TOPICS IN CONTEMPORARY ASTROPHYSICS

A.C. Fabian

Institute of Astronomy, Madingley Road, Cambridge CB3 0HA

1 Introduction

Studies of the Universe test our understanding of physics. They range over the most extremes of temperature, density, pressure, magnetic fields, distances etc. and involve all branches of physics. In the past, astronomy and astrophysics have led to the discovery of new physics and they should continue to do so in the future. In this book we concentrate on gravity, which although weak compared with the other forces, dominates on scales exceeding those of the Earthbound laboratory. Unlike electric charges, there are no repulsive masses – only attractive ones.

Gravity is responsible for binding the approximately 10^{51} particles in the Earth and the $\sim 10^{57}$ particles in the Sun into closely spherical shapes. In both cases the atomic nuclei are separated by similar distances to those of matter on the surface of Earth, i.e. about 10^5 times the diameters of the nuclei. In a collapsed star such as a neutron star, the number of particles is similar to that of the Sun, yet the nuclei are touching – the object is effectively a giant atomic nucleus of atomic number 10^{57}. The size of such an object is thus only 10 km, comparable to the size of city. At the surface of such an object the force of gravity is so strong that relativistic effects are important. Indeed the radius of a neutron star only exceeds the radius of a black hole of similar mass by a factor of a few.

Gravity also affects the most distant parts of the Universe and is the controlling force in cosmology. At large distances the gravitational forces are so large (given the enormous enclosed mass) that relativity again plays a major role.

We shall explore these objects and effects using relatively simple approaches. The basics rather than the details are the important issues, although we shall introduce those details necessary for understanding and maintaining interest.

1.1 Brief overview

The simple distribution of matter in space is unstable to gravitational collapse (Newton). The expansion of the Universe counteracts this in the mean but overdense regions have nevertheless collapsed to form galaxies and stars. In stable objects collapse is opposed by thermal pressure, radiation pressure, degeneracy pressure, electrostatic forces and angular momentum.

Thermal and radiation pressure (such as maintains the size of stars) imply an outward flow of energy, which continues so long as power is available. In stars this is due to thermonuclear fusion, which lasts as long as light elements (e.g. H) are abundant in the core. Later, gravitational collapse takes over. The details for stars are complex (the end of the H-burning phase in the core of a star leads to collapse of the core and expansion of the outer envelope, known as the red giant phase) but eventually the cores of all stars form a collapsed object, some with catastrophic consequences (supernovae). From a gravitational perspective, a star can be thought of as a temporary halt in the collapse of a gas cloud as it rids itself of nuclear energy.

The power of a star is released in its core (where $T > 10^7$ K) and it is highly thermalized by its passage through the overlying layers. It emerges as quasi-blackbody radiation; very little power is
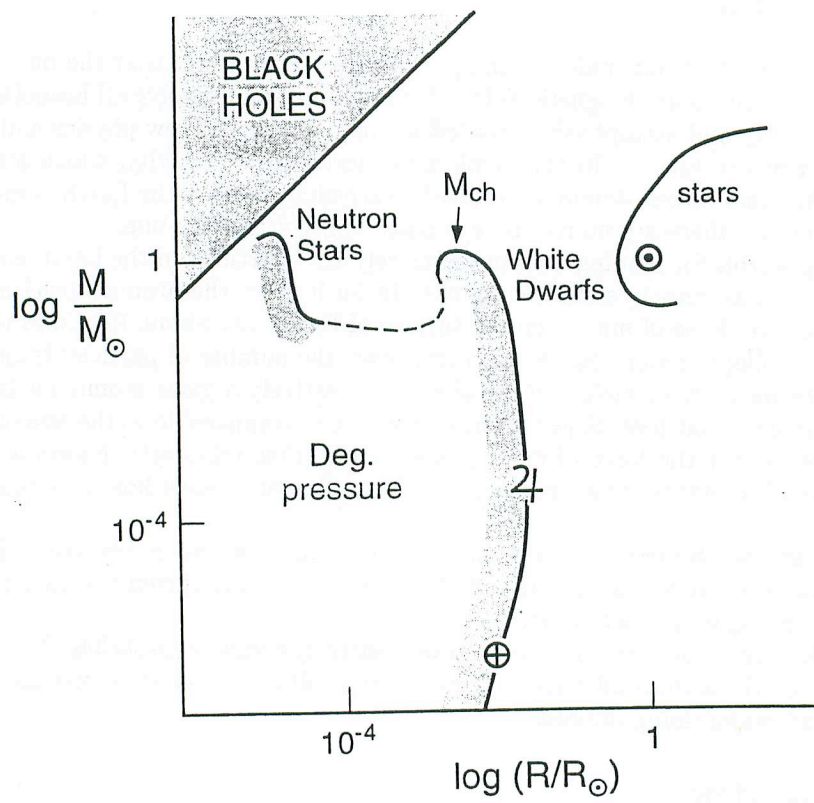
Figure 1: Mass-radius diagram for objects of planetary and stellar mass.

either at high – X-ray, or low – radio, frequencies. In solar-type stars (our Sun is a dwarf G star) the emission peaks in the green part of visible spectrum.

Accretion and other processes to be discussed later that occur around compact objects (White Dwarfs, Neutron Stars or Black Holes) can be much more efficient in converting mass to energy and involve much less surrounding matter, so allowing the direct power source to be seen. The emission can be spread widely across much of the electromagnetic spectrum. Note that if

$$h\nu = kT = \frac{1}{2}m_p v^2$$

then as $v \to c$ the spectral peak shifts to hard X-rays and the temperatures to 10^9 K and above. Electrons with relativistic speeds are generated which radiate over wide spectral bands. Such sources involve the most powerful (highest L), most compact (highest L/R) and most variable (smallest R for large L) objects known, testing physics to its extremes.

Gravity is an essential process in the formation and/or operation of such objects – pulsars, quasars, and accreting black hole sources in binaries and active galaxies and quasars. Much can be understood without recourse to detailed General Relativity, although the limitations of Newtonian gravity must be recognized.

We shall also look at some of the largest gravitationally-bound objects in the Universe: the clusters of galaxies. Most of the binding matter here is dark and of unknown composition.

2   STARS AND STELLAR EVOLUTION

2.1   Timescales

The collapse time, t_{ff} on which a star would fall in on itself if only gravity acts, is given approximately by the radius divided by the free-fall velocity at that radius;

$$t_{ff} \approx \left(\frac{R^3}{GM}\right)^{\frac{1}{2}} \approx \frac{R}{v_{ff}} \sim 27\left(\frac{R}{R_\odot}\right)^{3/2}\left(\frac{M}{M_\odot}\right)^{-1/2} \text{ min.}$$

The time to release the gravitational binding energy of a star, at luminosity L, is known as the Kelvin timescale;

$$t_{Kelvin} \approx \frac{GM^2}{RL} \approx 3 \times 10^7 \left(\frac{M}{M_\odot}\right)^2 \left(\frac{R}{R_\odot}\right)^{-1} \left(\frac{L}{L_\odot}\right)^{-1} \text{ yr.}$$

2.2   Equations of Stellar Structure

Since $t_{ff} \ll t_{Kelvin}$, the Sun must be very close to hydrostatic equilibrium, *i.e.*

$$\frac{dP}{dr} = -\rho\frac{GM(r)}{r^2}, \tag{S1}$$

$$\text{where} \quad M(r) = \int_0^r \rho(r)4\pi r^2 dr. \tag{S2}$$

S1 and S2 are the first two equations of stellar structure.

Note that

$$\frac{dP}{dr} = -\frac{GM}{4\pi r^4}\frac{dM}{dr}$$

$$\text{and} \quad \frac{d}{dr}\left(P + \frac{GM^2}{8\pi r^4}\right) = \left(\frac{dP}{dr} + \frac{GM}{4\pi r^4}\frac{dM}{dr}\right) - \frac{GM^2}{2\pi r^5} < 0.$$

$P + \frac{GM}{8\pi r^4}$ must be greatest in the centre of a star and therefore must exceed the value at the surface of the star, *i.e.*

$$P_c > \frac{GM^2}{8\pi R^4} \approx 4.4 \times 10^{13} \left(\frac{M}{M_\odot}\right)^2 \left(\frac{R}{R_\odot}\right)^{-4} \, \text{Nm}^{-2}.$$

For the sun, this estimate falls short by a factor of about 100 since it does not take into account the central density concentration.

If we use the simple dimensional estimate $P_c \sim GM^2/R^4$ (pressure is energy density or binding energy over volume) and the ideal gas law $P_g = k\rho T/\mu m$, where μm is the mean molecular weight, then the central temperature of a star is

$$T_c \sim \frac{GM^2}{R^4} \frac{\mu m}{k\rho} \sim 6 \times 10^7 \, \text{K}.$$

Since kT is much greater than the binding energy of an atom of hydrogen (which dominates the composition of the Sun), the gas is ionized (and $\mu \sim 0.5$).

Note that

$$kT \approx \frac{GM\mu m}{R},$$

a result which can be generalized in the *Virial Theorem*: A self-gravitating closed system of point masses bound by inverse square forces has (time-averaged)

$$\bar{E}_{kin} = -\bar{E}_{pot}/2.$$

For a star,

$$E_{pot} = E_{grav} = -\int_0^R \frac{GM(r)}{r} \rho 4\pi r^2 dr$$

which from the equation of hydrostatic support gives

$$E_{grav} = \int_0^R \frac{dP}{dr} 4\pi r^3 dr.$$

Integration by parts (noting that $P = 0$ when $r = R$) then gives

$$E_{grav} = -3 \int_0^R P dV$$

where V is volume. For a perfect gas, $P = nkT$, and $E_{kin} = \frac{3}{2}nkT$, so

$$E_{grav} = -2E_{kin},$$

thus proving the Virial Theorem.

Note that gravitational collapse leads to increased thermal energy ($T \propto 1/r$). In stars this leads to self-regulation of the nuclear burning processes in regions where pressure is dominated by thermal pressure (slight collapse raises the core temperature which in turn raises the energy generation rate and thus the pressure resisting collapse).

If we now include all forms of energy (internal energy in degrees of freedom etc. characterized by the ratio of specific heats γ). Per kg the thermal energy

$$E_T = \frac{3}{2} \frac{kT}{\mu m},$$

and the internal energy

$$U = C_V T.$$

4

Since $C_P - C_V = k/\mu m$ and $\gamma = C_P/C_V$ so $C_V(\gamma - 1) = k/\mu m$ and $U = \frac{kT}{\mu m(\gamma - 1)}$, then

$$E_T = \frac{3}{2}(\gamma - 1)U = -E_{grav}/2.$$
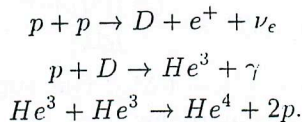
The total energy of the star,

$$E_{Tot} = U + E_{grav} = -\frac{E_{grav}}{3(\gamma - 1)} + E_{grav} = -E_{grav}\frac{(3\gamma - 4)}{3(\gamma - 1)}.$$

A star is therefore unstable to collapse or infinite expansion (E_{Tot} positive) if $\gamma < 4/3$.

The basis of the above results can be seen by consideration of a static cloud of gas. If it is squeezed inward then it will bounce back unless it loses some of the kinetic energy of its gas particles. It can do that by radiating energy away, which is the usual reason for a cloud collapsing, or by some of the gravitational energy released in the collapse going into molecular rotation, say, or ionization of the gas. In the last two examples the ratio of specific heats is certainly less than 5/3 and may be less than 4/3, in which case too little kinetic energy remains to support the cloud which continues to collapse. Another important situation for which the effective $\gamma = 4/3$ occurs if the pressure is dominated by a relativistic gas or photons. This is relevant for stars supported by either radiation pressure or relativistic degeneracy pressure. Such stars are inherently unstable.

When a star forms by gravitational collapse from a diffuse gas cloud most of the gravitational energy released is radiated away and the central pressure and temperature rise until thermonuclear fusion takes place. Since the most abundant element in the Universe and thus of interstellar gas clouds is hydrogen, fusion of hydrogen to form helium is the most important process. The strong electrostatic repulsion between protons means that high temperatures are needed, of about ten million K (at which point the gas will be highly ionized). Quantum mechanical tunnelling is important for the fusing of protons (which means a steep temperature dependence for the reaction).

For the Sun, the most important reactions is the *p-p chain*;

$$p + p \rightarrow D + e^+ + \nu_e$$
$$p + D \rightarrow He^3 + \gamma$$
$$He^3 + He^3 \rightarrow He^4 + 2p.$$

the total energy release is about 26.2 MeV per He^4 formed (*i.e.* $\sim 0.007mc^2$, where m is here the mass of the 4 protons used in the chain). The rate of energy generation $\epsilon \propto \rho T^4$ and is about 10^{-4} W kg^{-1} in the Sun.

Further reactions can take place, the CNO cycle being important in massive stars ($\epsilon_{CNO} \propto \rho T^{15}$). About 2% of the Solar energy is emitted in the form of neutrinos, which readily escape from the core of the Sun. Experiments are in operation to detect these neutrinos (the observed rates do not agree with the predicted ones!)

Energy generation gives us the third equation of stellar structure

$$\frac{dL}{dr} = 4\pi r^2 \rho \epsilon. \tag{S3}$$

Energy escapes from the core through either radiation or convection (both are important at different radii in the Sun; radiation in the core and convection throughout much of the envelope).

The fourth equation of stellar structure relates to the energy transfer in the star; for radiative transport

$$\text{Flux} = \frac{L(r)}{4\pi r^2} = -D\frac{dU(r)}{dr},$$

where $U(r)$ is the energy density in radiation (aT^4) at radius r. Now from kinetic theory, the diffusion coefficient $D = \frac{1}{3}\lambda c$ and $\lambda = 1/(n\sigma) = 1/(\kappa\rho)$, where σ is the cross section for interaction of matter and radiation and κ is the absorption coefficient. So

$$\frac{L(r)}{4\pi r^2} = -\frac{4}{3}\frac{acT^3}{\kappa\rho}\frac{dT}{dr}. \tag{S4}$$

5

2.3 Stellar Models

To make a stellar model, the equations of stellar structure (S1 – S4) are solved together with some equation for the absorption coefficient κ, matching the surface properties and mass with the observed values. Usually this can only be done numerically. Some scaling laws, as functions of M, R or L can be obtained by noting that

$$\frac{L}{R^2} \propto \frac{T^4}{\kappa \rho R}; T \propto \frac{M}{R}; \rho \propto \frac{M}{R^3},$$

assuming say that electron scattering dominates the radiative diffusion (as it does in massive stars), so $\kappa =$ a constant, gas pressure dominates the pressure and that $L = \epsilon M$ with $\epsilon \propto \rho T^n$. Roughly, $L \propto M^3$. Typically, about 12% of the mass of a star undergoes $H \to He$ fusion, yielding a hydrogen burning lifetime for the Sun of about 10^{10} yr, scaling as mass^{-2} or so for more massive stars.

When the hydrogen in the core of a star is consumed the core must contract in order for the temperature to rise for He-burning to occur. The outer envelope then swells by a factor of 100 or more. There is no simple description of why the envelope expands so much, but numerical (computer) models of stars and observation of real stars are in good agreement with such a dramatic expansion of the envelope. The star is then in its red giant phase which last a much shorter time than the hydrogen burning phase due to the much higher core temperature. The Earth and its orbit will be engulfed by the Sun in 5 billion years time when it becomes a red giant.

Radiation pressure is more important than gas pressure in the most massive stars. To assess when this happens, assume that $P_g > P_{rad}$, then

$$P_g \sim nkT \sim \frac{GM^2}{R^4},$$

$$P_{rad} \sim aT^4 \sim \frac{aG^4 M^4 (\mu m)^4}{R^4 k^4},$$

use page 4, 3rd equation from top

so $P_{rad} > P_g$ when $M > k^2 a^{1/2} (\mu m)^2 G^{3/2}) \sim 6\,M_\odot$. Gas supported by radiation pressure has $\gamma \sim 4/3$ so massive stars are prone to catastrophic mass loss. This limits the mass of stars to $\lesssim 100\,M_\odot$ (see also later discussion on Eddington limit).

3 Degeneracy Pressure

The low mass limit on stars is determined by *degeneracy pressure*, which also supports the cores of some stars and particularly *White Dwarfs*. It is due to the compression of fermions which obey the Pauli Exclusion principle (only 2 fermions are allowed per energy level, with opposing spins). When matter is compressed the zero-point (or Fermi) energy increases, which generates a pressure (energy per unit volume).

An approximate (but qualitatively accurate) description of degeneracy pressure can be obtained from the Uncertainty Principle, $\Delta p \Delta x \approx \hbar$. The zero-point energy per particle

$$E_{deg} = \frac{p^2}{2m} = \frac{\hbar^2}{2m\Delta x^2}$$

increases as the size available to a particle decreases. The consequent pressure $P_{deg} = \frac{E_{deg}N}{V}$, for N particles in volume V. Using density $n = N/V$ and $\Delta x \approx (V/N)^{1/3}$,

$$P_{deg} \approx \frac{N}{V} \frac{p^2}{2m} = \frac{\hbar^2}{2m} n^{5/3}.$$

Note that $P \propto \rho^{5/3}$ so it behaves like an ideal gas with $\gamma = 5/3$. Also note that $P_{deg} \propto m^{-1}$ so electron degeneracy pressure is about 2000 times more important that that for protons.

So far it has been assumed that the zero-point motion of the electrons is non-relativistic ($v \ll c$). As the density rises however the speed increases and the electrons go relativistic, where the energy $E = pc$, not $p^2/2m$ so

$$P_{deg,rel} = \frac{N}{V}\frac{\hbar^2}{2m}\left(\frac{N}{V}\right)^{1/3} = \hbar c n^{4/3}.$$

The effective γ is thus $4/3$ and the object is unstable to gravitational collapse.

A rough limit to the mass of a H-burning star is obtained by noting that a central temperature $T_c \approx \frac{GMm}{Rk} \gtrsim 10^7$ K is required for the p-p chain to operate. This can occur in the collapse of a gas cloud provided that the thermal pressure (nkT_c) exceeds the degeneracy pressure ($\frac{\hbar^2}{2m_e}n^{5/3}$ where the subscript e denotes electrons) as the core heats up to T_c. This corresponds to a mass exceeding (fundamental constants times T_c)$^3/4$ or about $0.05\,M_\odot$. Collapsed objects of lower mass, say $0.01\,M_\odot$, are supported by degeneracy pressure before they have collapsed sufficiently for T_c to exceed 10^7 K. Object of mass $0.001\,M_\odot$ or less (i.e. the mass of Jupiter or less) are supported by a combination of degeneracy pressure and electrostatic forces (in a sense the Pauli and Uncertainty Principles applied within individual atoms).

In summary, we have found that stars, shining by the light of self-generated nuclear fusion, have masses lying between 0.05 and $100\,M_\odot$. The most massive stars are the most luminous by far and also the most short lived.

4 White Dwarfs

Collapsed objects supported by electron degeneracy pressure are called White Dwarfs. Their properties can be determined simply by considering the total energy of the star (ignoring geometrical factors) (i) when the electrons are non-relativisitic

$$E = NE_{deg} + E_{grav} = \frac{\hbar^2}{2m_e}\left(\frac{N}{V}\right)^{2/3}N - \frac{GM^2}{R}.$$

Using $N = M/m_p$ and $V = R^3$, gives

$$E = \frac{\hbar^2}{2m_e}\left(\frac{M}{m_p}\right)^{5/3}\frac{1}{R^2} - \frac{GM^2}{R}.$$

Equilibrium occurs when E is minimized, yielding

$$M \approx \left(\frac{\hbar^6}{m_e^3 G^3 m_p^5}\right)R^{-3}.$$

Note that the radius *decreases* as the mass increases. To understand this recall the Uncertainty Principle which requires that the space available to the electrons be reduced in order for the momentum to increase. When the electrons become relativistic, the energy equation is

$$E = hc\left(\frac{M}{m_p}\right)^{4/3}\frac{1}{R} - \frac{GM^2}{R}.$$

Both terms on the RHS have the same radius dependence and there is no minimum. The star is unstable to collapse (or expansion to infinity). The mass at which this occurs can be obtained by setting $E = 0$ above (*i.e.* the maximum mass for which $E \leq 0$), giving

$$M_{Ch} = \left(\frac{\hbar c}{G}\right)^{3/2}\frac{1}{m_p^2},$$

7

which is known as the *Chandrasekhar mass*. Note that it depends only on fundamental constants. The exact value of M_{Ch} depends on the composition of the White Dwarf. The zero-point energy of the ions in massive cold white dwarfs is sufficient to allow nuclear fusion to proceed (pyconuclear fusion) and the equilibrium mixture is mainly C and O. M_{Ch} is then about $1.4\,M_\odot$.

The electron mean-free-path in degenerate matter is large (cf. metals, because of a filled Fermi sea) and heat conduction is efficient. A small layer of non-degenerate matter on the surface of white dwarfs acts as a blanket keeping the heat (from the collapse etc.) in and they take $\sim 10^{10}\,\mathrm{yr}$ to cool, so appear hot and white in a telescope. The radius of a white dwarf is similar to that of the Earth, or about $5 \times 10^6\,\mathrm{m}$.

5 Neutron Stars

Another solution to that of the white dwarf exists for dense matter, the neutron star. When the degeneracy energy of an electron exceeds the mass-energy difference between the neutron and proton;

$$E_{deg} > (m_n - m_p)c^2 = 1.29\,\mathrm{MeV}$$

(at which point the electrons are mildly relativistic) then inverse β-decay is favourable

$$e^- + p \to n + \nu.$$
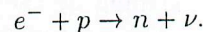
(The reverse process becomes blocked by a lack of free electron states.) The density at which this occurs is about $10^{10}\,\mathrm{kg\,m^{-3}}$. At still higher densities the matter become neutron rich (with sufficient electrons and ions to fill all free energy states and keep the matter electrically neutral).

Neutron degeneracy pressure is then important; from the mass-radius relation for white dwarfs

$$M \propto \frac{1}{(m_{deg}R)^3},$$

where m_{deg} is the mass of the degenerate particles. So for $M_\odot$ and comparing electrons and neutrons

$$R_{ns} \approx R_{wd}\left(\frac{m_e}{m_p}\right) \approx 10^4\,\mathrm{m}.$$

At this radius, the interparticle separation $\Delta x \approx \left(\frac{N}{V}\right)^{-1/3} \approx 10^{-15}\,\mathrm{m}$, which is similar to the range of the strong nuclear force (*i.e.* the Compton wavelength of the proton: the particles 'touch'). The matter density in a neutron star is thus comparable to nuclear matter density (ρ_{nuc}). The equation of state of matter at even $2\rho_{nuc}$ is not known from laboratory experiments so the precise structure of massive neutron stars is unclear. They may (or may not) have solid cores. The radius of a neutron star is about 10 km.

To make a model of a neutron star requires that General Relativity be included since

$$\frac{GM^2}{R} > 0.1Mc^2,$$

(*i.e.* R_{ns} is a few R_S). The pressure needed to oppose gravity is an energy density and this can be seen as an extra mass density.

Note that there is a lower mass to a neutron star (of about $0.1\,M_\odot$) below which the neutrons are unstable to β-decay (the radius increases as the mass decreases for degenerate matter). There is also an upper mass of about $3\,M_\odot$. This is derived using models for neutron stars and an equation of state (P, ρ) which begins (for the outer layers) with known properties of nuclear matter and is extrapolated in assuming causality ($dP/d\rho < c$).

Gathering together the structures of stars, white dwarfs and neutron stars, gives us the mass-radius relation for cosmic objects;
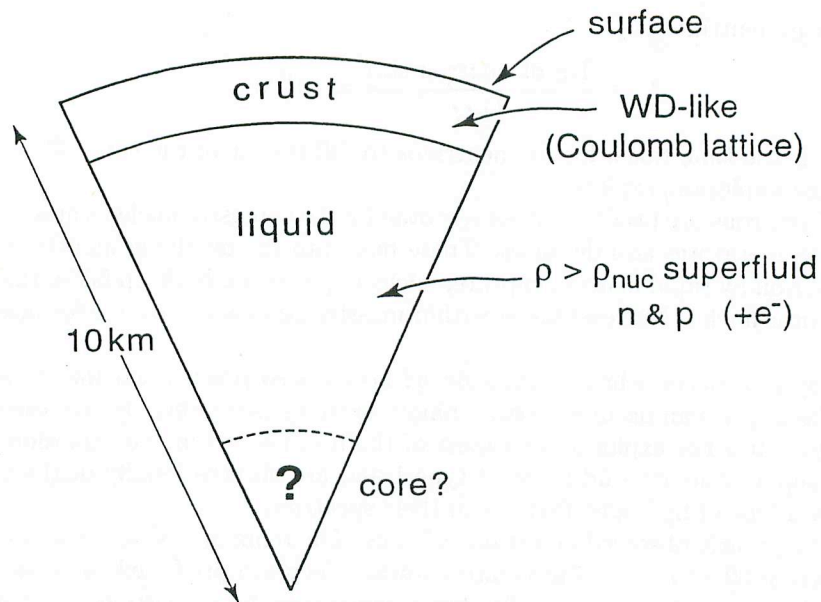
Figure 2: The structure of a neutron star. The outer crust is similar to the core of a white dwarf and has a crystal lattice structure. The bulk of the star consists of superfluid neutrons. The structure of any core is unknown.

Coulomb electrostatic energy becomes important for low mass objects such as planets (and us), where matter may form into a lattice. At lowest masses this dominates and the object need not be spherical, with mountains of sizes comparable to the 'planet' itself occurring in asteroids.

White dwarfs form from the cores of stars with masses $M \lesssim 8 \, M_\odot$. Sufficient mass is shed in the form of a wind, especially when the star evolves to become a red giant, that the resulting core is less than M_{Ch}.

6 Supernovae

When a massive star ($M > 8 \, M_\odot$) reaches the end of its life, its core collapses and may form a neutron star. In such massive stars, core nuclear burning can take place all the way to iron, above which further nuclear fusion is endothermic (it requires energy rather than yields it). The resulting iron core, partly supported by electron degeneracy pressure, then collapses. The gamma-rays produced by the enormous release of gravitational energy cause photodissociation of the iron nuclei, which disintegrate back into protons.

Very high temperatures are reached in the collapse and a large fraction of the energy is released as neutrinos. These slow the collapse of the core from the free-fall time to a few seconds, since at densities as high as 3×10^{14} kg m^{-3}, they scatter off the matter (weak neutral currents are important). The core bounces when it reaches nuclear densities (i.e. the nuclei touch) and the overlying matter, which is most of the star, is ejected at high velocities of about $10,000 - 20,000$ km s^{-1}. The sudden release of the energy in this envelope and in its subsequent recombination gives rise to a highly luminous outburst, with a visual luminosity of up to $10^9 \, L_\odot$. Such an event is known as a Supernova. The light drops rapidly over the next few months as the remaining sources of power (principally decay of radioactive species produced in the explosion such as Ni^{56} $\rightarrow$ Co^{56} $\rightarrow$ Fe^{56}) diminish. The cores of such supernovae (known as Type II supernovae) remain as neutron stars, or black holes if the mass of the progenitor star was high enough (this is uncertain, perhaps above $30 \, M_\odot$).

The luminosity in neutrinos is

$$L_\nu \approx \frac{\text{BE of neutron star}}{t_{diff}} \approx 10^{45}\text{W}$$

where $t_{diff} \sim 10\,\text{s}$ is the time taken for the neutrinos to diffuse out of the core. This is the main energy loss from the implosion/explosion.

Large fluxes of neutrons are produced in supernovae by the explosive nucleosynthesis that takes place in the high temperatures and densities. These pass into the overlying mantle and build up the elements above iron by rapid neutron capture; nuclear species are built up faster than β-decays destroy the progenitor nuclei. Nucleosythesis within massive stars is the source for most elements beyond helium.

Supernovae may also occur when a white dwarf accretes so much mass that it exceeds the Chandrasekhar mass. This can occur in certain binary systems (see later). In this case, no stellar remnant is expected after the explosion and most of the iron formed in the explosion is expelled into space. Such supernovae are said to be of type I and are observationally distinguished from those of type II by a lack of hydrogen features in their spectrum.

Supernovae are regularly observed in distant galaxies. The remnants of about 5 supernova that occurred in the last 1000 yr in our Galaxy are known. They are predicted to occur at a mean interval of about 30 yr; most are not readily found since they lie in parts of the Galactic disk that are obscured by dust clouds. The nearest 'recent' supernova was in 1006, when the object appeared as bright as the Moon and could be seen in the daytime for a month or so.

6.1 SN 1987A

The nearest supernova observed since the telescope has been used for astronomical observation (i.e. about 1610) occurred in the Large Magellanic Cloud (about 50 kpc away) in February 1987. The progenitor star was massive ($\sim 15\,\text{M}_\odot$) and at maximum the supernova was visible to the naked eye. As a supernova it was underluminous because it was not a red (super)giant when it exploded but a blue one, either due to its lower metallicity (typical of the Clouds) or because it had lost much mass, perhaps through having a close binary companion.

Neutrinos were detected by large underground detectors on Earth (designed to look for proton decay). The flux detected (about 20 neutrinos with energies of between 5 - 10 MeV), when combined with the interaction cross section of the detectors, is consistent with a total energy of about 10^{46} J or the binding energy of a neutron star. It is likely that a neutron star (which could later have collapsed to become a black hole) was formed in the event. The fact that all the detected neutrinos arrived within 20s after travelling a distance of about 150 thousand ly enables a mass limit to be placed on the neutrino (of about 15eV) through the detected energy $E = mc^2$ and $m = m_0/\sqrt{1 - v^2/c^2}$, i.e.

$$E = \frac{m_0 c^2}{\sqrt{1 - v^2/c^2}}.$$

Differentiating this with respect to v and then approximating it to $\Delta E/\Delta v$, where $\Delta v \sim c^2 \Delta t/R$ and R is the distance to the supernova, leads to the rest mass $m_0 c^2 \approx E^{3/2}\sqrt{\left(\frac{c\Delta t}{\Delta E R}\right)}$. For SN 1987A the limit on the rest mass corresponds to about 10 eV.

The decay of the light from the supernova has followed an exponential law due to the main power source being the radioactive decay of Co→Fe. Any additional power due to a young neutron star (for example) must be less than about 10^{30} W, which is less than expected.

7 Supernova Remnants

The envelope of the progenitor star is ejected into space at about $10^4\,\text{km}\,\text{s}^{-1}$. It interacts there with the stellar wind ejected much earlier (and at lower velocity) from the progenitor star, together
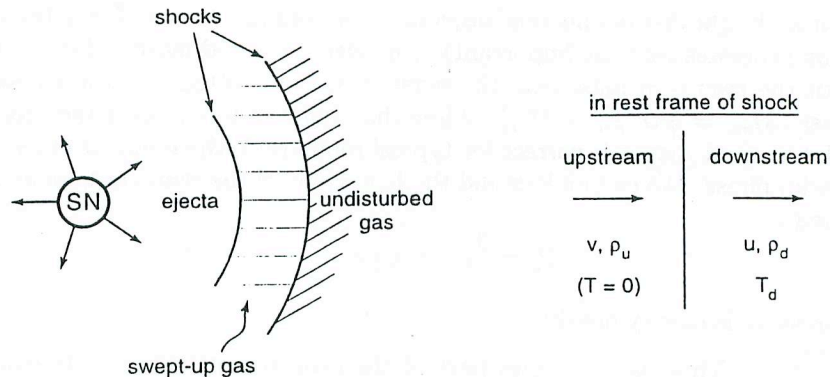
Figure 3: Supernova blast wave. In the early phases of the life of a supernova remnant there is a shock wave between the freely expanding ejecta and the swept-up gas and another between that gas and the undisturbed interstellar gas. In the rest-frame of the shock, matter streams in at low temperature and flows out at temperature T_{d}.

with any other surrounding or interstellar gas. In doing so it causes a strong shock to occur since the ejecta is moving highly supersonically compared with the surrounding matter (*i.e.* it is moving faster than the velocity of sound in the interstellar gas).

To deduce what happens in a strong shock, consider the situation at the interface between shocked and unshocked gas. The transition is abrupt there since the matter upstream of the shock, which is flowing into the shock, can have no knowledge of it.

Now consider the equations of mass, momentum (pressure) and energy conservation;

$$\text{Mass}: \ \rho_u v = \rho_d u,$$

$$\text{Momentum} \ \rho_u v^2 = P_d + \rho_d u^2,$$

$$\text{Energy} \ v(\frac{1}{2}\rho_u v^2) - u(\frac{1}{2}\rho_d u^2 + \frac{3}{2}P_d) = uP_d.$$

Note that the these equations are carried out for unit area at the shock, ρv^2 is the ram pressure (rate of change of momentum) and that uP represents the rate at which PdV work is done at the shock.

Rearranging, we get

$$\frac{\rho_d^2}{\rho_u^2}u^3 - 5\frac{\rho_d}{\rho_u}u^3 + 4u^3 = 0$$

$$\text{or} \ \left(\frac{\rho_d}{\rho_u} - 4\right)\left(\frac{\rho_d}{\rho_u} - 1\right) = 0.$$

So valid solutions are either that nothing changes (which must be one solution) or that

$$\frac{\rho_d}{\rho_u} = 4; \ \frac{v}{u} = 4.$$

From the ideal gas law, we obtain $T_d = P_d \mu m / \rho_d k$ or

$$T_d = \frac{3}{16}\frac{\mu m v^2}{k}.$$

use $*$, $**$ and solve for P_δ

This means that a significant fraction of the kinetic energy upstream of the shock is thermalized at the shock where it appears downstream as heat.

Much of the kinetic energy in the expanding ejecta from a supernova is therefore converted into heat. The temperatures that result exceed 10^7 K in the first few 1000 yr and expanding supernova

11

remnants are seen as bright X-ray emitting objects (very roughly $h\nu \approx kT$ for bremsstrahlung and other radiation processes that are important). The ejecta expand more or less freely (*i.e.* $v =$ const, $r \propto t$) until the swept-up mass (*i.e.* the mass it has interacted with) has a similar mass to the ejected mass ($M_{sw} = \frac{4}{3}\pi r^3 \rho_u = M_{ej}$). After that (and assuming that radiative cooling is unimportant in the shocked gas, as is correct for typical remnants), the remnant enters the Sedov-Taylor (or Adiabatic) phase. We can understand the behaviour of the remnant here by noting that the pressure driving it

$$P_d = \frac{3}{4}\rho_u v^2 \propto \rho \dot{r}^2$$

is also $\propto E/r^3$ (pressure is energy density).

Thus $r \propto \left(\frac{E}{\rho_u}\right)^{1/5} t^{2/5}$. (This also describes part of the evolution of atom bomb explosions, for which it was first applied.)

Later the remnant moves slowly, radiative cooling of the post-shock gas is important and it merges with the rest of the interstellar gas. Its size may then by about 30 pc. Supernova remnants are a major source of energy and the source of 'metals' in the interstellar medium. Most of the elements of which the Earth is composed were forged in supernova explosions and were expelled through supernova remnants.

8 OBSERVATIONS OF COMPACT OBJECTS

8.1 Discovery of White Dwarfs

Sirius A (which is the brightest star in the Sky) is perturbed by a faint companion – Sirius B, inferred from the 'wobble' in its proper motion in the 1830s and first seen in 1865.

Sirius A and B form a binary with an orbital period of 49.9 yr. $2a \approx b$, where a and b are the semi-major axes of the orbits of A and B. This, with an estimated mass of Sirius A of $2\,M_\odot$, enables us to deduce that $M_B \approx 1\,M_\odot$.

Both stars have similar (white) colour, so have similar surface temperatures $T_s \sim 10^4\,\mathrm{K}$, but their luminosities are observed to differ by a large factor, $\sim 10^4$. So, using $L = 4\pi R^2 \sigma T^4$ (Stefan–Boltzmann law), $R_B \sim R_A/100$ and we conclude that $\rho_B \sim 10^6 \rho_A$. Assuming that A has a mean density similar to that of the Sun (which is about the density of water) then $\rho_B \sim 10^9\ \mathrm{kg\ m^{-3}}$. This implies that it is a white dwarf.

White dwarfs were not fully understood until quantum mechanics and degeneracy pressure were introduced in the 1920s (initially for white dwarfs by R. Fowler). The surface gravitational redshift ($z \sim 10^{-4}$) was measured from the shift of known spectral lines in the 1960s (claims in the 1930s were erroneous). Helium is absent from the surface layers of many white dwarfs due to settling in the large gravitational field.

White dwarfs are common in the Galaxy, in wide binaries (as in Sirius), in close binaries (see later) and as lone objects. The Sun will become one in another 5 billion yr or so.

9 Lone Neutron stars - Pulsars

Neutron stars initially cool rapidly after formation due to neutrino emission (they initially cool from the inside out). T_{core} drops to $\lesssim 10^8\,\mathrm{K}$ (from $10^{11}\,\mathrm{K}$) in 10^3 yr. $T_{surface}$ is then about $10^6\,\mathrm{K}$. The radius is so small that the emitted luminosity is also small. The emission is predominantly in the EUV band where photoelectric absorption by Galactic hydrogen can drastically reduce the flux to below any observable level. There is no clear detection yet of thermal radiation from the surface of a cooling neutron star (but see pulsars).

Although neutron stars were predicted to occur by Landau, Zwicky and others in the 1930s (soon after the neutron was discovered) they were considered to be undetectable until cosmic X-ray

sources were discovered in the early 1960s. Zwicky and Baade did predict that neutron stars could be formed in supernovae and even suspected the supernova remnant known as the Crab Nebula of containing one.

Neutron stars were discovered serendipitously by Bell, Hewish and others in 1967. They found radio sources emitting flashes with a period of about 1 s. Many such sources, called PULSARS, are now known, with periods

$$1.6 \text{ ms} \lesssim P \lesssim 5 \text{ s}.$$

P is found to increase with time (apart from so-called glitches), $i.e.$ $\dot{P}$ is positive. Pulsars can be very good clocks.

Immediately after the discovery of pulsars it was realized that some form of compact object was involved. Rotation or pulsation of any self-gravitating object must give

$$\Omega^2 \lesssim \frac{GM}{R^3} \sim G\rho,$$

($i.e.$ it cannot spin faster than breakup speed or oscillate faster than the escape velocity), so values of the pulse frequency Ω limit the density ρ of the object. $P < 1\,\text{s} \Rightarrow$ a neutron star.

The generally accepted and tested model of a pulsar is the rotating, magnetized neutron star (first put forward by Gold).

9.1   The magnetic dipole model

Consider a spinning sphere which has a magnetic dipole moment and surface magnetic field (at the pole) m and B_p, respectively, spinning at angular frequency Ω. This emits dipole radiation (cf. Larmor's formula) at a power

$$\mathcal{P} = \frac{2}{3c^3} \left(\frac{\mu_0}{4\pi} \right) |\ddot{m}|^2$$

$$\text{where} \quad m = \frac{B_p R^3}{2} \left(\frac{4\pi}{\mu_0} \right).$$

Therefore, since $\ddot{m} = m \sin \alpha \Omega^2$,

$$\mathcal{P} = \frac{2}{3c^3} \left(\frac{\mu_0}{4\pi} \right) m^2 \Omega^4 \sin^2 \alpha = -\frac{dE_{rot}}{dt}$$

$$= \frac{4\pi}{\mu_0} \frac{2 B_p^2 R^6 \Omega^4 \sin^2 \alpha}{3c^3}.$$

The power is emitted at the expense of the rotational energy of the star, $E_{rot} = \frac{1}{2} I \Omega^2$, where I is the moment of inertia of the star. Now

$$\dot{E}_{rot} = I \Omega \dot{\Omega} = -\frac{8\pi}{\mu_o} \frac{B_p^2 R^6 \Omega^4}{3c^3} \sin^2 \alpha.$$

Thus $\dot{\Omega} < 0$ ($\dot{P} > 0$) so the pulsar slow down due to the loss of rotational energy, as observed.

[Note that the B-field corotates with the neutron star out to a radius R_c at which the speed of a corotation particle equals c – the speed of light cylinder. $R_c = c/\Omega$. The field there (assuming a dipole all the way) $B_c \sim B_p (R/R_c)^3$ and the energy density $\varepsilon_c \approx B_c^2 \mu_O^{-1}$. This is being 'stripped off' and is flowing away at c as radiation so

$$\mathcal{P} \sim \varepsilon_c R_c^2 c \Rightarrow \frac{B_p^2 R^6 \Omega^4}{\mu_o c^3}.$$

Thus giving a simple physical explanation of the above result.]

Observationally, a Characteristic Age, τ = 'spin-down time,' can be determined with which the surface magnetic field can be deduced (assuming theoretical properties of a neutron star and some value for α).

$$\tau = -\left(\frac{\Omega}{\dot{\Omega}}\right)_0 = \frac{6Ic^3\mu_0}{B_p^2 R^6 \sin^2 \alpha \Omega_0^2 4\pi}.$$

The basic dipole theory can also be tested in a few cases from the *braking index*, defined by $\dot{\Omega} \propto -\Omega^n$. Observationally,

$$n = -\frac{\Omega\ddot{\Omega}}{\dot{\Omega}^2}$$

and should be 3 for magnetic dipole radiation. For the few cases where it is measured (*e.g.* the Crab pulsar) it is about 2.5.

The convincing early success for the magnetic dipole model came from the discovery of the pulsar in the Crab Nebula. The Crab is associated with the supernova observed in 1054 AD in the constellation Taurus. The nebula was found in the 1950s to emit synchrotron radiation (it is strongly polarized) in the radio and optical bands. It was also found to be an X-ray source in the 1960s. The radiative lifetime of the electrons emitting the optical and X-ray emission is shorter than 10 yr and so requires some continuous injection of power, at a rate $\mathcal{P} \sim 5 \times 10^{31}$ W. The pulsar in the centre of the Crab has a period of $P = 33$ ms and $\dot{P} = 4.23 \times 10^{-13}$ s s^{-1}. Assuming that it is a neutron star of mass $M = 1.4\,M_\odot$ and radius $R = 12$ km gives $I \approx 1.4 \times 10^{38}$ kg m^2, which has to emit energy at a rate 6×10^{31} W to give the observed $\dot{P}$. Such good agreement strengthens the pulsar model. The Nebula is in a sense acting as a calorimeter for the energy lost by the pulsar. We can deduce that $B_p \approx 5 \times 10^8$ T (or about $10^{13} \times B(\text{Earth})$).

Note that the high spin frequency of pulsars is due to conservation of angular momentum of the core of the progenitor star during its collapse. The same may be true of the magnetic moment, although this is debatable. [Note that the magnetic moment of pulsars is not particularly high; Jupiter has a larger magnetic moment than many pulsars.]

Most of the power emitted is in the form of low frequency waves (the dipole frequency) which cannot travel far through the ionized interstellar medium. Their energy goes into accelerating fast particles which produce the blue nebula in the case of the Crab. A *small* fraction is emitted as the radio pulses (the precise location of the emission source around the pulsar is still disputed – the mechanism is coherent curvature radiation and synchrotron emission) and in some young pulsars a much large fraction of the power is emitted as gamma-rays. The nearest pulsar to the Earth (less than 100 pc away), which is the second brightest gamma-ray source in the Sky (after the Crab) and is known as Geminga, appears to emit mostly gamma-rays; its X-ray pulses have about 1% of the gamma-ray power. The optical source at its position has in turn a luminosity only 1% of that. The supernova that created Geminga must have been one of the most spectacular events to have been witnessed by early people (about 100,000 yr ago) and may have created the 'Local Bubble', a 100 pc size region of million K gas around the Sun.

Electric fields develop around spinning neutron stars which can be powerful enough to 'rip' electric charges from the surface so filling the surrounding *magnetosphere* with matter.

The electrodynamics of this have been solved only for the aligned case (*i.e.* when the magnetic dipole and spin axes are aligned, which can be thought of as an uninteresting case from an observational point of view!). There are open field lines extending to infinity near the poles, otherwise the field lines are closed (they do not extend beyond the light cylinder). To a poor approximation, looking in along an open field line we see a surface field B_p moving with velocity $v = \Omega R$. The consequent electric field,

$$E \approx v \times B \approx \Omega R B$$

yields a voltage of order $\Omega R^2 B$ which is about $10^{11} P^{-1}$ V for a typical value of B (say for the Crab pulsar). This pulls charges from the surface and causes a current flow in the magnetosphere (therefore modifying the field; the solution is therefore non-linear and thus difficult). In the general
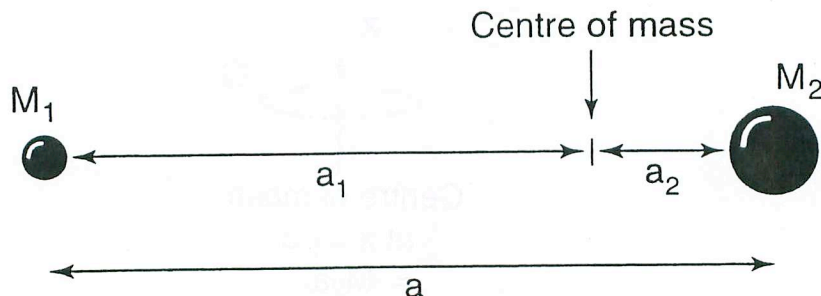
Figure 4: Definition of parameters for a binary star system.

case it is believed that a voltage gap exists above the surface of the neutron star within which the subsequently radiating charges are accelerated.

Note that the voltage decreases as the period increases. At a period of a few seconds the voltage is to low to create the situation necessary for radio pulse emission and the pulsar becomes undetectable (the pulsar 'dies').

Superimposed on the steady spindown of fast pulsars may be sudden changes in period, known as glitches. For a glitch in the Crab pulsar $\frac{|\Delta P|}{P} \sim 10^{-8}$, and for the Vela pulsar it may be $\sim 2 \times 10^6$. Glitches are probably due to crustquakes (release of stored elastic energy as the crust relaxes during spindown – the centrifugal force at the equator varies as P^{-2}). The period changes because the moment of inertia is altered by the rearrangement of crust in the quake, whereas the angular momentum of course remains unchanged. The slow recovery of the spin of the star after a glitch is because the sudden behaviour happens only in the crust and requires days to be transmitted to the superfluid interior, *i.e.* it shows the superfluid nature of the core.

Few pulsars are actually associated with supernova remnants. The youngest (*i.e.* smallest τ) are associated, although not all young supernova remnants have a central pulsar (perhaps due to a) a large proper motion of the pulsar, b) the supernova being of Type I – a white dwarf pushed over M_{Ch}), c) a black hole having formed, or d) the pulsar beam does not pass Earth). No pulsar has yet been seen in SN1987A; the luminosity of the expanding remnant excludes an early version of the Crab pulsar unless the magnetic field does not form until later.

10 BINARY STARS

About 70% of all stars in the Galaxy are members of multiple star systems (binaries being the dominant class). When a compact object occurs in a close binary system, mass transfer from the companion star can make it very luminous due to gravitational energy release as it accretes that matter. To understand this process and the interpretation of observations of binary systems involving compact objects we first outline some of the basic properties of binary systems.

10.1 *Mass Determination*

The inclination of the orbit is defined such that $i = 0$ when the orbit is on the plane of the Sky. The stars orbit about the centre of mass of the system; $a = a_1 + a_2$ and $M_1 a_1 = M_2 a_2$. Kepler's law gives

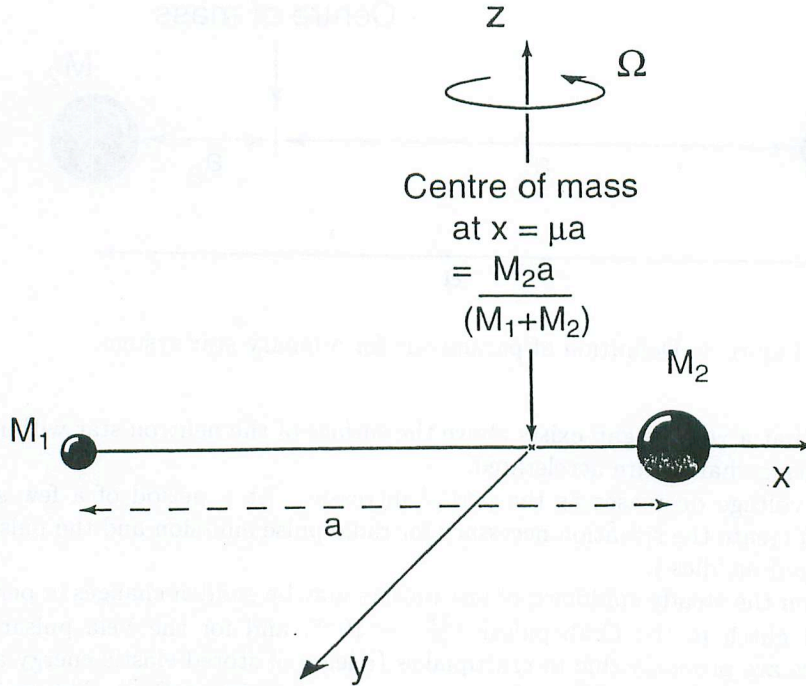$$\frac{G(M_1 + M_2)}{a^3} = \left(\frac{2\pi}{P}\right)^2 .$$

Figure 5: Definition of 3D parameters of a binary system.

Now if the projected orbital velocity of star 1 is measureable, say by a periodic doppler shift of its spectral lines, or of its pulse period (if it is a pulsar), with semi-amplitude

$$K_1 = \frac{2\pi}{P} a_1 \sin i$$

then the *mass function*

$$f_1 = \frac{(M_2 \sin i)^3}{(M_1 + M_2)^2} = \frac{P K_1^3}{2\pi G}.$$

If both K_1 and K_2 are observable, then the mass ratio M_1/M_2 can be obtained and separate masses as function of inclination. Further information (say from eclipses, if observed) is required to obtain the separate masses.

The process of *mass transfer* from one star to another was discovered through the *Algol Paradox*. The bright (naked eye) star Algol (β Per) is eclipsed every 2.9 days (discovered and interpreted over 200 yr ago by J. Goodricke). The brighter star is a typical hydrogen burning star of about $4\,M_\odot$, whereas the eclipsing one is an evolved (sub)giant star of about $0.8\,M_\odot$. The paradox arises because more massive stars evolve faster than less massive ones: so how can the more massive star be the less evolved one in the Algol system? The problem was resolved by Crawford and others in the 1960s. When the more massive star evolves and swells as it completes core hydrogen burning, its less massive companion tidally distorts and captures its envelope, with the consequence that the mass ratio is later reversed.

To understand this in more detail, consider the equipotential surfaces of a pair of orbiting point masses, including both gravity and rotation, from the point of view of a system of coordinates corotating with the binary.

The total potential at point (x, y, z), $\phi(x, y, z)$ = gravitational force from star 1 + gravitational force from star 2 + centrigugal potential, *i.e.*

$$\phi = -\frac{GM_1}{(x^2 + y^2 + z^2)^{1/2}} - \frac{GM_1}{((x - a)^2 + y^2 + z^2)^{1/2}} - \frac{\Omega^2((x - \mu a)^2 + y^2)}{2}.$$
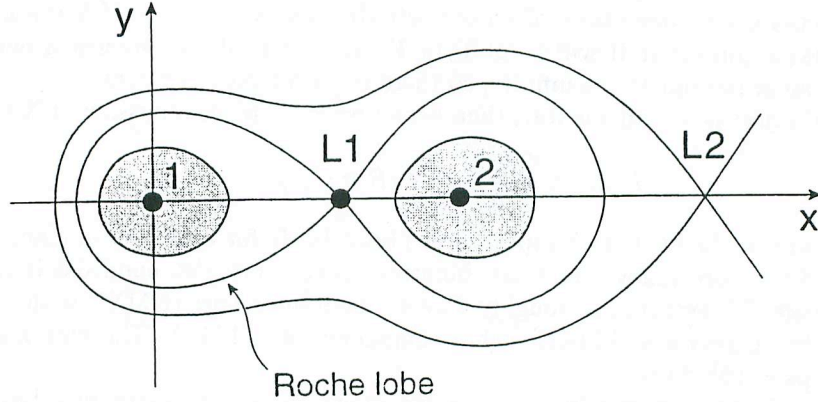
16

Figure 6: Equipotentials in the corotating frame of a binary system. The first equipotential common to both stars is called the Roche lobe and the common point between the stars, from which mass transfer occurs as one star swells to fill its Roche lobe, is called the inner Lagrangian or L_1 point.

The surface of a star will follow an equipotential, so as, say, star 1 evolves, it spills mass onto star 2 through the inner Lagrangian point L_1 when it expands to touch the common equipotential, called the Roche lobe. This process (sometimes called Roche lobe overflow) is one way in which mass transfer can proceed. (It can also occur if star 1 has a strong wind some of which is captured by star 2.)

There are 3 classes of close (interacting) binaries;

In Roche lobe overflow, angular momentum conservation causes the separation of the stars, a, to change. The total angular momentum $J = (M_1 a_1^2 + M_2 a_2^2)\Omega$ can be rewritten as

$$ J = \frac{M_1 M_2}{(M_1 + M_2)} a^2 \Omega. $$

So if the mass transfer is conservative ($\dot{J} = 0$, $\dot{M}_{total} = 0$ – both are unlikely in practice!) then

$$ \dot{J} = (\dot{M}_1 M_2 a^2 \Omega + M_1 \dot{M}_2 a^2 \Omega + 2 M_1 M_2 a \dot{a} \Omega + M_1 M_2 a^2 \dot{\Omega})/(M_1 + M_2) = 0. $$

Since from Kepler's law we find $\dot{a}/a = -2\dot{\Omega}/3\Omega$ then

$$ \frac{3\dot{M}_1(M_1 - M_2)}{M_1 M_2} = -\frac{\dot{\Omega}}{\Omega} = \frac{\dot{P}}{P}. $$

This means that $|\dot{P}| \propto \dot{M}$ and the sign depends upon whether $M_1 > M_2$ or not, e.g. if $\dot{M}_1 < 0$ and $M_1 < M_2$, then $\dot{P} > 0$ and the mass transfer is stable, but if $M_1 > M_2$, $\dot{P} < 0, \dot{a} < 0$, the orbit shrinks as mass is lost and $\dot{M}$ can increase $\Rightarrow$ rapid mass transfer. Mass is usually first transferred from the more massive star in a binary system and once started happens rapidly until the masses are equalized, after which it proceeds at a more stable rate (the details also depend on how the radius of that star responds to the mass loss). Mass loss and mass transfer in a binary can significantly change the orbital period, say from 100s of days to hours!

If one star in a close binary has evolved to become a white dwarf and the orbital period is a few hours ($a \lesssim R_\odot$!) then mass transfer can lead to the system appearing as a cataclysmic variable such as a dwarf nova. These are 'stars' that flare up optically by about a factor of 100 every few weeks. The mass transfer process is unstable and leads to 'waves' of matter accreting onto the white dwarf. Much of the power in such objects is gravitational.

Sometimes the accreted matter, which is hydrogen rich, can burn explosively and a bright outburst is seen. Such an object is known as a nova. If the matter blown off by nova explosions is

less than the accreted mass, then eventually the white dwarf may exceed the Chandrasekhar limit and a supernova explosion occur It will be of Type I). The released thermonuclear energy here is sufficient to blast all of the matter to infinity, so there is no compact remnant.

If the compact object is a neutron star, then accretion leads to the emission of X-rays.

$$L_{grav} = \frac{GM\dot{M}}{R} = 4\pi R^2 \sigma T_{surface}^4$$

where the accretion rate $\dot{M}$ leads to temperatures above 10^6 K for accretion luminosities greater than $L_\odot$. Such objects are known as X-ray binaries. X-rays are also emitted if the compact object is a black hole. These occur in roughly 2 types; high mass ones (HMXB with a companion of $\gtrsim 10\,M_\odot$) and low mass ones (LMXB with a companion of $< 1\,M_\odot$). The high mass ones are short lived (perhaps $< 10^{4-5}$ yr).

Note that X-ray binaries generally require a supernova to have occurred in a binary system. The exploding star has to be the less massive one if the binary is to remain bound (*i.e.* mass transfer must have occurred). Otherwise a runaway star (+ a high velocity pulsar) are formed. We shall soon discuss X-ray binaries in more detail.

11 Accretion

Accretion by compact objects accounts for some of the most spectacular, variable, powerful and energetically efficient objects known. The luminosity is usually just due to the release of gravitational energy. To accrete, matter must become bound to the star, which if it can be treated as a fluid leads to the concept of the accretion radius;

$$R_A = \frac{2GM}{v^2},$$

where v is the velocity of the gas with respect to the compact object, either due to its motion, or it represents the sound speed of the gas (accretion studies were pioneered by Bondi, Hoyle and others).

The accreting matter may fall in radially within R_A if there is little angular momentum (say from a stellar wind) or in the form of an *accretion disk* if there is much angular momentum (as happens in the case where Roche lobe overflow occurs). In the first case there may be a shock above the surface of the compact object at which the matter is decelerated from the free-fall velocity ($v_{ff} = \sqrt{2GM/r}$) and density

$$\rho = \frac{\dot{M}}{4\pi r^2 v_{ff}} \propto r^{-3/2}.$$

The matter behind the shock cools by emitting radiation and settles onto the surface. On a neutron star, $v_{ff} \sim c/3$, and the matter has to slam into the surface (or be decelerated by the outgoing radiation). For a black hole, spherical accretion can be inefficient since the hole can swallow matter and its kinetic energy and even, at high $\dot{M}$, the emitted radiation – there is no need for the matter to decelerate as happens when there is a solid surface.

In the case with angular momentum, the accreting matter may orbit the compact object. Friction between the inner parts of the resulting ring and its outer parts, which are at different velocities (Kepler's law) leads to the angular momentum being transported outward and the matter inward. It thus spreads into a disk around the object, with matter at the inner edge falling onto the object. The resultant luminosity is reduced from $L = GM\dot{M}/R$ by the kinetic energy that the matter has when it is accreted onto the star (at most a factor of 2 for a star corotating with the disc; for a black hole it corresponds to half the binding energy of the innermost stable orbit).

We can now summarize some properties of accretion onto compact objects in terms of
i) the mass⇒energy conversion efficiency

$$\epsilon = \frac{GM}{Rc^2} = \frac{R_S}{2R}$$

18

where R_S is the Schwarzchild radius of the compact object, *i.e.* $\sim 3 \times 10^{-4}$ and 0.15 for a white dwarf (WD) and neutron star/black hole (NS/BH), respectively, (possibly modified by the last consideration above for an accretion disk – $1/12$ for an accretion disc around a 'Newtonian' black hole; General Relativity makes a small correction necessary to this estimate).

ii) Minimum temperature of the emitting gas (equals the blackbody temperature)

$$T_{BB} = \left(\frac{GM\dot{M}}{4\pi R^3 \sigma} \right)^{1/4}.$$

For luminosities $L > 1000\,L_\odot$, $T_{BB} = 2 \times 10^5$ and 6×10^6 K for a WD and NS/BH, respectively. The black hole is assumed to be of Solar mass (note that $T_{BB} \propto T^{1/4}$).

iii) The Shock temperature

$$T_s = \frac{3}{16} \frac{GMm_p}{kR},$$

which is about 2×10^8 and 2×10^{11} K, for a WD and NS/BH, respectively.

The observed radiation will peak between photon energies of $T_B B$ and T_s, the distinction depending on how much matter is available to interact with the photons and 'degrade' the spectrum into a blackbody shape.

11.1 Accretion onto Magnetized Neutron Stars or White Dwarfs

Neutron stars and white dwarfs often have magnetic fields strong enough to channel an accretion flow when it is close to the surface. To understand what happens, we begin by assuming that the accretion flow is spherically symmetric, and that the compact object has a dipolar magnetic field. The magnetic field of the star then dominates the flow within a region of radius R_M, known as the magnetosphere, where the magnetic pressure (or stress) equals the ram pressure of the inflowing gas; $P_B > P_{flow}$. (This is the reason that the Earth's magnetic field causes a Solar-wind-free cavity to exist around the Earth.)

$$P_B = \frac{B^2}{2\mu_0} = \frac{m^2}{32\pi^2 R_M^6} = P_{flow} = \rho v_{ff}^2 = \frac{\dot{M}}{4\pi R_m^2} \sqrt{\frac{2GM}{R_M}}.$$

Therefore $R_M = m^{4/7} \dot{M}^{-2/7} (2G)^{-1/7} M^{-1/7} (8\pi)^{-2/7}$,

which, since $L = GM\dot{M}/R_\star$ (where $R_\star$ is the radius of the compact object) gives

$$R_M \propto L^{-2/7} m^{4/7}.$$

The powers of 2/7 are due to the magnetic pressure scaling as R_M^{-6} whereas the ram pressure scales $R_M^{-5/2}$, the balance thus yielding a $R_M^{7/2}$ dependence.

If the magnetized object is accreting from a disk, calculations show that $R_M(\text{disk}) \sim \frac{1}{2} R_M(\text{spherical})$. Also, until the object corotates with the inner edge of the disk, it accretes angular momentum, *i.e.* the object *spins up*. The sense of $\dot{P}$ is thus opposite from that of a radio pulsar, which spins down.

Many HMXB systems contain an X-ray pulsar with period ranging from 10s of ms to minutes. Observations reveal orbital periodic shifts in the spin period P, generally superimposed on a gradual spin up ($\dot{P} < 0$). These systems consist of a magnetized neutron star orbiting a massive star from which it is accreting matter, either from the star's wind or via Roche-lobe overflow. The pulses are due to the strong magnetic field funnelling the inflowing gas onto the polar caps of the neutron star, which are not aligned with the the spin axis. The pulse period is just the rotation period of the neutron star.

The specific angular momentum (*i.e.* per kg) of matter in a Keplerian orbit is $\sqrt{GMR}$, so
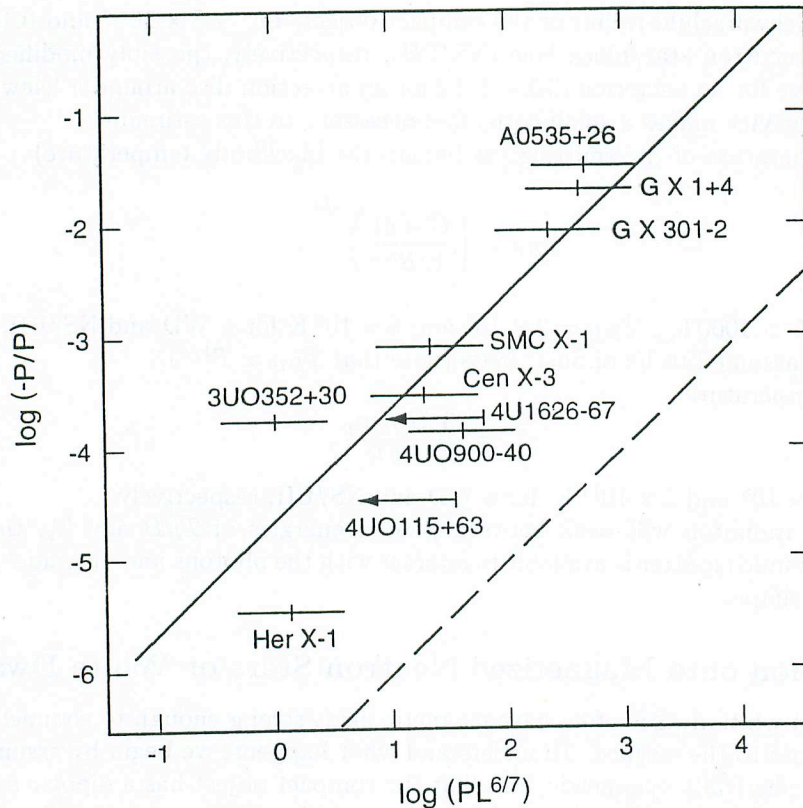
$$I\dot{\Omega} = \dot{M}\sqrt{GMR_M},$$

Figure 7: $-\dot{P}/P$ plotted against $PL^{6/7}$ for a number of accreting X-ray pulsars. The solid line is the expected relation for an object with the magnetic moment of a pulsar and moment of inertia of a neutron star; the dashed line is the similar relation for a white dwarf. Her X-1 lies below the neutron star line because it has spun up close to equilibrium (i.e. the magnetosphere is spinning at a rate close to the Kepler velocity of the inner disc).

$$\text{so} \quad \dot{P} = -P^2 \frac{\dot{M}\sqrt{GMR_M}}{2\pi I}.$$

Using the above scaling for R_M, we finally obtain

$$\frac{\dot{P}}{P} \propto \frac{-PL^{6/7}m^{2/7}}{I}.$$

The good agreement between data for binary X-ray pulsars and the above theory, assuming pulsar-like magnetic fields, shows that the accreting object is a neutron star, not a white dwarf. (White dwarfs have a moment of inertia about m_p/m_e times larger than that of a neutron star and so could not spin up as fast as observed.)

A good example of an HMXB is Centaurus X-3, which is in a circular orbit of period 2.087 day about a supergiant star. Pulses are observed with a period of 4.8 s and with $\dot{P}/P = -2.8 \times 10^{-4}$ yr^{-1}. Using observations of the eclipses, the estimated mass of the supergiant (from its luminosity), the orbital period and the doppler shift of the pulses, the mass of the neutron star can be obtained. From its rate of spin-up we can also obtain an estimate of its moment of inertia.

Another binary system with an X-ray pulsar is Hercules X-1, which in addition to showing pulses, orbital doppler variations and eclipses also has cyclotron lines in its X-ray spectrum which show that the polar magnetic field is about 3×10^8 T. As with many accreting objects, the X-ray emission is not very steady and shows many variations, some roughly periodic (perhaps due to

occultation by the accretion disk) and some chaotic in nature. The companion star to Her X-1 is about $2\,M_\odot$ and receives more power from the accreting X-ray source than it generates by its own thermonuclear burning. The face of the star closest to the X-ray source is therefore heated by X-ray irradiation to a temperature about 50% higher than the opposite face, leading to a well-observed orbital variation of the optical light curve (which was known before the X-ray source was first detected).

Most LMXB do not pulse. Also eclipses are rare (the companion star is usually small, of about $0.5\,M_\odot$, and subtends a much smaller solid angle at the X-ray source than is the case for the massive stars in a HMXB). They appear to be undergoing relatively stable mass transfer over rather long timescales (probably $> 10^8\,$yr). The lack of pulses is presumed due to a low surface magnetic field. Pulsar magnetic fields appear to decay on timescales greater than a million years, either intrinsically (as may happen with radio pulsars) or, if mass transfer has occurred, due to some effect of accretion (the details of magnetic field decay are uncertain). Many LMXB do show complex chaotic behaviour in their X-ray flux with the appearance of complex quasi-periodic oscillations (QPO), perhaps due in some cases to a beat between the spin frequency of the neutron star with the orbital frequency of the inner radius of the disk (*i.e.* at R_M).

11.2 X-ray Bursts

nuclear burning

Many LMXB show *X-ray bursts* in which the persistent X-ray flux rises by factors of 100s in a few seconds then decays back to the original level over the next few 10s of seconds. They often recur every few hours. The bursts are due to a thermonuclear flash triggered by explosive[1] Degeneracy pressure is important in the accreted matter just under the surface. Since this is independent of temperature, the temperature rise due to nuclear burning is not accompanied by a density decrease, as happens if thermal pressure dominates. The burning then takes place ever more rapidly. He-burning (H-burning involves the weak force and so cannot give rise to a 1 s flash) of the accreted matter. Nuclear burning gives about 6 MeV/nucleon whereas accretion gives about 200 MeV/nucleon. So we expect the ratio

$$\alpha = \frac{\text{ave. persistent flux}}{\text{time} - \text{ave burst flux}} \gtrsim \frac{200}{6} \sim 30.$$

The observed values of $\alpha \sim 50 - 200$ so the agreement is fair (especially if nuclear burning is incomplete in a burst). Bursts require that the polar field $B_p \lesssim 10^8\,$T, otherwise matter is channelled onto the poles at high density and burns quietly and continuously.

The surface is observed to cool after a burst as a blackbody. Using $L = 4\pi R^2 \sigma T^4$ and obtaining L and T from the observed flux and spectrum (and some knowledge of the distance to the LMXB) enables the radius to be found. Observations show that R is constant while L changes by a factor of about 100. (Note that General Relativistic corrections need to be applied for a precise determination of the radius.) If spectral features are observed (say some absorption line) then the surface gravitational redshift could be measured and thus the mass to radius ratio M/R and some constraints on the equation of state of neutron stars. So far no convincing spectral feature has been found. Bursts demonstrate that LMXB often contain neutron stars.

11.3 The Eddington Limit

The peak luminosity of X-ray bursts appears to be limited by radiation pressure. This can be so high that the surface layers are blown into space. The limiting luminosity at which the force due to radiation just balances the attraction due to gravity on matter is known as the *Eddington limit*, L_{Edd}. It is important in providing a mass-dependent limit to the luminosity of objects (both massive stars and especially to accretion flows). Any phenomenon lasting more than a few

---

[1]*

dynamical times (thus excluding supernovae) should have $L < L_{edd}$. The basic limit is obtained by comparing the force of radiation acting on an electron (through Thomson scattering of cross-section σ_T) at the surface of the object with the force of gravity on a proton there. The electron and proton, although assumed free, are bound together by electrostatic attraction.

$$F_{grav} = \frac{GMm_p}{R^2} = F_{rad} = \frac{L}{4\pi R^2 h\nu}\sigma_T\frac{h\nu}{c},$$

where the radiative force term is the flux of photons (of typical energy $h\nu$) times the cross-section times the momentum of a photon. This gives

$$L_{Edd} = \frac{4\pi GMm_p c}{\sigma_T},$$

which is about 10^{31} W per Solar mass. (Compare with the previous discussion on the mass limit of stars.)

12 Accreting Black Holes

Some luminous X-ray binaries, *e.g.* Cygnus X-1, LMC X-3, probably contain black holes accreting from the companion star. The basic evidence for a black hole in these systems is a high mass for the compact X-ray emitting object (usually based on spectroscopic evidence including doppler shifts from the companion) and a small size, based on rapid X-ray variability (say on timescales down to ms). Unless relativistic motions are involved, an object cannot appear to vary on a timescale shorter than its light crossing time, so if it appears to vary by a large factor (say 2) on a timescale shorter than Δt it must be smaller than $c\Delta t$.

12.0.1 Cygnus X-1

is the X-ray emitting companion to the supergiant star HDE226868. Its orbital period (from optical spectroscopy and some dips in the X-ray data) is 5.6d. $K_1 \sim 70\,\mathrm{km\,s^{-1}}$ and $f = 0.252\,\mathrm{M_\odot}$. Studies of stars similar to HDE226868 indicate that it must have a mass greater than $8.5\,\mathrm{M_\odot}$ so, from the mass function, the mass of the X-ray component $M_X > 3.3\,\mathrm{M_\odot}$. Assuming that no eclipses occur (so $\sin i$ is not 1), that HDE226868 does not fill its Roche lobe and that the distance to the system is greater than 6000ly leads to an improved estimate of $M_X \gtrsim 6\,\mathrm{M_\odot}$.

Cygnus X-1 shows a hard X-ray spectrum with a chaotic flickering light curve, unlike that seen from neutron star X-ray Binaries (and similar to Active Galactic Nuclei apart from timescale). Other objects, like the so-called hard transients, flare to very luminous levels then decay over the next few months. Later studies of their low-mass companions reveal large radial velocity amplitudes and imply a large masse for the X-ray emitting object in them. These objects are very strong candidates for being black holes.

Clear evidence that black holes have been observed in the Galaxy has not yet been obtained. Future X-ray spectroscopic studies of matter swirling in the accretion disc close to the central object in systems such as described above offer the best opportunity for now.

13 The Evolution of X-ray Binaries

Massive X-ray Binaries originate as a binary consisting of 2 normal massive stars closely orbiting each other. The more massive one uses up its core hydrogen first (in a few million yr), expands as it becomes a giant, transfers mass to the companion then becomes a supernova leaving a neutron star or black hole. Provided that it was the less massive star before the SN, the system remains bound, although the orbit may be eccentric. The neutron star may become a radio pulsar (which will be difficult to detect if the companion has a wind) which will spin down. Later the companion
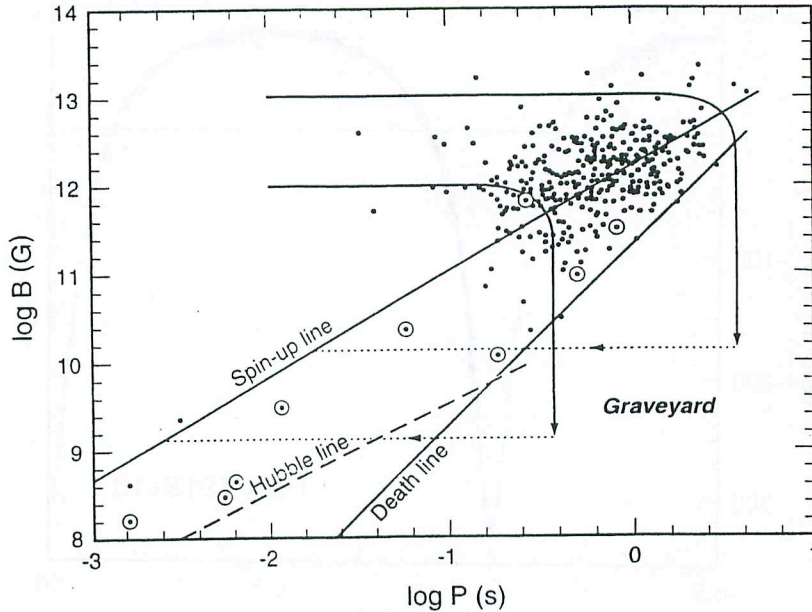
Figure 8: The evolution of magnetic field and pulse period for radio pulsars. See text for detailed description.

reaches the end of core H-burning and begins to expand and transfer mass back to the compact object. This is the HMXB phase. It does not last long since rapid mass transfer probably ensues and the compact object is smothered. The neutron star then spirals into the core of the bloated companion (due to viscous drag on its orbit). After the companion goes supernova, a binary pulsar may remain (*i.e.* a young radio pulsar orbiting an old one, or even a black hole). (The final result depends upon how much mass from the companion is lost during the spiral-in phase.)

The origin of LMXB is less clear. They are relatively plentiful in globular clusters where there may be about 100,000 stars within a radius of a few pc (the Galaxy has over 100 such clusters orbiting it). At such high star densities, close interactions between stars and even collisions are probable. A lone neutron star (say an old radio pulsar which has spun down to the point where the voltages at the poles are insufficient to drag charges into the magnetosphere) may tidally capture a low-mass star that strays close. Accretion in LMXB is so long-lived that over $0.1\,M_\odot$ can be accreted, spinning the neutron star up to very high frequencies (periods P close to 1 ms). This can only happen if the magnetic field is weak, due either to age or accretion, so $R_M \sim R_\star$ (most HMXB have a strong B which prevents spin up to P much below 1s since the inner edge of the disk is then in corotation with the star). When the neutron star has been spun up and the mass transfer rate starts to diminish, the emission of magnetic dipole and high energy radiation by the neutron star can now again become important. It may even cause the companion to evaporate. The object now becomes either a lone, or binary, *millisecond pulsar*. Several globular clusters contain many such millisecond pulsars, which have been essentially recycled by tidal capture and accretion (some have called it 'life after death' for pulsars!). Millisecond pulsars, and LMXB, also occur in the Galactic disk.

The points in the figure represent 403 observed pulsars ($B \propto \sqrt{P\dot{P}}$), the arrowed lines indicate possible evolution with the field decaying exponentially on a 5 million year timescale (note that this may only occur if accretion takes place). The 'death line' corresponds to a voltage at the pulsar surface below which detectable pulsar activity does not occur. The 'spin-up line' corresponds to the minimum period to which spin-up can occur in Eddington-limited accretion. The 'Hubble line' represents a pulsar spin-down age, $\tau = P/2\dot{P}$ of 10^{10} yr. The horizontal dashed lines show the
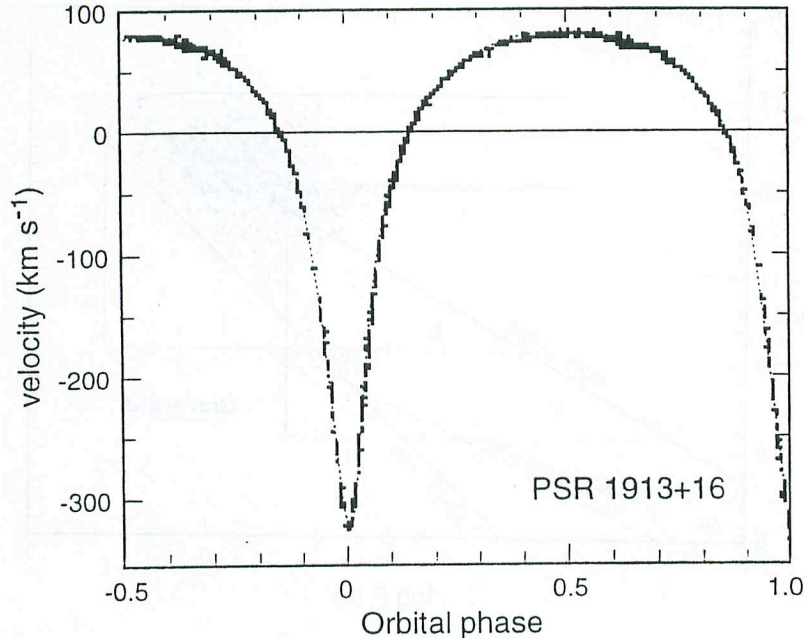
Figure 9: Radial velocity curve of the binary pulsar PSR 1913+16.

spin-up of a dead pulsar due to accretion. An example of evolution might be that a lone pulsar in a globular cluster spins is born with a strong field and spins down to a period of several seconds. It later captures a normal star into a close orbit about itself and subsequent evolution of that star causes mass transfer to ensue and accretion to take place. The field decays and the pulsar spins up to a period of a few milliseconds.

14 The Binary Pulsar

Many binary pulsars are now known. Some pulsars (*e.g.* the ms pulsars) also are very good clocks. ($\dot{P} \sim 10^{-19}\,\mathrm{s\,s^{-1}}$ – which gives an accuracy of better than 10^{-6} s over a timespan of 10 yr. This is substantially better than individual atomic clocks – but not if several clocks are bootstrapped together). The first binary pulsar, discovered in 1974 by Hulse & Taylor, is both. It is a radio pulsar which shows large doppler shifts on an 8 hr period. It is essentially a high precision clock orbiting with a velocity of $v \sim 300\,\mathrm{km\,s^{-1}}$ in an eccentric orbit close to a $\sim 1\,\mathrm{M_\odot}$ companion (thought to be an old neutron star). It provides the best test yet of several aspects of General Relativity.

The mass function is determined very accurately and since terms of the order of v^2/c^2 can be measured, the transverse Doppler shift and gravitational redshift (due to the companion) are also obtained. These lead to a new function relating M_1, M_2, a and the eccentricity e. The precession of periastron ($\dot{\omega} \sim 43$" per century for Mercury's orbit about the Sun) is 4.2 deg per year. This is related to the orbital parameters by

$$\dot{\omega} = \frac{6\pi G M_2}{a_1(1 - e^2)Pc^2}$$

and allows a complete solution of the binary parameters to be obtained. Both M_1 and M_2 appear to be close to $M_{Ch} = 1.4\,\mathrm{M_\odot}$.

The orbit is steadily tightening at a rate consistent with the emission of gravitational waves. To understand the emission of such waves, first consider electromagnetic waves where the dipole power
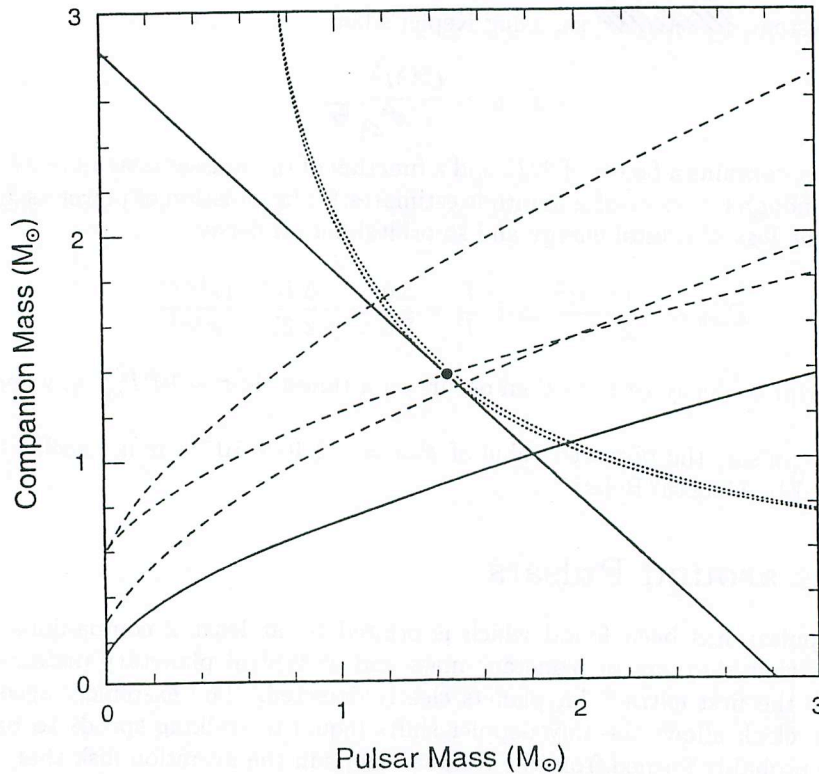
Figure 10: Binary pulsar mass.

emitted is proportional to the square of the acceleration of the emitting particles. Accelerating masses emit gravitational waves, but conservation of momentum precludes dipole radiation in this case (the acceleration of a mass in one direction must equal that of another mass in the opposite direction and the effects cancel – it works for electric charges since they can be of opposite sign). The simplest emission of gravitational radiation is quadrupole, with a power

$$L_{GW} = \frac{1}{5}\frac{G}{c^5} \left\langle \overset{\cdots}{I}\;\overset{\cdots}{I} \right\rangle$$

where $\overset{\cdots}{I}$ is the rate of rate of rate of change of the moment of inertia, approximately $MR^2/T^3 \sim Mv^3/R$ (from a dimensional argument). Thus

$$L_{GW} \sim \frac{c^5}{G}\left(\frac{R_S}{R}\right)^2 \left(\frac{v}{c}\right)^6.$$

Note that c^5/G is a very large number! Power of that order can however it can only be released by aspherical objects moving very fast within a region close to their Schwarzschild radii, so compact objects are required. For example, if

$$v \sim v_{ff}, \quad L_{GW} \sim (c^5/G)(R_S/R)^5.$$

A supernova explosion at the centre of the Galaxy, if the core collapse in an assymmetric manner, emits a large luminosity L_{GW}. However at Earth this would cause only a tiny displacement ($\lesssim 10^{-18}$ m or $10^{-3}\times$ the size of an atomic nucleus on a 1m bar, i.e. $\Delta\ell/\ell \sim 10^{-18}$). Gravitational Wave astronomy is difficult, but not impossible; such displacements are now measurable and much smaller ones are within the reach of instruments under development.

25

$$\ddot{I} \sim M a^2 \Omega^3$$

For a binary system, ~~$\ddot{I} \sim M a^2 \Omega^3$~~ so, using Kepler's law,

$$L_{GW} \sim \frac{G^4 M^5}{c^5 a^5}.$$

(The correct answer contains a factor of 32/5 and a function of the masses instead of M – the above result is accurate enough for order-of-magnitude estimates.) The emission of power as gravitational radiation caused the loss of orbital energy and so orbital period decay;

$$E_{orb} \sim -\frac{1}{2}\frac{GM^2}{a} \quad \text{and} \quad \frac{\dot{P}}{P} = \frac{3}{2}\frac{\dot{a}}{a} = -\frac{3}{2}\frac{\dot{E}}{E} \propto \frac{G^3 M^3}{c^5 a^4}.$$

The time for an orbit to decay $(a \Rightarrow 0)$ then occurs on a timescale $\tau \sim 10^5 P_{orb}^{8/3}$ s, where P_{orb} is in seconds.

For the binary pulsar, the observed value of $\dot{P}_{orb} = -2.40 \times 10^{-12}$ is in excellent agreement with that predicted by General Relativity.

15 Planets around Pulsars

One millisecond pulsar has been found which is orbited by at least 2 companions. The mass function indicates that they are of planetary mass and at typical planetary distances from the object. These are the first extra-Solar planets clearly detected. The extremely accurate nature of the pulsar as a clock allows the tiny doppler shifts (equal to walking speed) to be measured. The planets have probably formed from the residual debris in the accretion disk that spun up the pulsar to a ms period. If planets can form in such a hostile environment then they probably form easily in protostellar disks.

$$L_x \approx 10^{37} \text{ erg/s}$$

16 A 'bizarre' binary system – SS433

$$L_{kinetic} \approx 10^{40} \text{ erg/s}$$

This object displays emission lines at optical and X-ray wavelengths that move around the spectrum, indicating simultaneous radial velocities of about $\pm 50,000 \text{ km s}^{-1}$, on a period of 164 days! The central system is a 13d binary system. The moving emission lines originate in jets of plasma ejected in opposite directions at a velocity of 0.26c from the system. Somehow the jets undergo precession on the 164d period. Radio observations show a 'corkscrew' picture of blobs emitted from the central object. The transverse Doppler effect (i.e. time dilatation) is obvious in the spectra. The mass loss rate in the jets is $\sim 10^{17} \text{ kg s}^1$ so the kinetic power is very high $\gtrsim 10^{32} \text{W}$. How and why the system operates is not known in detail. Other jet emitting systems have been found in our Galaxy.

17 Gamma-Ray Bursts

The US Vela satellites launched in the 1960s to monitor the treaty banning nuclear tests in space carried X-ray and gamma-ray detectors designed to detect flashes of radiation. Using the data from these satellites, R. Klebesadel and his colleagues discovered brief energetic gamma-ray flashes which originate beyond the Solar system. Rough positions and a minimum distance to the source of a burst could be deduced from the time at which it was detected at separate satellites.

The flux of gamma-rays in a burst can be very high (comparable to the optical flux received from Venus when at its brightest). Bursts are often highly structured in time, perhaps with many peaks and on scales down to ms or less. They are presumed to originate from compact objects, probably neutron stars. Recent observations with the Compton Gamma-Ray Observatory (in a near Earth orbit) shows at least one burst per day and an overall brightness distribution suggesting that the
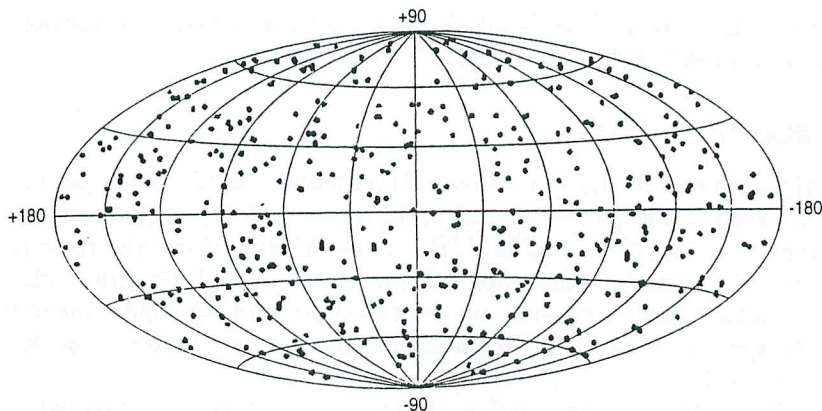
Figure 11: Positions of gamma-ray bursts detected by the BATSE instruments on NASA's Compton Gamma-Ray Observatory. Galactic coordinates are used in which the Galactic Centre is at (0,0) and the Galactic Plane (Milky Way) lies along the line of zero latitude. Note the uniform distribution of the bursts.

source locations in space are inhomogeneous – there are too few faint ones for a homogeneous distribution in space. Curiously they also appear to be isotropic, with no obvious concentration of bursts to any particular part of the Sky. This appears to rule out simple models in which the bursts are from old Galactic neutron stars, where a concentration towards the Galactic Centre is expected. Currently favoured models suggest that they are at cosmological distances (say billions of light years) and are due to the merger of a pair of binary neutron stars (this should eventually happen to the Binary Pulsar). Such events should create a fireball with a suitable power. Details of the gamma-ray flickering and spectrum within a burst have not yet been worked out. The merger of 2 ten km size objects may create a flash easily detectable across the whole Universe!

18 ACTIVE GALACTIC NUCLEI

18.1 Basic observations

In some galaxies, a compact central region outshines all the $\sim 10^{11}$ stars in the surrounding galaxy. Luminosities observed are $10^{36} - 10^{40} W$ (i.e. up to at least $2.10^{13} L_\odot$). The observed radiation from these galactic nuclei spans many different wavebands, and is not from ordinary stars. The luminosity sometimes varies on timescales shorter than a day, and this sets an upper limit to the size of the 'central engine'. The radiation in the optical and infrared bands does not all come directly from the 'central engine' – it may have been absorbed and re-emitted by surrounding clouds of gas or dust occupying a larger volume (but still very small compared with the entire galaxy). Active galactic nuclei (AGNs) span a wide range of properties, and astronomers have classified them in rather confusing ways. The lower-luminosity categories include the so-called Seyfert galaxies; the more powerful objects are called 'quasars'. Quasars are so bright that many have been observed at such great distances that their host galaxy cannot be seen. Some (but not all) AGNs are strong radio sources.

Quasars are specially important for cosmology because they serve as probes for the high-redshift universe – some have such large redshifts that the Lyman alpha line of hydrogen (121.6 nm), normally in the far UV, is shifted into the red end of the visible spectrum (around 700 nm).

Strong emission lines in the optical and UV are common in many AGN due to photoionization of surrounding cool gas. Much of this gas lie within a pc of the central engine and moves at high velocities ($> 1000\,km\,s^{-1}$) and so doppler broadens the emission lines – the region is known as the

Broad Line Region or BLR. Strong broad emission lines were one of the first means for detecting AGN and still remain a powerful detection tool.

18.2   Interpretations of AGNs

Many details of AGNs are still poorly understood. Most models attribute the power output from AGNs (or at least from the more powerful ones) to accretion onto a black hole of $\sim 10^8 \, \mathrm{M}_\odot$.

Accretion luminosity is simply $L = GM\dot{M}/R$,, where M and R are the mass and radius of the central object and the mass accretion rate is $\dot{M}$, provided that all the gravitational energy is released. This is not so for a black hole, since much of the kinetic energy and some of the radiation emitted by the infalling gas is swallowed by the hole. In this case calculations show that $L = \epsilon \dot{M} c^2$, where the efficiency $\epsilon \sim 0.1$.

The way in which this power is radiated is, however, still uncertain. The physical situation may resemble Cygnus X1 (a black hole candidate in our Galaxy, with mass around $10 \, \mathrm{M}_\odot$) but the luminosity is higher by 10^7 and variations are slower by a similar factor.

There are some simple scaling laws which relate some of the processes around supermassive holes to those around stellar-mass holes (eg Cygnus X-1). One important relation concerns the *Eddington limit*, L_{Edd}. It is important in providing a mass-dependent limit to the luminosity of objects (both massive stars and especially to accretion flows). Any phenomenon lasting more than a few dynamical times (thus excluding supernovae) should have $L < L_{edd}$.

Suppose the accretion rate (and hence, for equal efficiencies ϵ, the luminosity) scales with M. The luminosity is then the same fraction of L_{Edd} in both cases, implying that (for instance) the relative importance of gravity and radiation pressure would be similar. The orbital timescale for material in the innermost stable orbit around a black hole is about

$$10^{-4} \left(\frac{M}{\mathrm{M}_\odot} \right) \text{ sec.}$$

(the exact value depends on whether the black hole is non-rotating (Schwarzschild) or spinning (Kerr)). This sets a characteristic timescale for the variability of radiation emitted near the hole, which scales with M.

Suppose there were an accretion disc around the hole (the infalling matter is likely to have enough angular momentum for it to go into orbit before reaching the black hole, viscosity in the disk then allows it to spiral in, losing energy to radiation as it does so). The surface area of the disc goes as M^2. The emission per unit area then scales as $\dot{M}/M^2 \propto M^{-1}$; since the power radiated per unit area goes as T^4, the characteristic black-body temperature then scales as $M^{-\frac{1}{4}}$. For supermassive holes, the primary thermal output would therefore be in the UV, rather than X-rays (though there is generally also strong X-ray emission from hotter optically thin gas, probably in a 'corona' surrounding the hole). Some of the UV is absorbed by clouds of gas further away from the hole (at typical distances $\gtrsim 10^4 r_s$), and reprocessed into emission lines, as in a gaseous nebula.

The motions in the inward-spiralling material near the hole involve speeds of $\gtrsim c/2$. Magnetic fields threading this material probably accelerate particles to relativistic speeds, rather as they do in pulsar magnetospheres. Characteristic field strengths are 1 T. These magnetic fields and relativistic particles are the dominant constituents of the relativistic jets and radio-emitting regions (discussed later).

[Historical note. John Michell (1784) pointed out that an object with the same density as the Sun but 10^8 times the Sun's mass would create such a strong gravitational attraction that not even light could escape from it.]
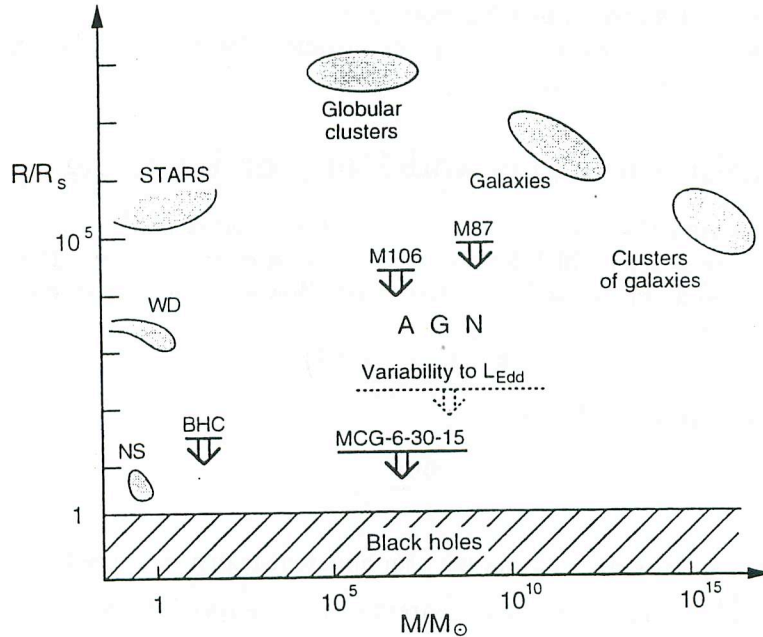
Figure 12: Radii of cosmic objects measured relative to the Schwarzschild radius for that mass (R_S. The downward pointing arrows are observational constraints obtained for the active nuclei of the galaxies M106 (from doppler shifts in water maser emission lines in a surrounding torus), M87 (from optical emission lines in a surrounding disc), MCG–6–30–15 (from X-ray iron line emission), Black Hole Candidates (e.g. Cygnus X-1, from variability measurements) and AGN in general (also from variability and Eddington limit arguments).

19 Why should we take the existence of massive black holes seriously?

Gas would naturally tend to accumulate in the centre of any galaxy (the bottom of potential well). A supermassive star may form, but this is unstable and quickly collapses to black hole. A massive star might alternatively build up by collisions and coalescence of stars in a dense central cluster. Any black hole that formed would subsequently grow by accretion of gas, or even of entire stars.

The quasar phenomenon requires a power-production mechanism with high efficiency. If spectral lies were emitted from the surface of a disc, we would expect a characteristic two-peaked profile. This is now detected for X-ray emission lines and shows the expected gravitational redshift as well as doppler shifts.

20 Jets and radio sources

Another distinctive feature of AGNs is that they often involve not just thermal radiation from hot gas, but also material in more exotic forms – magnetic fields and relativistic electrons which emit synchrotron radiation. This is channelled into jets, which squirt from the centre at close to the speed of light. (These radio loud objects are about 10% of all AGN and generally occur in elliptical galaxies – radio-quiet objects are generally in spiral hosts.)

It is not clear exactly how the jets are produced. Possibilities include (i) 'nozzles' (ii) channelling along evacuated vortex near rotation axis of a thick disc. (iii) collimation by 'coiled-up' magnetic fields carried by wind from disc.

In some cases, the jets penetrate to very great distances (larger even than the entire host galaxy), to create an extended radio source.

21 Superluminal motions and Doppler beaming

There is good evidence that the jets flow out at close to the speed of light.

Suppose that, at a time zero a blob leaves the centre of a distant galaxy. Then after time t_0 the blob will have moved a transverse distance $vt_0 \sin\theta$. However, the corresponding time in the observer's frame is

$$\Delta t = t_0 \left(1 - \frac{v}{c}\cos\theta\right).$$

The apparent transverse speed is therefore

$$\frac{v\sin\theta}{1 - \frac{v}{c}\cos\theta}.$$

So a blob moving at a speed close to c can have an apparent transverse speed that is as large as γv, where $\gamma = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}}$. The maximum apparent transverse speed occurs when $\sin\theta = \gamma^{-1}$.

For example, if $\gamma = 10$, and the blobs are moving at 6 degrees (0.1 radian) to the line of sight, the apparent transverse speed would be almost 10c. Motions of this kind have been discovered in many radio quasars, using the technique of Very Long Baseline Interferometry (VLBI) – this technique allows angular resolutions of much better than a milli-arc-second to be achieved, corresponding to only tens of lightyears even at the Hubble distance.

Extreme superluminal motions are only expected if the jet is almost pointing at us (to within an angle $\sim \gamma^{-1}$). This does not mean that we would expect the phenomenon to be only rarely observable, since the jets that are pointing towards us would emit much more radiation towards us ('Doppler boosting'). If a blob emits radiation with a spectrum $P(\nu) \propto \nu^{-\alpha}$, the observed intensity at ν scales as $D^{3+\alpha}$, when

$$D = \frac{1}{\gamma\left(1 - \frac{v}{c}\cos\theta\right)}$$

is the doppler shift. Since the doppler boosted sources (which are precisely the ones oriented in the way required for superluminal effects) appear much more intense, superluminal effects can therefore occur in a large fraction of a radio source sample selected from any survey that is complete down to a given apparent radio intensity. It is therefore not surprising that a substantial fraction of known radio quasars should display superluminal motions.

22 Beaming and 'unified models' of AGNs

The observed one-sidedness of jets in double radio sources is probably a consequence of doppler favouritism rather than intrinsic asymmetry — the approaching side would be much stronger than the receding side. (The extended diffuse radio lobes, due to plasma which has been shocked where the jets encounter the external medium, should be more symmetrical because they are not moving so fast.) Some astronomers have proposed more elaborate unified models, where there is bright optical emission directed along jet. The well-known jet in the elliptical galaxy M87 would look very bright if the jet were pointing towards us. (Objects know as BL Lacs are thought to be in this class. Intense gamma-ray blazars are an extreme example of looking down a jet.)

23 Quasar lifetimes and demography.

The characteristic timescale for doubling the mass of an object radiating at the Eddington luminosity is $L_{\text{Edd}} = 4\pi GM \, m_p c/\sigma_T$ is

$$M/\dot{M} = M/\left(L_{\text{Ed}}/\epsilon c^2\right) = \frac{c\sigma_T\epsilon}{4\pi Gm_p} = 4\times10^8\epsilon \quad \text{years.}$$

For an efficiency $\epsilon \simeq 0.1$ this is $\sim 4\times10^7$ years.
[Note: this lifetime would be even shorter if efficiency were as (relatively) low as it is for nuclear fusion, with $\epsilon = 0.007$ (cf the lifetime of massive main-sequence stars where radiation pressure is important and $L \approx L_{\text{Edd}}$).] This characteristic lifetime for powerful AGNs is 100 times shorter than the Hubble time, a fact that has important implications for the number of dead quasars.

The number of quasars observed is about 1 per cent of number of galaxies. But if the active lifetime of a quasar were indeed only 1 per cent of the Hubble time, there would be 100 generations of quasars. It is therefore possible that every galaxy goes through a 'quasar phase' after it forms. (Warning: this is just a 'plausibility argument' — the issue of quasar lifetimes is still a matter of lively controversy!) Dead quasars would be massive black holes lurking in the centres of galaxies, quiescent because they are not being fuelled by accretion. These might therefore be expected even in nearby galaxies. The orbital speed of a star at distance $r \, (\gg r_s)$ from a black hole is

$$v_{orb} \simeq c \left(\frac{r}{r_s}\right)^{-\frac{1}{2}}$$

The characteristic speeds of stars in the central bulge of a galaxy are 100-300 km/sec (i.e. $3.10^{-4} - 10^{-3}$ c). Therefore a central black hole would substantially modify the motions of any stars within $10^6 - 10^7 r_s$ of the hole.

Studies of the Andromeda Galaxy (M31) indicate that the stars in its central regions are 'feeling' the gravitational pull of a dark mass of $3\times10^7 \, M_\odot$. Occasionally a star would pass so close to the hole that it was tidally disrupted. The resultant "flare" when the debris was swallowed would be a good diagnostic of the hole's properties. This requires it to get within the "tidal radius" r_T, where

$$\frac{GM}{r_T^3} \simeq \frac{Gm_*}{r_*^3} \simeq G\rho_*$$

(this is a 'Newtonian' argument which of course breaks down if the predicted value of r_T is $< r_s$. If $M \gg 10^8 \, M_\odot$, the hole would swallow stars like the Sun without disrupting them.)

24 CONSTITUENTS OF THE UNIVERSE

24.1 Galaxies

Galaxies are typically classified as either Spiral, Elliptical or Irregular. The stars in a galaxy are in orbit around the common centre of mass; in a spiral galaxy the orbits are ordered and all in the same sense whereas they tend to be more randomly oriented in ellipticals. The Milky Way and our nearest large neighbour, M31, are spiral galaxies; our two small satellite galaxies, the Magellanic Clouds are irregulars. The largest galaxies are ellipticals (there is also a class of dwarf ellipticals) and the nearest to us lie in the Virgo Cluster, about 15 Mpc away (about 20 times the distance to M31).

Typical radii (say, within which half the light is contained) of galaxies lie in the range of $1 - 10$ kpc,
masses lie in the range of $10^8 - 10^{12} \, M_\odot$,
and luminosities $\sim 10^8 - \text{few} \times 10^{11} \, L_\odot$,

so that the mass-to-light ratio *for the optically visible part* $M/L \sim 1 - 10 \, M_\odot / L_\odot$.

Low mass galaxies are the commonest but most of the total mass of galaxies is in the rarer luminous ones. The mean smoothed out density of stars in galaxies is less than 1% of the closure density of the Universe. (Note that it is possible that there is a significant population of galaxies which either have surface brightnesses too low to have been detected or so bright and compact that they have been mistaken for stars.)

Irregular galaxies are the most numerous and often the smallest galaxies. S0 (lenticular) galaxies are like spirals but without cold gas and are common in clusters.

The variation of the surface density of galaxies with brightness (called number counts) does not at faint visual magnitudes agree with the expectations from simple cosmology. In a Euclidean Universe the number of galaxies increases with volume as r^3 and the brightess decreases as r^{-2}, making the number N brighter than a given apparent flux S vary as $S^{-3/2}$. In an expanding Universe we expect a flatter dependence of N on S due to the effect of redshift etc. The observed counts maintain a steep relation, suggesting that a) there are more, or b) they were brighter, at higher redshifts. There are problems with both a and b since in a) where have the galaxies gone? (there are problems with mergers) and b) we expect to see the galaxies to greater distances, but do not.

Galaxies do interact and merge with each other, but just how much and to what extent this is a serious influence in the evolution of an average galaxy is unknown. The thin disk of our galaxy indicates that any major merger (ie with an object exceeding 10% of its mass) happened a very long time ago during the likely formation phase of the galaxy.

24.2   Clustering of galaxies

Most galaxies are found near other galaxies. About one half of all galaxies are in groups of 2 – 10 large members (the Milky Way is part of the Local Group which has about 12 members of various types and sizes; the MW, M31 and M33 being the dominant ones) and about 1% are in rich clusters with 100s of members. The typical radius of a group or cluster is about 1 Mpc.

One measure of the clustering of galaxies is the correlation length, which is defined as the radius from a galaxy within which the probability of finding another galaxy is double that expected if galaxies were entirely randomly distributed in space. It is about 5 Mpc and the probability varies with radius as a power law of index 1.7.

The large scale structure (i.e. the spatial distribution of galaxies) of the Universe out to several 100 Mpc has been mapped through measurements of the velocities of many galaxies. Assuming that the redshifts translate straight into distance through the Hubble relation (i.e. the galaxies are part of the 'Hubble flow') the results convert into a 3D picture of the galaxy distribution. Various structures apart from the obvious well-known clusters are apparent, such as the Local Void. The process has been taken further by using the various distance estimation methods to obtain relative distances to a fair numbers of galaxies, thereby determining departures from a smooth Hubble flow. Such departures, or peculiar velocities, are presumably due to the gravitational effect of the underlying mass distribution, which can thereby be mapped (with some assumptions).

About 1% of all bright galaxies occur in rich clusters of galaxies, with 30 or more members. Clusters and groups of galaxies contain enormous quantities of diffuse gas. It consitutes about 25% of the totals mass of a cluster – 5 times the mass of all the detectable stars ~~into~~ the member galaxies. The sound speed of the gas is similar to the velocity dispersion of the cluster which is typically 500 to 1200 $\mathrm{km\,s^{-1}}$ and so the gas predominantly radiates X-rays. Diffuse X-radiation is the principal source of information on the intracluster medium (ICM). There is further indirect evidence for the gas in 'head-tail' radio sources and from theories of the propagation of double-lobe radio sources.

Most of the observed intracluster gas has an electron density, n_e, in the range of $10^{-4} - 10^{-2} \, \mathrm{cm^{-3}}$ and a temperature $T \sim 10^7 - 10^8 \, \mathrm{K}$, and is contained within a radius of 1 to 2 Mpc. The total bremsstrahlung luminosity is $\sim 10^{42} - 10^{46} \, \mathrm{erg\,s^{-1}}$. The 6.7 keV iron emission line is
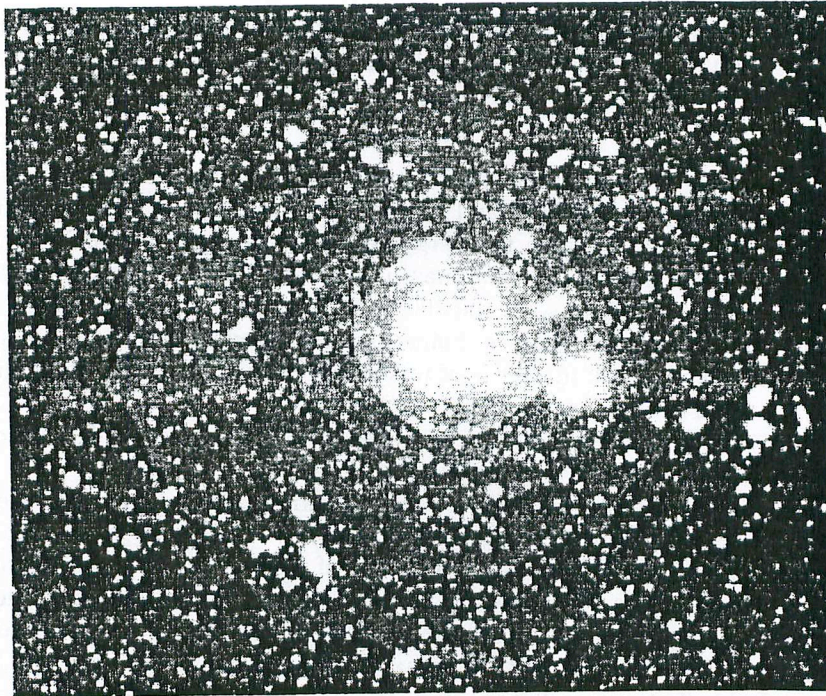
Figure 13: Superimposed X-ray and optical images of the Perseus cluster of galaxies.

commonly observed and on average the gas has ~ 0.4 times solar abundance in iron. X-ray emission lines from other elements (e.g. Ne, Si, S) indicate that the enrichment was primarily from Type II supernovae.

Some cluster properties show correlations which are fairly good if just X-ray quantities are used (e.g. X-ray luminosity L_X versus gas temperature T_X). The temperature correlates with optically-determined velocity dispersion (which is a notoriously uncertain quantity) and with the central galaxy number density. Generally, the deeper the potential well, the more gas and galaxies it contains and the more luminous it is in X-rays. The gas is at a temperature close to that given by the Virial Theorem, i.e.

$$kT_{gas} \approx \frac{GM_{cluster}m_p}{R}.$$

24.3   DARK MATTER

Estimates of the mass of stars in a cluster, from the product of the total visible luminosity and the mass-to-light ratio of a typical galaxy yield values of $M_\star \sim 10^{13} - 10^{14} \, M_\odot$. This conflicts with dynamical mass estimates obtained from the Virial Theorem;

$$M \approx v^2 R/G,$$

where v is assumed to be the dispersion in the radial velocities of the member galaxies,

$$M_{dyn} \sim 10^{14} - 10^{15} \, M_\odot \approx 10 M_\star.$$

$$i.e. \ M_{cluster} \gg \Sigma M_{galaxies}.$$

The implication is of *dark matter* in between the galaxies in a cluster (first suggested by Zwicky in the 1930s). (It is sometimes called 'missing mass' but a better description is 'missing light'.)

Some (but not all) of the dark matter in clusters is detectable in X-rays since there is an intracluster medium (ICM) of hot gas spread throughout a cluster. The temperature of the ICM

$$T \sim 10^8 \, \text{K} \approx v^2/k,$$

i.e. the sound speed in the gass is comparable to the velocity dispersion of the member galaxies. This is a consequence of the Virial Theorem and suggests that the high sound speed and velocity dispersion (typically $1000 \, \text{km s}^{-1}$) are due to energy release in the gravitational collapse during cluster formation. Up to about 30% of the total mass in clusters is in the hot ICM when averaged over the whole cluster, but only 10% or so in the 500 kpc radius core, where the gas is insufficient to gravitational bind itself and/or the galaxies. Something else, as yet undetected, must dominate the mass there. The bulk of this dark matter is of unknown origin. It could be some fundamental particle(s) which is a relic of the very early universe.

M/L for rich clusters is 100–200; the total mean density of such clusters is only comparable with that of galaxies (due to the rarity of clusters).

Dark matter is also inferred to occur in many large galaxies. Measurements of the rotational velocity of stars (and at large galaxy radii, of gas) about the centre of spiral galaxies do not always drop with galaxy radius as expected from the application of Kepler's law and if most of the mass is distributed like the visible light. The curves often continue at roughly constant velocity which implies that the mass is rising roughly in proportion to the galaxy radius, whereas the total light within that radius does not. In large elliptical galaxies, the mass at large radii can be inferred from the temperature of the hot interstellar medium seen in X-rays and again the mass is seen to be much higher than can be accounted for by the detectable stars alone. In all cases it seems that large galaxies are surrounded by massive haloes of dark matter.

The mean mass-to-light ratio of the Universe depends on the mass density. If it equals the closure density, *i.e.* $\rho = \rho_{crit}$ so $\rho/\rho_{crit} = \Omega = 1$, then $M/L \sim 1500 \, \text{M}_\odot/\text{L}_\odot$. Most of the mass of the Universe is then dark matter and what we actually see with telescopes is only a few per cent.

A comparison of the abundance of helium, deuterium etc. that are formed early in the evolution of the Universe (within the first 3 min) by cosmic nucleosynthesis (see later) with the observed abundances indicates that the mean baryon density is about 5 per cent of the closure density. The baryons observed in galaxies and clusters account for only about 1% of the closure density meaning that there are also a lot of 'missing baryons'. The most likely place for them to lie is in an intergalactic medium.

It is curious that the fractional mass of a cluster which is in baryons – gas – in the ICM is about 0.3, much higher than in any other large objects (*e.g.* galaxies) and is a higher ratio than expected from cosmic nucleosynthesis and a closure density (~ 0.05).

25 GRAVITATIONAL LENSES

The distribution of mass – and therefore of dark matter – can sometimes be mapped through its effect on bending light from an object lying behind the mass. Masses can thereby act as 'gravitational lenses' which distort our view of background objects, occasionally magnifying them by large factors so that an otherwise undetectable galaxy, say, becomes easily visible.

In a gravitational field, light is bent by twice the 'Newtonian' value (which was first estimated by Söldner in 1801);

$$\alpha = \frac{4GM}{bc^2} = \frac{2R_S}{b}.$$

$\alpha = 1.75(M/\text{M}_\odot)(b/\text{R}_\odot)^{-1}$ arcsec, first measure by Dyson, Eddington et al. in 1919 (they compared photographs of the star field close to the Sun during a total eclipse with one taken at night 6 months previously). Parallel rays of light grazing the Sun are therefore brought to a focus at the edge (1000 AU) of the Solar System (note that rays with different impact parameters, b, have different foci, so the 'lens' is non-linear).
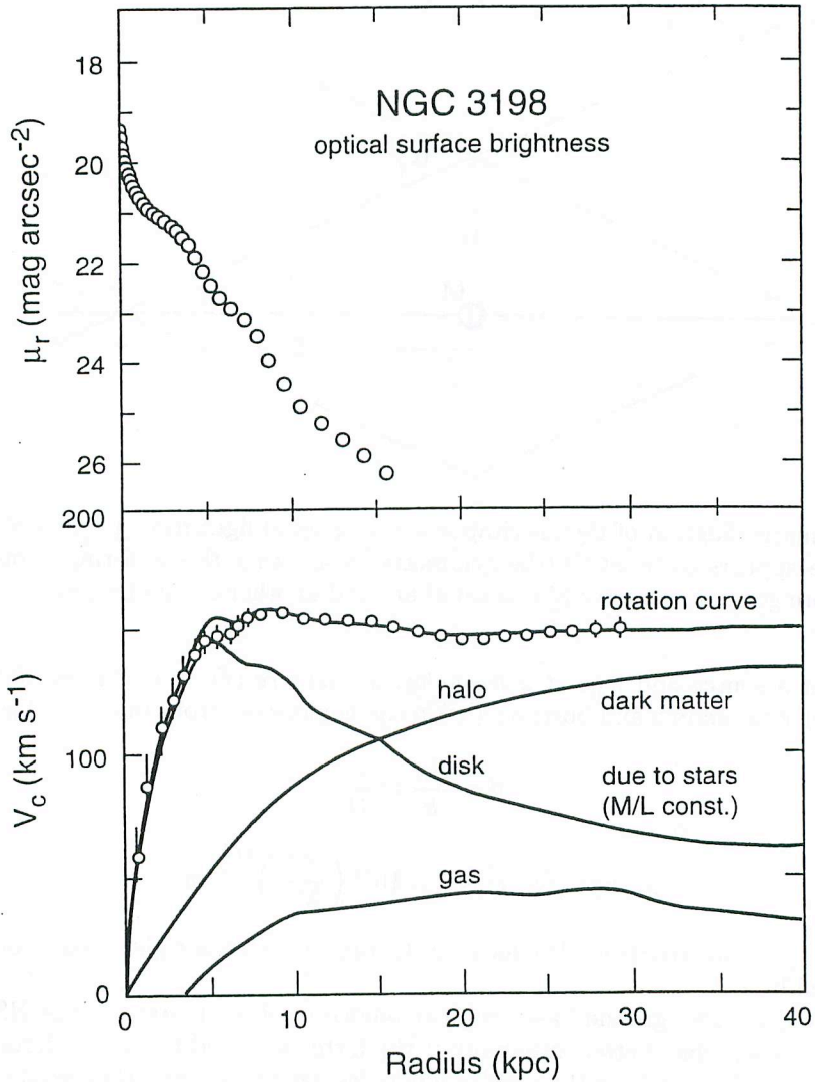
Figure 14: Evidence for the presence of dark matter in spiral galaxies. The upper panel shows the optical surface brightness of the spiral galaxy NGC 3198 (magnitudes are a logarithmic measure of the flux of light). The lower panel shows the observed radial velocity curve (dots) and the expected curve due to stars, assuming a constant mass-to-light ratio. Also shown is the contribution due to gas (note that this adds quadratically to that from the stars). A dark matter distribution as indicated is required.
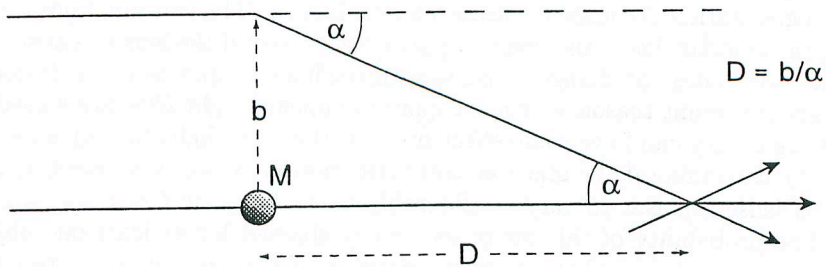

Figure 15: Schematic diagram of the gravitational focussing effect of mass M on a parallel beam of light.
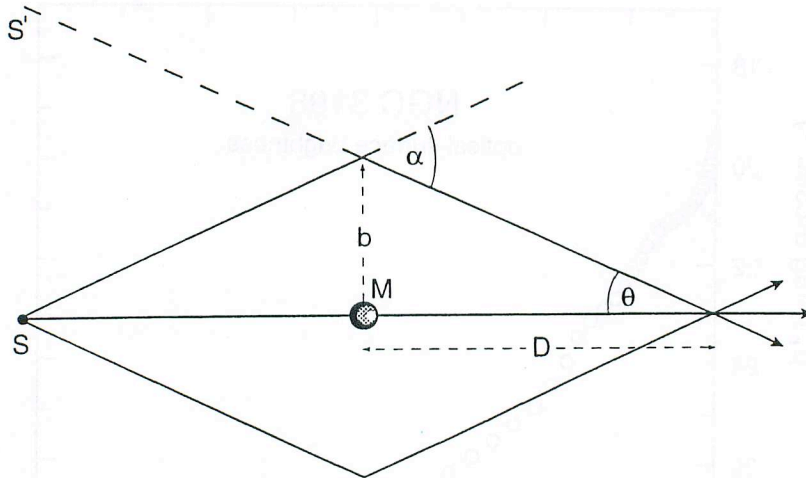
Figure 16: Schematic diagram of the gravitational focussing of light from a source S at a distance $2D$. The source appears to be at S' (the symmetry means that this is a ring about the lensing object; for a more general geometry S' is a small arc and another occurs below).

Now consider a source and lens at a *cosmological distance* ($D \sim 10^{10}$ ly $= 10^{26}$ m) and for simplicity consider the source and observer (us) to be equidistant from the lens. Then

$$\theta = \frac{R_S}{b} = \frac{b}{D}$$

$$\text{so} \quad b = \sqrt{R_S D} \sim 5 \times 10^{14} \left(\frac{M}{M_\odot} \right)^{1/2} \text{m}$$

and $\theta \sim 1(M/M_\odot)^{1/2}$ microarcsec. The lens can (if the geometry is right) then produce a pair of images separated by $\sim \theta$.

To be resolved by any ground-based optical telescope, $\theta \sim 1$ arcsec, (the HST and radio telescopes can do somewhat better, especially if the latter are used as an interferometer), which means that $M \gtrsim 10^{12} M_\odot$, at least the mass of a massive galaxy. We also then need the mass to lie within a radius $b < 10^{20}$ m at the lens! Extensive dark haloes are not therefore so good at making lenses, what is needed is a large mass within a relatively small radius (in galaxy terms).

The first gravitational lens was discovered by Walsh et al in 1979. They found 2 quasars separated by 6 arcsec with identical spectra – the Double Quasar. Studies show that these are 2 images of the same quasar at $z = 1.4$ lensed by a massive galaxy in a (small) cluster along the line of sight. Only if the lens is symmetric and aligned with the object would we actually see a ring (the 'Einstein ring') – one radio source has this structure.

Over a dozen gravitationally-lensed quasars are now known. The lens amplifies the total brightness by making the angular size of the images appear larger than if the lens was absent (the surface brightness of an object does not change on passing through a transparent lens). It does not appear that lensing is an important reason why most quasars appear bright (few are lensed). Individual stars in the lensing galaxy can have some effect by perturbing the light rays by a few microarcsec. This effect is only detectable if the star has tranverse motion (velocity v) when it can cause the brightness of the lensed quasar to vary considerably on a timescale $t \sim b/v$ – this is known as microlensing. The probability of this being seen (it is claimed for at least one object) depends on the star density in the lens and the mass distribution of the stars, it is one test for dark matter composed of low-mass stars (brown dwarfs that are not massive enough to ignite hydrogen burning).
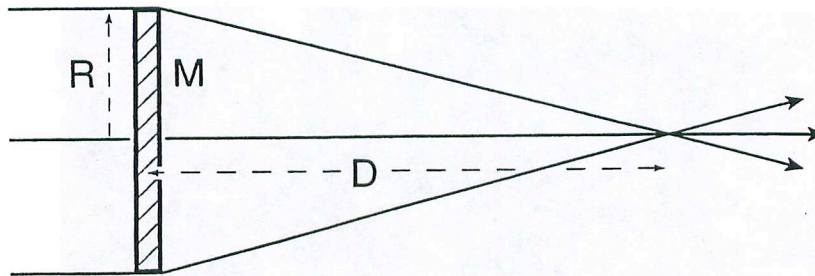
Figure 17: Lensing by a distributed mass. Only that mass M within the focussed beam causes the focussing.

Microlensing does not require the objects to be at cosmic distances to be detectable. Redoing the above expressions for an object of mass M half the way to the Large Magellanic Cloud (10^5 ly) gives $\theta \sim 2 \times 10^{-9}(M/M_\odot)^{1/2} \sim 400(M/M_\odot)^{1/2}$ microarcsec. This means that the fraction of one square degree of the LMC being lensed is $\sim \theta^2 \sim 10^{-14}(M/M_\odot)$. Now if the halo of our Galaxy is composed of $10^{12} M_\odot$ in objects each of mass M then in one square degree that will roughly be $10^7(M/M_\odot)^{-1}$ of them so the probability of any particular star in the LMC being lensed is the product of these quantities or 10^{-7} and since there are about 10^7 stars per square degree in the LMC there should be one star being lensed at any given time. This is detected because the halo object should be moving and thus the LMC star will brighten then dim symmetrically (and achromatically) on a timescale $\sim b/v \sim 10^7(M/M_\odot)^{1/2}$ s. Several such events have been seen by the MACHO and EROS collaborations now monitoring the LMC. Curiously they see many lensing events towards the Galactic Centre, suggesting that there is aelongation or bar there pointing in our direction.

The different light paths to the separate images are of different lengths so any intrinsic variation in the brightness of a lensed quasar will be seen first in one image and then in the other. If the lens can be accurately modelled then the difference in light paths gives an estimate of the distance to the lens independent of the Hubble constant. The redshift of the lens can then be used to measure H_0.

25.1 Extended gravitational lenses

Consider a self-gravitating object of size R causing a deflection $\alpha = 4GM/Rc^2$. From the Virial Theorem $v^2 \sim GM/R$, so $\alpha \approx 4v^2/c^2$. Large angular deflections thus require objects with large internal velocities ($> 300\,\mathrm{km\,s^{-1}}$ for 1 arcsec). These are relatively rare objects. Rich clusters of galaxies have $v \sim 1000\,\mathrm{km\,s^{-1}}$ and $R \sim 3 \times 10^6$ ly, so can we expect $\alpha > 10$ arcsec? What is needed is that $R > b$, so for a cluster to act as a lens, it must have a compact core.

We require

$$R^2 < b^2 = \frac{4GM}{c^2}D$$

so surface mass density in a lens has to exceed

$$\frac{M}{\pi R^2} = \Sigma_{lens} \gtrsim \frac{c^2}{4\pi GD}.$$

For cosmological distances (say $D \sim c/H_0$) this is $\sim 1\,\mathrm{g\,cm^{-2}} \sim 10\mathrm{kg\,m^{-2}}$. It is similar to the surface density of your hand or of a typical paperback book.

(Note also that it is similar to the mean surface density in a universe of closure density, $\rho_{crit}c/H_0$. An Einstein-de-Sitter Universe is essentially the ultimate gravitational lens. The probability that a class of objects gives rise to lensing along a random line of sight is proportional to their fractional contribution to the closure density, $\rho_\ell/\rho_{crit} = \Omega_\ell$. Thus since galaxies have
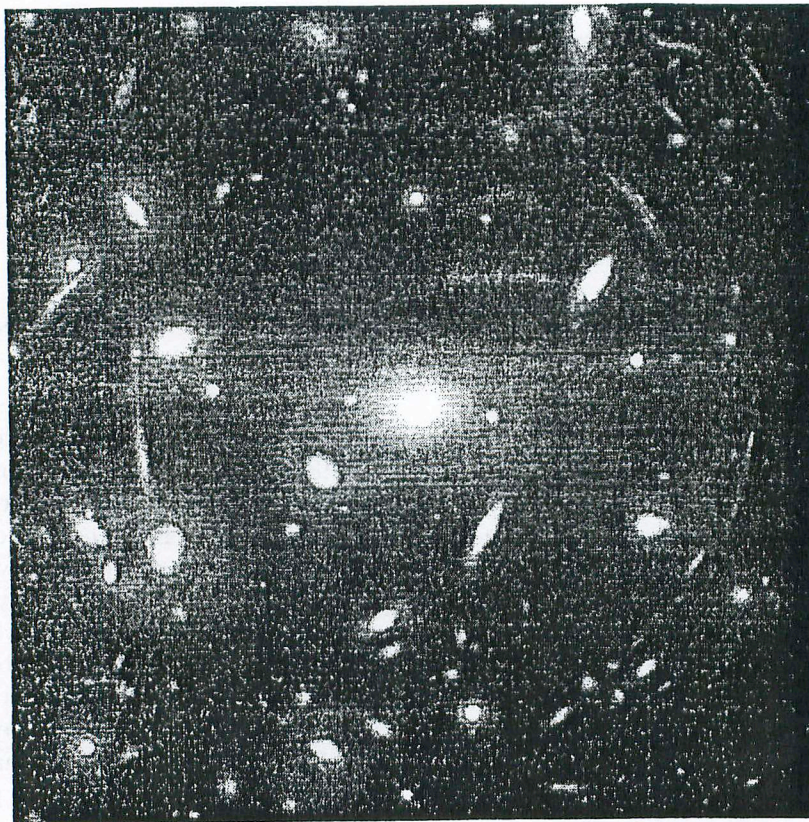
Figure 18: An optical image of the core of the cluster A2218 from the Hubble Space Telescope. Many fine arcs due to gravitational lensing of distant galaxies are seen.

$\Omega_{gal} \sim 0.01$. there is somewhat less than a 1% chance that lensing occurs due to a galaxies along a random l.o.s.)

Clusters which have such high central surface densities produce 'arcs' of light, first discovered in 1986 by Soucail et al. and by Lynds & Petrosian. These are images of distant galaxies that have been magnified by large amounts and so become visible (spectra of the arcs show much higher redshifts than the cluster. The cluster cores are acting as 'giant telescopes' with apertures of 300,000 ly! (In some sense, the Universe itself acts as the ultimate gravitational telescope.)

As noted previously, gravitational lenses are highly non-linear and do not give simple images like a convex spectacle lens does. The effect from a complex gravitational field is more similar to that produced by the surface of a swimming pool on sunlight illuminating the bottom of the pool, with complex bands of light and dark (caustics) seen. We, the observer on the bottom, see an bright image if we lie on a bright band. (To see caustics, look at the bottom of the inside of a cylindrical mug illuminated by a strong light.)

TOPICS IN CONTEMPORARY ASTROPHYSICS – APPENDIX 1

January 23, 1996

1 BLACK HOLES

The Schwarzschild solution for the metric around a point mass is

$$ds^2 = -(1 - \frac{2GM}{c^2 r})c^2 dt^2 + (1 - \frac{2GM}{c^2 r})^{-1} dr^2 + r^2(d\theta^2 + \sin^2 d\phi^2).$$

(r is here defined so that $2\pi r$ is the circumference of a circle centred on the mass.)
A static observer (constant r, θ, ϕ) measures proper time

$$c^2 d\tau^2 = -ds^2 = (1 - \frac{2GM}{c^2 r})c^2 dt^2$$

so

$$\frac{d\tau}{dt} = (1 - \frac{2GM}{c^2 r})^{1/2}.$$

Thus gravitational redshift

$$(1 + z_g) = (1 - \frac{2GM}{c^2 r})^{1/2}.$$

This approximates to $GM/c^2 r$ at large radii where $r \gg GM/c^2$.
Note that $z_g \to \infty$ as $r \to \frac{2GM}{c^2 r} = R_S$.
$R_S = \frac{2GM}{c^2 r}$ is known as the Schwarzschild radius and has the value $3M/M_\odot$ km.
Note that the same result can be obtained from Newtonian gravity assuming that R_S is the radius for a body of mass M such that the escape velocity from the surface equals the velocity of light (first discussed by Michel 1784 and Laplace 1798).

A collapsing object would not experience anything peculiar in falling through the radius R_S, although it would meet a singularity at the centre where conventional physics breaks down. From the outside it would appear to take an infinite time to fall within R_S (the redshift becomes infinite). It does however rapidly go black since the object only emits a finite amount of light in falling within R_S.

Orbits about a black hole are calculated from geodesics. A simple approach (justified only since it gives correct answers) is to calculate orbits in a Newtonian manner except use a pseudo-Newtonian potential

$$\Phi = \frac{GM}{R - R_S}.$$

This shows that innermost bound orbit is at $2R_S$ and the innermost stable orbit is at $3R_S$.

Black holes have limited complexity (to the outside observer). They can have only mass, charge and spin. The solution for the spinning black hole is due to Kerr. The maximum spin occurs when the angular momentum per unit mass $a = GM/c$. Bound orbits about a maximally spinning black hole occur all the way down to $R = GM/c^2$ if they are corotating.

All inward going photons emitted from within $R = 3GM/c^2$ are captured by a non-spinning black hole.

2 ACCRETION

The behaviour of the gravitational infall, or accretion, of gas onto a mass, M, can only be calculated analytically in a few specialized cases. They are pure spherical accretion in which the mass is at rest in an infinite gas at temperature T: line accretion where the mass is moving at velocity v through gas of negligible temperature and disk accretion where the gas has some initial angular momentum and accretes in a plane. In all cases the thermal or bulk motion acts to stem or channel the accretion.

2.1 Spherical accretion

Assume that the gas acts as a fluid so that we can write down the force (Euler's) equation at radius r from the mass;

$$u\frac{du}{dr} + \frac{1}{\rho}\frac{dP}{dr} = -\frac{GM}{r^2}$$

where u and P are the velocity and pressure. Now the sound speed $c_s^2 = dP/d\rho$, so this can be rewritten as

$$u\frac{du}{dr} + \frac{c_s^2}{\rho}\frac{d\rho}{dr} + \frac{GM}{r^2} = 0.$$

Continuity (conservation of mass) means that at any radius

$$\frac{1}{r^2}\frac{d}{dr}(r^2\rho u) = 0.$$

so

$$\frac{2}{r} + \frac{1}{u}\frac{du}{dr} + \frac{1}{\rho}\frac{d\rho}{dr} = 0.$$

Eliminating $d\rho/dr$ gives

$$\frac{du}{dr} = \frac{(2c_s^2/r - GM/r^2)}{(u - c_s^2/u)}.$$

This becomes infinite when $u = c_s$ unless at that radius $2c_s^2/r = GM/r^2$. The radius $r_s = GM/2c_s^2$ marks the transition to supersonic flow and is known as the sonic radius.

To proceed further we need to integrate the force equation

$$\int u\,du + \int \frac{dP}{\rho} + \int \frac{GM}{r^2}\,dr = 0$$

and adopt an equation of state $P \propto \rho^\gamma$;

$$\frac{1}{2}u^2 + \frac{c_s^2}{\gamma - 1} - \frac{GM}{r} = \text{const} = \frac{c_\infty^2}{\gamma - 1}.$$

This is the energy equation, KE + Enthalpy - PE = constant. The last step is found by setting $u = 0$ when $r = \infty$.

At the sonic radius we now find

$$r_s = \frac{GM}{4c_\infty^2}(5 - 3\gamma).$$

Using $(\rho_s/\rho_\infty)^{(\gamma-1)} = (c_s/c_\infty)^2$ in the mass flow rate equation (from continuity) $\dot{M} = 4\pi r_s^2 \rho_s u_s$, we obtain

$$\dot{M} = 4\pi \frac{(GM)^2}{c_\infty^3} \rho_\infty \frac{1}{4} \left(\frac{5-3\gamma}{2}\right)^{\frac{(5-3\gamma)}{2(\gamma-1)}}.$$

The factor after ρ_∞ above is approximately unity so $\dot{M} \approx 4\pi R_A^2 \rho_\infty c_\infty$ where the accretion radius

$$R_A = \frac{GM}{c_\infty^2}.$$

Note that the above rate in which we have assumed that collisions make the gas act as a fluid exceeds that for a collisionless cloud of particles on pure ballistic orbits by a factor of $\sim (c/c\infty)^2$. This can be an enormous factor for interstellar gas accreting onto a black hole. Collisions limit the increase in tangential motion of particles without changing much their radial velocities, since they fall in together.

Provided that the accretion radius exceeds the radius of the central mass, within R_A, $u \approx u_{ff} = \sqrt{2GM/r}$, and from continuity $\rho \propto r^{-3/2}$.

2.2 Accretion Discs

Here we assume that the gas has significant angular momentum so that it goes into orbit about the mass. Particles at neighbouring radii have different speeds, so friction occurs and will act to transport angular momentum outward and matter inward. In a steady state the matter will form a disk stretching down to the central mass. We shall assume that the orbits are Keplerian (tangential velocity v_ϕ much greater than radial velocity v_r).

The angular momentum transfer rate $\dot{J} = \dot{M} \times$ (specific angular momentum in Keplerian orbit) $= \dot{M}\sqrt{GMr}$.

The tangential velocity $v_\phi = r\Omega = \sqrt{GM/r}$.

The tangential stress s_ϕ due to viscosity,

$$s_\phi = \eta \times \text{(velocity gradient acted on by viscous force)}.$$

and the velocity gradient

$$\frac{dv_\phi}{dr} = \Omega + r\frac{d\Omega}{dr}.$$

Only the last term is relevant here since the Ω term represents uniform rotation and is necessary to prevent slippage.

So

$$s_\phi = \eta r \frac{d\Omega}{dr} = -\frac{3}{2}\eta\Omega$$

in a Keplerian disk.

The torque due to the viscous force is the rate of change of angular momentum, so taking the total torque at radius r

$$s_\phi (2\pi r 2h)r = \dot{j}^+ - \dot{j}^- = \dot{M}(\sqrt{GMr} - \beta\sqrt{GMr_*}),$$

where the $+$ and $-$ superscripts represent the flow in and out, h is the semithickness of the disk, and allowance has been made for a fraction β of the angular momentum at r_* being absorbed by the star (for a black hole $\beta \approx 1$). The last 2 equations enable us to eliminate s_ϕ from the relation between η and $\dot{M}$ giving the velocity gradient

$$\frac{3}{2}\Omega = \frac{\dot{M}(\sqrt{GMr} - \beta\sqrt{GMr_*})}{4\pi r^2 h\eta}.$$

The energy production rate per unit volume is the product of the force per unit area (i.e. stress) and the velocity gradient acted on by that force,

$$\dot{Q} = \frac{3}{2}\Omega \times \frac{\dot{M}(\sqrt{GMr} - 3\sqrt{GMr_*})}{4\pi r^2 h\eta},$$

independent of η.

If the energy is radiated locally (assuming that the disk is then so $h \ll r$) then the surface flux

$$F(r) = \frac{1}{2}2h\dot{Q} = \frac{3\dot{M}}{8\pi r^2}\frac{GM}{r}\left(1 - \beta\left(\frac{r}{r_*}\right)\right),$$

taking into account the fact that the disk has 2 sides.

The total luminosity $L = \int 2F \times 2\pi r\, dr = (\frac{3}{2} - \beta)\frac{GM\dot{M}}{r}$, as required by the Newtonian binding energy. If the central mass is a black hole so that $\beta = 1$ then half the total gravitational energy is liberated, the remainder is swallowed by the hole (the matter falls in with its orbital energy).

If the disk is optically thick to radiation and acts as a blackbody then

$$T(r) = (F(r)/\sigma)^{1/4}.$$

It can be about 10^7 K for an accreting stellar mass black hole.

Approximating blackbody emission as a δ-function at frequency ν given by $h\nu = kT$, then the spectral luminosity

$$\frac{dl}{d\nu} = 2\pi r F(r)\frac{dr}{d\nu} \propto \frac{r}{r^3}\frac{dr}{dT} \propto \frac{\nu^{-7/3}}{r^2}.$$

Since $r \propto \nu^{-4/3}$ then

$$\frac{dL}{d\nu} \propto \nu^{1/3},$$

up to ν_{max} corresponding to the temperature of the innermost radius of the disk.

So much can be deduced without knowing the origin of the viscosity. Molecular viscosity such as is relevant in the air we breathe is too small to account for the large accretion rates inferred for many cosmic objects. Turbulence and magnetic fields are likely to be involved there. Since the dimensions of viscosity are density, velocity and length, then a turbulent disk probably has $\eta \propto \rho v_{\text{turb}}\ell$, where v_{turb} and ℓ are some characteristic turbulent velocity and length scale. Now $v_{\text{turb}} < c_s$ or the gas will shock, and $\ell < h$, so $\eta \sim \rho c_s h$.

Perpendicular to the disk, hydrostatic equilibrium gives

$$\frac{1}{\rho}\frac{dP}{dz} = -\frac{GM}{r^2}\frac{z}{r}$$

so approximating $dP = P$ and $dz = h$ we have

$$h \sim \left(\frac{P}{\rho}\right)^{1/2}\left(\frac{GM}{r^3}\right)^{-1/2} = \frac{c_s}{\Omega}.$$

Thus $\eta \approx \rho c_s^2 \Omega \approx P/\Omega$ and $s_\phi \sim P$. Since this is a maximum, it is generally written as $s_\phi = \alpha P$ with $\alpha < 1$, giving what is known as an 'α- disk'.

Disks may be subject to many instabilities and overall it does appear that accretion is a highly unsteady process. It is not yet understood in detail.