right]\frac{d\phi}{2\pi} \\ &= \frac{3}{8}\left[1+\mu^2\mu'^2+\frac{1}{2}(1-\mu^2)(1-\mu'^2)\right] \\ &= \frac{3}{8}\left[\frac{3}{2}(1+\mu^2\mu'^2)-\frac{1}{2}(\mu^2+\mu'^2)\right] \end{aligned} \tag{111}$$

The (3/8) factor normalizes the pdf:

$$\int p'(\mu';\mu)d\mu' = \frac{3}{8}\left[3+\mu^2-\mu^2-\frac{1}{3}\right] = 1 \tag{112}$$

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

$$p(\mu,\mu') = p(\mu)p'(\mu';\mu) \tag{113}$$

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:

$$\begin{aligned} \left\langle\frac{\nu'''}{\nu}\right\rangle &= \int\int\gamma^2(1+\beta\mu)(1+\beta\mu')p(\mu)p'(\mu';(\beta+\mu)/(1+\beta\mu))d\mu d\mu' \\ &= \gamma^2\int\frac{(1+\beta\mu)^2}{2}d\mu \\ &\times \int\frac{3}{8}\left[\frac{3}{2}\left(1+\left[\frac{\beta+\mu}{1+\beta\mu}\right]^2\mu'^2\right)-\frac{1}{2}\left(\left[\frac{\beta+\mu}{1+\beta\mu}\right]^2+\mu'^2\right)\right](1+\beta\mu')d\mu' \\ &= \gamma^2\left(1+\frac{\beta^2}{3}\right) \end{aligned} \tag{114}$$

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

$$P_{IC} = \frac{4}{3}\gamma^2 c\, \sigma_T\, u_{ph} \tag{115}$$

where u_{ph} is the energy density of the photons the electron is colliding with. The typical output frequency is

$$\nu_{IC} = \frac{4}{3}\gamma^2 \nu_{in} \tag{116}$$

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

$$\nu L_\nu = \left(\frac{\gamma N(\gamma)}{2}\, \frac{4}{3}\gamma^2 c \sigma_T u_{ph} \right) \Bigg|_{\gamma = \sqrt{\nu/\nu_1}} \tag{117}$$

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

$$u_{CMBR} = aT_\circ^4 = \frac{4\sigma}{c} T_\circ^4 = 4 \times 10^{-13}\ \mathrm{erg/cm^3} \tag{118}$$

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:

$$\frac{(\nu L_\nu)_{IC}}{(\nu L_\nu)_S} = \left(\frac{3.3\ \mu\mathrm{G}}{B} \right)^2 \tag{119}$$

and

$$\frac{\nu_{IC}}{\nu_S} = \frac{6 \times 10^4\ \mathrm{G}}{B} \tag{120}$$

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

$$p_R(R) = \int\int \delta(R - \gamma^2(1+\beta\mu)(1+\beta\mu'))p(\mu)p'(\mu';\mu)d\mu d\mu' \qquad (121)$$

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

$$\begin{aligned}
\left\langle \left(\frac{\nu'''}{\nu}\right)^2 \right\rangle &= \int\int \gamma^4(1+\beta\mu)^2(1+\beta\mu')^2 p(\mu)p'(\mu';\mu)d\mu d\mu' \\
&= \gamma^4 \int \frac{(1+\beta\mu)^3}{2} \int \frac{(1+\beta\mu')^2}{2} d\mu' d\mu \\
&= \gamma^4 \int \frac{(1+\beta\mu)^3}{2} \left[1+\frac{\beta^2}{3}\right] d\mu \\
&= \gamma^4 \left[1+\beta^2\right] \left[1+\frac{\beta^2}{3}\right] \\
&= \gamma^4 \left[1+\frac{4}{3}\beta^2+\frac{1}{3}\beta^4\right] \qquad (122)
\end{aligned}$$

This differs from the exact result only in the β^4 term.

15. Shocks

See §11.3 in Longair.

In the frame of reference of the shock front, the upstream gas flows into the shock at $x = 0$ with speed $v_1 = U$. It flows out into the post-schock region with speed v_2. Mass conservation gives

$$\rho_1 v_1 = \rho_2 v_2 \qquad (123)$$

which looks like a momentum equation but is not. Then momentum convservation gives

$$\rho_1 v_1^2 + P_1 = \rho_2 v_2^2 + P_2 \qquad (124)$$

which looks like an energy equation but is not.

Finally the energy equation is

$$v_1(\tfrac{1}{2}\rho_1 v_1^2 + u_1 + P_1) = v_2(\tfrac{1}{2}\rho_2 v_2^2 + u_2 + P_2) \qquad (125)$$

If we take $u_1 = P_1 = 0$ for a high Mach number shock, and $P_2 = \frac{2}{3}u_2$ for a monoatomic gas behind the shock, and let $f = \rho_2/\rho_1$, then

$$\begin{aligned}
v_2 &= v_1/f \\
(1-f^{-1}) &= \frac{P_2}{\rho_1 v_1^2} \\
\frac{1}{2} &= \frac{1}{2f^2} + \frac{(u_2+P_2)v_2}{\rho_1 v_1^3} \\
&= \frac{1}{2f^2} + \frac{1}{f}\frac{u_2+P_2}{\rho_1 v_1^2} \qquad (126)
\end{aligned}$$

These have the solution $f = 4$, $P_2 = \frac{3}{4}\rho_1 v_1^2$, $u_2 = \frac{3}{2}P_2$, so $u_2 = \frac{9}{8}\rho_1 v_1^2$, since $1/2 = 1/32 + (1/4)(9+6)/8$.

15.1. Sedov Phase of a Supernova Remnant

Since the density behind the shock is $4\rho_1$, the pressure behind the shock is given by $P = \frac{3}{16}\rho_1 v_1^2$, but also $P = nkT$. So we find the temperature is

$$kT = \frac{3\mu m_p v_1^2}{16} \tag{127}$$

The internal energy per gram is then $\frac{3}{2}nkT/\rho$ and $n = \rho/\mu m_p$ so the energy per gram is $u/\rho = (3/32)v_1^2$. The velocity of the material behind the shock is $\frac{3}{4}v_1$ in the lab frame, so the kinetic energy per gram is $(9/32)v_1^2$. Thus the total energy can be written as $\frac{3}{8}v_1^2$ per gram of swept up material. The mass of the swept up material is $M_{swept} = (4\pi/3)R^3\rho_1$.

The energy contained within a expanding spherical shock wave is constant except for the loss of energy due to radiation, so one can write $\rho_1(4\pi/3)R^3\rho_1(3/8)\dot{R}^2 = E = \text{const}$. As a result, one gets a differential equation stating $\dot{R} \propto R^{-3/2}$, so $R \propto t^{2/5}$. Thus at age t, $v_1 t = (2/5)R$.

For example, Kepler's supernova remnant, with a distance of 4 kpc, is about $4'$ in diameter. It has been expanding since 1604. With a linear radius of $R = 7.2 \times 10^{18}$ cm, and $t = 1.29 \times 10^{10}$ sec, the shock velocity should be 2200 km/sec. The temperature for $\mu = 0.62$ (90% H by number, 10% He) is $kT = 6$ keV. The total energy in the supernova remnant is then given by $E \approx (4\pi/3)R^3\rho_1\ \frac{3}{8}(0.4R/t)^2 = 5(n_H \text{ in cm}^{-3}) \times 10^{49}$ ergs. A typical energy in a supernova remnant is 10^{51} ergs, also called a *foe* for 10 to the Fifty One Ergs.

15.2. Supernovae

See §13.1 of Longair

Supernovae are stellar explosions with peak luminosities as large as an entire galaxy. The energy seen in supernova remnants is typically 10^{51} ergs, but the luminous output during the supernova is about 10^{49} ergs. A typical duration within 1 magnitude of peak brightness is 2×10^6 seconds and the peak brightness is about 5×10^{42} erg/sec. This is more than 10^9 $L_\odot$ or $M_V < -18$.

Supernova models are divided into two classes:

1. Exploding white dwarfs, or

2. Core collapse.

Observed supernovae are divided into two classes as well:

1. Hydrogen-free spectra: Type I,
2. Hydrogen-rich spectra, Type II.

Many subdivisions of these types have been made, but for the most part Type I supernovae with no hydrogen in their spectra are due to exploding white dwarfs that did not have any hydrogen to start with. These supernovae are the Type Ia supernovae used to make measurements of the expansion of the Universe. Type Ia supernovae are the most otpically luminous supernovae. Type II supernovae are usually the result of the collapse of the iron core of a red supergiant star. With about 10 $M_\odot$ of hydrogen in the envelope, these naturally have a hydrogen-rich spectrum. Some supergiants manage to eject their hydrogen envelope before exploding, leading to Type Ibc supernovae which are a contaminant in cosmological studies.

The observed energy in a supernova remnant can be produced by the sudden nuclear burning of about 0.5 $M_\odot$ of carbon or oxygen into iron. This releases about 0.1% of the rest mass into energy, or $0.001 \times 0.5 \times 2 \times 10^{33} \times 9 \times 10^{20} \approx 10^{51}$ ergs. A sudden release of this energy can occur in a white dwarf, where the carbon or oxygen exists at high density but is too cool for fusion to occur. A triggering event can start fusion at one location, leading to either a deflagration (sub-sonic flame front) or a detonation (super-sonic burning in a shock front) of most of the white dwarf. The nature of the triggering event is not certain. The most accepted model in accretion-induced collapse (AIC) where slow accretion of material from a companion pushes the white dwarf past the Chandrasekhar limit, leading to higher densities which trigger fusion. An alternate model assumes a binary with two white dwarfs that spirals in due to gravitational radiation until a merger occurs. The merger leads to heating which triggers fusion. This is the double-degenerate (DD) model. Exploding white dwarf models need to use almost all the available mass as fuel to produce the observed energy in the supernova remnant.

On the other hand, core-collapse supernovae have an energy source that easily makes 100 times more energy than that seen in the supernova remnant. This energy source is the gravitational potential energy of a neutron star. With a mass of 1.4 $M_\odot$, and a radius of 10 km, the surface potential is $\phi \approx -GM/R = -0.2c^2$. This potential is deep enough that accurate calculations of neutron star structure require general relativity. But for our purposes, we can just take ϕ at the surface times the mass giving an energy of 5×10^{53} ergs. Much of this energy is radiated as neutrinos by the hot neutron star, which cools in seconds. The existence of this radiation was confirmed in SN 1987A by the detection of 20 neutrinos

in two different underground detectors from this Type II supernova in the Large Magellanic Cloud, about 50 kpc from Earth.

Based on stellar evolution models, stars lighter than 8 $M_\odot$ lose their envelopes and become white dwarfs, while stars heavier than 8 $M_\odot$ develop into core collapse supernovae. since stars lighter than 0.8 $M_\odot$ are still on the main sequence, one epxects about 20 times more white dwarfs have been formed than neutron stars for a Salpeter initial mass function, $dn/dM \propto M^{-2.35}$. But the frequency of Type Ia and Type II supernovae are comparable, and this is compatible with the requirement that a white dwarf be in a rather specific kind of binary system before it can become a supernova.

Hydrodynamic models of core collapse supernovae have had a hard time driving a strong enough shock into the envelope to lift it off leaving a neutron star and supernova remnant. $99 \pm 1\%$ of the energy comes off as neutrinos. Processes which capture a percent of the neutrino momentum flux and transfer it to the envelope are needed. A second important case is when the pressure behind the shock is dominated by radiation pressure. Then $u_2 = aT^4$ and $P_2 = \frac{1}{3}u_2$. These have a solution $f = 7$, $P_2 = \frac{6}{7}\rho_1 v_1^2$, and $\frac{1}{2} = \frac{1}{98} + 4\frac{1}{7}\frac{6}{7} = (1+48)/98$. For this case the internal energy density after the shock is $u_2 = 3P_2 = \frac{18}{7}\rho_1 v_1^2$. The kinetic energy density is $\frac{1}{2}\rho_2(\frac{6}{7}v_1)^2 = \frac{18}{7}\rho_1 v_1^2$. Thus the kinetic energy and the energy in radiation are equal after a strong radiation pressure dominated shock.

Strong radiation pressure dominated shocks are relevant during supernova explosions. So when the shock wave hits the surface of the star, it leaves behind a ball of expanding gas with a kinetic energy equal to the radiation energy. We need to understand what happens to 99 % of the radiation, since the observed light curve integrates to only 1% of the kinetic energy. We need to understand the time scale over which the light is emitted. We wil consider a simple model consisting of a rapidly expanding uniform density sphere of ionized gas, where electron scattering provides the dominant opacity.

The reason that only 1% of the energy comes out as light is that the expansion of the supernova converts the radiative energy into kinetic energy of expansion. This gives

$$-\frac{dE}{dt} = -\frac{v}{R}E \tag{128}$$

If we ignore the first doubling of the size of the supernova, we can consider the expansion velocity v to be nearly constant.

15.3. 3-D Radiative Transfer

In order to calculate the light emitted at various times, we need to define two quantities: the intensity, which is the energy per unit area-time-frequency-solid angle. Thus the intensity

I_ν has units of ergs/cm^2/sec/Hz/sr. The more familar quantity is the flux, the energy per unit area-time-frequency. This is the intensity, weighted by $\cos\theta$, integrated over solid angle. Thus the flux F_ν has units erg/cm^2/sec/Hz.

The flux follows an inverse square law, but the intensity is constant along a ray unless there is absorption or emission. The absorption per unit length is given by α_ν with units of per cm, and the emissivity is given by j_ν with units of ergs/cm^3/sec/Hz/sr. The basic equation of radiative transfer is

$$\frac{\partial I_\nu}{ds} = -\alpha_\nu I_\nu + j_\nu \tag{129}$$

which has the the solution $I_\nu = j_\nu/\alpha_\nu = S_\nu$ if α_ν and j_ν are constant as a function of path length s. The quantity $S_\nu = j_\nu/\alpha_\nu$ is called the *source function*.

Another integral over directions is the mean intensity J_ν. This is the average over solid angle of I_ν without the $\cos\theta$ factor, so $J_\nu = \int I_\nu d\Omega/4\pi$. The quantity H_ν is the flux F_ν divided by 4π.

The flux is really a vector quantity, and the scalar F used in plane-parallel problems is just the z component of it. The vector flux is given by

$$\vec{F_\nu} = \int I_\nu \hat{n} d\Omega \tag{130}$$

where $\hat{n}$ is the unit vector for the direction of the intensity. In a situation where the point at which the flux is being found is far from any boundaries, we can determine the flux from the gradient of the source function. The formal solution of the RT equations gives us

$$I_\nu(\vec{r},\hat{n}) = \int S_\nu(\vec{r}-\hat{n}s)\exp(-\alpha_\nu s)\alpha_\nu ds \approx S_\nu(\vec{r}) - \ (\alpha_\nu)^{-1}\hat{n}\cdot\vec{\nabla}S_\nu \tag{131}$$

so the flux is

$$\vec{F_\nu} = -\frac{4\pi}{3\alpha_\nu}\vec{\nabla}S_\nu \tag{132}$$

Note that for this I_ν the mean intensity is $J_\nu = S_\nu$. In most cases far from boundaries *local thermodynamic equilibrium* (LTE) is a good assumption, and this gives $S_\nu = B_\nu(T)$. If we ask for a mean opacity α_R that gives the bolometric flux

$$\vec{F} = -\frac{4\pi}{3\alpha_R}\vec{\nabla}S = -\frac{4\pi}{3\alpha_R}\frac{4\sigma T^3}{\pi}\vec{\nabla}T \tag{133}$$

then we find that α_R has to be the Rosseland mean opacity:

$$\frac{1}{\alpha_R} = \frac{\int \alpha_\nu^{-1}\frac{\partial B}{\partial T}d\nu}{\int \frac{\partial B}{\partial T}d\nu} \tag{134}$$

The condition of radiative equilibrium can be stated as $\vec{\nabla}\cdot\vec{F} = 0$.

15.4. Time-dependent Radiative Transfer

If radiative equilibrium does not apply, then the divergence of the flux determines the change in the energy density:

$$\frac{\partial u}{\partial t} = \frac{4\pi}{c}\frac{\partial J}{\partial t} = -\vec{\nabla}\cdot\vec{F} \approx \frac{4\pi}{3}\vec{\nabla}\cdot\left(\frac{1}{\alpha_R}\vec{\nabla}J\right) \tag{135}$$

Let us consider radiation scattering inside a uniform density sphere. For uniform density we can move the α through the divergence giving

$$\frac{\partial J}{\partial t} = \frac{c}{3\alpha_R}\nabla^2 J \tag{136}$$

Separation of variables gives us solutions of the form $J = \exp(-\lambda t)\mathrm{sinc}(kr)$ with $\mathrm{sinc}(x) = \sin(x)/x$. k and λ have to satisfy $\lambda = k^2c/(3\alpha)$, but what are the allowed values of k? Eddington (1926, *The Internal Constitution of Stars*) proposed a very useful approximate boundary condition for radiative transfer problems. He assumed that I_ν was constant for outgoing rays and zero for rays coming into a star. This gives the Eddington boundary condition $J = 2H = (2/3)dJ/d\tau$. This becomes

$$\frac{2}{3}\left(\frac{\cos(kr)}{r} - \frac{\sin(kr)}{kr^2}\right)\frac{1}{\alpha_R} = \frac{\sin(kr)}{kr} \tag{137}$$

at the surface of the sphere $r = R$. Let the optical depth to the center be $\tau_c = \alpha_R R$. Then we get

$$\cos(kR) - \mathrm{sinc}(kR) = 1.5\tau_c\mathrm{sinc}(kR) \tag{138}$$

As long as τ_c is large, the solutions are approximately $kR = n\pi$ for $n = 1, 2, \ldots$. Then the decay time constant of the scattering sphere is

$$\frac{1}{\lambda} = \frac{3\alpha_R}{k^2c} = \frac{3\alpha_R R^2}{n^2\pi^2 c} = \frac{3\tau_c}{n^2\pi^2}\frac{R}{c} \tag{139}$$

which is about $0.3\tau_c$ times the light travel time for the slowest decaying solution.

The energy lost to the escaping light is also the luminosity of the supernova, giving

$$-\frac{dE}{dt} = L = \frac{\pi^2}{3\tau_c}\frac{c}{R}E \tag{140}$$

The combined energy loss equation with both expansion loss and radiative loss is

$$-\frac{dE}{dt} = \left(\frac{v}{R} + \frac{\pi^2}{3\tau_c}\frac{c}{R}\right)E \tag{141}$$

The optical depth along the radius is given by $\tau_c = \frac{3M\sigma_T R}{4\pi R^3 \mu_e m_p}$ so equation reduces to

$$\begin{aligned} -\frac{dE}{dt} &= \left(\frac{v}{R} + \frac{4\pi^3\mu_e m_p c R}{9M\sigma_T}\right)E \\ &= \left(\frac{1}{t} + \frac{4\pi^3\mu_e m_p c v}{9M\sigma_T}t\right)E \end{aligned} \tag{142}$$

We can define a t_1 when the energy loss due to expansion is equal to the energy loss due to radiation. It is given by

$$t_1 = \sqrt{\left(\frac{9M\sigma_T}{4\pi^3\mu_e m_p c v}\right)} \tag{143}$$

this can be compared to the time t_0 given by the initial size of the supernova divided by the expansion velocity (after most of the radiation energy has been converted to kinetic energy), $t_0 = R_0/v$.

For a white dwarf in a Type Ia supernova, with $v = 10^4$ km/sec, and $R_0 = 10^4$ km, $t_0 = 1$ sec, and with $M = 1.4$ $M_\odot$, and $\mu_e = 2$, $t_1 = 1.2 \times 10^6$ sec. This means that only 1 part per million of the initial energy will be radiated, so Type Ia supernovae need another source of energy. For core collapse Type II supernovae, with $M = 10$ $M_\odot$, $\mu_e = 1.2$ and $R_0 = 1.5 \times 10^{13}$ cm, $t_0 = 1.5 \times 10^4$ sec, and $t_1 = 4 \times 10^6$ sec. Then a fraction about 0.003 of the initial energy gets radiated. Since the initial energy is $\frac{1}{2}Mv^2 = 10^{52}$ ergs, this is plenty to power the observed lightcurve.

The differential equation in Eq(142) can be simplified by changing variables to $y = t/t_1$. This gives

$$-\frac{dE}{dy} = \left(\frac{1}{y} + y\right) E \tag{144}$$

If we then let $E = z/y$, we get

$$-\frac{dz}{ydy} + \frac{z}{y^2} = \left(\frac{1}{y} + y\right)\frac{z}{y} \tag{145}$$

Canceling the z/y^2 from both sides gives

$$-\frac{dz}{z} = ydy \tag{146}$$

So

$$z = z_0 \exp(-\tfrac{1}{2}y^2) \tag{147}$$

Then $E = z/y = z_0 \exp(-\frac{1}{2}(t/t_1)^2)(t_1/t)$ and the luminosity is

$$L = z_0 \exp(-\tfrac{1}{2}(t/t_1)^2)[t_1/t](t/t_1^2) = (z_0/t_1)\exp(-\tfrac{1}{2}(t/t_1)^2) \tag{148}$$

Thus the luminosity is a half-Gaussian. The peak luminosity is at $t = 0$, and it decays to $1/\sqrt{e}$ of the peak in a time t_1. So the luminosity jumps abruptly from zero to the maximum luminosity. This abrupt onset of the luminosity in a core collapse supernova has been observed by Swift, which was observing a different supernova in the galaxy NGC 2770 when a new supernova went off (Soderberg *et al.*, 2008, Nature, 453, 469. arXiv:0802.1712). A bright X-ray transient was observed at shock break-out. This supernova was 2008D, a Type Ibc.

For Type Ia supernovae, an additional source of energy is needed to power the lightcurve, since energy from the shockwave is transferred to expansion within the first few seconds. This additional source is the radioactive decay of ^{56}Ni to ^{56}Co with a $1/e$ lifetime of 9 days and an energy release of 2.1 MeV, and then ^{56}Co decays to ^{56}Fe with a $1/e$ lifetime of 111 days and an energy release of 4.6 MeV. (Some of this energy is lost as neutrinos however.) Note that ^{56}Ni is formed by 14 α particles, so it can be made very rapidly by fusion involving only strong nuclear interaction. To actually make the stable isotope with $A = 56$, which is ^{56}Fe, takes two β^+ decays or electron captures requiring weak interactions.

The radioactive decays add a new energy source $P(t)$ to Eq(142), leading to

$$\begin{aligned} \frac{dE}{dt} &= -\left(\frac{1}{t} + \frac{4\pi^3 \mu_e m_p c v}{9 M \sigma_T} t\right) E + P(t) \\ \frac{dE}{dy} &= -\left(\frac{1}{y} + y\right) E + t_1 P(t) \end{aligned} \tag{149}$$

The radioactive decay power $P(t)$ starts at $t = 0$ with a large value P_0 and then falls slowly over days. The initial energy from the explosion is completely expanded away within seconds. Thus one can approximate $E = tP_0$ for times from 1 minute to a few days. The radiative loss term goes like tE, so the luminosity of a Type Ia supernova should go like t^2. This has actually been observed in SN 2011fe in M101, which was discovered by the Palomar Transient Factory on 24 Aug 2011. Nugent *et al.* (2011, Nature, 480, 344. arXiv:1110.6201) show the flux increasing like t^2 and an explosion time only 11 hours prior to the first observation.

The total radioactive decay energy in 0.5 $M_\odot$ of ^{56}Ni is 1.2×10^{50} ergs, so this energy source can easily power the observed lightcurve of Type Ia supernovae.

There is a correlation between the decay rate of a Type Ia supernova and the peak luminosity, with fast decays occuring in faint supernovae, while slowly decaying supernovae are bright. After correcting for the decay rate, one can measure the distance of a Type Ia supernova to $\pm 9\%$ using optical lightcurve data, or to $\pm 5\%$ using lightcurve data measured in the restframe near-IR. With several hundred Type Ia supernovae already measured, this is a powerful tool for measuring the expansion of the Universe.

16. Compact Objects

See §13.2 in Longair

Compact objects in astronomy are white dwarfs, neutron stars and black holes. These objects are not supported by thermal pressure, so one can talk about *cold, catalyzed matter*. This is matter at zero temperature, but also matter where the nucleons are arranged in the nuclei with the highest packing fraction, meaning that all possible fusion or fission energy

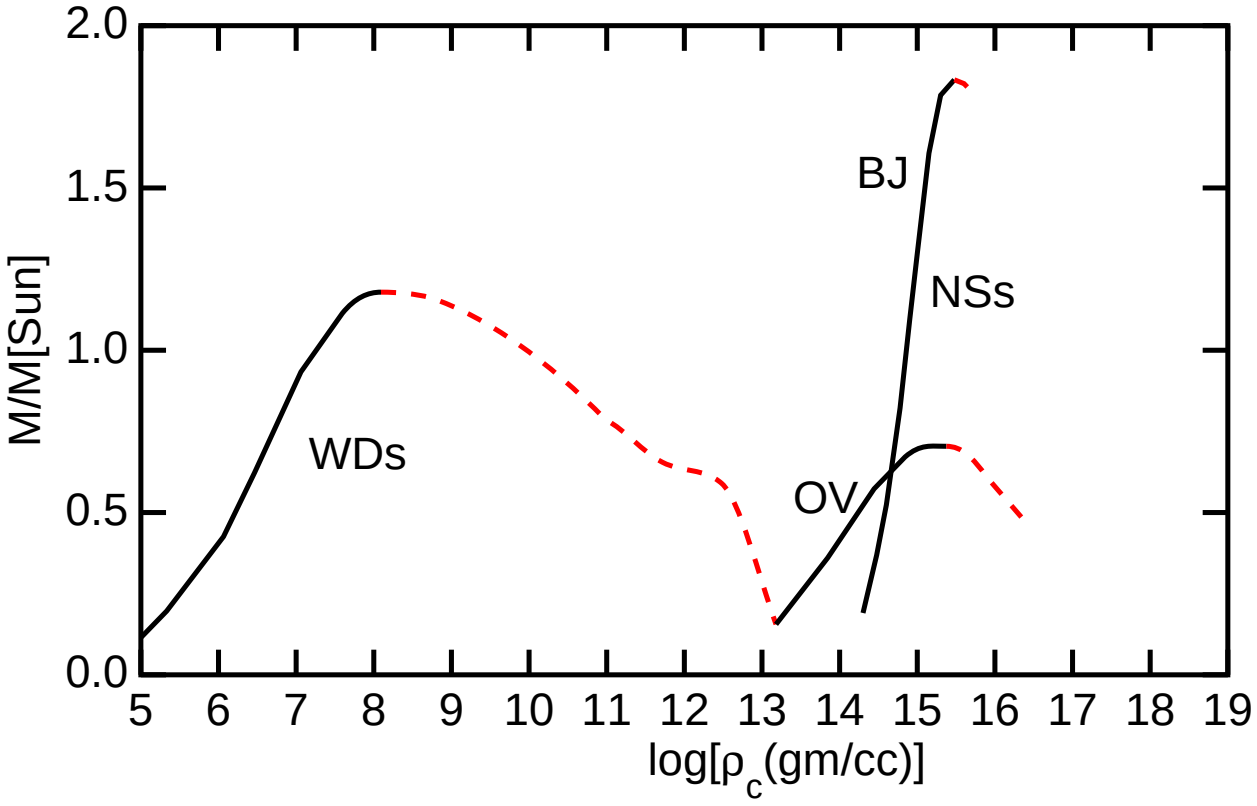

Fig. 8.— Masses of cold, catalyzed objects *vs.* central density. Adapted from Shapiro & Teukolsky (1983, *Black Holes, White Dwarfs, and Neutron Stars*). Black portions of the curve show stable objects, while the red portions are unstable. The OV (Oppenheimer-Volkoff) curve for neutron stars assumed the equation of state for free neutrons which is too soft, leading to a maximum neutron star mass that is too low. The BJ (Bethe-Johnson) curve uses a stiffer equation of state and accomodates the observed mass range for neutron stars.

has been extracted. At low densities this means ^{56}Fe, but at high enough densities the Fermi energy of the degenerate electrons gets high enough to make the electron capture $e^- + {}^AZ \rightarrow {}^A(Z-1) + \nu_e$ energetically favorable so the nuclei become more neutron rich. Ultimately the nuclei are so neutron rich that free neutrons are released and the matter changes over into neutrons giving a neutron star.

Instead of considering all these nuclear reactions we will consider two limiting cases: a degenerate electron plasma with nuclei like He, C or O that all have $Z = A/2$, and a pure degenerate neutron gas. In both cases the degenerate species has spin $\frac{1}{2}$, so the density of states up to momentum p is given by dividing phase space into cells of size h^3, and then multiplying by $g_s = 2$ to allow for the spin states:

$$n = \frac{2 \times 4\pi p^3}{3h^3} \tag{150}$$

We can write the momentum for particles of rest mass m as $p = mc \sinh\psi$, where ψ is the

rapidity, and $\beta = \tanh\psi$, $\gamma = \cosh\psi$. Thus

$$\sinh\psi_{max} = \frac{1}{mc}\left(\frac{3nh^3}{8\pi}\right)^{1/3} \tag{151}$$

Thus for large n one has $\psi_{max} \approx \frac{1}{3}\ln n +$ const. The density of the degenerate particles is given by

$$\begin{aligned}
\rho &= m\int \gamma g_s dp^3/h^3 \\
&= m(mc/h)^3 4\pi g_s \int_0^{\psi_{max}} \cosh^2\psi \sinh^2\psi d\psi \\
&= \frac{8\pi m^4c^3(\sinh(4\psi_{max}) - 4\psi_{max})}{32h^3} \\
&= \frac{8\pi m^4c^3(\psi_{max}^3 + (4/5)\psi_{max}^5 + \ldots)}{3h^3}
\end{aligned} \tag{152}$$

The density goes like ψ_{max}^3 for small ψ_{max}, but $\exp(4\psi_{max})$ for large ψ_{max}. In the case of electron degeneracy the density of the nuclei is much bigger than this, and goes like $\sinh^3\psi_{max}$. This gives a density of

$$\rho = (A/Z)m_p \frac{8\pi(m_e c\sinh\psi_{max})^3}{3h^3} \tag{153}$$

The pressure of the degenerate particles is given by

$$\begin{aligned}
P &= \int (mc\gamma\beta\cos\theta)(c\beta\cos\theta) g_s dp^3/h^3 \\
&= mc^2(mc/h)^3(4\pi g_s/3)\int_0^{\psi_{max}} \cosh^2\psi \tanh^2\psi \sinh^2\psi d\psi \\
&= mc^2(mc/h)^3(4\pi g_s/3)\int_0^{\psi_{max}} \sinh^4\psi d\psi \\
&= \frac{8\pi m^4c^5(\sinh(4\psi_{max}) - 8\sinh(2\psi_{max}) + 12\psi_{max})}{96h^3} \\
&= \frac{8\pi m^4c^5(\psi_{max}^5 + (10/21)\psi_{max}^7 + \ldots)}{15h^3}
\end{aligned} \tag{154}$$

This pressure goes like ψ_{max}^5 for small ψ_{max} but $\exp(4\psi_{max})$ for large ψ_{max}.

Thus for electron degeneracy, $P \propto \rho^{5/3}$ at low densities but switches over to $P \propto \rho^{4/3}$ when the electrons become relativistic but still do not dominate the density. For neutron degeneracy, $P \propto \rho^{5/3}$ at low densities but switches over to $P = \rho c^2/3$ when the neutrons become relativistic.

We can compare these to order of magnitude estimates of the central density and pressure, $\rho_c \propto M/R^3$ and $P_c \propto M^2/R^4$, so $P_c/\rho_c^{4/3} \propto M^{2/3}$. So when $P \propto \rho^{4/3}$, the mass is

independent of the central density. The relativistic electron equation of state leads to a constant mass, the Chandrasekhar mass limit for white dwarfs with $A/Z = 2$ at $M = 1.4$ M$_\odot$. If $A/Z = 2.15$ as for iron nuclei, then ρ_c goes up by 7.4% without changing P_c so the mass limit goes down by 15% to 1.2 M$_\odot$ for iron white dwarfs. As long as $dM/d\rho_c > 0$, the configurations are stable, since a slight compression leads to a higher density which could support a higher mass. Fig 8 shows the mass *vs.* central density plot, with the stable and unstable branches indicated. This means that the radius of stable white dwarfs goes down as the mass increases, with $R \propto M^{-1/3}$ for non-relativistic degeneracy. Since low mass neutron stars are also non-relativistic degenerate configurations, they also show $R \propto M^{-1/3}$.

For neutron stars, the mass *vs.* central density function goes through a maximum when $d \ln P / d \ln \rho \approx 4/3$. The density normalization factor $\rho_\circ = 8\pi m(mc/h)^3 = 1.8 \times 10^{16}$ gm/cc for neutrons, and the peak of $P/\rho^{4/3}$ occurs at $\rho = 1.1 \times 10^{16}$ gm/cc. General relativity also affects the maximum mass. Equation (10.1) in Longair gives a nuclear radius of $R = 1.2 \times 10^{-13} A^{1/3}$ cm, implying a nuclear density of $\rho_{nuc} = 3Am_p/4\pi R^3 = 2.3 \times 10^{14}$ gm/cc. Thus neutron stars are denser than nuclei.

For the free neutron equation of state we have used, which was also used by Oppenheimer & Volkoff (1939, PhysRev, 55, 374) in the first paper about neutron stars, this maximum occurs at $M \approx 0.7$ M$_\odot$ which is known to be too low. But we also know that nucleons interact quite strongly, leading to a much stiffer equation of state and to the liquid drop model for nuclei. The BJ curve in Figure 8 is one such stiffer equation of state. For this EoS, the radius of a 1.5 M$_\odot$ neutron star is 11.6 km.

Note that for $R = 2GM/c^2$, the order-of-magnitude density and pressure estimates are $\rho_c \approx M/R^3 \propto (c^2/G)^3/M^2$ and $P_c = GM^2/R^4 \propto (c^8/G^3)M^2$ so the pressure is of order ρc^2 but smaller densities lead to larger masses. If the equation of state is assumed to be known for densities less that 2 times nuclear densities, then has a sound speed $c_s = \sqrt{dP/d\rho}$ equal to the speed of light after that, the maximum possible neutron star mass is 2.9 M$_\odot$ (Kalogera & Baym, 1996, ApJL, 470, L61).

At lower densities, with $P \propto \rho^{5/3}$, we find that $M \propto \rho_c^{1/2}$, so lower central densities give lower masses. Thus neutron stars with low central densities are stable. Neutron stars with central densities higher than the ρ_c that gives the maximum mass are unstable on dynamical timescales, which for this density is only tens of microseconds.

More accurate masses can be found by choosing a ρ_c and then integrating the equation of hydrostatic equilibrium,

$$\begin{aligned} \frac{dP}{dR} &= -\frac{GM}{R^2}\rho \\ \frac{dM}{dR} &= 4\pi R^2 \rho \end{aligned} \tag{155}$$

from $R = 0$ with $M = 0$ until the pressure goes to zero, which defines the surface of the

object and gives the mass and radius. For neutron stars general relativity is significant so these Newtonian equations need to be modified.

16.1. Black Holes

See §13.11 in Longair

Black holes were predicted by Laplace based on Newtonian gravity, since the escape velocity is $\sqrt{2GM/R}$ and if this is greater than the speed of light then light could not escape. This defines a radius $R = 2GM/c^2$ which is 3 km for a solar mass. But this analysis does not make sense in special relativity, since light always travels at the speed of light. Thus if light could rise up to some apoapsis and then had to fall back down, the light at the apopasis would be stationary, contradicting the constancy of the speed of light.

So in general relativity, if an object is a black hole, then light cannot escape even a little bit beyond the surface of the black hole. In GR the geometry of space-time is described by a metric, such as the Schwarzschild metric

$$ds^2 = \left(1 - \frac{2GM}{rc^2}\right) c^2 dt^2 - \left[\left(1 - \frac{2GM}{rc^2}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right] \tag{156}$$

Note that the Schwarzschild metric has a coordinate singularity at $r = 2GM/c^2$, but this is due to the particular choice of coordinates and not to any fundamental singularity. The only fundamental singularity is at $r = 0$.

In GR light travels on null geodesics with $ds = 0$. One sees that in the Schwarzschild metric the path $r = 2GM/c^2, dr = d\phi = d\theta = 0$ is a null geodesic since the coefficient multiplying dt^2 vanishes, so light at just hangs at $r = 2GM/c^2$ without making any outward progress. The modern definition of a black hole is the set of points in space-time from which light cannot escape to infinity. The boundary of this set is called the *event horizon* of the black hole. Outside the event horizon there is always some directions for a light ray that will allow it to escape to infinity. Note that even outside the event horizon there are directions for light rays which do fall into the black hole.

The Schwarzschild metric has the annoying property that the space-like coordinate dr becomes time-like inside the event horizon. The Kerr (1963, PRL, 11, 237) metric avoids this problem. The Kerr metric is justly famous for applying to rotating objects, but if the spin is zero, it becomes

$$ds^2 = c^2 dt^2 - d\vec{x}^2 - \frac{2GM}{c^2|\vec{x}|}[\hat{x} \cdot d\vec{x} + cdt]^2 \tag{157}$$

In this metric the inward null geodesic always follows $dr/cdt = -1$ while the outward null

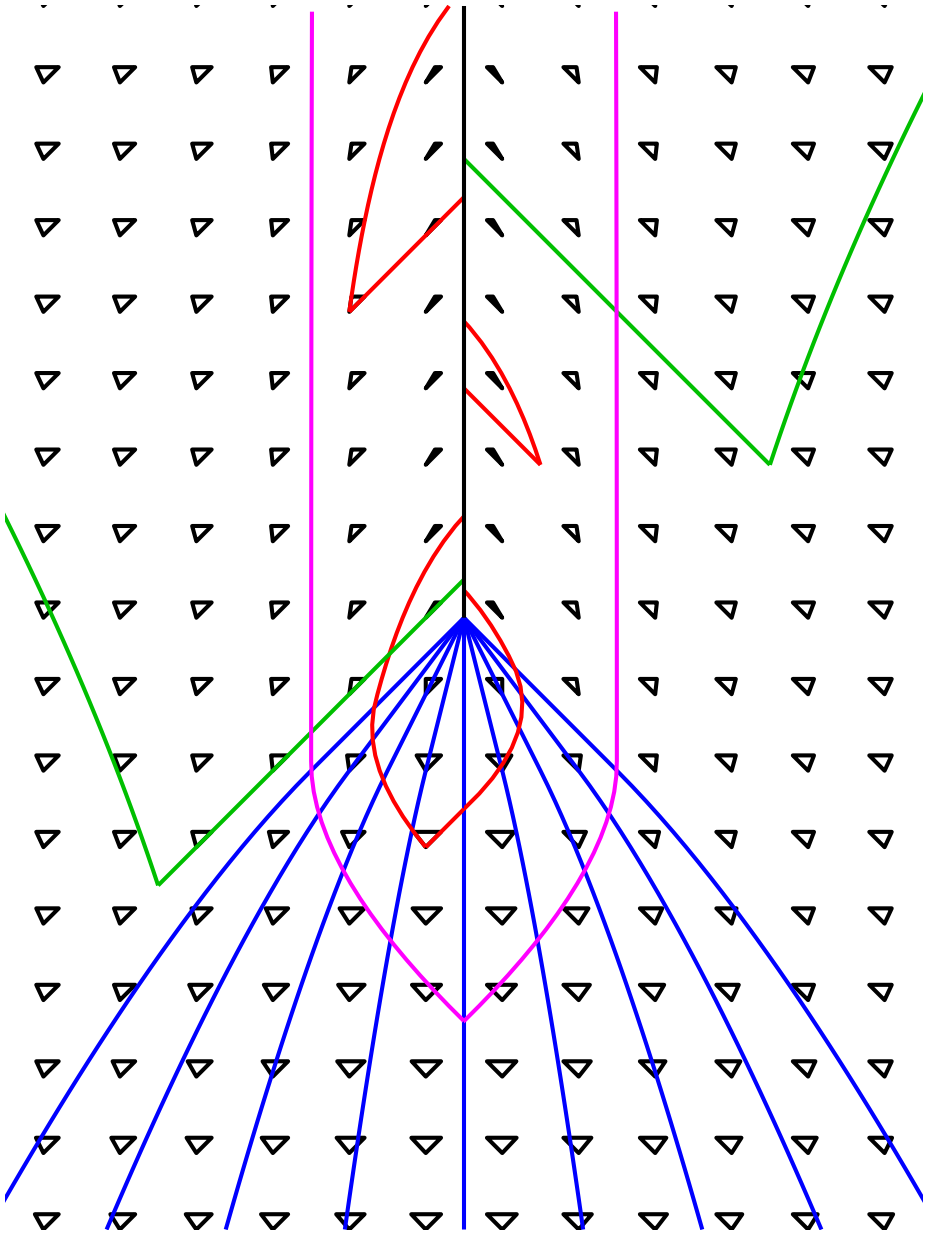

Fig. 9.— Space-time diagram showing the formation of a black hole. Blue world lines are matter collapsing to form the black hole. Magenta shows the event horizon, which forms before the central singularity. The little black triangles show infinitesimal light cones for many events, while the green light cones show outward rays escaping to infinity. The red light cones show that light emitted inside the event horizon cannot escape. Uses Kerr metric coordinates.

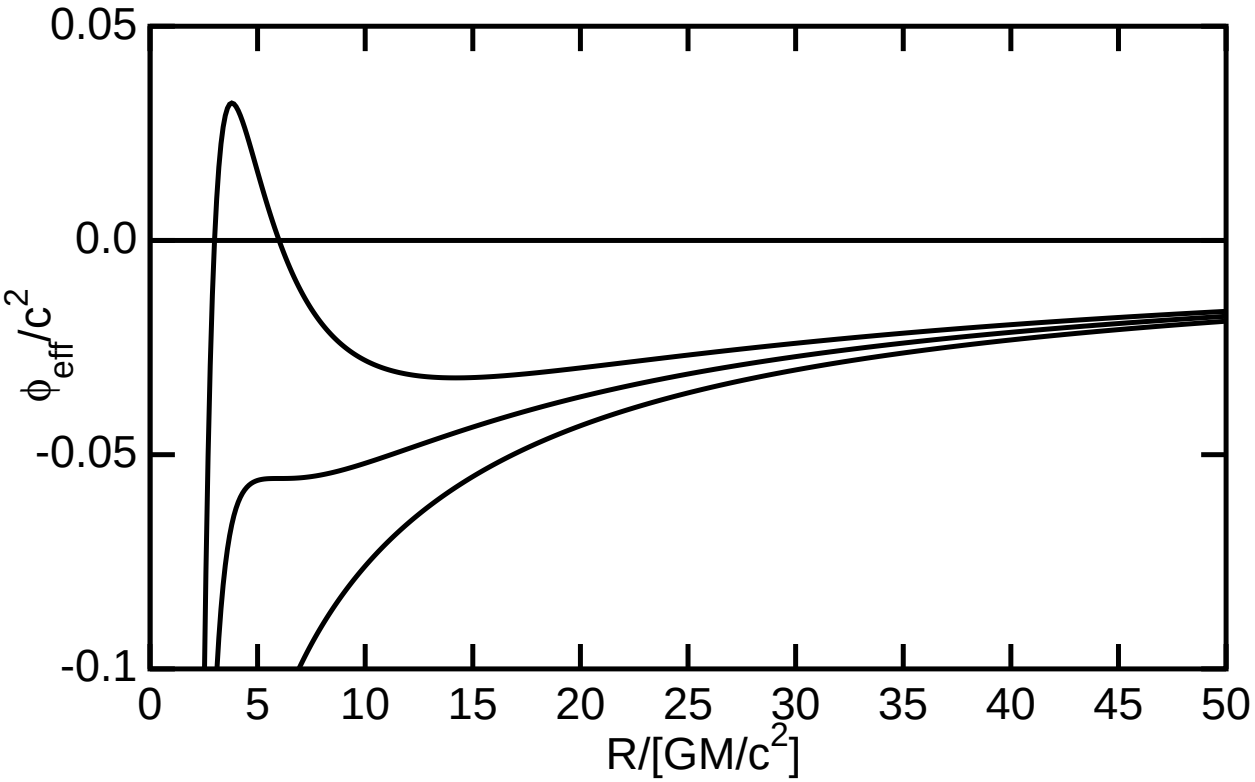

Fig. 10.— Plot of ϕ_{eff} *vs.* radius in a non-rotating black hole for values of $\ell^2 = 6$, 12 & 18 times $(GM)^2/c^2$. The top curve show a stable circular orbit, while the middle curve shows the innermost stable circular orbit at $R = 6GM/c^2$.

geodesic follows $dr/cdt = (1 - R_s/r)/(1 + R_s/r)$. There is no singularity in the metric coefficients at the Schwarzschild radius, but at the Schwarzschild radius the outward null geodesic becomes stationary, and inside the Schwarzschild radius both the inward and outward null geodesics are heading for the central singularity.

One of the significant differences between GR and Newtonian gravity is a strong attraction near the center that leads to black holes, but also causes the excess precession of the perihelion of Mercury that provided the first evidence for GR. The equation of motion for an object orbiting a mass in GR is given by

$$C_3 = \dot{r}^2 + \frac{\ell^2}{r^2} - \frac{2GM}{r} - \frac{2GM\ell^2}{c^2 r^3} \tag{158}$$

where ℓ and C_3 are constants of integration corresponding to the angular momentum per unit mass and the velocity at infinity squared. The last term on the right hand side is the difference introduced by GR. It reduces the effectiveness of angular momentum for preventing infall. In other words, it reduces the centrifugal potential. Note that C_3 can be negative for closed orbits. $C_3/2$ is the energy integral of the motion, and the effective potential for radial motion after the angular motion has been integrated to give the angular momentum is given

by

$$\phi_{eff} = \frac{\ell^2}{2r^2} - \frac{GM}{r} - \frac{GM\ell^2}{c^2 r^3} \tag{159}$$

If $\dot{r} = 0$ then

$$C_3 = \left(\frac{\ell^2}{r^2}\right)\left(1 - \frac{2GM}{c^2 r}\right) - \frac{2GM}{r} \tag{160}$$

For any value of ℓ, one can find the radius r that gives the minimum value of C_3 by solving $dC_3/dr = 0$ giving

$$\ell^2 \left(\frac{-2}{r^3} + \frac{6GM}{c^2 r^4}\right) + \frac{2GM}{r^2} = 0 \tag{161}$$

This gives

$$\ell^2 = \frac{2GM/r^2}{2/r^3 - 6GM/c^2 r^4} = \frac{GMr}{1 - 3GM/c^2 r} \tag{162}$$

Even with infinite angular momentum, circular orbits are impossible inside $r = 3GM/c^2$. But circular orbits become unstable at a larger radius, $r_{ISCO} = 6GM/c^2 = 3R_s$. ISCO stands for Innermost Stable Circular Orbit.

At this radius, $\ell^2 = 12(GM)^2 c^2$, and putting that into the expression for the effective potential gives

$$C_3 = \frac{12(GM)^2}{c^2}\left(\frac{1}{r^2} - \frac{2GM}{c^2 r^3}\right) - \frac{2GM}{r} \tag{163}$$

which has first derivative

$$\frac{dC_3}{dr} = \frac{12(GM)^2}{c^2}\left(\frac{-2}{r^3} + \frac{6GM}{c^2 r^4}\right) + \frac{2GM}{r^2} \tag{164}$$

This vanishes at $r = 6GM/c^2$ as expected.

Note that ℓ^2 is kept constant when varying r to take derivatives, even though we have a formula for ℓ^2 that is a function of r. ℓ^2 is a constant of integration so it does not vary for a given orbit.

Now we evaluate the second derivative of C_3 with respect to r, which determines the radial oscillation frequency. This is

$$\frac{d^2 C_3}{dr^2} = \frac{12(GM)^2}{c^2}\left(\frac{6}{r^4} - \frac{24GM}{c^2 r^5}\right) - \frac{4GM}{r^3} \tag{165}$$

This also vanishes at $r = 6GM/c^2$. Thus the small radial oscillations about a circular orbit, which in a Keplerian orbit have the same frequency as the angular motion, now have a zero effective restoring force, and the slightest perturbation will lead to infall into the black hole.

For a non-spinning one solar mass object, $r_{ISCO} = 9$ km, while the orbit of Mercury is at 0.4 AU or 60 million km. The excess precession of Mercury is $43''$ per century or $43/206265/(2\pi \times 36525/88) = 8 \times 10^{-8}$ radians per radian, or $\mathcal{O}(r/r_{ISCO})$.

16.2. Stationary Limit

The Kerr metric exhibits a stationary limit outside the event horizon. The stationary limit marks where a stationary world-line with $dt = 0$ becomes space-like, or faster than the speed of light. Between the stationary limit and the event horizon, objects can travel and also escape from the black hole, but only if they are orbiting around the black hole in the same sense as the spin of the black hole. The radius of the event horizon shrinks as the spin increases, with

$$r_{EH} = \frac{r_s + \sqrt{r_s^2 - 4(J/Mc)^2}}{2} \tag{166}$$

so if the angular momentum $J = 0.5Mcr_s = GM^2/c$, for a maximally spinning Kerr black hole, the radius is one-half the Schwarzchild radius. The stationary limit occurs at a radius given by

$$r_{SL} = \frac{r_s + \sqrt{r_s^2 - 4(J/Mc)^2 \cos^2\theta}}{2} \tag{167}$$

so it touches the event horizon at the poles but is further out near the equator. The region between the event horzion and the stationary limit is called the *ergosphere* because energy can be extracted from the black hole by processes that occur there. These are called Penrose processes, and they can reduce the mass of the black hole, but not below the irreducible mass given by

$$M_{ir}^2 = \frac{M^2 + \sqrt{M^4 - (Jc/G)^2}}{2} \tag{168}$$

So for a non-spinning black hole, $M_{ir} = M$. For a maximally spinning black hole, $M_{ir} = M/\sqrt{2}$. Finally, the area of the event horizon is given by $A_{EH} = 16\pi(GM_{ir}/c^2)^2$.

16.3. Beckenstein Entropy, Hawking Radiation

Since the area of the event horizon for a black hole is non-decreasing, Beckenstein put forward the hypothesis that black holes have entropy which is proportional to the event horizon area. Black holes certainly have an energy, so if they have an entropy, then they have a temperature. Hawking (1974, Nature, 248, 30) proposed that black holes actually radiate at this temperature. The temperature is proportional to the acceleration of gravity at the surface of the black hole. Unruh (1976, PRD, 14, 870) showed that one sees a temperature in an accelerating frame of reference even in flat space-time, with $kT_U = \hbar * a/2\pi c$. A non-spinning black hole has $g = GM/(2GM/c^2)^2$ at the event horizon, so

$$kT = \frac{\hbar c^3}{8\pi GM} \tag{169}$$

and $T = 61(M_\odot/M)$ nK. Using $dQ = c^2 dM$, and

$$S/k = \int dQ/kT = \int \frac{8\pi GM dM}{\hbar c} = 4\pi \left(\frac{M}{M_{pl}}\right)^2 \tag{170}$$

where the Planck mass $M_{pl} = \sqrt{\hbar c/G} = 22\ \mu$g. This entropy of a black hole is very large for super-massive black holes. Defining the Planck length $\lambda_{pl} = \hbar/M_{pl}c = 1.6 \times 10^{-33}$ cm, we find that the ratio of the event horizon area to the entropy is given by $A_{EH}/(S/k) = 4\lambda_{pl}^2$.

Since black holes less massive than the maximum neutron star mass are not going to be formed under present conditions, and any black with a low enough mass to have a moderate temperature will be much less massive than this, only *primordial black holes* can produce much energy by Hawking radiation. For a mass of 10^{14} grams, the temperature is $T = 1.2 \times 10^{12}$ K or $kT = 100$ MeV. For a mass of 10^{15} grams $kT = 10$ MeV. At this T, a blackbody will radiate photons, electrons, positrons and neutrinos. Electrons have the same number of spin states as photons, while neutrinos only have one spin state but there are 3 kinds all with antiparticles. Fermion species contribute $\frac{7}{8}$ as much as photons, so the luminosity at 10^{15} grams is

$$L = (43/8)4\pi R_s^2 \sigma_{SB} T^4 = 1.86 \left(\frac{10^{15}\ \text{gm}}{M}\right)^2 \times 10^{16}\ \text{erg/sec} \tag{171}$$

Now $dM/dt = -L/c^2 \propto M^{-2}$ so the mass *vs.* time goes like $M^3 = M_\circ^3 - \alpha t$. The final evolution is an explosion as M goes to zero.

Using the values for $M = 10^{15}$ grams gives $\alpha = 6.2 \times 10^{25}$ gm^3/sec. Putting in the age of the Universe, 13.7 Gyr, and setting the current M to zero, we find that $M_\circ = 3 \times 10^{14}$ gram primordial black holes will be exploding now. But since $kT = 33$ Mev for this mass, hot enough to radiate blackbody muons, the mass limit is slightly increased. Primoridal black hole explosions have been proposed as the sources for short GRBs but the energy released in the last second is only about 10^{30} ergs so the sources would have to be closer than 0.03 pc which is unlikely. In addition, the temperature in the last second is much hotter than the observed $E_{pk} \approx 260$ keV for GRBs.

17. 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 $E = 0.5\, I\omega^2 \approx 10^{48}(140\ \text{ms}/P)^2$ 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

$$L = -I\omega\dot{\omega} = 4\pi^2 I\dot{P}/P^3. \tag{172}$$

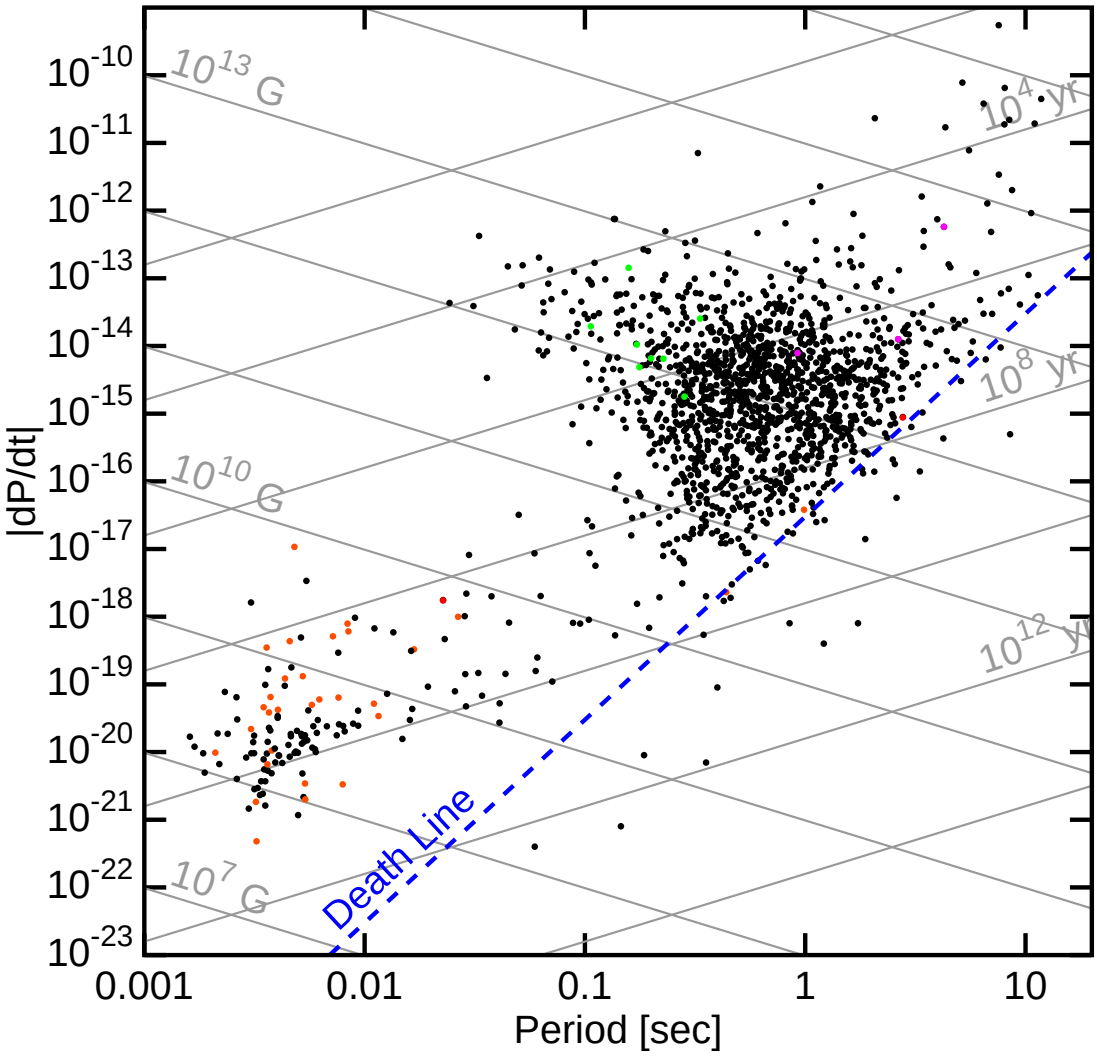

Fig. 11.— The distribution of 1743 known pulsars in the P-$\dot{P}$ plane. The orange dots have negative $\dot{P}$, probably due to gravitational acceleration in the galaxy or a globular cluster. The magenta dots are recently discovered occasional pulsars that only emit a pulse several times per day, 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 empical "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.

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 $\mu_e = 0.5\, E_p R^3$. 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

$$\frac{B_p^2\, R^6\, \omega^4}{6c^3} = -I\omega\dot{\omega} \tag{173}$$

so an estimate for the surface magnetic field is given by

$$B_p = \left(\frac{6c^3 I}{4\pi^2 R^6}\right)^{0.5} \sqrt{P\dot{P}} = 2\times 10^{12}\ \text{gauss} \left(\frac{I}{10^{45}}\right)^{0.5} \left(\frac{R}{10\ \text{km}}\right)^{-3} \left(\frac{P}{1\ \text{sec}}\ \frac{\dot{P}}{10^{-15}}\right)^{0.5} \tag{174}$$

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

$$\tau = \frac{P}{2\dot{P}}. \tag{175}$$

For the Crab nebula pulsar the characteristic age is 1240 years while the actual age is 950 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 11. 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

$$T_e = \frac{I_\nu}{2k(\nu/c)^2} = \frac{10^{-23} \times 90000 \times (3.08 \times 10^{21})^2}{2 \times \pi \times (10^6)^2 \times 1.38 \times 10^{-16}} = 10^{28} \text{ K} \tag{176}$$

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

$$\omega_c = 1.5\gamma^2 \frac{eB_{eff}}{mc} = 1.5\gamma^3\beta\frac{c}{R} \tag{177}$$

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

$$P = \frac{2e^2a^2}{3c^3} = \frac{2\gamma^4\beta^4e^2c}{3R^2} \tag{178}$$

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.

17.1. 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.

18. Accretion Powered Sources

X-ray binary stars have a mass-losing star and a compact object. The compact object can be a white dwarf, a neutron star or a black hole. The important thing about the compact object is that its gravitational potential well is very deep, so ϕ/c^2 is large (in magnitude, since ϕ itself is negative). Here ϕ is the Newtonian gravitational potential, $\phi \approx -\sum GM/R$.

18.1. Mass functions

Consider a system with masses M_1 and M_2. Let the separation between the stars be a. In the center of mass frame the stars are $a_1 = aM_2/(M_1+M_2)$ and $a_2 = aM_1/(M_1+M_2)$ away from the CoM. The orbital frequency is found using $\omega^2 a_1 = GM_2/a^2$ and $\omega^2 a_2 = GM_1/a^2$. Luckily these both give the same orbital frequency, $\omega^2 = G(M_1 + M_2)/a^3$.
Thus $a = [2\pi/P]^{-2/3}[G(M_1 + M_2)]^{1/3}$.
The deprojected radial velocity amplitude $K_2/\sin i$ of star 2 is given by

$$K_2/\sin i = \omega a_2 = \frac{aM_1}{M_1 + M_2}\frac{2\pi}{P} \tag{179}$$

were P is the period. Substituting for a and writing $M_2 = qM_1$ gives

$$K_2/\sin i = \left(\frac{2\pi}{P}\right)^{1/3} (GM_1)^{1/3}(1 + q)^{-2/3}. \tag{180}$$

This leads to the *mass function* which provides information about the mass M_1:

$$\frac{(M_1 \sin i)^3}{(M_1 + M_2)^2} = \frac{M_1 \sin^3 i}{(1 + q)^2} = \frac{K_2^3 P}{2\pi G}. \tag{181}$$

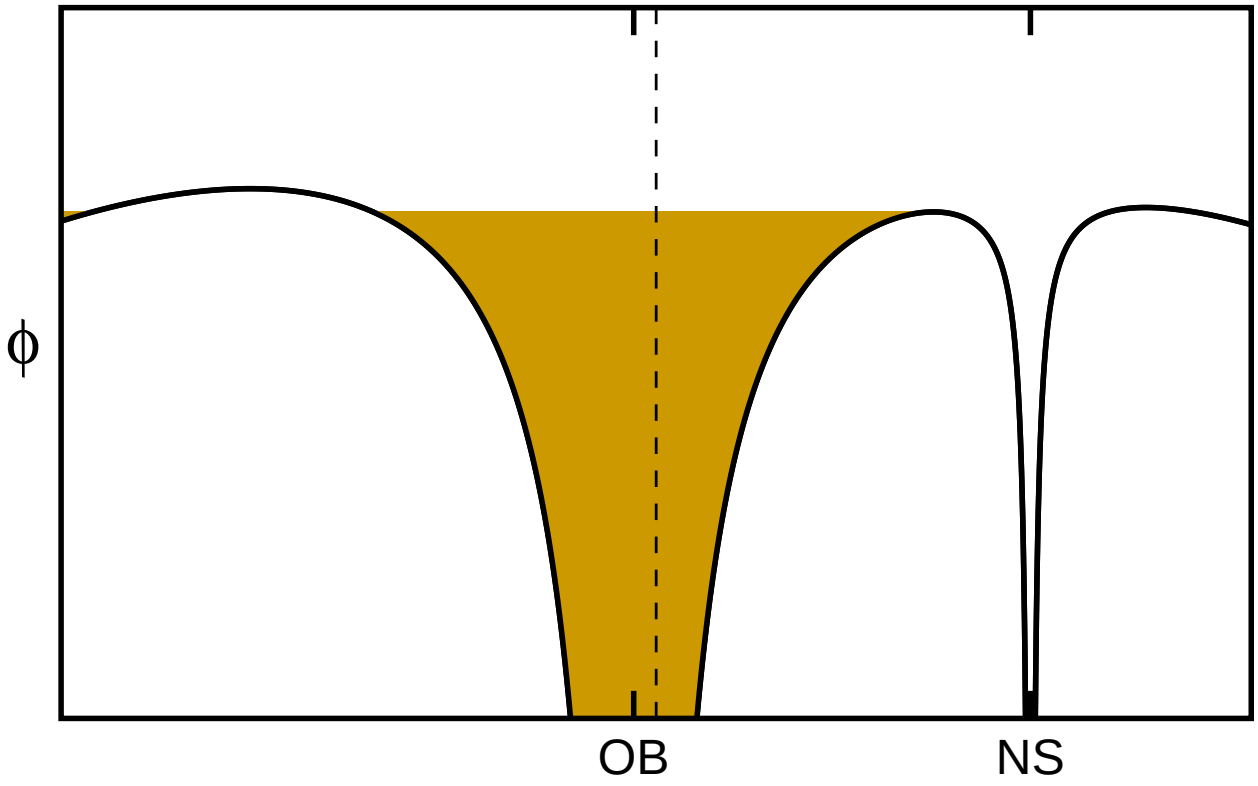

Fig. 12.— Plot of the effective potential in a rotating frame along the line between the two stars in the Cen X-3 system. A mass ratio of $q = 0.06$ is used, with the neutron star at $M = 1.21\ M_\odot$ and the mass-losing star having $20.5\ M_\odot$. This star has spectral type OB II.

If the velocity amplitude of star 1 can also be measured, then the symmetric relation

$$\frac{(M_2 \sin i)^3}{(M_1 + M_2)^2} = \frac{K_1^3 P}{2\pi G} \tag{182}$$

gives more information. For a pulsating source the timing residuals give $\omega a_n \sin i = K_n$, usually with excellent accuracy. For the high mass X-ray binary (HMXB) Cen X-3, discovered by the UHURU satellite, the accreting neutron star has a mass of of $1.21 \pm 0.21\ M_\odot$, while the mass-losing companion has a mass of $20.5 \pm 0.7\ M_\odot$. With a period of 2.087 days, the orbit radius $a = [G(M_1 + M_2)P^2/4\pi^2]^{1/3}$ is 0.089 AU.

The Roche lobe is the volume a star can fill before it starts losing mass across the saddle point at the L_1 Lagrange point. Thus the Roche lobe radius is about 10^{12} cm, considerably larger than the Sun.

For Sco X-1, neither star has a well-measured radial velocity curve, and the system does not eclipse. The best estimates are the mass of donor star $M_1 = 0.42\ M_\odot$, accreting onto a neutron star with mass $M_2 = 1.4\ M_\odot$. This is a low-mass X-ray binary (LMXB). The Roche lobe radius is about $0.32a$, and $a \approx 0.02$ AU. Thus the Roche lobe radius is 10^{11} cm, slightly larger than the Sun.

In a corotating frame revolving with frequency ω, the effective potential per unit mass on a zero velocity object at position r toward M_2 is $\phi_e = -0.5\omega^2 r^2 - GM_2/(a_2 - r) - GM_1/(a_1 + r)$. This potential is minus infinity at $r = -\infty$, $-a_1$, a_2 & $+\infty$. In between these infinities are three maxima corresponding to the L_3, L_1 and L_2 Lagrange points. We want to find the point between the two stars where $d\phi_e/dr = 0$. This defines the inner Lagrange point L_1 between the two stars and the radii of the Roche lobes, $(a_1 + r)$ around M_1 and $(a_2 - r)$ around M_2. Since $\omega^2 = G(M_1 + M_2)/a^3$, one gets

$$\frac{G(M_1 + M_2)r}{a^3} + \frac{GM_2}{(a_2 - r)^2} - GM_1(a_1 + r)^2 = 0. \tag{183}$$

Letting $x = r/a$ one gets

$$(1 + q)x + \frac{q}{[(1 + q)^{-1} - x]^2} - \left[\frac{q}{1 + q} + x\right]^{-2} = 0. \tag{184}$$

which becomes

$$(1 + q)^3 x + \frac{q(1 + q)^2}{[1 - (1 + q)x]^2} - \frac{(1 + q)^2}{[q + (1 + q)x]^2} = 0. \tag{185}$$

18.2. Post-Keplerian Parameters

When the binary contains a pulsar, extremely precise data on the orbit becomes available through pusle timing. A Keplerian orbit has 6 parameters, and the mass ratio, but usually

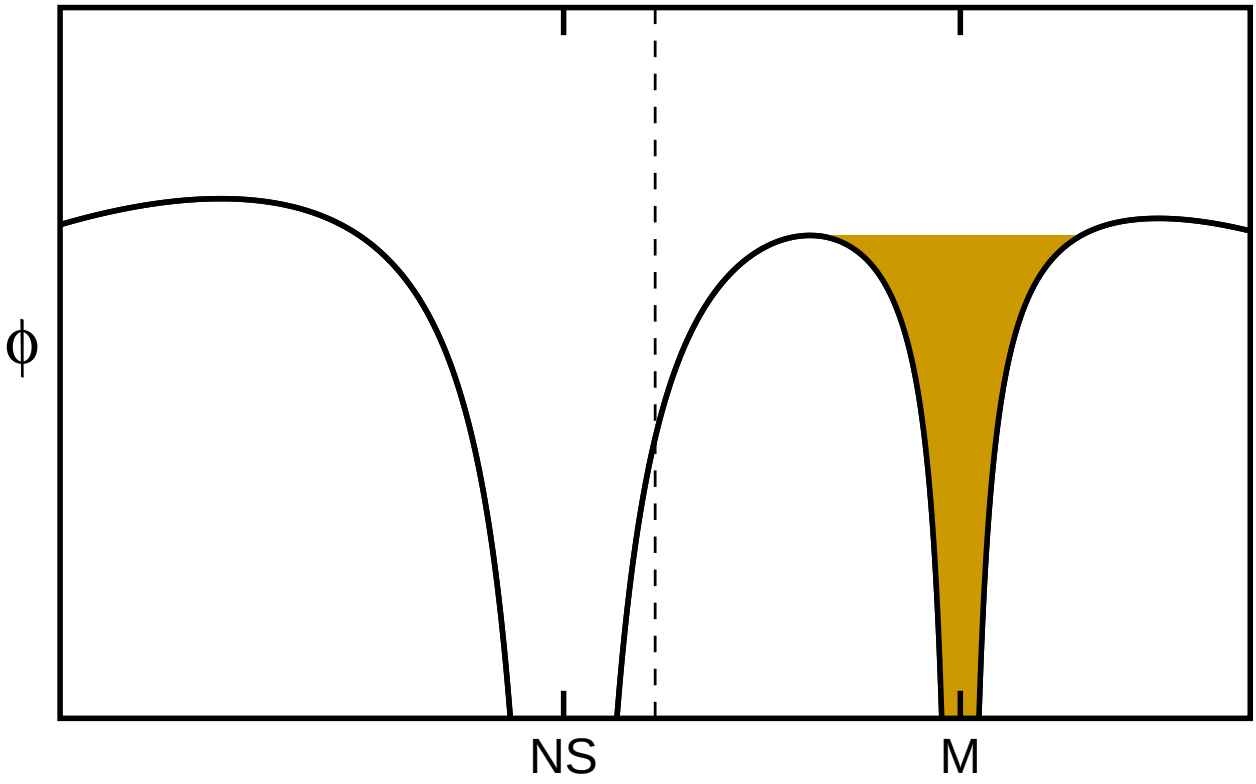

Fig. 13.— Plot of the effective potential in a rotating frame along the line between the two stars in the Sco X-1 system. A mass ratio of $q = 0.3$ is used, with the neutron star at $M = 1.4\ M_\odot$ and the mass-losing star having $0.42\ M_\odot$. This would be an M dwarf if on the main sequence, but the light from the star is all due to the hot side heated by the X-ray source.

one cannot measure all of these parameters. For non-eclipsing systems, it is hard to measure the inclination. But when the pulsar is behind the companion star, one can often measure the Shapiro time delay, which is the analog of the gravitational deflection of light. This gives a delay that goes like the logarithm of the impact parameter, so the width of the bump in the delay can give the inclination.

Measurement of the precession of the periapsis gives the depth of the potential. The precession in radians per orbit gives GM/rc^2. Given the Keplerian parameters, this gives a precise value for the sum of the masses, $M_1 + M_2$. The formula for the precession of the periapsis is

$$\frac{d\phi}{d\theta} = \frac{3G(M_1 + M_2)}{c^2 a(1 - e^2)} \tag{186}$$

radians/radian. With observations of either star one also has e and the period P. The period is a different function of $M_1 + M_2$ and a: $4\pi^2/P^2 = G(M_1 + M_2)/a^3$, so the addition of the periapsis precession to any spectroscopic binary establishes both the sum of the masses and the semi-major axis. The addition of either the mass ratio if both stars can be observed or the inclination if the system eclipses or shows the Shapiro delay then establishes the individual masses of the stars.

18.3. Gravitational Radiation

Compact binaries also radiate gravitational waves. The formula for radiation from a binary star is

$$-\frac{dE}{dt} = \frac{32G^4 M_1^2 M_2^2 (M_1 + M_2)}{5c^5 a^5} \frac{1 + \frac{73}{24}e^2 - \frac{37}{96}e^4}{(1 - e^2)^{1/2}} \tag{187}$$

For $e = 0$ this can written in terms of the quadrupole moment $Q^{ij} = \sum M x^i x^j - \frac{1}{3}\delta^{ij} \sum M r^2$. This is a traceless version of the moment of inertia tensor. For convenience put the stars along the x^1 axis, then

$$\begin{aligned} Q^{11} &= \tfrac{2}{3}\left(a^2(M_1 M_2^2 + M_2 M_1^2)/(M_1 + M_2)^2\right) = \tfrac{2}{3}\left(a^2 M_1 M_2/(M_1 + M_2)\right) \\ Q^{22} &= Q^{33} = -\tfrac{1}{2}Q^{11} \end{aligned} \tag{188}$$

and the non-diagonal terms are zero. When the binary is rotated by 45° we get $Q^{12} = \frac{1}{2}(a^2 M_1 M_2/(M_1 + M_2))$. 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 $\frac{1}{2}(a^2 M_1 M_2/(M_1 + M_2))$. Thus

$$\left\langle \sum |d^3 Q^{ij}/dt^3|^2 \right\rangle = (4 \times \tfrac{1}{2} \times 64\omega^6 \times \tfrac{1}{4}) \left(\frac{a^2 M_1 M_2}{M_1 + M_2}\right)^2 = 32\omega^6 \left(\frac{a^2 M_1 M_2}{M_1 + M_2}\right)^2 \tag{189}$$

where $\langle\rangle$ is a time average over the period. Therefore

$$-\frac{dE}{dt} = \frac{G\left\langle \sum |d^3Q^{ij}/dt^3|^2 \right\rangle}{5c^5}. \tag{190}$$

This is the gravitational wave equivalent of the Larmor formula for electric dipole radiation. 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 $M_{pl} = \sqrt{\hbar c/G} = 22$ μg 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

$$\text{Flux} = \frac{c^3}{16\pi G}\omega^2(h_+^2 + h_\times^2) \tag{191}$$

which is the correct formula for the gravitational wave flux. This flux has yet to be measured by any detector.

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 $h = 1.2\times10^{-18}$. LIGO could detect this signal from a NS-NS merger out to 10 Mpc.

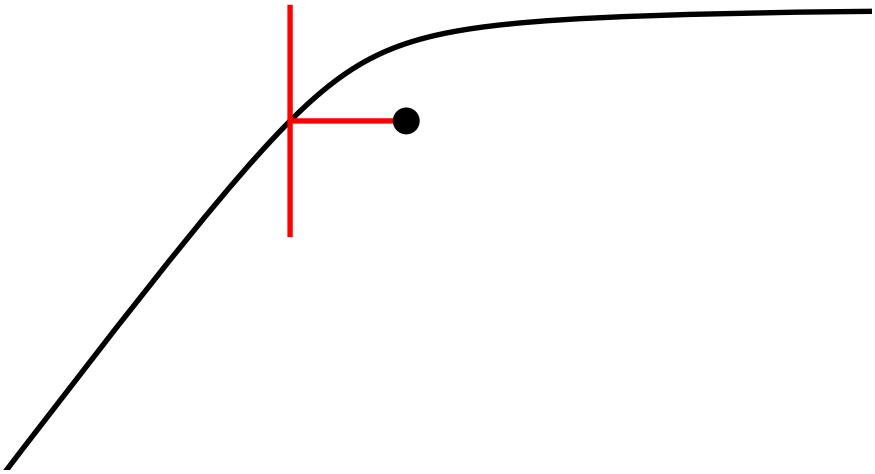

Fig. 14.— Orbit of a particle with impact parameter equal to the accretion radius, $b = 2GM/v_\infty^2$. The red line is set a distance b behind the object, showing that the wake crossing occurs at $r = b$ for this b.

19. Accretion

Black holes are pretty black but their surrounding are not. Compact objects including neutron stars and black holes usually radiate much more from the accretion disks that surround them than from their surface. For a neutron star with $M = 1.4\ M_\odot$ and $R = 10$ km, the gravitational potential is $\phi = -GM/R = -0.21c^2$. Thus material accreting onto the NS can in principle convert as much as 21% of its rest mass into radiated energy. This is much greater than the 0.7% converted into energy by fusing hydrogen to helium in a star.

There are two main modes of accretion: accretion from a wind or flow, and the accretion disk. In a wind, gas flows by a compact object on a hyperbolic orbit, and its path is bent by gravity. In a wake directly behind the object, the gas collides with gas coming from the other side, and the transverse velocity is dissipated by collisions. If the gas is now bound to the compact object it is accreted. Gas with velocity at infinity v_∞ and impact parameter b has angular momentum per unit mass $\ell = bv_\infty$, and energy per unit mass of $v_\infty^2/2$. The kinetic energy due to the transverse velocity at radius r is $[(b/r)v_\infty]^2/2$. Thus if the wake crossing radius is $< b$ the kinetic energy due to transverse motion is greater than the excess energy at infinity, so the gas is then bound after the transverse velocity is eliminated and accretes. The equation for a hyperbolic orbit in polar coordinates is

$$r = \frac{a(1-e^2)}{1+e\cos\theta} \tag{192}$$

with $v_\infty^2/2 = -GM/2a$ so $a = -GM/v_\infty^2$ and is negative. But the eccentricity e is greater than unity so the radius r is positive. The angular momentum per unit mass can be evaluated

at the periapsis with $\theta = 0$, so $r_p = a(1-e)$. The velocity is given using the energy,

$$v_p^2/2 = v_\infty^2/2 + GM/[a(1-e)] = \frac{v_\infty^2}{2}\left(1 + \frac{2}{e-1}\right) \tag{193}$$

so $v_p = v_\infty\sqrt{1+2/(e-1)}$ and $v_p r_p = b v_\infty$. Thus $b = r_p\sqrt{1+2/(e-1)}$ and

$$b = \frac{GM}{v_\infty^2}\sqrt{(e-1)^2 + 2(e-1)} = \frac{GM}{v_\infty^2}\sqrt{e^2 - 2e + 1 + 2e - 2} = \frac{GM}{v_\infty^2}\sqrt{e^2-1} \tag{194}$$

Thus $e = \sqrt{1 + (bv_\infty^2/GM)^2}$.

When $b = 2GM/v_\infty^2$, we get $e = \sqrt{5}$, the asymptotes of the hyperbola are at $\pm 116.6^\circ$. The wake crossing occurs at $\theta = 63.4^\circ$, which is 180° from the incoming asymptote, and $\cos(63.4^\circ) = 1/\sqrt{5}$ so $r = b$. Thus the cross-section for accreting material is $\pi b^2 = 4\pi(GM/v_\infty^2)^2$. If the density of the undisturbed wind is ρ, the accretion rate is $\dot{M} = 4\pi\rho v(GM/v^2)^2 = 4\pi\rho(GM)^2/v_\infty^3$.

Note that in the impulse approximation, $\Delta v = \int GMb(b^2 + v^2t^2)^{-3/2}dt = 2GM/bv = v$ if $b = 2GM/v^2$. Also the $b = 2GM/v^2$ is the distance where the velocity at infinity is the escape velocity.

We have ignored any internal pressure in the material being accreted. Internal pressure will lead to a sound speed c_s in the material. Allowing for the sound speed and pressure, the accretion rate is modified to

$$\dot{M} = \frac{4\pi\rho(GM)^2}{(c_s^2 + v_\infty^2)^{3/2}}. \tag{195}$$

The accretion process gives a luminosity $L = \eta\dot{M}c^2$, and the radiative efficiency η can be as large as 0.06 for a non-rotating black hole, 0.2 for a neutron star, or 0.4 for a rotating black hole.

19.1. Eddington Limit

The accretion process will stop if the accreting material experiences a repulsive radiation pressure larger than the attractive force of gravity. An element of material with cross-section σ and mass m at a distance R from a source of luminosity L feels a radiative force $F_r = L\sigma/(4\pi R^2 c)$. The force of gravity is $F_g = GMm/R^2$. For gravity to be larger than the radiation pressure requires that

$$L < L_{edd} = 4\pi GM(m/\sigma)c \tag{196}$$

where L_{edd} is the Eddington limit. This limit is normally evaluated for electorn scattering opacity in fully ionized gas. One gram of cosmic abundance stuff has 4.5×10^{23} hydrogen

atoms and 3.75×10^{22} helium atoms, with 5.25×10^{23} electrons having a total cross section of 0.35 cm^2, so $\sigma/m = 0.35$ cm^2/gm, making $L_{edd}/M = 3.7 \times 10^4\ L_\odot/M_\odot$. If one assumes instead that the accreting material is all helium (or any material with $A/Z = 2$) then the limit is $7 \times 10^4\ L_\odot/M_\odot$.

19.2. Eddington Time

Since accretion produces luminosity which is limited by L_{edd}, there is a maximum allowable accretion rate for a given radiative efficiency factor η. This is given by $\eta \dot{M} c^2 < 4\pi GM(m/\sigma)c$. Recognizing that only a fraction $(1-\eta)$ of $\dot{M}$ is actually added to the mass of the accreting body, we get a minimum time scale for adding mass by accretion given by

$$\tau = \frac{M}{\dot{M}} = \left(\frac{\eta}{1-\eta}\right)\frac{(\sigma/m)c}{4\pi G} = 4 \times 10^8 \left(\frac{\eta}{1-\eta}\right)\left(\frac{\sigma/m}{0.35\ \mathrm{cm}^2/\mathrm{gm}}\right)\ \mathrm{years} \qquad (197)$$

which is considered embarassingly long considering that very luminous quasars, and hence very massive black holes, are observed less that 10^9 years after the Big Bang. But since η can be small the time scale can be shorter. If $\eta = 0.06$ corresponding to the maximum efficiency for a non-rotating black hole, $\tau = 25 \times 10^6$ years allowing 30 e-foldings before redshift $z = 7.085$, the highest known quasar redshift. This is probably enough, but even faster mass growth is possible with radiatively inefficient accretion flow (RIAF) models. Note that there is no requirement that matter falling into a black hole radiate at all.

19.3. Accretion Disks

Material in an X-ray binary star that falls over the L_1 point will have angular momentum that prevents a direct infall into the compact object. Material in a disk galaxy will have angular momentum that keeps it from falling into the central supermassive black hole. As a result, the accreting material goes into orbit around the compact object. Elliptical or inclined orbits lead to collisions between gas elements in the accreting material, so it will settle into a fairly flat disk with material in circular orbits. This is called an accretion disk. The angular momentum per unit mass in a circular orbit is $\ell = \sqrt{GMr}$ at radius r, so in order for the material to accrete there must be frictional torques in the accretion disk, leading to heating and radiation.

A very simple but approximate way to think about an accretion disk is to consider the energy lost in a ring between r and $r + dr$, which is $dE = (GM\dot{M}/r^2)drdt$. But half of this energy goes into the kinetic of a circular orbit. If the other half is radiated as a blackbody with area $dA = 4\pi r dr$ (note this includes both sides of the disk) then one gets $\sigma_{SB}T^4 = GM\dot{M}/(8\pi r^3)$. Thus the temperature of the disk goes like $T \propto r^{-3/4}$.

Since $\dot{M} \propto M$ due to the Eddington limit, and the minimum radius $r_{min} \propto M/\eta$, we find that the maximum temperature $T_{max} \propto M^{-1/4}$. We have $r_{min} = 0.5\eta^{-1}GM/c^2$, and $\dot{M} = (\eta c^2)^{-1}(L/L_{edd})4\pi GM(m/\sigma)c$ so one gets

$$T_{max} = \left(8\frac{(L/L_{edd})\eta^2(m/\sigma)c^5}{2GM\sigma_{SB}}\right)^{1/4} = 2.4\times 10^7 \left(\frac{\eta}{0.1}\right)^{1/2}\left(\frac{M_\odot}{M}\right)^{1/4} \text{ K} \qquad (198)$$

Thus stellar mass black holes in X-ray binaries get hot enough to radiate thermal X-rays from the inner edge of the accretion disk, but super-massive black holes in quasars will primarily radiate ultraviolet light.

The spectrum radiated by an optically thick accretion disk can be calculated exactly by adding up appropriately weighted blackbodies. But a quick order of magnitude approximation comes from noting that an energy $\nu L_\nu \propto GM\dot{M}/r$ is radiated at $\nu \approx 4kT/h \propto r^{-3/4}$. Thus $\nu L_\nu \propto \nu^{4/3}$ and the spectrum follows $L_\nu \propto \nu^{1/3}$ up to a cutoff at $\nu \approx 4kT_{max}/h$.

19.4. Geometrically Thin

If the disk is radiating like a blackbody, then one can show that it is geometrically thin. Take the potential energy per hydrogen atom near the disk,
$-GMm_p/\sqrt{r^2+z^2} \approx -GMm_p/r + 0.5GMm_pz^2/r^3$.
Setting the z^2 term to $kT(r) \propto r^{-3/4}$ gives $z^2 \propto r^{9/4}$, so $z \propto r^{9/8}$. Hence the disk flares out slowly, with z/r gradually increasing with r. The actual value for a $1M_\odot$ object at $r = 10GM/c^2$, with T as given above, is given $(m_pc^2/20)(z/r)^2 = kT_{max}$ so $(z/r) = \sqrt{(2\text{ keV})/(47\text{ MeV})} = 0.007$.

19.5. Hot or Two Temperature Disks

There is no requirement that the disk radiate like a blackbody. If the disk is optically thin, it can be much hotter. Furthermore, the electrons in the disk do the radiating, while the protons get most of the energy input, and the coupling between the proton temperature and the electron temperature is often quite weak.

Often accreting binary sources have two states, a cool thermal state with a spectrum like the optically thick disk, and a non-thermal hard state that is associated with relativistic jets, synchrotron and inverse Compton emission.

19.6. α-Disks

The simple blackbody analysis above gives a reasonable estimate for temperature of an accretion disk but can tell us nothing about the mass of the disk or how long it takes material to spiral in to the compact object. To estimate these quantities one needs to consider the viscosity of the disk material. The molecular viscosity is tiny, leading to extremely large Reynolds numbers. The Reynolds number is a ratio of dynamical forces to viscous forces. If the Reynolds number is large, the flow becomes turbulent, which creates a large effective viscosity. If the disk is ionized, the shearing velocity in the disk will lead to the amplification of magnetic fields leading to a magnetohydrodynamic (MHD) instability and a large effective viscosity. Shakura & Sunyaev (1973) proposed a simple model for these effective viscosities where the shear stress (the off-diagonal components) of the stress tensor) are a constant fraction α of the pressure (the diagonal component of the stress tensor).

At the inner edge of the accretion disk at r_*, the material falls freely into the compact object. This is especially true for black holes where r_* is the innermost stable circular orbit (ISCO). The torque carried in the accretion disk thus vanishes at r_*. Moving outward, the torque increases as $\Delta\tau = \dot{M}\sqrt{GMr}$, so the torque at radius r is $\tau(r) = \dot{M}\sqrt{GM}(\sqrt{r} - \sqrt{r_*})$ since the angular momentum per unit mass is $\ell = \sqrt{GMr}$. The power dissipated into a ring is given by $\tau\Delta\Omega$ where $\Omega = \sqrt{GM/r^3}$. There is no dissipation for solid-body rotation. Thus $dP = (3/2)GM\dot{M}r^{-2}(1 - \sqrt{r_*/r})dr$. The integrated power from r_* to ∞ is $0.5GM\dot{M}/r_*$, which is the same as the simple model discussed above, but the inner edge is cooler and the outer disk is hotter. The effective temperature of the disk is

$$T_{eff} = \left[\left(\frac{3GM\dot{M}}{8\pi\sigma_{SB}r^3}\right)\left(1 - \sqrt{\frac{r_*}{r}}\right)\right]^{1/4} \tag{199}$$

The maximum dissipated power per unit area occurs for $r = (49/36)r_*$ and is a flux of $6.8 \times 10^{-3} \times GM\dot{M}/r_*^3$.

If one assumes an isothermal equation of state with $P = (kT/\mu m_p)\rho$, then the vertical integral of the pressure is related to Σ, the surface density of the disk by $\int P dz = (kT/\mu m_p)\Sigma$. The torque at a given radius is related to the shear stress by $\tau(r) = r \times 2\pi r \int P_{shear} dz$ so in an α-disk we get $\tau(r) = 2\pi r^2 \alpha (kT/\mu m_p)\Sigma$. Thus the product $\alpha\Sigma$ is specified, and low α corresponds to a more massive disk and vice versa. If we take $\tau \propto r^{1/2}$, and the blackbody temperature $T \propto r^{-3/4}$, we get $\alpha\Sigma \propto r^{-3/4}$. The radial velocity v_r is set by $-v_r\Sigma(2\pi r) = \dot{M}$. To have a constant $\dot{M}$ one needs $v_r \propto r^{-1/4}$. This gives a residence time that goes like $r/v_r \propto r^{5/4}$. Material spends the most time at large radius and then passes quickly through the small radii on its way to colliding with the compact object. For a given $\dot{M}$, a disk with high α and lower Σ will have a shorter residence time.

19.7. QPOs

Quasi-Periodic Oscillations (QPOs) have been observed in many sources. These differ from true pulsations by having a varying frequency and phase, leading to a low Q. If the period of these is the orbital period in the ISCO, then the frequency is $\nu = 2700(M_\odot/M)$ for a Schwarzschild black hole. Sco X-1 has shown QPOs at 1100 Hz. Sco X-1 is unlikely to be a 2.5 $M_\odot$ black hole, and it is probably a 1.4 $M_\odot$ neutron star with an inner edge of the accretion disk at about $R = 16$ km.

19.8. Fate of the Accreted Material

Material accreted onto a black hole just disappears through the event horizon. But material accreted onto white dwarfs or neutron stars can lead to explosive nuclear burning. Hydrogen rich layers on the surface of a white dwarf will burn explosively in a *nova* explosion if the accretion rate is correct. Novae do not lead to the total destruction of the white dwarf, and appear to be recurrent with periods of hundreds to thousands of years. Burnable nuclear fuel on the surface of neutron star leads to *X-ray bursts*.

20. Gamma Ray Bursts

See §22.7 in Longair

Gamma ray bursts (GRBs) are short bursts of "γ-rays", where the quotes indicate that the energy is closer to 100 keV which is a dentist's X-ray, and much lower than the energies observed by γ-ray telescopes like Fermi, HESS or VERITAS.

GRBs were discovered by the Vela satellites that were launched to monitor the nuclear test-ban treaty, since nuclear explosions in space would produce short bursts of hard X-rays. While the Vela satellites were classified, the existence of GRBs was announced by Klebesadel, Strong & Olson (1973, ApJL, 182,L85). There were four Vela satellites almost equally spaced in a circle with a geocentric radius of 1.2×10^5 km (0.4 light-seconds) and by using the light travel time across the constellation of satellites approximate locations could be derived. The durations were 0.1 to 30 seconds, and the time integrated flux or *fluence* $\mathcal{F} = \int F dt$ in 0.2-1.5 MeV was $10^{-5} - 2 \times 10^{-4}$ erg/cm^2. Klebesadel *et al.* reported 16 bursts in 3 years. The Vela GRBs appear to follow a Euclidean source count law, with the number of sources brighter than a fluence $\mathcal{F}$ going like $N(> \mathcal{F}) \propto \mathcal{F}^{-3/2}$. This would apply to sources uniformly distributed in Euclidean space. They could be extragalactic or they could be in the solar neighborhood.

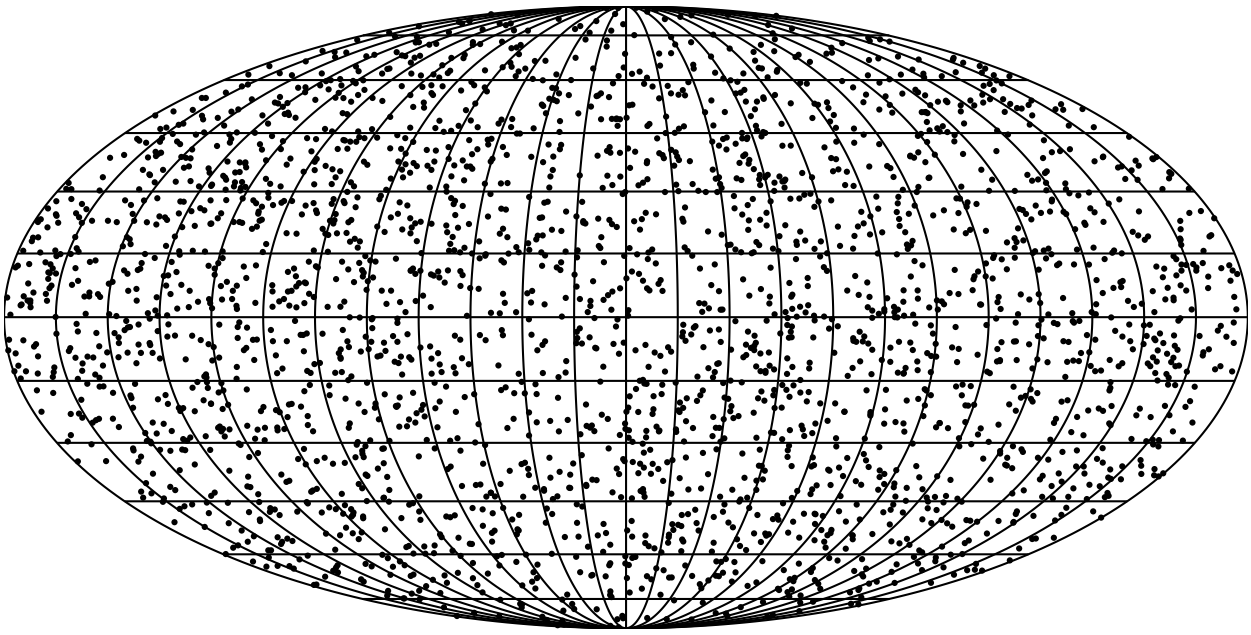

Fig. 15.— Map in galactic coordinates of the positions of 2702 GRBs observed by the BATSE experiment on the Compton Gamma Ray Observatory.

Starting in 1977, GRB detectors flew on several interplanetary spacecraft. The long baselines allowed the determination of very accurate positions using time delays. These data showed that the bright GRBs are isotropically distributed. No obvious progenitors or remnants were seen in the error boxes for the positions. But it took weeks to months to collect the data from all the spacecraft before the accurate position could be calculated.

Also in the 1970s, balloon-borne detectors showed that bursts with small fluence like 10^{-7} erg/cm^2 were much less frequent that the Euclidean source count model prediction. This break meant that an edge had been reached. It could be the transition from a uniform distribution in th solar neighborhood to a disk distribution in the Milky Way. In a disk, the volume within D grows like $D^2 \propto 1/S$, so $N(>\mathcal{F}) \propto \mathcal{F}^{-1}$. But if the faint GRBs were coming from a flat disk, they would not be isotropic.

Over 150 theoretical models for GRBs were created. The one with the least energetic mechanism is probably the relativistic BB model by Grindlay & Fazio (1974, ApJL, 287, L93) which proposed that an iron grain of radius 0.1 cm approaching the Earth with $\gamma = 1000$ will see Sun light blueshifted to keV X-rays, and will break up 100's of AU from the Sun. The cloud of iron atoms will scatter sunlight back toward the solar system with energies increased by a factor $\approx \gamma^2$ giving MeV γ-rays. The angular spread in $1/\gamma$, so the total energy for a burst with fluence 10^{-5} erg/cm^2 is only $10^{-5} \times (300 \text{ AU} \times 1.5 \times 10^{13} \text{ cm/AU}/\gamma)^2 = 2 \times 10^{20}$ ergs. The kinetic energy of a 0.1 cm radius iron sphere with $\gamma = 1000$ is 2×10^{22} ergs.

The most energetic models considered objects at cosmological distances where the flux

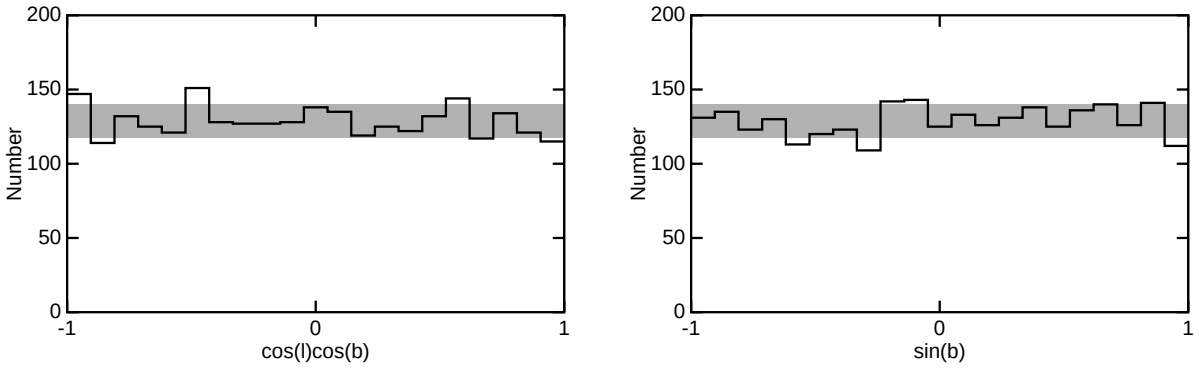

Fig. 16.— Histograms showing the concentration toward the Galactic Center on the left, and toward the galactic plane on the right. In both cases, there is not any concentration. The grey band shows the expected mean ± one standard deviation for each histogram. In both cases, 6 out of 21 bins are outside the ± 1 σ range, consistent with the expected 6.7 ± 2.6 bins.

is given by $F = L/(4\pi D_L^2)$ where D_L is the luminosity distance. Since $dt_{obs} = (1+z)dt_{em}$ the total energy radiated assuming isotropic emissions is given by

$$\mathcal{F} = \int F dt_{obs} = \frac{(1+z)\int L dt_{em}}{4\pi D_L^2} = \frac{(1+z)E_{iso}}{4\pi D_L^2} \tag{200}$$

Cosmological models require energies $E_{iso} \approx 4\pi(c/H_\circ)^2 \times 10^{-5} = 5 \times 10^{51}$ ergs. This is a huge amount of energy to come out in 300 keV photons.

The launch of the Compton Gamma Ray Observatory in 1991 provided crucial data to cut down most of these models. Its BATSE instrument had 8 large area flat panels arranged like the corners of a cube. Using the flux ratios among the panels simultaneously detecting a burst, an approximate position accurate to $\pm 5^\circ$ could be derived. The area was large enough to detect faint bursts, and BATSE clearly showed both the change in $N(>\mathcal{F})$ from $\mathcal{F}^{-3/2}$ for bright bursts to a shallower slope for faint bursts, and the isotropy of the faint bursts. This information provided strong evidence for cosmological models for GRBs.

The BATSE GRB burst catalog has positions for 2702 GRBs, and they are consistent with isotropy. There is a slight quadrupole caused by the CGRO observing pattern. Otherwise there is neither a large scale anisotropy associated by the Milky Way nor small scale anisotropy associated with source clustering.

If one computes the dipole component toward the Galactic Center for the BATSE GRB catalog, which is given by $D_x = \langle \cos\theta \rangle = [\sum \cos(l)\cos(b)]/N$, one gets -0.0104. The standard deviation for one burst would be $\sqrt{0.5}$, so for N bursts it is smaller by a factor of $\sqrt{N}$, giving $\sigma = 0.0136$. So if the GRBs came from a spherical halo centered on the Milky Way, the size would have to be greater than twenty times the 8 kpc distance of the Sun from the GC, or a radius bigger than 170 kpc. If there is a halo around the Milky Way, there should also be

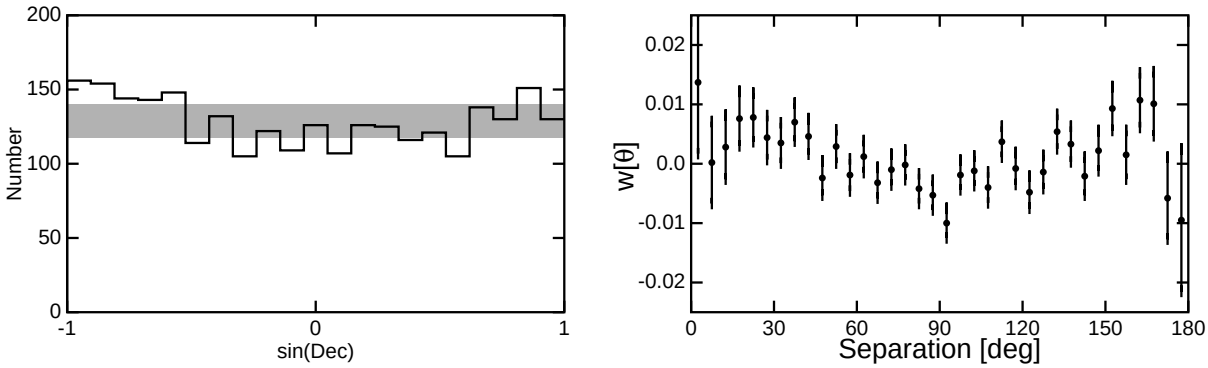

Fig. 17.— Left: histogram showing number of BATSE GRBs vs declination. There is significant effect caused by the scan pattern. Right: the angular correlation function of the BATSE bursts. The only significant effect is due to the scan pattern.

one around M31, but no concentration of GRBs is seen toward M31. So the isotropy of the BATSE GRB catalog requires a cosmological origin for the GRBs.

Figure 19 shows the number of GRBs *vs.* fluence from the BATSE catalog. Cosmological models predict that the break in $N(>\mathcal{F})$ occurs at about $z = 0.25$. There were about 50 GRBs in 10 years at the break, and the volume out to $z = 0.25$ contains about 35×10^6 L_* galaxies. Thus there are approximately 10^{-7} GRBs per L_* galaxy per year.

20.1. General Description of Anisotropy

Looking for concentrations toward specific axes is not a completely general procedure. But a general description is possible. If the number density of sources on the sky is $n(l, b)$ with average $\overline{n}$ then one can write the *two-point correlation function* as

$$w(\theta) = \left\langle \left(\frac{n(l,b)}{\overline{n}} - 1 \right) \left(\frac{n(l',b')}{\overline{n}} - 1 \right) \right\rangle \tag{201}$$

with

$$\cos\theta = \sin b \sin b' + \cos b \cos b' \cos(l - l') \tag{202}$$

so θ is the angle between (l, b) and (l', b').

For point sources like GRBs, the number density is a sum of delta functions, $n(l, b) = (4\pi)^{-1} \sum \delta^2((l, b) - (l', b'))$. Then $w(\theta)$ can be calculated by dividing the observed number of pairs in a given bin of θ by the expected number of pairs, and then subtracting one. This is shown on the right side of Figure 17 where with $N = 2702$ bursts, the $N(N-1)/2$ burst separations were binned into 5° wide bins, and then divided by the expected number which is $N(N-1)\Delta\cos\theta/4$ to give $1 + w$. The errorbars one the plot are computed as $1/\sqrt{N_{pair}}$.

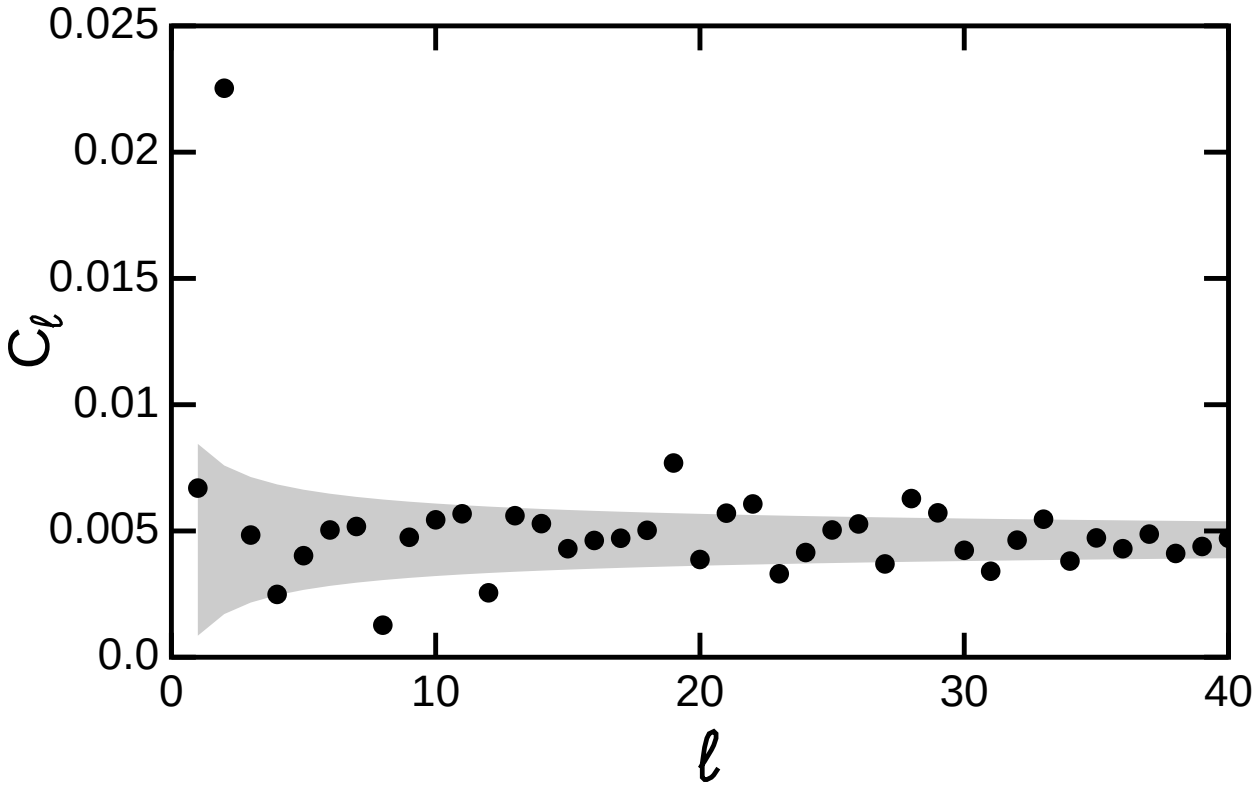

Fig. 18.— The angular power spectrum of the BATSE GRB catalog.

One can also expand the density contrast into spherical harmonics,

$$\frac{n(l,b)}{\overline{n}} - 1 = \sum a_{\ell m} Y_{\ell m}(l,b) \tag{203}$$

By the orthonormality of the spherical harmonics we can write

$$a_{\ell m} = \int \left(\frac{n(l,b)}{\overline{n}} - 1 \right) Y^*_{\ell m}(l,b) d\Omega \tag{204}$$

Because the density contrast has zero mean by construction, $a_{00} = 0$.

If we are looking for a cosmological effect that predicts a statistically isotropic Universe, then

$$C_\ell = \langle |a_{\ell m}|^2 \rangle \tag{205}$$

is independent of the azimuthal quantum number m. $a_{\ell m}$ has zero mean, and the standard deviation can be found by realizing that the $Y_{\ell m}$, sampled at a random location, has a variance of $1/4\pi$. The mean density $\overline{n}$ is $N_b/4\pi$. So $a_{\ell m}$ is the sum of N_b samples each with variance $(4\pi/N_b)^2/4\pi = 4\pi/N_b^2$. Hence $a_{\ell m}$ has variance $4\pi/N_b$ and this variance is the expected value of C_ℓ, and the standard deviation of C_ℓ is given by $\sigma(C_\ell)/C_\ell = \sqrt{2/(2\ell+1)}$. Figure 18 shows that the angular power spectrum of the GRB catalog is consistent with isotropy except for $\ell = 2$ which is affected by the anisotropic observation time pattern with more coverage at the celestial poles.

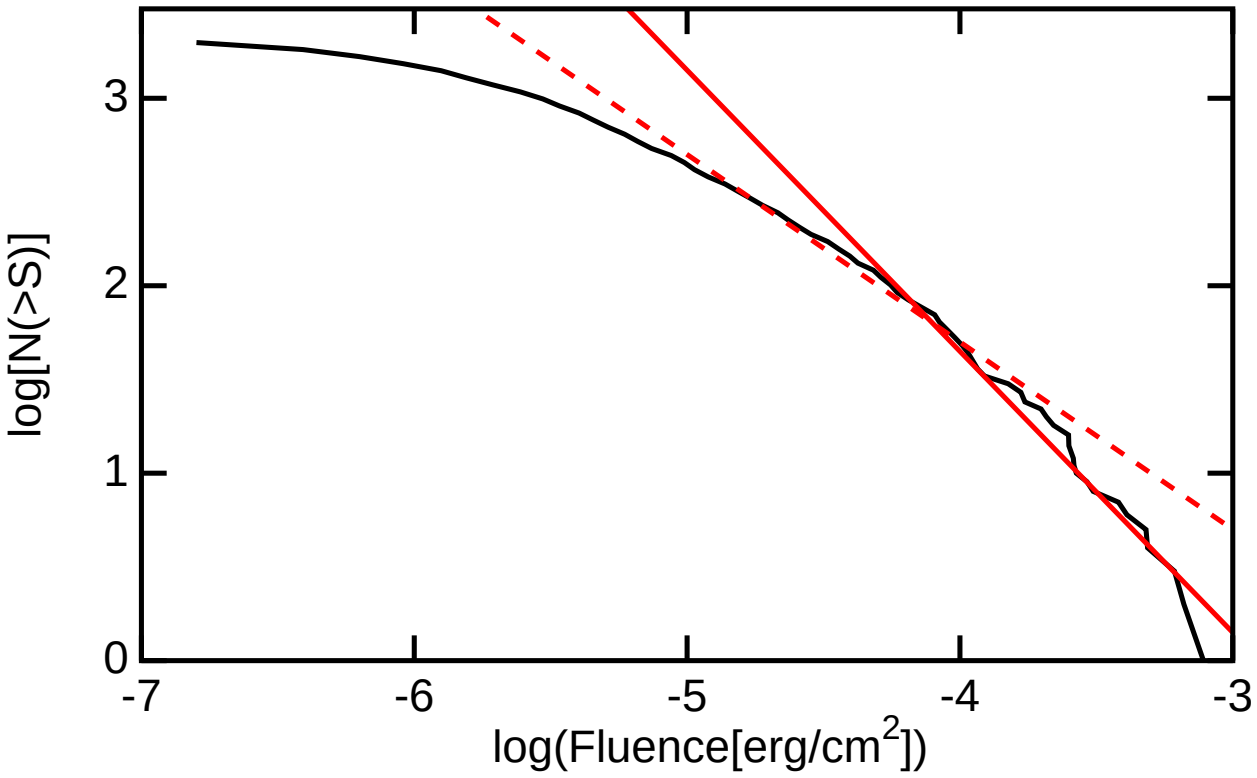

Fig. 19.— The number of GRBs *vs.* fluence for BATSE GRBs. The solid red line shows a Euclidean source count with $N(>\mathcal{F}) \propto \mathcal{F}^{-3/2}$, while the dashed line shows $N(>\mathcal{F}) \propto \mathcal{F}^{-1}$.

20.2. Discovery of Counterparts

The BeppoSAX satellite had an X-ray telescope which it could point at GRB positions. It also carried a Wide-Field Camera (WFC) with a coded-aperture system covering $40^\circ \times 40^\circ$ on the sky. When the WFC saw a GRB, the X-ray telescope was pointed at the position, leading to the discovery of fading X-ray transients (FXT) following GRBs. GRB 970228 was seen optically as well by a large optical telescope on the ground. After the fading optical transient (FOT) had faded completely, a very faint galaxy was seen at its position. Then GRB 970508 was localized rapidly, enabling a spectrum to be taken of the fading optical transient, which revealed a redshift of $z = 0.835$.

The GRB 980425 was followed by a coincident core collapse supernova, SN 1998bw, showing that GRBs comprise a subset of Type II supernovae.

GRB 090423 was seen as a fading infrared transient, and the location of a break in the spectrum gives a redshift of $z = 8.2$ if it is due to Lyman-α.

In 2011, a claim was made that GRB 090439B had a redshift of $z \approx 9.4$ (Cucchiara *et al.*, arXiv:1105.4915). The fading infrared transient was much brighter at 1.6 and 2.2 μm

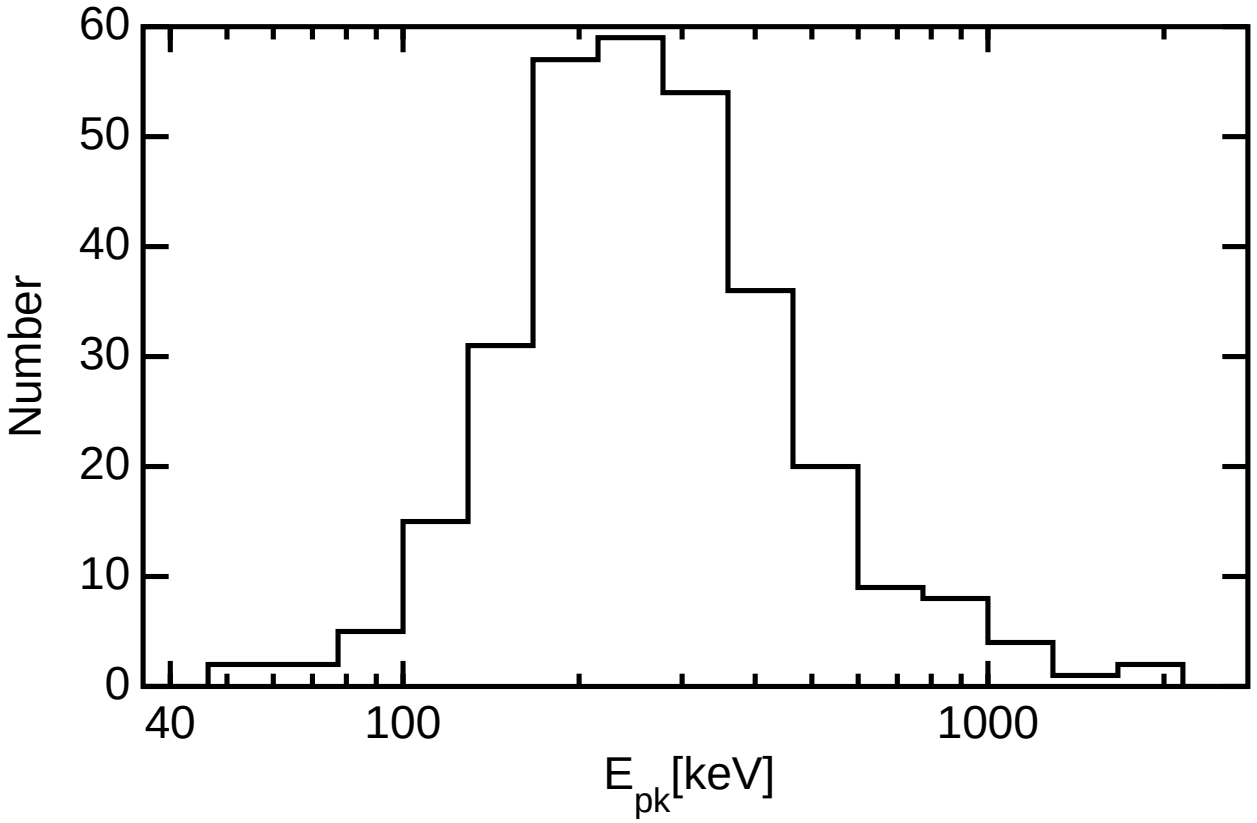

Fig. 20.— Histogram of E_{pk} from the fits by Kaneko *et al.* (2006).

than at 1.25 μm, suggesting that the Lyman-α break was redward of 1.25 μm, leading to the claimed high redshift.

20.3. Spectra of GRBs

Band *et al.* (1993, ApJ, 413, 281) found that GRB spectra are well fit by a function that at low energies was a power with an exponential cutoff, but at high energies was just a steeper power law. This Band function for the photon flux, $N_E(E)$ in photons/cm^2/sec/keV, is specified by 4 parameters, the low energy power law index α, the exponential cutoff energy $E_\circ$, the high energy power law β, and the overal flux normalization A. The function is

$$N_E(E) = \begin{cases} A\left(\frac{E}{100\ \mathrm{keV}}\right)^{\alpha} \exp\left(-\frac{E}{E_\circ}\right), & E < (\alpha-\beta)E_\circ; \\ A\left[\frac{(\alpha-\beta)E_\circ}{100\ \mathrm{keV}}\right]^{\alpha-\beta} \exp(\beta-\alpha)\left(\frac{E}{100\ \mathrm{keV}}\right)^{\beta}, & E \geq (\alpha-\beta)E_\circ. \end{cases} \tag{206}$$

This function is constructed to be continuous with continuous derivative at the break energy $E_{break} = (\alpha-\beta)E_\circ$. Kaneko *et al.* (2006, ApJS, 166, 298; astro-ph/0605427) has plots of spectral fits to BATSE bursts and histograms of the parameters. The median low energy photon spectral index α is -1.07, while the median high energy photon spectral index β is

≈ -2.48. Therefore the peak energy per octave, $\nu F_\nu \propto E^2 N_E(E)$, usually occurs in the part of the spectrum below the break at $E_{pk} = E_\circ/(2+\alpha)$. The median of the distribution of E_{pk} shown in Figure 20 is 263 keV.

20.4. Beaming and Spectrum

The very large inferred GRB luminosity for cosmological distances is a good argument for beaming. If the material emitting the γ-rays is moving with bulk Lorentz factor γ then the radiation will be beamed with a HWHM of about $1/\gamma$ radians. Then the total energy emitted in γ-rays during the burst is not E_{iso} but γ^2 smaller. This makes the energy comparable to the energies associated with a supernova. There are approximately 10^{-2} supernovae per L_* galaxy per year, so if the bulk $\gamma = 300$, there could be a γ-ray burst associated with every core collapse supernova. It is more likely that only a subset of core collapse supernovae emit GRBs. One reasonable model is that very massive stars collapse to a Kerr black hole instead of a neutron star, and it could rapidly form an accretion disk, which could create relativistic jets. Observers situated along the jet axis would see a GRB associated with a supernova, while observers over the other 99.999% of possible directions would only see the supernova.

20.5. Short bursts

A distinct subset of GRBs are short, with durations less than 2 seconds. These bursts appear to be less luminous, but sitll cosmological, and may be due to merging NS-NS binaries.

21. Cluster of Galaxies

An important class of X-ray sources are clusters of galaxies like the Coma cluster, or the members of the Abell Catalog. These vast assemblages of matter have very deep potential wells, and the galaxies in the potential well move with very high velocities through the cluster. The velocity dispersion of the galaxies is mirrored by the velocity dispersion of ions in a hot plasma that fills the cluster.

When we count the stars in the solar neighborhood, and add up both their masses and their luminosities, we get a mass to luminosity ratio of 3 in solar units in the V band. The mass to luminosity ratio of clusters of galaxies is about 300. This is mainly dark matter. The ratio of dark matter to baryonic matter is well measured by CMB experiments like WMAP to be 4.97 ± 0.17. Thus 250 out of the 300 is dark matter. Of the remaining 50, only 5 is

in stars. Since the galaxies in clusters have older stars and little star formation, the M/L ratio will be higher. That still leaves 90% of baryons unaccounted for. These baryons are contained in the intracluster medium between the galaxies.

When a patch with a density excess collapses, the linear density contrast is small. It is exactly 1.69 for an Einstein-de Sitter model. After the collapse, a very simple model gives a density contrast of 177. Thus clusters are usually defined by a radius with a given density contrast, such as R_{200} which is defined by

$$N\rho_\circ = \frac{\int_0^{R_N} \rho(r) 4\pi r^2 dr}{(4\pi/3)R_N^3} \tag{207}$$

where $\rho_\circ$ is the average density of the Universe. If we consider different clusters with different radii, then we know that the density within R_N is always the same, so the mass scales likes $M \propto R_N^3$, so the potential depth is $GM/R \propto R_N^2$. This potential depth is proportional to $\mu m_p kT$, so $T \propto R_N^2 \propto M^{2/3}$.

The X-ray luminosity is proportional to $Mn_e\sqrt{T}$ when integrated over energy or $Mn_e \exp(-E/kT)/\sqrt{T}$ for a given band. Since the density contrast and hence the density are similar for all clusters, we expect to see $L_X \propto M^{4/3}$.

One of the good things about X-ray emitting clusters of galaxies is that all of the baryons are radiating and thus visible. The baryons in stars provide the light from the galaxies, while the other 90% of the baryons are hot enough to radiate X-rays.

Masses of clusters of galaxies can be measured many ways. The mass of baryons can be measured because all of the baryons are visible in clusters of galaxies: the stars in the galaxies plus the hot gas in the intracluster medium can both be observed. To observe the hot gas one uses X-ray observations. Here one observes an angular size θ and an X-ray flux F_X and electron temperature T_e. The distance of the cluster is given by $D = cz/H_\circ$. The X-ray flux is given by

$$F_X \propto \frac{(\theta D)^3 n_e^2 \sqrt{T_e}}{D^2} \tag{208}$$

We can solve this for

$$n_e \propto F_X^{1/2} D^{-1/2} T_e^{-1/4} \theta^{-3/2} \tag{209}$$

and then compute the baryonic mass

$$M_b \propto n_e (\theta D)^3 \propto F_X^{1/2} D^{5/2} T_e^{-1/4} \theta^{3/2} \tag{210}$$

In most analyses to date, the electron density is assumed to follow an *isothermal β-model*, with

$$n_e(r) = n_e(0) \left(1 + \frac{r^2}{r_c^2}\right)^{-3\beta/2}. \tag{211}$$

This is a fairly crude approximation, and it is definitely not the gas density that would result in hydrostatic equilibrium in a Navarro, White & Frenk (1997, ApJ, 490, 493-508) density profile for the dark matter. The NFW profile is

$$\rho(r) = \frac{\delta_c \rho_{crit}}{(r/r_s)(1 + r/r_s)^2} \tag{212}$$

Here r_s is the radius at which the density variation changes from a r^{-1} dependence to a r^{-3} dependence. In this density law the enclosed mass varies like r^2 for small r, so the gravitational acceleration g is constant. The potential then is shaped like a conical pit, and the density of isothermal gas in hydrostatic equilibirum would also show a conical peak instead of a smooth quadratic variation with radius r. Inaccuracies like this have in the past led to factor of two discrepancies among various cluster mass determinations but more careful analyses have reduced these problems.

The X-ray data can also be used to compute the total mass of the cluster. Hydrostatic equilibrium says that $dP/dr = -\rho g$ which gives

$$\frac{n_e k T_e}{\theta D} \propto \frac{n_e \mu G M_t}{(\theta D)^2} \tag{213}$$

where μ is the mass per electron. This determines the total mass

$$M_t \propto (kT_e/\mu)\theta D. \tag{214}$$

Then the baryon fraction is given by M_b/M_t or

$$f_b \propto F_X^{1/2} D^{3/2} \theta^{1/2} T_e^{-5/4}. \tag{215}$$

LaRoque, Bonamente, Carlstrom, Joy, Nagai, Reese[1] & Dawson (2006, ApJ, 652, 917-936) find the X-ray determined baryon fraction is $f_b = 0.119(70/H_\circ)^{3/2}$.

The Sunyaev-Zeldovich effect allows one to measure $y = \tau_e k T_e / m_e c^2$. This gives another way to solve for n_e:

$$n_e \propto \frac{y}{T_e \theta D}. \tag{216}$$

and

$$M_b \propto n_e (\theta D)^3 \propto= y^1 T_e^{-1} (\theta D)^2. \tag{217}$$

Finally

$$f_b \propto y^1 T_e^{-2} \theta^1 D^1. \tag{218}$$

LaRoque *et al.* find the S-Z effect gives $f_b = 0.121(70/H_\circ)$. Note that these values are determined for the volume inside r_{2500}, where r_{2500} is defined as the radius within which

[1]UCLA grad

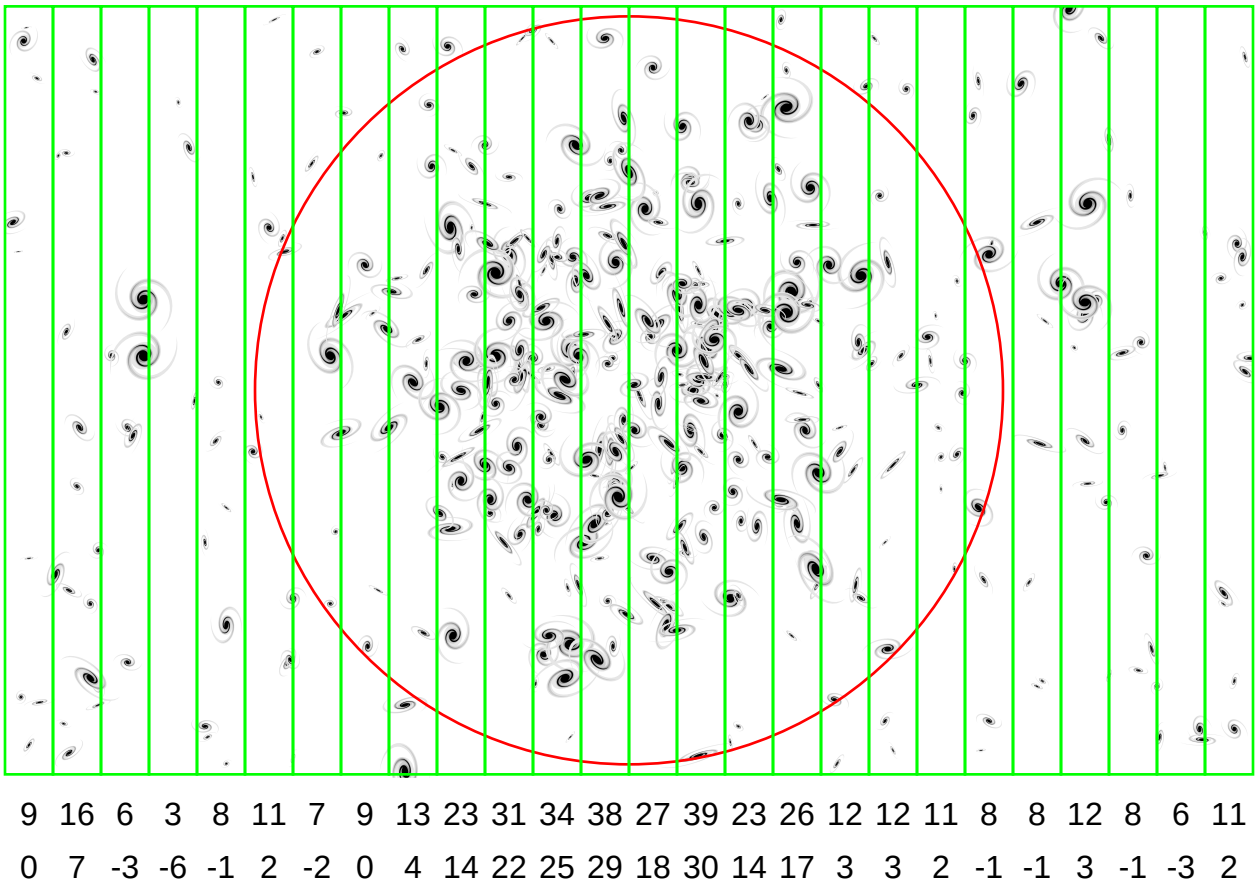

Fig. 21.— A schematic cluster with strip counts, background subtracted strip counts, and R_e computed using Eqn(252) shown by the circle.

the mean density is $\overline{\rho} = 2500 \times \rho_{crit}$. Recall that in the Press-Schechter model a cluster is virialized when $\overline{\rho} = 177 \times \rho_{crit}$ so this is just the dense central core. Within this core f_b is only 0.68 ± 0.10 times the cosmic Ω_b/Ω_m from the CMB.

With two ways to find n_e that scale differently with distance D, one can find D by requiring consistency:

$$\begin{aligned} n_e &\propto y^1 T_e^{-1} \theta^{-1} D^{-1} \\ &\propto F_X^{1/2} D^{-1/2} T_e^{-1/4} \theta^{-3/2} \\ D &\propto y^2 F_X^{-1} T_e^{-3/2} \theta. \end{aligned} \tag{219}$$

Given the distance and the redshift one can find the Hubble constant from the S-Z effect data, and Bonamente, Joy, LaRoque, Carlstrom, Reese & Dawson (astro-ph/0512349) determine $H_\circ = 77 \pm 10$ km/sec/Mpc.

The ESA mission *Planck*, launched in 2009, should provide a wealth of S-Z data on thousands of clusters.

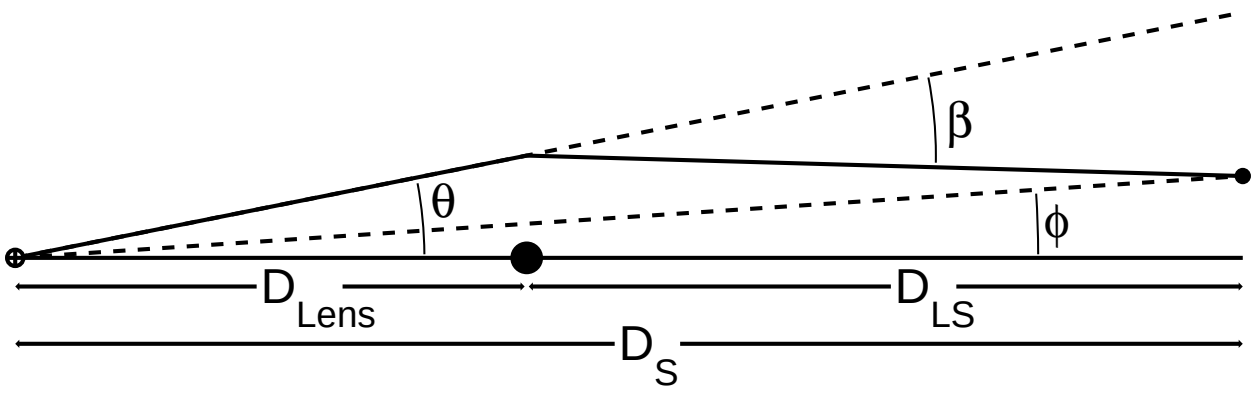

Fig. 22.— Definitions of the apparent position θ, the deflection β, and the undeflected positions ϕ.

One of the most useful methods for obtaining masses of systems is the *Virial* Theorem. This states that the total kinetic energy of a bound system is minus one-half of the total potential energy:

$$\mathrm{KE} = -\frac{1}{2}\mathrm{PE} \tag{220}$$

Remember that the potential energy of a bound system is negative so the minus sign gives a positive kinetic energy. This gives

$$(3/2)M_t\sigma^2 \propto (1/2)\frac{GM_t^2}{R_e} \tag{221}$$

where σ is the radial velocity dispersion and R_e is the effective radius that gives PE $= -GM^2/R_e$. This assumes an isotropic velocity distribution, so $\langle v^2\rangle = 3\sigma^2$. Therefore

$$M_t = \frac{3\sigma^2\theta_e D}{G}. \tag{222}$$

This is equivalent to the total mass equation from X-ray data with $\sigma^2 \propto kT_e/\mu$.

Note that the effective radius R_e is quite large. For a cluster with a Gaussian profile, R_e is $\sqrt{2\pi/\ln 2} = 3.01$ times the half-density radius.

For the Plummer model with $\rho \propto [1+(r/b)^2]^{-5/2}$ the effective radius is $R_e = (32/3\pi)b = 3.395b$. In this model the half-density radius is $0.565b$, so R_e is 6.01 times the half-density radius. Fig. 21 shows how R_e is considerably bigger than the half-density radius of a cluster.

Gravitational lensing also provides a way to measure masses. The deflection of light passing by a mass at impact parameter b is given by $\beta = 4GM/bc^2$. If the deflections are small, then it is not important where the mass is along the line of sight, so the mass of a cluster can be collapsed into a sheet, giving a surface mass density $\Sigma(x, y)$. This is the *thin lens approximation* and it is very accurate for clusters of galaxies.

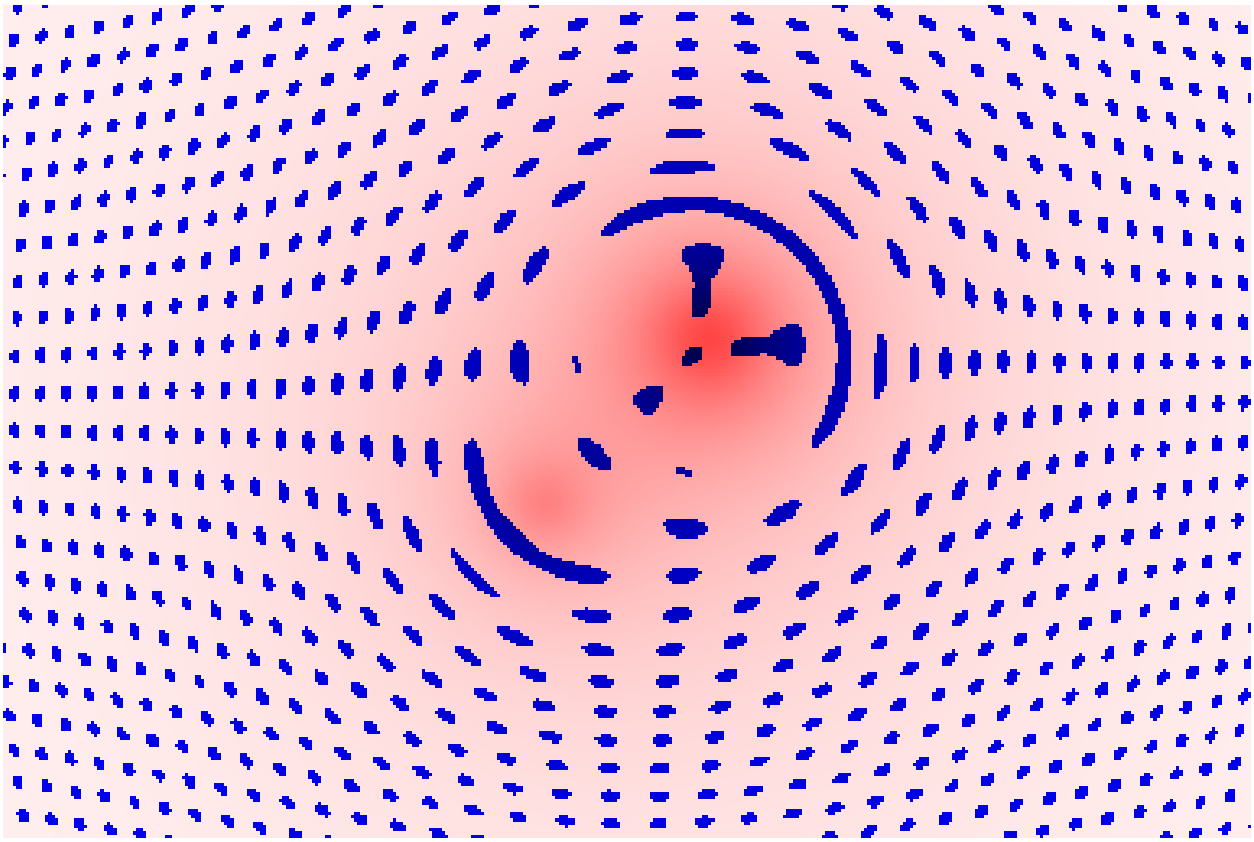

Fig. 23.— An asymmetric foreground cluster with surface density shown by the pink shading, and a regular array of circular background galaxies in blue.

The deflection produced by a uniform ring of mass vanishes for a light ray passing through the ring. The closer distance to the near side of the ring is exactly canceled by the greater amount of mass on the far side of the ring. In addition, for a ray passing outside the ring the deflection is the same as it would be for all the mass collapsed to a point in the center. Thus if the mass distribution collapsed into a sheet is circularly symmetric, one gets an exact formula for the mass enclosed within a cylinder of radius b around the line of sight:

$$M(<b) = \frac{\beta(b) b c^2}{4G} \tag{223}$$

The deflection β occurs at the lens and is not the same as the image displacement observed at the Earth. The impact parameter b is clearly given by $b = D_A(z_L)\theta$ where θ is the observed angular separation between the lens and the apparent position of the source. I will use D_{Lens} as a shorthand for the angular size distance to the lens, $D_A(z_L)$. The actual radius of the source at the source plane is given by $D_S\theta - \beta(D_{Lens}\theta)D_{LS}$, or an undeflected angular position of $\phi = \theta - \beta(D_{Lens}\theta)D_{LS}/D_S$, where $D_S = D_A(z_S)$ is the angular size distance to the source, and D_{LS} is the angular size distance of source seen from the lens. This is computable from the usual formulae for D_{Lens} and D_S using

$$D_{LS} = [R_\circ/(1+z_S)]S[S^{-1}([1+z_S]D_S/R_\circ) - S^{-1}([1+z_L]D_{Lens}/R_\circ)], \tag{224}$$

where $S() = \sin()$ for closed Universes with $k = 1$ or $S() = \sinh()$ for open Universes with $k = -1$, and $R_\circ = (c/H_\circ)/\sqrt{|1 - \Omega_{tot}|}$. This can also be written as

$$D_{LS} = \frac{R_\circ}{1 + z_S} S\left(\int_{1/(1+z_S)}^{1/(1+z_L)} \frac{cda}{R_\circ a\dot{a}} \right). \tag{225}$$

In non-flat Universes $D_S \neq D_{Lens} + D_{LS}$. This violation of Euclidean triangle rules can be used to measure the curvature of space, even with the very long skinny triangles involved in lensing. Figure 22 shows the definitions of these angles.

For a given lens and source redshift there is an *Einstein ring radius* θ_E such that the actual source position is on-axis when the apparent position is at θ_E. This is found by solving

$$D_S\theta_E - \beta(D_{Lens}\theta_E)D_{LS} = 0. \tag{226}$$

Clearly $\beta = \theta_E D_S/D_{LS}$ for the Einstein ring.

In *strong* lensing one identifies multiple images of a source which are separated by angles comparable to θ_E. Using other data such as flux ratios it is easy to determine θ_E. Then the mass contained within a cylinder of radius $b = D_{Lens}\theta_E$ is

$$M(< D_{Lens}\theta_E) = \beta bc^2/4G = \theta_E^2(D_{Lens}D_S/D_{LS})c^2/4G. \tag{227}$$

In *weak* lensing one observes the shapes of background galaxies more distant from the lens than the Einstein ring radius. The tangential dimension of a background source is increased by a factor θ/ϕ, the radial dimension is multiplied by $\partial\theta/\partial\phi$. But since the deflection usually decreases for sources more distant from the lens, $\partial\theta/\partial\phi$ is usually less than one and sources are compressed radially. Thus the average ratio of tangential size to radial size is greater than one. This ratio is given by $\langle a/b\rangle = 1 + 2\langle\gamma_T\rangle$ where γ_T is called the tangential shear. For a point mass lens, $\phi = \theta - \theta_E^2/\theta$, so

$$\begin{aligned} \langle a/b\rangle &= \frac{\theta/\phi}{\partial\theta/\partial\phi} \\ &= \frac{\theta}{\theta - \theta_E^2/\theta}\frac{d\theta + \theta_E^2/\theta^2 d\theta}{d\theta} \\ &= \frac{1 + \theta_E^2/\theta^2}{1 - \theta_E^2/\theta^2} \approx \frac{2\theta_E^2}{\theta^2} \end{aligned} \tag{228}$$

Thus $\theta_E^2 = \langle\gamma_T\rangle\theta^2$ and

$$M = \frac{c^2}{4G}\frac{D_{Lens}D_S}{D_{LS}}\langle\gamma_T\rangle\theta^2 \tag{229}$$

For a *Singular Isothermal Sphere* (SIS) lens, with $\rho = (\sigma^2/2\pi G)r^{-2}$, where σ is the velocity dispersion in each axis, the deflection angle is constant since $M(< b) \propto b$. Then $\phi = \theta - \theta_E$, and $\partial\theta/\partial\phi = 1$, so $\gamma_T \approx 0.5\theta_E/\theta$. In this case

$$M(< D_{Lens}\theta) = \frac{c^2}{8G}\frac{D_{Lens}D_S}{D_{LS}}\langle\gamma_T\rangle\theta^2 \tag{230}$$

There is a strong correlation between the temperature of the hot gas in X-ray luminous clusters and the total mass M_{2500} within a sphere of radius r_{2500}. Hoekstra (arXiv:0705.0358) uses weak lensing to find M_{2500} and gets

$$M_{2500} = (1.4 \pm 0.2) \times 10^{14} h^{-1} M_\odot \left(\frac{kT_e}{5 \text{ keV}} \right)^{1.34 \pm 0.29} . \tag{231}$$

Note that since $kT \approx GM/r$ and $(3/4\pi)M_{2500}/r_{2500}^3 = 2500\rho_{crit}$, we can derive $M \propto r^3$, and $T \propto r^2$, so $M \propto T^{3/2}$ is the expected scaling. The observations are consistent with this expectation. Also, for $H_\circ = 71$ and $kT = 10$ keV, the mass is 5×10^{14} $M_\odot = 10^{48}$ grams, and $\overline{\rho} = 2.36 \times 10^{-26}$ gm/cc, giving a radius $r_{2500} = 0.7$ Mpc.

To prove the virial theorem consider the moment of inertia

$$I = \sum m_i \vec{r}_i^2 \tag{232}$$

If the system has settled into a steady state, then the time-averaged value of the second time derivative of I will be zero. So

$$\dot{I} = \sum 2 m_i \vec{r}_i \cdot \dot{\vec{r}}_i \tag{233}$$

and

$$\ddot{I} = \sum 2 m_i \left(\dot{\vec{r}}_i^2 + \vec{r}_i \cdot \ddot{\vec{r}}_i \right) . \tag{234}$$

Thus

$$\ddot{I} = 4 \sum \frac{1}{2} m_i \vec{v}_i^2 + 2 \sum_i m_i \vec{r}_i \cdot \left(\sum_{j \neq i} \frac{G m_j (\vec{r}_j - \vec{r}_i)}{|\vec{r}_j - \vec{r}_i|^3} \right) . \tag{235}$$

Now

$$\sum_i \sum_{j \neq i} \frac{G m_i m_j \vec{r}_i \cdot (\vec{r}_j - \vec{r}_i)}{|\vec{r}_j - \vec{r}_i|^3} = \sum_j \sum_{i \neq j} \frac{G m_i m_j \vec{r}_j \cdot (\vec{r}_i - \vec{r}_j)}{|\vec{r}_j - \vec{r}_i|^3} \tag{236}$$

because the RHS is obtained by just interchanging the indices i and j, and the names of the indices don't matter when they are summed over. But if these two quantities are equal, we can add them together and divide by two and get the same value again. This gives

$$\begin{aligned} \sum_i \sum_{j \neq i} \frac{G m_i m_j \vec{r}_i \cdot (\vec{r}_j - \vec{r}_i)}{|\vec{r}_j - \vec{r}_i|^3} &= \frac{1}{2} \sum_i \sum_{j \neq i} \frac{G m_i m_j (\vec{r}_i - \vec{r}_j) \cdot (\vec{r}_j - \vec{r}_i)}{|\vec{r}_j - \vec{r}_i|^3} \\ &= -\frac{1}{2} \sum_{i,j \neq i} \frac{G m_i m_j}{|\vec{r}_j - \vec{r}_i|} = \text{PE} \end{aligned} \tag{237}$$

Thus

$$\ddot{I} = 4(\text{KE}) + 2(\text{PE}) = 0 \tag{238}$$

in a steady state. This needs to be taken as a time average, since in many bound systems such as a elliptical binary the ratio of the instantaneous kinetic energy to the instantaneous potential energy varies with phase. But the time-averaged values satisfy

$$\langle \text{KE} \rangle = -\frac{1}{2} \langle \text{PE} \rangle \tag{239}$$

Usually we apply the virial theorem to a distant cluster where only radial velocities can be measured. If the radial velocity dispersion is $\sigma(v_r)$ then the kinetic energy is

$$\mathrm{KE} = \frac{3}{2} M \sigma(v_r)^2 \tag{240}$$

where the "3" comes from assuming isotropy in the velocity distribution. The potential energy is

$$\mathrm{PE} = -\frac{GM^2}{R_e} \tag{241}$$

where R_e, the effective radius, is found from

$$-\frac{1}{2} \sum_{i,j \neq i} \frac{G m_i m_j}{|\vec{r}_j - \vec{r}_i|} = \mathrm{PE} = -\frac{GM^2}{R_e} \tag{242}$$

so

$$\frac{1}{R_e} = \frac{1}{2} \left(\sum_{i,j \neq i} \frac{(m_i/M)(m_j/M)}{|\vec{r}_j - \vec{r}_i|} \right). \tag{243}$$

A stable evaluation of the potential energy can be found if a model for the density law of the cluster, $\rho(r)$, is found. Then

$$\mathrm{PE} = -16\pi^2 G \int_0^\infty \rho(r) r \left(\int_0^r \rho(r') r'^2 dr' \right) dr \tag{244}$$

Peebles in "Physical Cosmology" (the 1971 book, not the later "Principles of Physical Cosmology") gives an interesting way to determine R_e from strip counts. Let $S(\delta)$ be the count of objects in a strip displaced by δ from the center of the cluster on the sky. This strip is really a plane in 3-D. Let x and y be coordinates in that plane. Then

$$S(\delta) = \int \int n(\sqrt{\delta^2 + x^2 + y^2}) dx dy = 2\pi \int_0^\infty \eta n(\sqrt{\delta^2 + \eta^2}) d\eta \tag{245}$$

where $\eta = \sqrt{x^2 + y^2}$. Now $r = \sqrt{\delta^2 + \eta^2}$ so $r dr = \eta d\eta$. Thus

$$S(\delta) = 2\pi \int_\delta^\infty r n(r) dr \tag{246}$$

and

$$\frac{dS}{d\delta} = -2\pi \delta n(\delta). \tag{247}$$

Now let m be the mass of an object and M be the mass of the cluster:

$$M = 4\pi m \int_0^\infty n(r) r^2 dr = 2m \int_0^\infty S(\delta) d\delta \tag{248}$$

The potential energy is

$$\begin{aligned} \mathrm{PE} &= -16\pi^2 G m^2 \int_0^\infty n(r) r \left(\int_0^r n(r') r'^2 dr' \right) dr \\ &= -4 G m^2 \int_0^\infty \frac{dS}{dr} \left(\int_0^r r' \frac{dS}{dr'} dr' \right) dr \end{aligned} \tag{249}$$

This can be integrated by parts with $q = S$ and $p = \int_0^r r'(dS/dr')dr'$ giving

$$\mathrm{PE} = 4 G m^2 \int_0^\infty S r \frac{dS}{dr} dr. \tag{250}$$

Once again integrate by parts with $q = S^2/2$ and $p = r$, giving

$$\mathrm{PE} = -2 G m^2 \int_0^\infty S^2 dr \tag{251}$$

Now

$$R_e = -\frac{GM^2}{\mathrm{PE}} = \frac{4Gm^2 \left[\int_0^\infty S d\delta\right]^2}{2Gm^2 \int_0^\infty S^2 d\delta} = \frac{2\left[\int_0^\infty S d\delta\right]^2}{\int_0^\infty S^2 d\delta} = \frac{\left[\int_{-\infty}^{+\infty} S d\delta\right]^2}{\int_{-\infty}^{+\infty} S^2 d\delta} \tag{252}$$

Applying the virial theorem gives

$$M = \frac{3\sigma(v_r)^2}{G} \frac{2\left[\int_0^\infty S d\delta\right]^2}{\int_0^\infty S^2 d\delta} \tag{253}$$

If the strip counts follow

$$S \propto \frac{1}{\delta^2 + \delta_\circ^2} \tag{254}$$

then

$$\begin{aligned} \int S(\delta) d\delta &= \frac{\pi}{2} \frac{1}{\delta_\circ} \\ \int S(\delta)^2 d\delta &= \frac{\pi}{4} \frac{1}{\delta_\circ^3} \\ R_e &= \frac{2(\pi/(2\delta_\circ))^2}{\pi/(4\delta_\circ^3)} = 2\pi\delta_\circ \end{aligned} \tag{255}$$