

Lecture notes in High Energy Astrophysics

Claes Fransson

January 21, 2007

Contents

1	What is high energy astrophysics?	5
2	Equations of stellar structure	5
3	The equation of state	7
4	The Chandrasekhar mass	10
5	Solar neutrinos	13
5.1	Predicted fluxes	13
5.2	Neutrino experiments	14
6	Pre-supernova Evolution of Massive Stars	21
6.1	Helium burning	21
6.2	Low mass versus high mass evolution	21
6.3	Advanced nuclear burning stages	23
6.3.1	Carbon burning	23
6.3.2	Neon burning	23
6.3.3	Oxygen burning	24
6.3.4	Silicon burning	24
6.4	Nuclear statistical equilibrium	25
6.5	Neutrino cooling	26
6.6	Duration of the burning stages	28
6.7	Mass loss	29
6.8	Evolution in the HR diagram	31
6.9	Structure before explosion	32
6.10	Triggering of collapse	34
6.11	Neutrino trapping	36
6.12	Collapse	37
6.13	Energy losses in the shock	38
6.14	Neutrino heating	39
6.15	Explosive nucleosynthesis	42
7	Observables of core collapse supernovae	46
7.1	Neutrinos from SN 1987A	46
7.2	Supernova classification	48
7.3	Radioactivity	53
7.4	Light curves	56
7.4.1	The diffusion phase of the light curve	56

7.4.2	The late light curve and radioactive isotopes	57
7.4.3	The bolometric light curve of SN 1987A	59
8	Type Ia supernovae	60
9	Pulsars	66
9.1	What are the pulsars?	66
9.2	Pulsar slowing down (Longair p. 100-104)	66
9.3	Electrodynamics of the pulsar	68
9.4	The magnetosphere. Longair p. 107-108	69
10	Neutron stars. Longair 15.3.3	71
10.1	Crust	71
10.2	Core	72
10.3	Binary pulsars	84
10.4	The binary mass function. (Longair p. 113)	87
11	Hydrodynamics	88
11.1	The equations of fluid dynamics	88
11.2	The shock jump conditions	91
12	Viscosity	95
12.1	The Navier-Stokes equation	96
12.2	The energy equation for a viscous fluid	97
13	Spherical accretion	99
14	Accretion efficiency	102
15	The Eddington luminosity	103
16	Accretion disks	104
16.1	Vertical structure	104
16.2	Radial structure	105
16.3	Energy loss from the disk	108
16.4	Viscosity	110
16.5	α -disks	112
16.6	Thick disks. Skip this section!	115
17	Magnetospheric accretion	118
18	Spin-up by accretion	121

19 Characteristics of X-ray Binaries	124
20 Wind accretion	127
21 Cyclotron radiation	129
22 Non-thermal emission	132
23 Energy losses of relativistic electrons	132
24 The equation for the diffusion losses	135
24.1 Example	136
24.2 The minimum energy argument	137
25 The expansion of a point explosion	139
26 Collisionally ionized plasmas	142
27 Emission processes	145
28 Supernova remnants	147
28.1 The adiabatic stage	147
28.2 The radiative stage	149
29 Hot gas in clusters of galaxies	151
29.1 The Sunyaev-Zeldovich effect	155
30 The Fermi mechanism	156
31 Gamma-ray Bursts	161
31.1 Historical overview	161
31.2 Summary of observations	163
31.2.1 Prompt phase	163
31.2.2 Afterglow phase	165
31.3 The necessity of relativistic expansion	167
31.4 General scenario for the prompt and afterglow emission.	169
31.5 Afterglow spectra	174
31.6 Jet steepening	176
32 GRB Progenitors	178
32.1 The supernova - GRB connection	178
32.2 The GRB environment	182

32.3	Collapsars	182
32.4	Neutron star mergers	186
33	Active Galactic Nuclei	189
33.1	Classification	189
33.2	Regions	189
33.3	Emission lines	189
33.4	Superluminal expansion	189
33.5	The broadline region	190
33.6	Jets	191
33.7	Evidence for black holes in AGNs	193
	33.7.1 Fe $K\alpha$ lines from accretion disks	193
	33.7.2 The Galactic Center	194
34	Pair production	197
A	Equations of relativistic hydrodynamics	198
B	Relativistic shocks	200
C	Relativistic blast waves	202

1 What is high energy astrophysics?

High energy astrophysics has many different interpretations. In the most narrow sense this is the type of observations involving high energy photons, primarily X-rays and gamma-rays. From a more physical point of view one usually means the study of objects which involve either extreme conditions, like high energies, temperature or densities, or photons and particles with high energies. Therefore, one includes such objects as cosmic rays, which traditionally was the first area of high energy astrophysics, high energy neutrinos, X-rays, gamma-rays, from the 'detector' point of view. From a more astrophysical point of view this includes supernovae, supernova remnants, neutron stars, black holes, binary X-ray sources, gamma-ray bursts, active galactic nuclei, radio jets, clusters of galaxies. In addition to these fairly exotic objects also more ordinary objects like ordinary stars and galaxies are also emitters of non-thermal radio emission and X-rays. Also the neutrinos from the sun are usually included. Some areas like high energy neutrino astronomy and gravitational waves are not mature enough to be useful as diagnostics of these objects, but will probably be extremely useful in a few years. This also applies to different types of dark matter detection, where e.g., the GLAST gamma-ray satellite will be extremely interesting for detection of gamma-rays resulting from e.g. dark matter annihilation in the Galaxy or from discrete sources like the sun or the galactic center.

Classes of objects: Stellar remnants, active galactic nuclei, clusters of galaxies,

Supplement to Longair Chap. 13 and 14

2 Equations of stellar structure

See e.g., Kippenhan and Weigert

Mass conservation

$$\frac{dM}{dr} = 4\pi r^2 \rho(r) \quad (2.1)$$

Hydrostatic equilibrium

$$dF = 4\pi r^2 [p(r + dr) - p(r)] = -4\pi r^2 \rho(r) \frac{GM(r)}{r^2} \quad (2.2)$$

or

$$\frac{dp}{dr} = -\frac{GM(r)\rho(r)}{r^2} \quad (2.3)$$

Energy conservation

$$\frac{dL}{dr} = -4\pi r^2 \epsilon(r) \quad (2.4)$$

where L is the luminosity (total energy loss per unit time) and ϵ is the energy generation rate per volume.

Finally we need an equation describing the transport of energy from the center to the surface. Suppose first that the energy transport is by diffusion of the photons. In the interior the gas is so optically thick that the mean free path, λ , is small compared to the dimensions of the system (e.g., the star). The photon will then perform a random walk, or diffusion. The radiation field will therefore be almost, but not completely, isotropic. We will now derive a diffusion equation describing the flow of radiation as a result of the small, but finite temperature gradient.

Consider an almost isotropic radiation field, as is the case in the interior of a star, and let us calculate the energy flowing through a surface area dA from a direction θ to the radial direction. The projected area is then $\cos\theta dA$. The photons will on the average come from a distance λ equal to the mean free path. The fraction of photons coming from this direction is $2\pi \sin\theta d\theta/4\pi$. The total energy through dA will then be

$$dE = 1/2 \sin\theta d\theta u(r + dr) \cos\theta dA c dt = 1/2 \sin\theta d\theta u(r - \lambda \cos\theta) \cos\theta dA c dt \quad (2.5)$$

Averaging over all directions

$$dE = c/2 \int_0^\pi \sin\theta \cos\theta u(r - \lambda \cos\theta) d\theta dA c dt \quad (2.6)$$

or making a Taylor expansion

$$dE = c/2 \int_0^\pi \sin\theta \cos\theta u(r) d\theta dA c dt - c/2\lambda \int_0^\pi \sin\theta \cos^2\theta \frac{du}{dr} d\theta dA c dt . \quad (2.7)$$

Since $u(r)$ is nearly isotropic (diffusion approximation!), the first term is zero and

$$dE = -\frac{c\lambda}{3} \frac{du}{dr} dA c dt . \quad (2.8)$$

In terms of the flux, $F = dE/dA c dt$, we have

$$F = -\frac{c\lambda}{3} \frac{du}{dr} \quad (2.9)$$

The mean free path is $\lambda = 1/\rho\kappa$ where κ is the opacity and ρ the density. The radiation density in an optically thick atmosphere is $u = aT^4$, so

$$F = -\frac{4acT^3}{3\kappa\rho} \frac{dT}{dr} \quad (2.10)$$

The luminosity is given by $L = 4\pi r^2 F$, so

$$\frac{dT}{dr} = -\frac{3\kappa\rho L(r)}{16\pi a c r^2 T^3} \quad (2.11)$$

Equations (2.1),(2.3),(2.4),and (2.11) define the structure of a star dominated by radiation transport. Under certain conditions (especially high temperature sensitivity of the energy generation or low temperature (high opacity)), resulting in very steep temperature gradients *if* the transport occurs only by radiation, the energy transport instead occurs by convection. A nearly adiabatic temperature gradient is then set up,

$$\frac{dT}{dr} = -\frac{\Gamma_2 - 1}{\Gamma_2} \frac{T}{p} \frac{dp}{dr} \quad (2.12)$$

where Γ_2 is an effective adiabatic index.

Most of the physics of the star and its evolution is determined by the opacity, $\kappa(\rho, T)$, the energy generation, $\epsilon(T)$, and the equation of state, $p(\rho, T)$.

3 The equation of state

The pressure, P , of a non-degenerate, perfect gas with temperature T is given by

$$P = knT \quad (3.1)$$

where n is the number of particles per volume and k is Boltzmann's constant, $k = 1.38 \times 10^{-16}$ ergs K⁻¹. In terms of the density this can be written as

$$P = \frac{k}{m_u \mu} \rho T \quad (3.2)$$

where μ is the mean mass per particle and m_u the atomic mass unit, 1.667×10^{-24} g.

For a gas of fermions the number density of particles is

$$n = \frac{8\pi}{h^3} \int_0^\infty f(p) p^2 dp \quad (3.3)$$

where p is the momentum, $f(p) = 1/[\exp(E - \mu)/kT + 1]$ is the Fermi-Dirac distribution, and μ is the chemical potential. The factor $4\pi p^2 dp/h^3$ is the phase space factor, and another factor of two comes from the spin of the

electrons. For a fully degenerate gas $f(p) = 1$ for $p < p_F$ and $f(p) = 0$ for $p > p_F$, allowing us to solve for p_F

$$p_F = \left(\frac{3h^3 n_e}{8\pi} \right)^{1/3} \quad (3.4)$$

This can be written in terms of the Fermi energy using $E_F = \sqrt{p_F^2 c^2 + m_e^2 c^4}$. For a non-relativistic gas $E_F = p_F^2/2m_e$, while for a relativistic gas $E_F = p_F c$.

The pressure, P , is given by

$$P = \frac{1}{3} \int_0^\infty v(p) p \frac{dn(p)}{dp} dp \quad (3.5)$$

where $dn/dp = 8\pi f(p) p^2/h^3$. For a fully degenerate gas we get

$$P = \frac{8\pi}{3h^3} \int_0^{p_F} v(p) p^3 dp . \quad (3.6)$$

Now $p = m_e v / \sqrt{1 - v^2/c^2}$, or

$$P = \frac{8\pi}{3h^3} \int_0^{p_F} \frac{p^4}{\sqrt{m_e^2 + p^2/c^2}} dp . \quad (3.7)$$

For simplicity we consider the non-relativistic and ultra-relativistic limits separately. The transition occurs when $p_F \approx m_e c$. Using $\rho = \mu_e m_p n_e$ this occurs at

$$\rho_r = 9.7 \times 10^5 \mu_e \text{ g cm}^{-3} \quad (3.8)$$

For a non-relativistic gas $p \ll mc$ and Eq. (3.7) shows that

$$P = \frac{8\pi}{15mh^3} p_F^5 . \quad (3.9)$$

With p_F from Eq. (3.4) we finally get

$$P = \frac{1}{20} \left(\frac{3}{\pi} \right)^{2/3} \frac{h^2}{m} n^{5/3} , \quad (3.10)$$

which is the equation of state for a non-relativistic, completely degenerate gas. In terms of the density we get in cgs units

$$P = 1.00 \times 10^{13} \mu_e^{-5/3} \rho^{5/3} . \quad (3.11)$$

In the opposite limit of an ultra-relativistic gas we obtain in the same way from Eq. (3.7)

$$P = \frac{2\pi c}{3h^3} p_F^4 . \quad (3.12)$$

and

$$P = \frac{1}{8} \left(\frac{3}{\pi} \right)^{1/3} hc n^{4/3} , \quad (3.13)$$

which is the equation of state for an ultra-relativistic, completely degenerate gas. Note that the adiabatic index in this case is 4/3, while in the non-relativistic case it is 5/3. In cgs units

$$P = 1.24 \times 10^{15} \mu_e^{-4/3} \rho^{4/3} . \quad (3.14)$$

For a non-relativistic gas the boundary between degeneracy and perfect gas equations of state is obtained by setting the non-degenerate pressure (Eq. (3.2)) equal to the degenerate, given by Eq. (3.10)

$$\frac{T}{\rho^{2/3}} = 1.2 \times 10^5 \frac{\mu}{\mu_e^{5/3}} \quad (3.15)$$

Similarly, if the gas is relativistic one finds that the boundary between degeneracy and perfect gas equations of state is given by

$$\frac{T}{\rho^{1/3}} = 1.5 \times 10^7 \frac{\mu}{\mu_e^{4/3}} . \quad (3.16)$$

The density when degeneracy sets in depends on the mass of the particle and temperature as $n_{deg} \propto m^{3/2} T^{3/2}$ in the non-relativistic case and $n_{deg} \propto m^3$ in the relativistic. Therefore, even if the electrons are degenerate, the ions are usually non-degenerate. The total pressure is then given by

$$P = P_e + P_{ion} \quad (3.17)$$

where P_e is given by either Eq. (3.10) or Eq. (3.13) and P_{ion} by Eq. (3.1). In the strongly degenerate case the ion pressure is much smaller than that of the degenerate electrons, and can usually be neglected.

Finally, the boundary between an ideal gas pressure and that of radiation dominated pressure is given by

$$\frac{RT\rho}{\mu} = \frac{aT^4}{3} \quad (3.18)$$

or

$$\frac{T}{\rho^{1/3}} = 3.2 \times 10^7 \mu^{-1/3} \quad (3.19)$$

In Fig. 1 we show the different regions defined by Eqns. (3.8), and (3.15) – (3.19) in the $\rho - T$ plane.

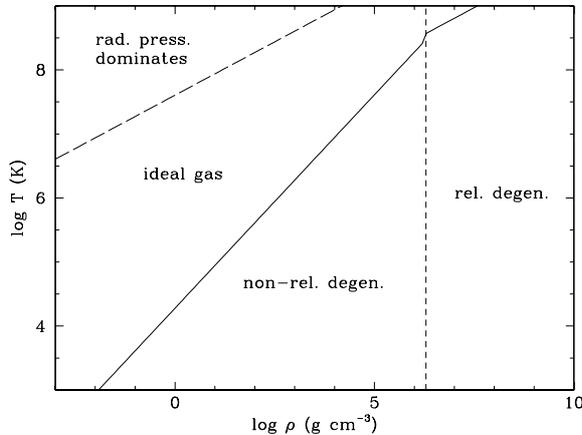


Figure 1: Regions of different equations of state in the $\rho - T$ plane.

4 The Chandrasekhar mass

In Longair a derivation of the Chandrasekhar mass is given in a rigorous way. Because of its importance it may be of interest also to derive it in a 'quick and dirty' way, without losing too much of the physics.

Let us consider the structure of a completely degenerate core. The equation of hydrostatic equilibrium gives

$$\frac{dP}{dr} = -\frac{Gm(r)\rho}{r^2} \quad (4.1)$$

or approximately

$$P \approx \frac{GM\rho}{r} \quad (4.2)$$

Now $\rho \approx 3M/(4\pi R^3)$, so

$$P \approx GM^{2/3}\rho^{4/3} \quad (4.3)$$

Here we can use the expression for the equation of state derived above. Alternatively, we note that the pressure is given approximately by

$$P \approx \frac{1}{3}n_e p v \quad (4.4)$$

where p is the 'typical' momentum of the electrons, and n_e the electron density and v the velocity. An estimate of p can be obtained from the uncertainty relation $\Delta p \Delta x \approx \hbar$, or since $n_e \approx 1/\Delta x^3$,

$$P \approx \frac{1}{3}\hbar n_e^{4/3} v \quad (4.5)$$

If the electrons are relativistic then $v = c$, and we have

$$P \approx \frac{1}{3} \hbar n_e^{4/3} c \quad (4.6)$$

which is apart from a numerical factor Eq. (3.13).

If the electrons are non-relativistic $v \approx \hbar n_e^{1/3} / m_e$, again using the uncertainty relation. Therefore,

$$P \approx \frac{1}{3} \frac{\hbar^2}{m_e} n_e^{5/3} c \quad (4.7)$$

To convert from electron density to total mass density we use $n_e = Z\rho/Am_p \approx 1/2\rho/m_p$ for a gas dominated by heavy elements. This together with Eq. (4.7) in Eq. (4.3) gives in the non-relativistic case

$$R \approx \frac{\hbar^2}{Gm_e M^{1/3}} \left(\frac{Z}{Am_p} \right)^{5/3} \quad (4.8)$$

This shows that as the mass of the star increases the radius *decreases*. Therefore, as the mass increases the density will increase as $\rho \propto R^{-6}$, and the electrons will therefore become relativistic at $\sim 10^6$ g cm⁻³.

The mass – radius above can be tested directly if masses and radii can be determined from e.g. binary motions and spectroscopy, including the gravitational redshift. In Figure 2 we show a comparison between observations of a sample of white dwarfs and the theoretical mass - radius relation for different interior compositions.

When the electrons are relativistic the equation of state is given by Eq. (4.6) which together with Eq. (4.3) gives an expression independent of radius,

$$M_{\text{Ch}} \approx \left(\frac{c\hbar}{G} \right)^{3/2} \left(\frac{Z}{Am_p} \right)^2 \quad (4.9)$$

This is the Chandrasekhar mass apart from numerical factors. A more accurate value is

$$M_{\text{Ch}} = 2.018 \frac{(3\pi)^{3/2}}{2} \left(\frac{c\hbar}{G} \right)^{3/2} \left(\frac{Z}{Am_p} \right)^2 = 5.836 \left(\frac{Z}{A} \right)^2 M_{\odot} \quad (4.10)$$

For $Z/A = 0.5$ we get $M_{\text{Ch}} = 1.46 M_{\odot}$.

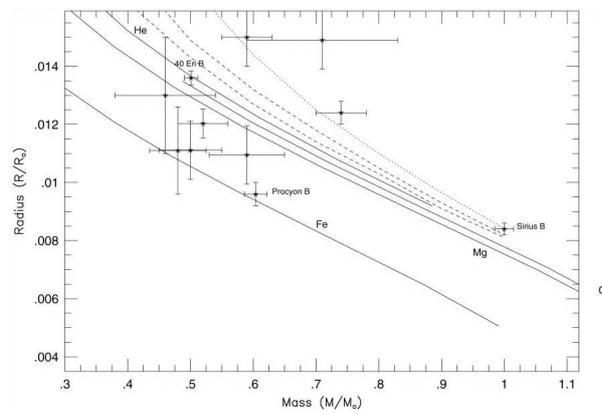
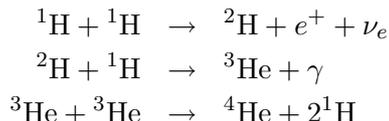


Figure 2: Comparison between observations and the theoretical mass - radius relation for different interior composition. (Provencal et al.(1998))

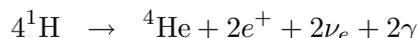
5 Solar neutrinos

5.1 Predicted fluxes

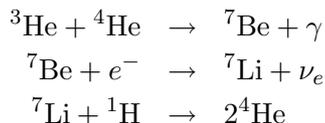
In the Sun nuclear hydrogen burning takes place at $\sim 1.5 \times 10^7$ K. The most important chain of reactions is the PP I chain



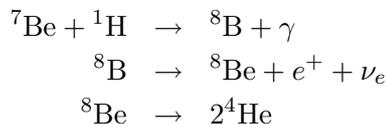
Summarizing



which takes place in 69 % of the cases. In half as many cases, 33 %, the last step is replaced by the so called PP II reaction



Finally, one very rare, but important reaction, PP III, takes place roughly once every thousand reactions,



The total reaction rate is moderately sensitive to the temperature, $\epsilon \propto T^6$.

In several of these steps neutrinos are emitted with different energies. Using standard assumptions about the solar structure, the fluxes of these can be calculated in what is known as the 'standard' solar model, which was pioneered by the late John Bahcall. In Fig. 3 the expected fluxes from this model are shown. An important feature is here the different energies of these. In particular, the easiest neutrinos to detect are the ones with the highest energies. This was the reason that historically the first solar neutrinos detected were from the PP III chain, in the reaction ${}^8\text{B} \rightarrow {}^8\text{Be} + e^+ + \nu_e$, in spite of the rare occurrence of this reaction.

We will now discuss the different experiments one by one. It is in this context especially important to note the different threshold energies in Fig. 3 of the different experiments.

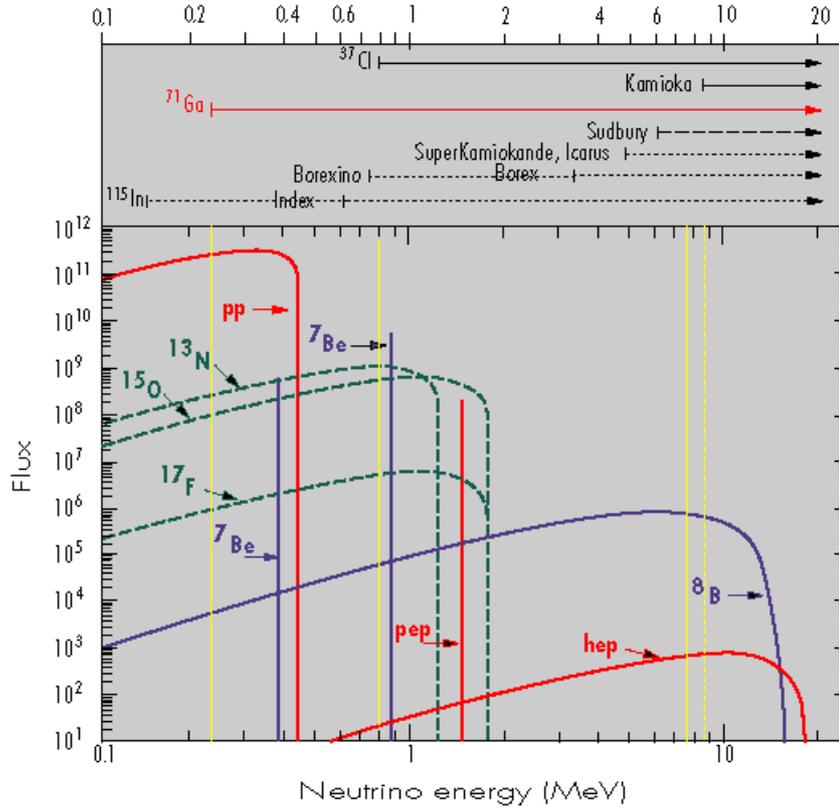


Figure 3: Predicted neutrino spectrum from the sun (Bahcall 2004)

5.2 Neutrino experiments

Chloride experiments: Homestake Gold Mine

This was the first neutrino experiment started in 1965 by Ray Davies. To protect it against cosmic rays it was located 1500 m underground in an old gold mine. As detector element it contained 615 tons of cleaning fluid, C_2Cl_4 . The important reaction was here



with a threshold of 0.8 MeV. Davies then detected the radioactive decay of ${}^{37}\text{Ar}$. Approximately 15 reactions were seen per month!

After several years of running the average measured rate of the Homestake experiment was 2.56 ± 0.25 SNU. The big surprise was now that this



Figure 4: The Homestake neutrino detector.

was a factor of roughly three lower than that predicted by the standard model, 8.1 ± 1.2 . This was known as the 'solar neutrino problem'. An important question is now if we can trust the 'standard model'. To make a long story short (see Longair for more details), in particular solar seismology models, which use the frequencies and amplitudes of different oscillations of the sun, can probe the structure of the sun, or more precisely the variation of the sound velocity as function of the radius. This in turn depends on the composition and temperature of the sun and therefore provides a probe of the solar interior, and a test of the solar model. The fact that one has a good agreement with the predicted and observed sound velocity provides good confidence in the solar models.

Water Cherenkov experiments: Superkamiokande

The most important of this class of experiments is the SuperKamiokande experiment, 1000 m underground. It contains 50,000 ton of pure water as detector element, and the reaction is the elastic scattering

$$\nu_e + e^- \rightarrow \nu_e + e^- \quad (5.2)$$

with a threshold of 5 MeV. The Cherenkov light from the scattered electrons are detected by thousands of PMTs. The measured rate with this detector is ~ 15 events/day, i.e. a factor 30 larger than Davies experiment. A

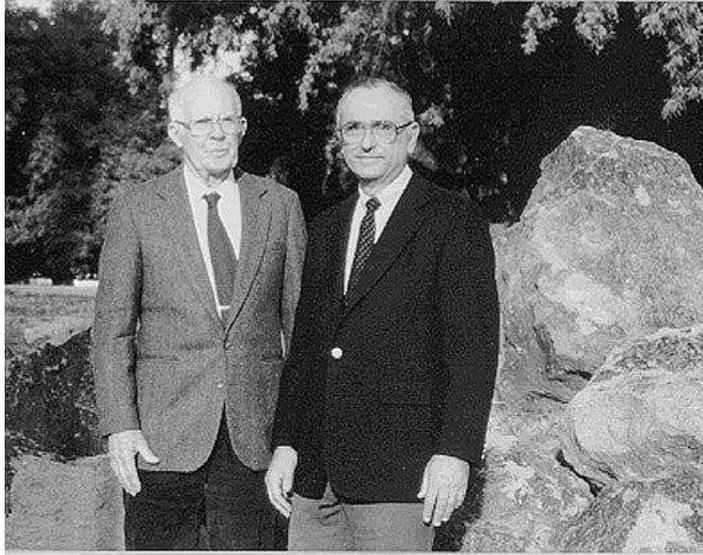


Figure 5: The pioneers of solar neutrino physics Ray Davies and John Bahcall.

most important property with this experiment is that since it is a scattering reaction it contains directional information about the neutrino. In fact, a neutrino image of the sun has been produced by this instrument!

Gallium detectors

An obvious drawback with the previous experiments is that they are only sensitive to the rare PP III cycle. In order to decrease the threshold to include the dominant PP I reaction one uses gallium in the reaction



The threshold of this is as low as 0.233 MeV.

There are two experiments of this type, the SAGE experiment with 50 tons of gallium, and the GALLEX experiment in the Gran Sasso laboratory in Italy with 30 tons of gallium. The former found a rate of 70.8 (+5.3, -5.2) (+3.7, -3.2) SNU, and the latter 77.5 +/- 8 SNU. This should be compared to the predicted rate of 129 +8/-6 SNU from the standard model, which is still a large factor above the measured.

This discrepancy is especially important since the neutrino rate of this reaction is directly proportional to the solar luminosity, and therefore less sensitive to details of the solar model.

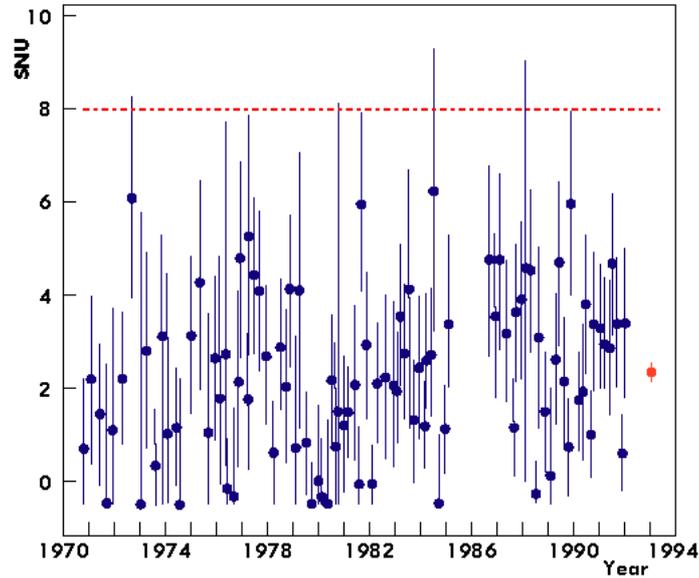


Figure 6: Neutrino flux as function of time for the Homestake chloride experiment

The SNO experiment

Since the book of Longair was published there has been considerable progress in terms of new experiments on solar neutrinos. The most spectacular of these is the SNO experiment (Sudbury Neutrino Observatory). While all the previous experiments are sensitive only to the electron neutrino, this is sensitive to all three types of neutrinos.

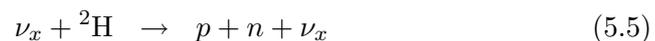
This experiment is based on a detector containing 1000 tons of heavy water, $^2\text{H}_2\text{O}$, nearly 2000 m underground close to Sudbury, Ontario. The incoming neutrinos can react in two different ways with the deuterium nuclei.

In the first reaction the ν_e can disintegrate the deuteron, resulting in a change of the charge of the nucleons



The electron can be detected from its Cherenkov radiation. This reaction only occurs for the electron neutrino.

The most interesting reaction is



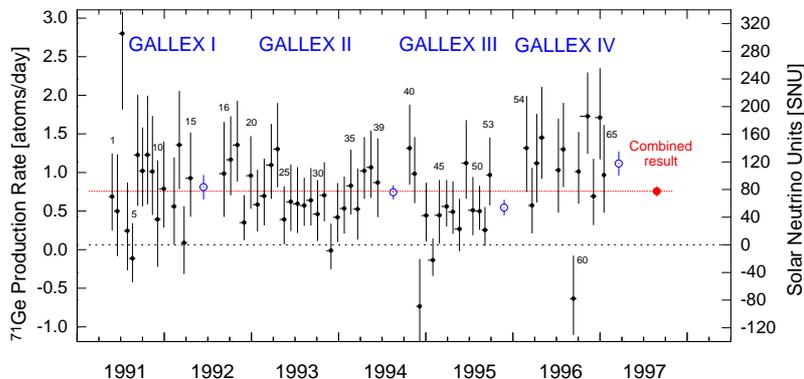


Figure 7: Reaction rates for the Gallex experiment.

which occurs by neutral currents. Here x stands for either e, μ or τ . This reaction, which only disintegrates the deuteron without any changes of the charges, occurs with all three neutrino types. Furthermore, the probability is the same for all three types. The neutrons liberated are captured by ^{35}Cl nuclei, which converts it to ^{36}Cl , with emission of gamma-rays.

The amazing result from this detector is that while the charged current reaction gave a similar result as the other neutrino experiments, $\sim 30\%$ of the expected rate, the neutral current reaction gave nearly exactly the rate predicted by the standard solar model! The situation is summarized in Fig. 9.

This result clearly demonstrated that the electron neutrinos changed flavor into the other two types on the way from the sun to us. Of the original electron neutrinos only about a third oscillated back to electron neutrinos, while the rest were converted to μ and τ neutrinos. Because the other experiments are not sensitive to these this explains the more than four decade long puzzle.

KamLAND

Besides the SNO experiment this is the most interesting. Here a detector in Kamioka, containing 1000 tons of liquid scintillator, detects neutrinos from nuclear reactors at different distances, 80–350 km, in Korea and Japan. The neutrino flux from these reactors passes through a long column density of mass. During this passage the neutrinos will mix by the MSW effect and only a fraction of the original emitted electron anti-neutrinos will be detected as such, the rest being in the tau- and mu- flavor states. The

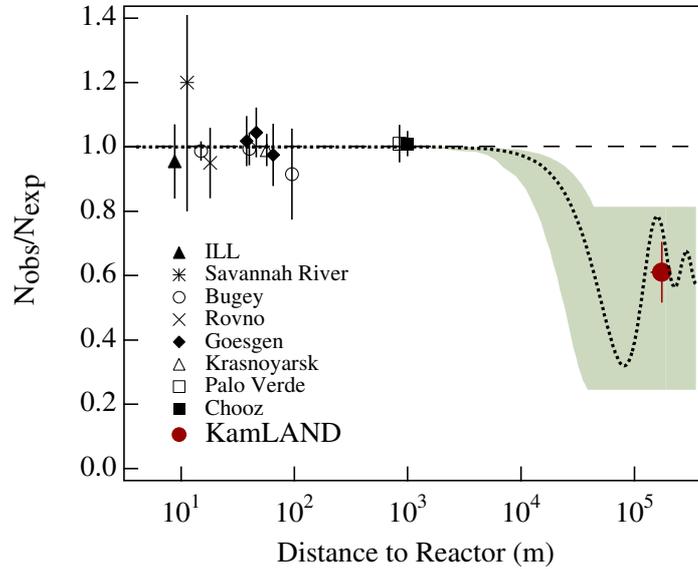


Figure 8: Neutrino flux as function of distance from the reactor. The line gives the expected rate, assuming that the neutrino oscillations have the mixing angles and mass differences derived from the solar observations.

neutrinos collide in the detector with protons, converting them to neutrons, which are inducing radioactive isotopes, which decay.



The experiment therefore checks the survival probability as function of distance from the reactor.

As shown in Fig. 8 the flux is indeed suppressed relative experiments done at short distances. The full line shows the predicted rate for mixing parameters determined from solar observations.

The fact that there is excellent agreement between the prediction and the experiment shows that neutrino oscillations are confirmed as an explanation of the 'neutrino problem'.

Summary

The neutrino fluxes *corrected for neutrino mixing* for the different reactions are summarized below relative to the predicted rates from the standard solar model:

Total Rates: Standard Model vs. Experiment
Bahcall–Serenelli 2005 [BS05(OP)]

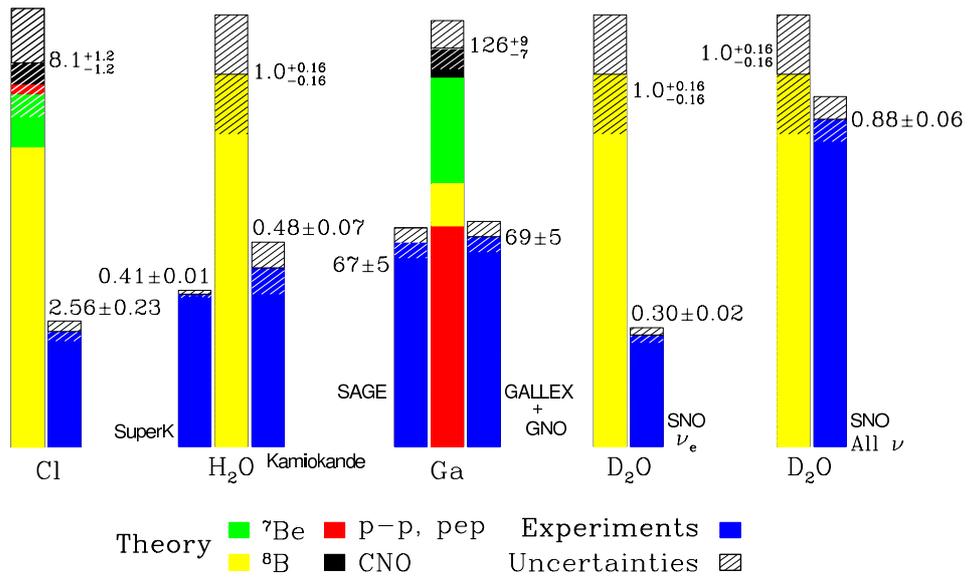


Figure 9: Neutrino fluxes as measured by different experiments. Note the agreement with the neutral current SNO measurements and the predicted value (Bahcall 2004)

$$\text{Flux}(pp) = 1.02 \pm 0.02 \pm 0.01 \times \text{theory}$$

$$\text{Flux}(^8\text{B}) = 0.88 \pm 0.04 \pm 0.23 \times \text{theory}$$

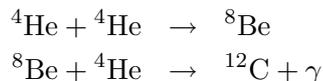
$$\text{Flux}(^7\text{Be}) = 0.91^{+0.24}_{-0.62} \pm 0.11 \times \text{theory}$$

The neutrino oscillations therefore explains the more than four decade long puzzle, and shows that we indeed understand why the sun shines!

6 Pre-supernova Evolution of Massive Stars

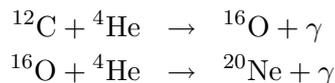
6.1 Helium burning

When the hydrogen in the core of a star has been consumed, two options are possible. Either hydrogen burning continues in a shell surrounding the helium core or the helium in the core itself is ignited. This occurs by the triple alpha process at a temperature of $\sim 10^8$ K. In detail the different steps are



The last step takes place by the famous Hoyle resonance reaction, which was predicted to take place by Hoyle before its experimental verification. The temperature sensitivity of this reaction is much higher than for hydrogen burning, $\epsilon \propto T^{41}$.

From the carbon produced further alpha captures may take place,



After helium burning a core of mainly carbon and oxygen therefore results. This is the endpoint of the evolution of low mass stars, like our sun. After the outer hydrogen rich layers have been expelled the result will then be a degenerate white dwarf consisting of roughly equal amounts of carbon and oxygen. Higher mass stars may, however, continue their nuclear burning to heavier elements.

6.2 Low mass versus high mass evolution

The evolution of the star is mainly the evolution of the core. This is in turn determined by the temperature and density. In particular, the final stages depend on whether or not the core becomes reaches a temperature high enough for carbon to ignite.

We can separate two qualitatively different behaviors, depending on the value of the core mass M_c . Core masses below M_{Ch} never reach a temperature larger than $T_{c,max}$, while those above can increase their temperature as the core contracts and ignite new fuels. For $M_c > M_{Ch}$ the temperature rises monotonically as $T_c \propto \rho_c^{1/3}$. The division of these two cases occur for $8 - 10 M_\odot$, Stars with mass smaller than this only evolves to the helium

burning stage, when they become red giants. At this point they lose a large fraction of their hydrogen envelope in a superwind lasting a few $\times 10^4$ years, forming a planetary nebula. The core, forming a white dwarf with a mass of less than the Chandrasekhar mass $1.46 M_{\odot}$, consists of mainly oxygen and carbon. Stars with larger mass than $8 - 10 M_{\odot}$, however, continue their evolution into the more advanced burning stages, and end their lives in a supernova explosion, and finally a neutron star or black hole.

In Fig. 10 this is shown in more detail from evolutionary calculations for stars of different masses. The $1 M_{\odot}$ and $2 M_{\odot}$ stars become degenerate before He-ignition, while the $7 M_{\odot}$ model ignites helium non-degenerately but then evolves into the degenerate regime. In all three cases does the cores not reach a temperature high enough for carbon burning to start. The $15 M_{\odot}$ model on the other hand continues to evolve in the partially degenerate regime and the temperature increases monotonically, and passes through all the nuclear burning stages.

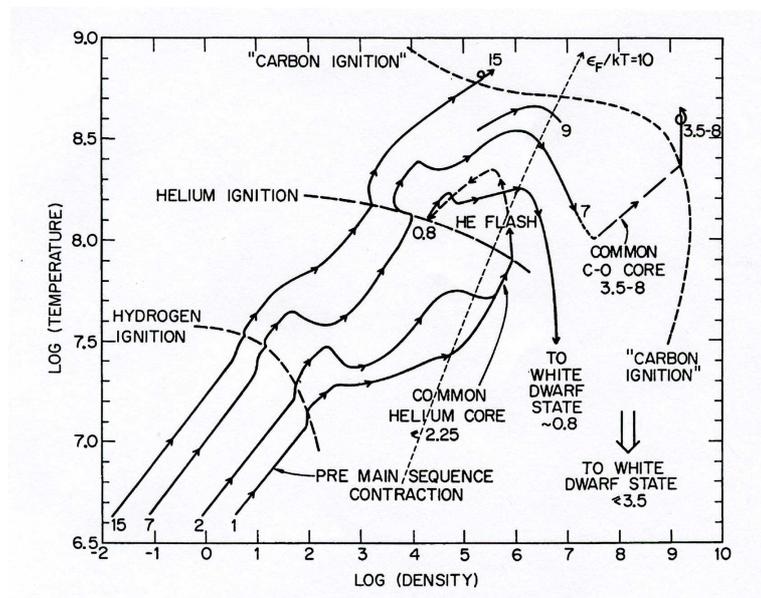


Figure 10: Evolution of the central density and temperature for stars of different masses. (Iben 1974)

In Fig. 11 a more recent calculation of the evolution of a $15 M_{\odot}$ and $25 M_{\odot}$ star is shown through all evolutionary stages up to silicon burning.

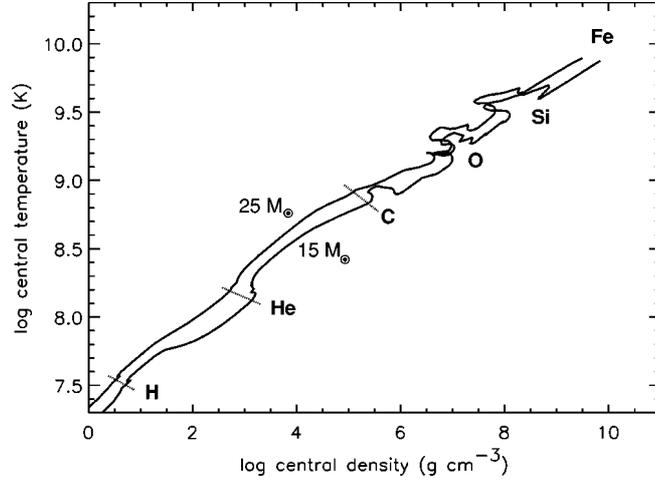


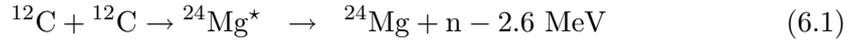
Figure 11: Evolution of the central density and temperature for a $15 M_{\odot}$ and $25 M_{\odot}$ star. (WHW02)

6.3 Advanced nuclear burning stages

In this section we include a brief discussion of the advanced burning stages. For a more complete account of the nuclear physics, as well as the hydrogen and helium stages, see especially Clayton (1967) or Arnett (1996).

6.3.1 Carbon burning

Carbon burning occurs at $(0.6 - 1.2) \times 10^9$ K. The principal reactions occur through the compound nucleus $^{24}\text{Mg}^*$, which decays as



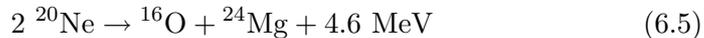
In the interesting temperature range the reaction rate depends on the temperature as $q \propto T^{29}$.

6.3.2 Neon burning

Neon burning occurs in a narrow range at $\sim 1.5 \times 10^9$ K. The first step is photo-disintegration



The next step is that the α particles are partly captured by ^{16}O to form ^{20}Ne , and partly by ^{20}Ne to produce ^{24}Mg , i.e., $^{20}\text{Ne} + \alpha \rightarrow ^{24}\text{Mg}$. The net result of each of these reactions can be summarized as

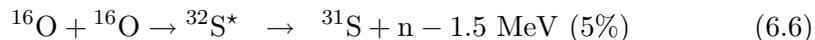


Note, however, that this is only symbolic, and is not a binary heavy ion reaction, like carbon burning. Because of the sensitivity of the α particle abundance to the temperature the reaction rate depends extremely sensitively on the temperature, as $q \propto T^{50}$.

One may ask why neon burning occurs before oxygen burning. The reason is that ^{16}O is a doubly magic nucleus, and has consequently a larger binding separation energy for α particles than neon, 7.2 and 4.7 MeV, respectively.

6.3.3 Oxygen burning

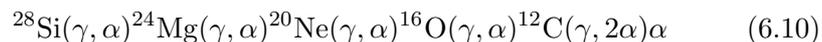
Oxygen can burn either as photo-disintegration, $^{16}\text{O}(\gamma, \alpha)^{12}\text{C}$, or as a fusion reaction, producing Si – S. During hydrostatic burning at $\sim 2 \times 10^9$ K the fusion reaction dominates, while in explosive oxygen burning in connection to the supernova explosion, photo-disintegration and fusion are equally important. The most important fusion reactions are



The main products of oxygen burning are ^{28}Si and ^{32}S . At $\sim 2 \times 10^9$ K the reaction rate depends on the temperature as $q \propto T^{33}$.

6.3.4 Silicon burning

Silicon does not fuse as a heavy ion reaction to ^{56}Ni , but instead melts at $\sim 3.5 \times 10^9$ K by photo-disintegration, due to the extremely energetic radiation density. The resulting nuclei in turn photo-disintegrate to lighter nuclei, etc. In summary,



The result is, however, not only lighter nuclei. The α particles produced by the melting will also be captured by the ^{28}Si to form ^{32}S , which may in

turn capture new α particles, etc. The result will be a quasi-equilibrium with successively heavier nuclei. The end result of this depends on the neutron excess η . For small values of $\eta \lesssim 6 \times 10^{-3}$, which is needed to produce the right abundances of the isotopes around the iron peak, the most abundant nucleus is also the most tightly bound nucleus ^{56}Ni (see next section). This radioactive isotope subsequently decays into ^{56}Fe . The fact that this radioactive isotope is main result of the silicon burning is important not only for the nucleosynthesis, but also for the observational properties of all kinds of supernovae.

The reaction rate at $\sim 3.5 \times 10^9$ K goes as $q \propto T^{49}$. The energy release is only one half of that of oxygen burning.

6.4 Nuclear statistical equilibrium

Because of the importance of nuclear statistical equilibrium (NSE) in the Si burning phase, as well as in several other contexts, we discuss it in some detail here.

The Saha equation, relating the number densities of two neighboring ionization stages in ionization balance through

$$n_i + \gamma \leftrightarrow n_{i+1} + e^-, \quad (6.11)$$

is

$$\frac{n_{i+1}n_e}{n_i} = \frac{G_{i+1}g_e}{G_i} \frac{(2\pi m_e kT)^{3/2}}{h^3} e^{-\chi_i/kT} \quad (6.12)$$

where G_i is the partition function of the ion i , g_e the statistical weight of the electron and χ_i the ionization potential.

In exact analogy with this, one can relate the equilibrium densities of two different isotopes in photodissociation balance

$$n_{Z,A} + \gamma \leftrightarrow n_{Z,A-1} + n \quad (6.13)$$

by

$$\frac{n_{Z,A-1}n_n}{n_{Z,A}} = \frac{2G_{Z,A-1}}{G_{Z,A}} \frac{(2\pi m_{Z,A-1}m_n kT)^{3/2}}{h^3 m_{Z,A}^{3/2}} e^{-Q_n/kT} \quad (6.14)$$

where we have used the fact that the statistical weight of the neutron is $g_n = 2$. Q_n is the binding energy of the neutron in the nucleus (Z,A) , $Q_n = (m_{Z,A-1} + m_n - m_{Z,A})c^2$. In more compact notation we can write this as

$$\frac{n_{Z,A-1}n_n}{n_{Z,A}} = \frac{2G_{Z,A-1}}{G_{Z,A}} \left(\frac{A-1}{A} \right)^{3/2} \theta e^{-Q_n/kT} \quad (6.15)$$

where $\theta \equiv (2\pi m_n kT)^{3/2}/h^3$.

Similarly, we can remove one proton from the nucleus we produced in reaction (6.13) by photodissociation according to

$$n_{Z,A-1} + \gamma \leftrightarrow n_{Z-1,A-2} + p \quad (6.16)$$

producing the next lighter element Z-1. As above we get for this balance

$$\frac{n_{Z-1,A-2} n_p}{n_{Z,A-1}} = \frac{2G_{Z-1,A-2}}{G_{Z,A-1}} \left(\frac{A-2}{A-1} \right)^{3/2} \theta e^{-Q_p/kT} \quad (6.17)$$

where $Q_p = (m_{Z-1,A-2} + m_p - m_{Z,A-1})c^2$.

This procedure can now be repeated until we have only protons and neutrons left. Putting these steps together we obtain

$$n_{Z,A} = G_{Z,A} \frac{A^{3/2} n_p^Z n_n^{A-Z}}{2^A} \theta^{1-A} e^{Q_{Z,A}/kT} \quad (6.18)$$

where now $Q_{Z,A} = (Zm_p + (A-Z)M_N - m_{Z,A})c^2$ is the total binding energy of the nucleus.

A particular composition is then characterized by a given value of Y_e or η . Once this is specified all other abundances can then be calculated from the NSE relation Eq. (6.18).

Which nucleus is most abundant in NSE depends on the value of η . In general, for $T \lesssim 10^{10}$ K the most tightly bound nucleus for a given value of η is favored. Consequently, for small neutron excesses one finds that ^{56}Ni , which is an even-even nucleus with $\eta = 0$, is the most abundant nucleus, while at $\eta \sim 0.07$ ^{56}Fe , with $\eta = (30 - 26)/56 = 0.071$, is most abundant. In Fig. 12 we show the isotopic abundances for a few different temperatures as function of η . We see that the abundances do not change appreciably as function of temperature, except for a general decrease in the iron peak abundances, reflecting the shift to ^4He , as photo-disintegration of ^{56}Fe becomes important.

6.5 Neutrino cooling

In the advanced burning stages cooling by neutrinos play an increasingly important role. At temperatures of the order of $m_e c^2/k \sim 5 \times 10^9$ K electron pair production by energetic photons becomes possible. In most cases these pairs annihilate into photons, but because the electrons and neutrinos couple through the weak interaction, occasionally a neutrino - anti-neutrino pair may be produced,

$$e^- + e^+ \rightarrow \nu + \bar{\nu} \quad (6.19)$$

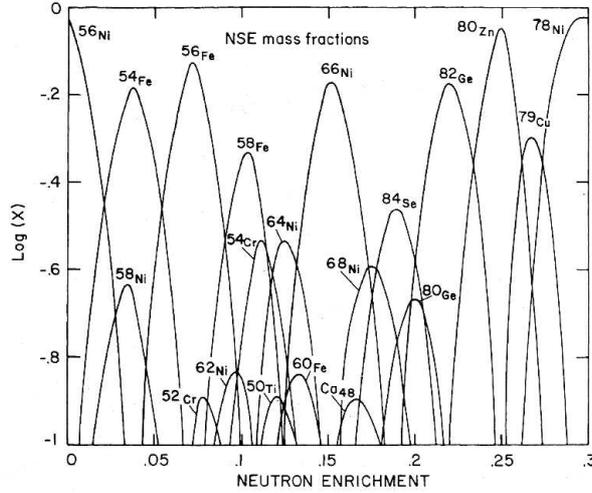


Figure 12: Abundances of iron peak elements at 3.5×10^9 K as function of the neutron excess. From left to right the isotopes are ^{56}Ni , ^{54}Fe , ^{56}Fe , ^{58}Fe . (Hartman, Woosley and el Eid 1985)

The typical neutrino cross section is

$$\sigma \approx 10^{-44} \left(\frac{E}{m_e c^2} \right)^2 \text{ cm}^{-2}. \quad (6.20)$$

The electron-positron annihilation cross section is of the order of the Thompson cross section, $\sigma_T = 0.665 \times 10^{-24} \text{ cm}^2$ (or rather Klein-Nishina at these energies). Therefore the probability for *neutrino pair production* is $\sim 10^{-20}$ of the electron pair annihilation rate.

The neutrino energy loss rate at $T < 10^9$ K is given by

$$\epsilon_\nu \approx 4.9 \times 10^{18} T_9^3 \exp^{-11.86/T_9} \text{ erg cm}^{-3} \quad (6.21)$$

and at $T > 3 \times 10^9$ K

$$\epsilon_\nu \approx 4.5 \times 10^{15} T_9^9 \text{ erg cm}^{-3} \quad (6.22)$$

This clearly illustrates the sensitivity of the neutrino losses to the temperature.

Except for pair production, also other neutrino cooling processes may be important. In particular, the plasma neutrino cooling process and photo-neutrino process are important in many circumstances.

Table 1: **Burning stages for a 15 M_{\odot} star (WHW02)**

Fuel	Ashes	T 10^8 K	ρ g cm $^{-3}$	M M_{\odot}	L $10^3 L_{\odot}$	R R_{\odot}	τ yrs
H	He, N	0.35	5.8	14.9	28.0	6.75	1.1×10^7
He	C,O	1.8	1.4×10^3	14.3	41.3	461.	2.0×10^6
C	Ne, Mg, O	8.3	2.4×10^5	12.6	83.3	803.	2.0×10^3
Ne	O, Mg, Si	16.3	7.2×10^6	12.6	86.5	821.	0.73
O	Si, S	19.4	6.7×10^6	12.6	86.6	821.	2.6
Si	Ni	33.4	4.3×10^7	12.6	86.5	821.	18 days

The photo-neutrino process is just pair production of a neutro-antneutrino pair

$$\gamma + e^- \rightarrow e^- + \nu + \bar{\nu}, \quad (6.23)$$

For most massive stars it is the pair annihilation cooling which accounts for most of the cooling, although plasma neutrino cooling is important in especially stars of lower mass.

6.6 Duration of the burning stages

Because of the increasing importance of neutrino losses as the temperature increases because of core contraction, and also the decreasing energy generation per mass, the durations of the burning stages decrease rapidly from thousands of years for carbon burning to days or less for silicon burning. In Tables 1 and 2 we give the duration, as well as the ignition temperature and other parameters, for a 15 M_{\odot} and a 25 M_{\odot} ZAMS star, including mass loss (from WHW02). These models include mass loss (see below), explaining the low masses in the table.

Note in this table the dramatic decrease in the duration of the advanced burning stages, because of increasing neutrino losses in the carbon burning stage and beyond. Also note, while there is only a small decrease in the total mass of the 15 M_{\odot} star, the 25 M_{\odot} star ends up with only half of the original ZAMS mass. We will discuss this further later.

After the carbon burning stage the neutrinos dominate the cooling of the core, as Fig. 13 clearly shows.

After carbon burning the diffusion time for the photons is much longer than the duration of these stages, and the core evolves independently from

Table 2: Same as above for a $25 M_{\odot}$ star

Fuel	Ashes	T 10^8 K	ρ g cm^{-3}	M M_{\odot}	L $10^3 L_{\odot}$	R R_{\odot}	τ yrs
H	He	0.38	3.8	24.5	110.	9.2	6.7×10^6
He	C,O	2.0	7.6×10^2	19.6	182.	1030.	8.4×10^5
C	Ne, Mg	8.4	1.3×10^5	12.5	245.	1390.	5.2×10^2
Ne	O, Mg	15.7	4.0×10^6	12.5	246.	1400.	0.89
O	Si, S	20.9	3.6×10^6	12.5	246.	1400.	0.40
Si	Ni	36.5	3.0×10^7	12.5	246.	1400.	0.73 days

the envelope. Unless some kind of shell flash or similar occurs, the envelope is essentially decoupled from the core.

6.7 Mass loss

Most massive stars experience mass loss to a varying degree. The properties of these winds, however, vary dramatically between the different evolutionary stages. In particular, the wind velocity scales roughly with the escape velocity of the star, which varies by a factor of about a hundred between the blue supergiant, red supergiant, and Wolf-Rayet phases.

In the blue supergiant (BSG) MS phase the winds are radiatively driven through momentum deposition from absorption of the photospheric radiation by the many resonance lines in especially the UV and far-UV. This is a fairly well understood process both theoretically and observationally. Typical mass loss rates are of the order of $10^{-6} M_{\odot} \text{ yr}^{-1}$ and the wind velocities are $1,000 - 3,000 \text{ km s}^{-1}$.

In the red supergiant (RSG) phase the winds are much less understood. Dust driving is believed to account for most of the momentum input. What initiates the wind (e.g., photospheric shocks connected to pulsations) is, however, not known. Further, it is likely that the star experiences a superwind phase, lasting $\sim 10^4$ years in the very last phases of the red supergiant stage. What drives this superwind is somewhat unclear, but pulsational instabilities may be particularly important (see e.g., Heger et al. 1997). Typical mass loss rates are in the general RSG phase of the order of $10^{-6} M_{\odot} \text{ yr}^{-1}$ and the wind velocities are $10 - 50 \text{ km s}^{-1}$. In the superwind phase mass loss rates as high as $10^{-4} - 10^{-3} M_{\odot} \text{ yr}^{-1}$ may occur. The duration of this phase must obviously be only of the order of a few times 10^4 years.

A useful formula which summarizes the mass loss rates on the main

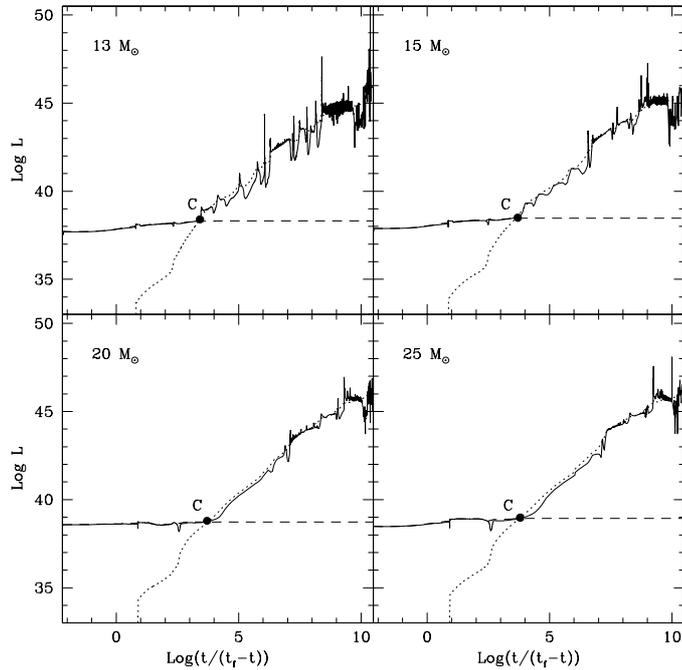


Figure 13: Neutrino luminosity of stars with mass 13 – 25 M_{\odot} compared to the nuclear energy generation and the photon luminosity (Limongi et al. 2002).

sequence and in the red supergiant stage is given by Nieuwenhuijzen & de Jager (1990),

$$\dot{M} = 9.6 \times 10^{-15} \left(\frac{L}{L_{\odot}} \right)^{1.42} \left(\frac{M}{M_{\odot}} \right)^{0.16} \left(\frac{R}{R_{\odot}} \right)^{0.81} M_{\odot} \text{ yr}^{-1} \quad (6.24)$$

Finally, in the Wolf-Rayet (WR) phase the wind velocities increase to 2,000 – 5,000 km s^{-1} , while the mass loss rate is $\sim 10^{-5} M_{\odot} \text{ yr}^{-1}$. The driving of the wind is here to a large extent by radiation on resonance lines, as in the OB star case. The initiation of the wind is, however, not clear, and pulsations may be important for this. Observationally, clumping of the wind is important, with a typical clumping factor of about two. Once this has been corrected for, the mass loss rates are fairly well determined.

The fact that the RSG phase last for $\sim 10^5$ years and the WR phase for a comparable period, and that the mass loss rates in these phases are $10^{-5} - 10^{-4} M_{\odot} \text{ yr}^{-1}$, means that stellar winds will have a major influence on the evolution.

6.8 Evolution in the HR diagram

Mass loss is crucial both for the observational properties the appearance of the star and for the internal structure. Because mass loss is increasingly important with mass, the effects increase strongly with mass. In Fig. 14 we show the evolution in the HR diagram of a $60 M_{\odot}$ star with and without mass loss. While both stars evolve to the RSG phase, the star without mass loss end its life in this phase. The star with mass loss, however, evolves back to the blue and becomes a hot star, now without any hydrogen envelope. It has become a helium star, or better known as a Wolf-Rayet star. This

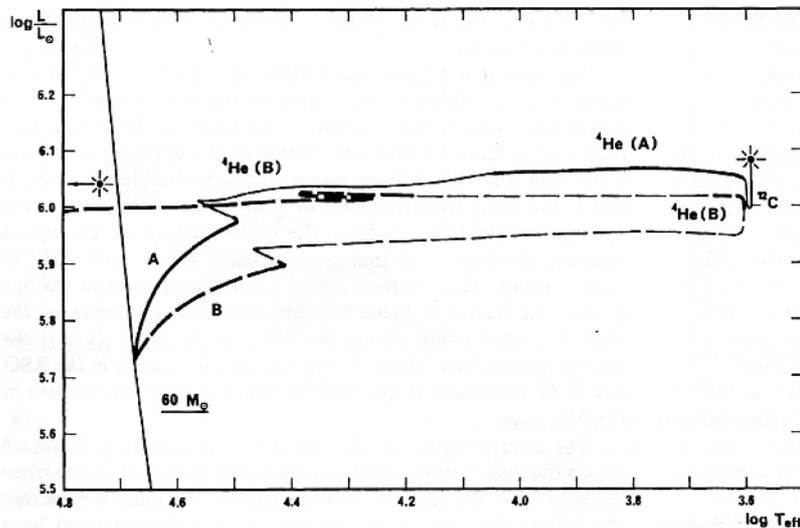


Figure 14: Evolution of a $60 M_{\odot}$ star with (dashed line) and without (solid line) mass loss (Maeder 1981)

evolutionary scenario is a general feature for massive stars above some limiting mass, M_{WR} , which is uncertain, but probably in the range $20 - 40 M_{\odot}$, depending on mass loss rates, rotation, metallicity etc. (see below). We therefore have the evolutionary sequence

$$O \rightarrow BSG \rightarrow RSG \rightarrow WR \quad (6.25)$$

The most massive stars may lose mass so fast that they never evolve to the RSG stage, but instead evolve as luminous blue variable (LBVs) and then directly to the WR stage.

$$O \rightarrow LBV \rightarrow WR \quad (6.26)$$

Because mass loss increases with luminosity and mass, $\dot{M} \propto L \propto M^{2-3}$, the effects are most important for the most massive stars. The result is that the final mass before the star collapses is nearly independent of the initial mass! When rotation is taken into account one finds a final mass of $10 - 15 M_{\odot}$ for all masses $\gtrsim 20 M_{\odot}$, as shown in Fig. 15.

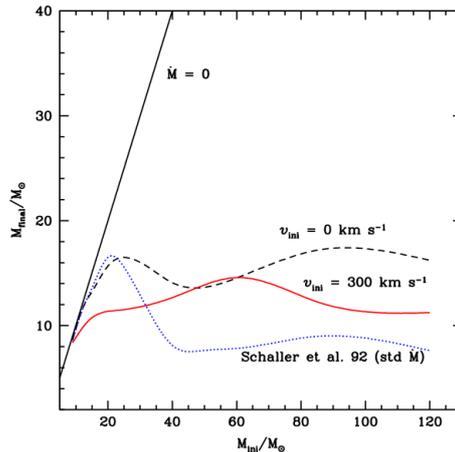


Figure 15: The effects of rotation and mass loss on the final mass of massive stars. The solid line gives the final mass for a ZAMS rotational velocity of 300 km s^{-1} , while the dashed line gives the zero velocity result, only including mass loss (Meynet & Maeder 2003)

6.9 Structure before explosion

The structure of the star just before collapse is extremely important for the outcome of the subsequent phases, including the supernova explosion. In Fig. 16 we show the abundance structure of a $15 M_{\odot}$ and a $25 M_{\odot}$ star shortly before core collapse. Because of mass loss on the MS and in the RSG phase, $\sim 3 M_{\odot}$ were lost before the explosion for the $15 M_{\odot}$ model, while the $25 M_{\odot}$ had lost $\sim 12 M_{\odot}$, illustrating the increasing importance of mass loss for the massive stars. Except for this, the general structure of the two models are similar.

The first thing to note is the pronounced onion shell structure with a number of distinct zones, reflecting the different burning stages at the time of core collapse. Concentrating first on the $25 M_{\odot}$ model, from the surface inwards to $\sim 8.2 M_{\odot}$ we have the unprocessed hydrogen envelope. We here note that the He mass fraction is roughly twice the original, and is

actually higher than that of H. Consistent with this is the high N abundance, depressed O abundance and nearly zero C abundance. This is a typical signature of CNO burning products, which have been brought to the surface by convective mixing of the envelope during the RSG phase. In the $15 M_{\odot}$ model this mixing is negligible, but this conclusion is sensitive to factors like rotation and convective treatment.

Inside the hydrogen envelope, the helium mantle has almost the same thickness in mass in both stars. The total mass of He produced is, however, different, because a large fraction of the He from the shell-burning has been mixed with the hydrogen envelope (see above). Besides He, the most important elements are Ne and C. Most of these are the result of He-burning, mixed throughout the He shell.

In terms of nucleosynthesis the most important region is the oxygen zone. The mass of the oxygen core varies strongly with ZAMS mass. For the $15 M_{\odot}$ model it is $3.1 M_{\odot}$, while it is $7.2 M_{\odot}$ for the $25 M_{\odot}$ model. The most abundant elements are O ($\sim 80\%$) and Ne ($\sim 20\%$). Also Mg has a substantial abundance in this region. The inner O zone has a large abundance of nuclei resulting from Ne and O burning, in particular Si, S, Ar, as a result of convective mixing from the O burning shell. These dominate completely in a thin region inside the O zone. The central $\sim 1.6 M_{\odot}$ consists of iron group elements from Si-burning. Between the iron core and oxygen core are products of oxygen burning and incomplete Si-burning.

The density structure before core collapse is important for the properties of the supernova explosion. Fig. 17 shows the density as function of mass from the center for stars of different masses. No mass loss has been included in these models, but the general structure including this is quite similar.

One can roughly distinguish three regions. Most of the volume consist of a very extended hydrogen envelope, with radius $\gtrsim 3 \times 10^{13}$ cm containing most of the mass. The mass of this is, not very surprising, sensitive to mass loss. The radius is, however, characteristic of a red supergiant as long, as there is at least $\sim 1 M_{\odot}$ of hydrogen left.

Next comes the helium mantle with radius $\sim 10^{11}$ cm and density $\sim 10^2$ g cm $^{-3}$. Inside this there is a gradual increase in the density in the oxygen core. Finally, in the inner $\sim 1.5 M_{\odot}$ we have the iron core with a radius of only $\sim 3 \times 10^8$ cm and a density $10^7 - 10^9$ g cm $^{-3}$.

When we compare the structure of the different models, the most apparent differences are the mass of the helium core, and the density gradient outside the iron core. In general the latter becomes less steep as the mass of the star increases. This will be important when we discuss the propagation of the shock wave after the bounce.

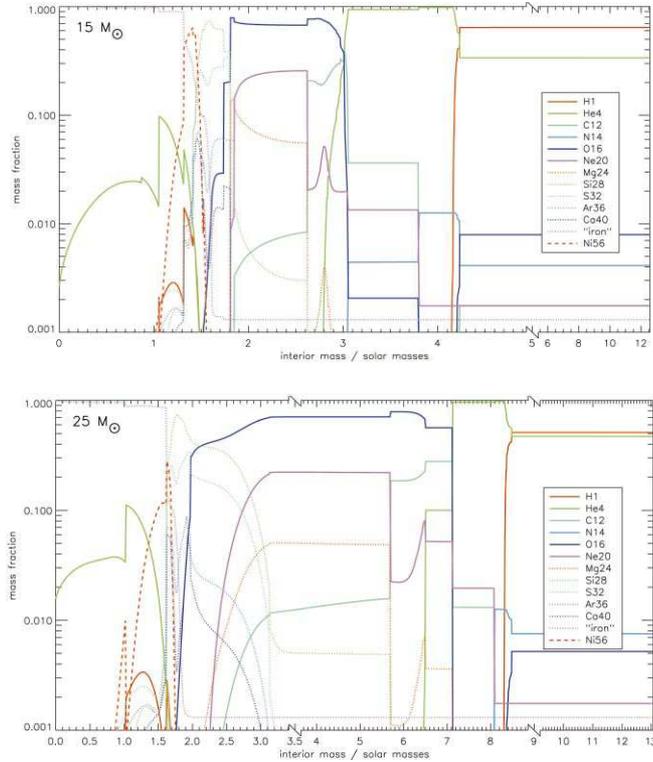


Figure 16: Abundance structure of a $15 M_{\odot}$ and $25 M_{\odot}$ star shortly before core collapse. (WHW02)

The mass of the Fe-core is crucial for whether the explosion will be successful or not. While this mass is only weakly dependent on the ZAMS mass in this interval, $13 - 25 M_{\odot}$, the small difference between $1.29 M_{\odot}$ and $1.53 M_{\odot}$ for the $13 M_{\odot}$ and $25 M_{\odot}$ models, respectively, may, however, be crucial for the outcome.

6.10 Triggering of collapse

A self-gravitating body is stable to perturbations as long as its adiabatic index is larger than $4/3$. As soon as $\gamma < 4/3$ the star (or core) will collapse.

There are several reasons why the core collapses. Photo-disintegration of Fe into α particles, or even nucleons, require ~ 8.8 MeV per nucleon. This takes away thermal energy from the core, and thus pressure support. A further reason for instability comes from electron capture on free protons

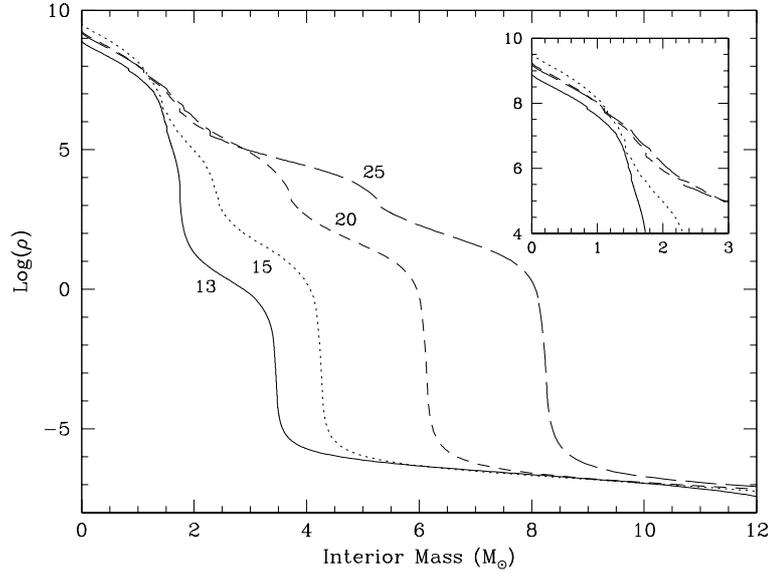


Figure 17: Density structure of four stars with ZAMS 13 – 25 M_{\odot} star. (Limongi et al. 2002)

and on bound protons in nuclei,



The removal of electrons again decreases the pressure in the core.

Once the core becomes unstable it will collapse on roughly a dynamical time scale, $t_{dyn} \sim R/v$, Assuming that the core collapses with the free-fall velocity, V_{ff} , we get

$$t_{dyn} = \frac{R}{V_{ff}} = \left(\frac{R^3}{2GM} \right)^{1/2} \quad (6.28)$$

In terms of the density this is

$$t_{dyn} = \left(\frac{3}{8\pi G\rho} \right)^{1/2} \quad (6.29)$$

At the edge of the iron core the density is $\sim 10^8 \text{ g cm}^{-3}$ when core collapse sets in, and at the center it is $\sim 3 \times 10^9 \text{ g cm}^{-3}$. Therefore,

$$t_{dyn} = 0.13 \rho_8^{-1/2} \text{ s} \quad (6.30)$$

and the collapse time scale is therefore of the order of milliseconds. Hydrodynamical models show that in reality the velocity is only 0.5 – 0.8 of the free-fall velocity, but as an order of magnitude estimate Eq. (6.30) is sufficient.

6.11 Neutrino trapping

During core collapse neutrinos produced in the core are absorbed and scattered by the nucleons. The most important elastic scattering processes are scattering by free neutrons and protons, and coherent scattering against bound neutrons and protons in nuclei,

$$\nu + n \rightarrow \nu + n \quad (6.31)$$

and

$$\nu + p \rightarrow \nu + p \quad (6.32)$$

and

$$\nu + (Z, A) \rightarrow \nu + (Z, A). \quad (6.33)$$

All three processes are mediated by neutral currents, and were before the Weinberg-Salam-Glashow electro-weak theory not considered. In fact, they make a crucial difference for the neutrino trapping, as we will see. Because of the coherence, the cross section of the last process is not only proportional to A , but to A^2 .

In addition to the scattering against the nucleons, *inelastic* scattering against electrons also take place,

$$\nu + e^- \rightarrow \nu + e^- . \quad (6.34)$$

The cross section of this is, however, only $\sim 1/600$ of that of the elastic nucleon scattering. Elastic scattering against the nucleons therefore dominate the inelastic scattering against the electrons by a large factor. When we estimate the mean free path to the scattering, we can therefore neglect the electrons. The inelastic scattering may nevertheless be important for thermalizing the neutrinos to the same temperature as the electrons.

For neutrino energies much less than $m_n c^2 \sim 1$ GeV the cross section for the nucleon scattering is

$$\sigma_\nu = \frac{1}{4} \sigma_0 \left(\frac{E_\nu}{m_e c^2} \right)^2 \quad (6.35)$$

where

$$\sigma_0 = \frac{4G_F^2 m_e^2}{\hbar^4} = 1.76 \times 10^{-44} \text{ cm}^2. \quad (6.36)$$

The mean free path for scattering is $\lambda_\nu = 1 / \langle n\sigma_\nu \rangle$, which is an average over the cross sections for these processes. An approximate expression for the mean free path is given by

$$\lambda_\nu \approx 2 \times 10^5 \left(\frac{E_\nu}{10 \text{ MeV}} \right)^{-2} \rho_{12}^{-1} \text{ cm} \quad (6.37)$$

The typical neutrino energy is ~ 20 MeV, so the mean free path is only $\sim 0.5 \rho_{12}^{-1}$ km.

Scattering is a diffusion process, and from the diffusion equation in spherical geometry one finds that the time for a neutrino to diffuse a radial distance R is

$$t_{diff} = \frac{R^2}{3\lambda_\nu c} \quad (6.38)$$

If we assume a uniform density sphere of mass $1.4 M_\odot$ and estimate the neutrino energy as the Fermi energy we get $E_\nu \approx E_F = 36.8 \rho_{12}^{1/3}$ MeV, and

$$t_{diff} = 5.2 \times 10^{-2} \rho_{12} \text{ s} \quad (6.39)$$

The diffusion time scale should be compared to the dynamical time scale, t_{dyn} , from Eq. (6.30),

$$\frac{t_{diff}}{t_{dyn}} = 40 \rho_{12}^{3/2} \quad (6.40)$$

Therefore, we find that *above a density of $\sim 10^{11}$ g cm $^{-3}$ the neutrinos become completely trapped in the core.* This has the consequence that the lepton number will be conserved in the core. If neutrino trapping would not set in, the lepton number would have decreased to a very low level because of beta decay and inverse beta decay (K-capture). This now instead happens **after** the collapse in the explosion phase when the density is low enough.

6.12 Collapse

Because of neutrino trapping, the collapse proceeds almost adiabatically. In Fig. 18 we show the velocity profile of the infalling core at different epochs during the collapse and the formation of the shock. The first curve, (a), corresponds to the final stages of the infall. The central density is $\sim 4.8 \times 10^{13}$ g cm $^{-3}$. At the next epoch the central density is $\sim 2.6 \times 10^{14}$ g cm $^{-3}$. The compression at the center generates sound waves which

propagate outwards. Curve (c) corresponds to maximum density, $\sim 9.7 \times 10^{14} \text{ g cm}^{-3}$. The inner 10 km is now at rest, and from the discontinuity in the velocity curve, it can be seen that the shock has just formed. In (d) the outgoing shock is very obvious. The central density is $\sim 6.9 \times 10^{14} \text{ g cm}^{-3}$, and the core is adjusting to its final density $\sim 4 \times 10^{14} \text{ g cm}^{-3}$. As we will

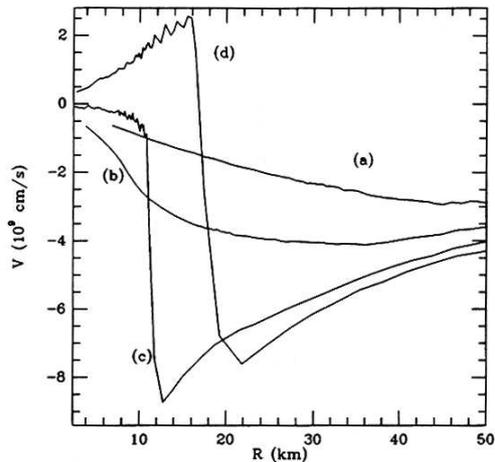


Figure 18: Velocity profile in the core at four epochs during collapse and explosion. Curve a) corresponds to the last epochs of the infall. b) The density at center is close to nuclear. The matter outside is falling in with increasing velocity. c) The infalling mass at the center has been brought to rest, while the accretion is occurring outside. d) The shock has been launched. (Cooperstein & Baron 1990)

see in next section, an important point for the survival of the shock is that it is not formed at the center of the star, but close to the outer edge of the core, at a mass of $\sim 0.5 M_{\odot}$ from the center, or $\sim 20 \text{ km}$.

6.13 Energy losses in the shock

The total energy of the shock, as it is launched outside the core, is roughly the binding energy of this, $\sim 5 \times 10^{51} \text{ ergs}$. While the initial energy of the shock is large enough to overcome gravity, the problem is that there are severe energy losses behind the shock, which takes away energy from it. These are mainly due to photo-disintegration of iron by the shock and due to neutrino losses in the hot gas behind the shock. Depending on the temperature, the photo-disintegration may proceed all the way to nucleons,

or for lower temperature to α particles. Total disintegration of an Fe nucleus to nucleons requires 8.8 MeV per nucleon. Therefore, for each $0.1 M_{\odot}$ of iron outside the core $\sim 1.5 \times 10^{51}$ ergs is lost by this process.

Success or failure depends on several factors. The mass of the iron core is crucial, to avoid disintegration losses. The smaller, the better. The maximum mass for an explosion is $\sim 1.2 M_{\odot}$. This is sensitive to factors like the $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$ reaction rate and the treatment of convection. Another important factor is the stiffness of the equation of state. A soft equation of state above nuclear density favors explosion. Unfortunately, this is the most uncertain regime of the equation of state.

Summarizing the current situation, the prevailing view is that this prompt explosion mechanism will probably not work without some additional energy input, or other ingredient, except possibly for stars of mass $\lesssim 12 M_{\odot}$.

6.14 Neutrino heating

In a now classical computer run, Jim Wilson let one of his apparently unsuccessful simulations run for a much longer time than before. When he looked at the result he saw, to his surprise, that the stalled shock had now become a successful one due to the late energy input from the neutrinos from the explosion and the newly formed hot proto-neutron star.

In Fig. 19 we show the resulting mass locations as function of time for this simulation. As we see, the collapse occurs on a time scale of a few tens of milliseconds. The shock forms at a distance of ~ 100 km from the center. This expands, but because of energy losses it loses speed, and after ~ 0.1 s it is almost stalled. Accretion continues, and normally one would consider this a failed explosion. However, because of the longer than normal simulation, we see that at ~ 0.5 s the shock suddenly gets new energy and rapidly expands out of the core. At that point the density at the shock is low enough for losses to be negligible, and the result is a healthy explosion.

Although successful, these simulations were based on an unrealistic equation of state above nuclear density, as well as other deficiencies. Nevertheless, what was most important was that they pointed out the importance of neutrino heating and the consequences of this on a long time scale. We will now discuss the details of this mechanism in more detail.

The proto-neutron star formed by the collapse cools by an enormous flux of neutrinos on a time scale given by the neutrino diffusion time scale t_{diff} , given by Eq. (6.39), but now with $\rho \sim 2 \times 10^{14}$ g cm $^{-3}$. Finally, as the density decreases at the boundary of the iron core, the density will be low enough for the neutrinos to escape freely. One can therefore, in analogy with

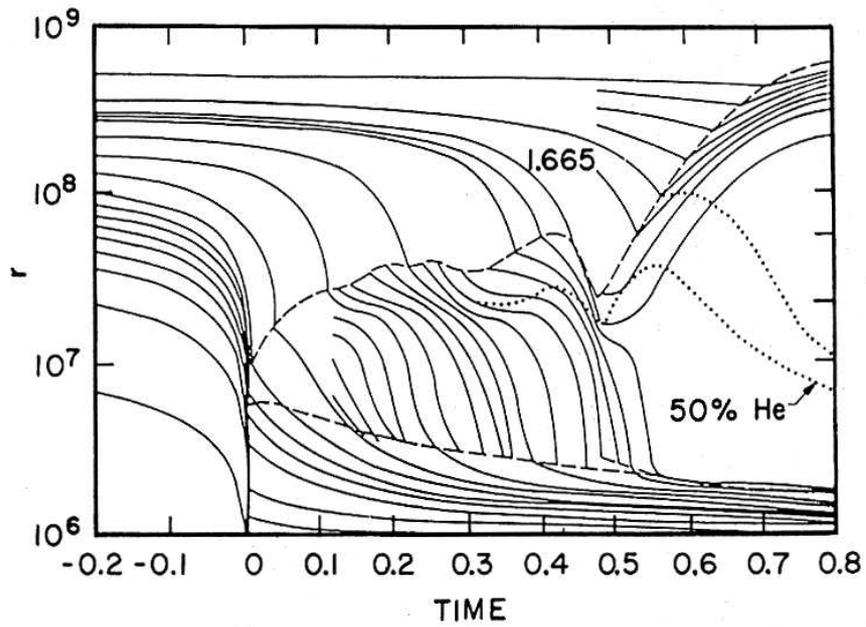


Figure 19: Trajectories of different mass shells for Jim Wilson's successful explosion model. The upper dashed line gives the position of the shock, while the lower is that of the neutrino photosphere. The dotted line marks the region where the abundance of He is 50%. Note the revival of the shock at 0.55 s due to neutrino heating (Bethe & Wilson 1985).

the photosphere, define a *neutrino-sphere*. As these propagate out through the shocked gas, they will scatter and be absorbed by the nucleons.

The most important heating processes are

$$\bar{\nu} + p \rightarrow e^+ + n \quad (6.41)$$

and

$$\nu + n \rightarrow e^- + p \quad (6.42)$$

The temperature behind the shock is set by the dis-integration of Fe. For each nucleon this costs a binding energy E_{bind} . Because the cooling decreases faster than the heating, there will be a radius where heating and cooling balance, often referred to as the gain radius. Inside of this there is a net cooling by the neutrinos, while outside there is a net heating. This is summarized in Fig. 20.

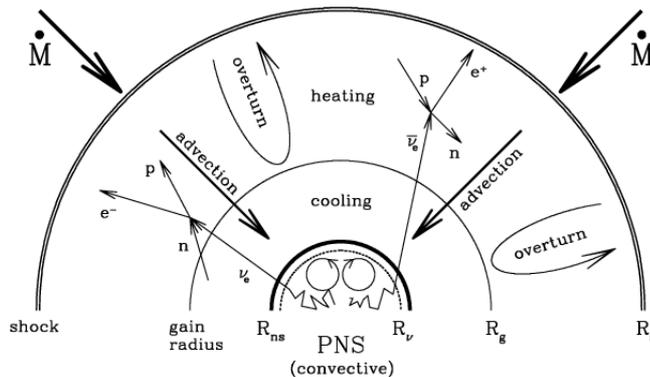


Figure 20: Schematic picture of the different regions close to the neutron star (Janka 2001)

The neutrino heat input is probably crucial for the outcome of the explosion. Just outside the gain radius the matter is heated to a high temperature, decreasing outwards. This therefore induces a strong entropy gradient outwards, leading to convection. The large scale convective motions transport entropy (hot gas) to the region close to the shock and can therefore re-energize the shock (see Fig. 19 and Fig. 21). At the same time low entropy (cool) gas sinks inward, inside the gain radius. There it will be heated by the neutrinos. This compensates the iron dis-integration losses, and if efficient enough, can cause the shock to survive through the whole iron core.

The efficiency of the late heat input mechanism depends naturally on the neutrino luminosity from the proto-neutron star. The neutrino emission is a

consequence of the de-leptonization of the proto-neutron star. The ν_e and $\bar{\nu}_e$ are produced by $e^- + p \rightarrow \nu + n$ and $e^+ + n \rightarrow \bar{\nu}_e + p$, while $e^+ + e^- \rightarrow \bar{\nu}_i + \nu_i$ are responsible for the production of μ and τ neutrinos, by roughly equal numbers.

Taking all these ingredients together, the increased neutrino luminosity from the proto-neutron star, the neutrino heating of the gas of the matter outside the neutron star and the convective motions behind the shock one has in some simulations been able to obtain an explosion. However, several of these successful attempts have weaknesses in the form of the treatment of the neutron star cooling, the radiative transfer of the neutrinos or the equation of state. It is therefore too early to make any conclusions of the success of this mechanism. In addition, there are other ingredients, like magnetic fields and rotation, which have only been included in simplified models. They may, however, be crucial for the outcome. In particular, the gamma-ray bursts may indicate the necessity to include these effects.

6.15 Explosive nucleosynthesis

If the shock has managed to escape the iron core, the density decreases rapidly, and with that the disintegration and neutrino losses. There will still be some explosive nucleosynthesis in the silicon and inner parts of the oxygen shell, but this only adds a small amount of energy to the shock. After this the shock will propagate through the whole star and disrupt this. As the shock reaches the surface the hot photons behind the shock are released and one observes the explosion. This is however several hours after the collapse of the core.

As the shock wave propagates through the silicon and oxygen rich gas close to the iron core, the temperature behind the shock will be high enough for explosive nucleosynthesis to take place in these regions. This burns most of the Si and S into nuclear statistical equilibrium, (see Fig. 22).

During the first seconds after the core bounce some of the most important isotopes are formed. Close to the border between the neutron star and the ejecta the explosive nucleosynthesis occurs in conditions close to NSE. As we saw in §6.4, it is therefore not surprising that the most abundant nucleus is ^{56}Ni . The exact mass of ^{56}Ni , which will be ejected, depends sensitively on where the split is between the matter falling into the neutron star and that expanding out. This is usually known as the 'mass cut'. Typically, the ^{56}Ni mass is $\sim 0.1 M_\odot$, but this can vary by a large factor, both upwards and downwards.

In addition to ^{56}Ni , substantial amounts of ^{57}Ni and ^{44}Ti are created.

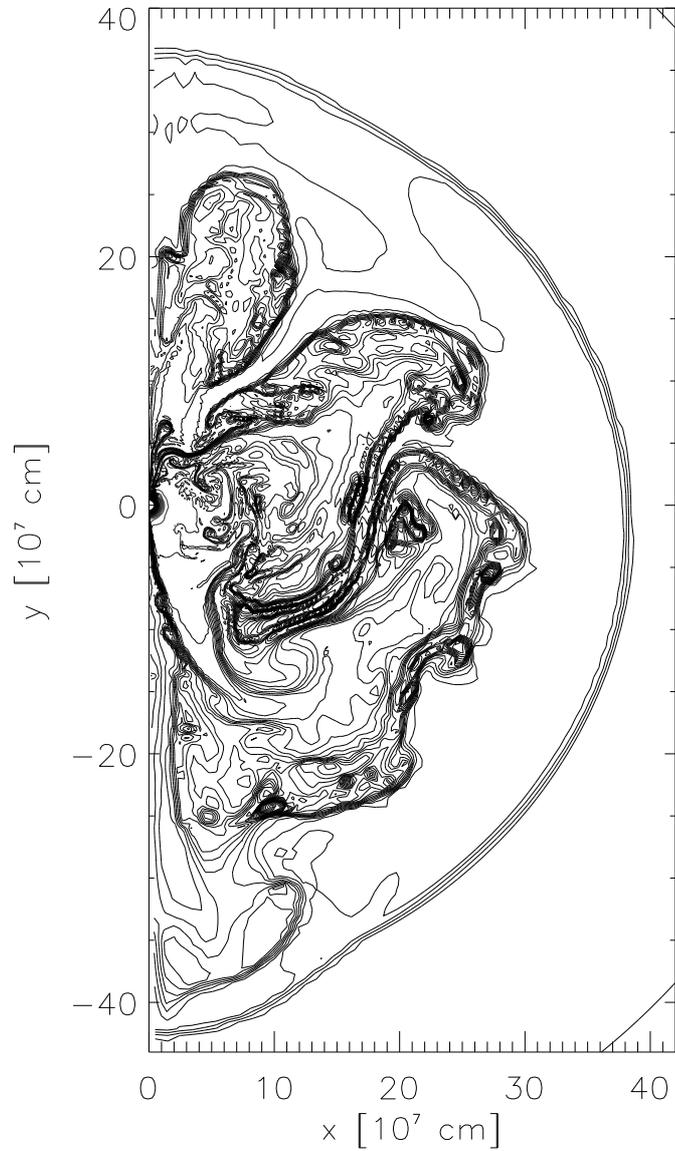


Figure 21: Two-dimensional simulation of the explosion at 0.38 s after core bounce. The contours show the entropy distribution. The shock is at 3800 km. Note the bubbles of neutrino heated gas, and the down drafts of sinking cooler gas. (Janka & Müller 1996)

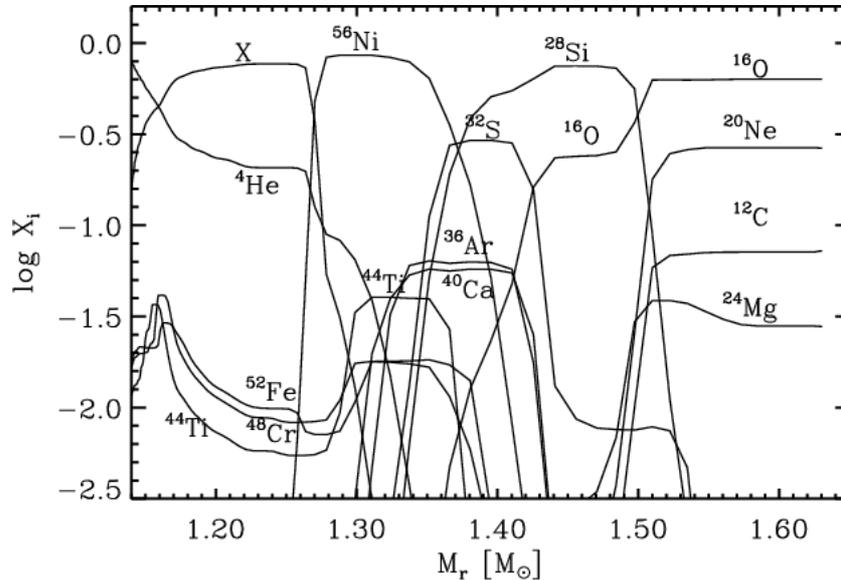


Figure 22: Explosive nucleosynthesis in the central region (Kifonidis et al. 2003).

The exact abundances of the three radioactive isotopes depend on the density, temperature and neutron excess. Therefore, a determination of these abundances provides a useful probe of the conditions at the time of the explosive nucleosynthesis, during the first seconds of the explosion.

Outside the silicon core, in the inner oxygen shell, the shock velocity and density are still high enough for the inner parts of the oxygen core to be transformed into Si/S. At this point the density becomes too low for any significant nucleosynthesis to take place. Outside of the inner oxygen shell, the composition just before core collapse is almost unaffected by the explosion. Summarizing the explosive nucleosynthesis, the most important elements affected by this is oxygen and elements heavier than this.

In addition to these elements, which account for most of the newly created elements in terms of mass, there is also a further process, which is extremely important for the elements beyond the iron peak. This relies on the fact that close to the mass cut, where the density and temperature is high, the abundance of free neutrons is also high. This high neutron flux can be absorbed by the different abundant iron group nuclei in this region, which leads to the build up of heavier, neutron-rich isotopes. These are usually unstable to especially beta decay, which creates more stable isotopes of

the same mass, but higher charge. This process is known as the r-process, r for rapid.

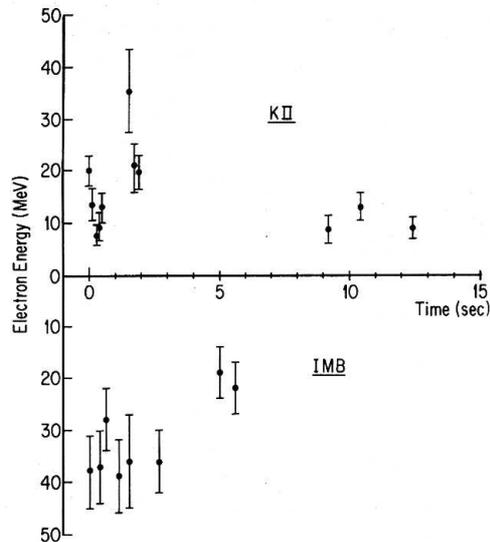


Figure 23: Energy versus arrival time for the detected neutrinos in the Kamiokande II and IMB detectors. (Burrows 1988).

7 Observables of core collapse supernovae

7.1 Neutrinos from SN 1987A

The most unique observation of SN 1987A is the first observation of neutrinos from outside of the solar system. Although already Chiu and Colgate & White in the 1960's had predicted that most of the gravitational energy in the collapse would emerge as neutrinos, the flux from supernovae at 'normal' distances is too low for the current (and probably next) generation of neutrino detectors. However, the small distance to SN 1987A, 50 kpc, meant the the flux was $\gtrsim 10^4$ larger than from a supernova in even the closest galaxies outside the Local Group.

As soon as the news of the discovery of SN 1987A came, the different teams looked at the registration journals of the most sensitive detectors, Kamiokande II in Japan and IMB (Irwine – Brookhaven – Michigan) in Ohio. To their satisfaction they saw a clear signal at exactly the same time, February 23 at 07:35:41 UT. Fig. 23 shows the energies of the individual electrons produced by the neutrinos, which approximately corresponds to a neutrino energy given by $E_\nu \approx E_e + 2$ MeV.

The total energy in the form of neutrinos is straightforward to calculate,

and was predicted long before SN 1987A. Because the kinetic energy of the shock is less than a percent of the total energy, what is emitted is just the binding energy of the neutron star formed. For a uniform density this is

$$E_b = \frac{3}{5} \frac{GM^2}{R} = 3.1 \times 10^{53} \left(\frac{M}{1.4 M_\odot} \right) \left(\frac{R}{10 \text{ km}} \right)^{-1} \text{ ergs} \quad (7.1)$$

Note that we here should use the radius of the cool neutron star 10–20 km. A more accurate calculation, taking the non-uniform density distribution into account, gives a similar result.

The duration of the burst is set by the diffusion time scale of the neutrinos as the proto-neutron star is deleptonized and is cooling down. The mean free path from Eq. (6.37) is therefore $\sim 10^6 \rho_{14}^{-1} (E_\nu/1 \text{ MeV})^{-2}$ cm. Using a constant density for the proto-neutron star with mass $\sim 1.4 M_\odot$, we have $\rho \approx 2.5 \times 10^{13} (R/30 \text{ km})^{-3}$. Using these expressions in the equation for the diffusion time, Eq. (6.38), we get

$$t_{diff} \approx 0.2 \left(\frac{R}{30 \text{ km}} \right)^{-1} \left(\frac{E_\nu}{100 \text{ MeV}} \right)^2 \text{ s.} \quad (7.2)$$

Typically, the neutrino energies are of the order of 100-200 MeV in the inner core.

In reality, the density in the center is higher than the mean density used above, and the neutrino energies also vary by a large factor, so this number should only be taken as indicative. The fact that it is much larger than the dynamical time scale, however, shows the importance of the neutrino diffusion. More accurate calculations show that the neutrinos diffuse out on a time scale of ~ 2 s.

Despite the high interior temperature the neutrinos which escape have a temperature of only $T_\nu \sim 4 - 5$ MeV, corresponding to a mean energy $\langle E_\nu \rangle \sim 3 T_\nu \sim 10 - 15$ MeV. Because of the high temperature and trapping an approximately equal number of all six neutrino species, $\nu_e, \bar{\nu}_e, \nu_\mu, \bar{\nu}_\mu, \nu_\tau,$ and $\bar{\nu}_\tau$ are produced by pair annihilation, the plasmon process and nucleon bremsstrahlung. The energy in each of the neutrinos is therefore $\sim 6 \times 10^{52}$ ergs.

The Kamiokande II and IMB detectors are both water Cherenkov detectors, shielded by several 1000'nds of meter of rock. The total amount of water in these are 2140 tons for Kamiokande II and 6800 tons for IMB. Only electron neutrinos are detected with these water detectors. This occurs through absorption on the protons in the water

$$\bar{\nu}_e + p \rightarrow e^+ + n \quad (7.3)$$

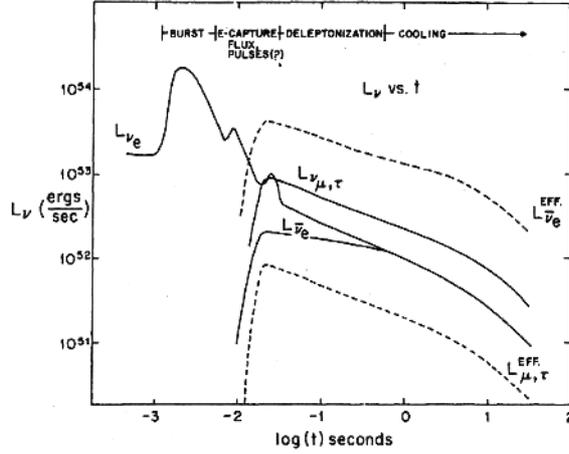


Figure 24: Approximate neutrino luminosities of the different neutron species. Note the logarithmic time scale. (Burrows 1984).

and elastic scattering including all neutrino species, $i = e, \mu, \tau$,

$$\nu_i + e^- \rightarrow \nu_i + e_i^- \quad (7.4)$$

The second reaction has a cross section which is only $\sim 5\%$ of that of the first reaction, which therefore dominates. The neutrino detectors are not sensitive to neutrinos below ~ 7 MeV for Kamiokande II and below ~ 19 MeV for IMB.

To compare the detected signal with the predicted one has to convolve the detector sensitivity with the calculated spectrum. Because of the small numbers (19 neutrinos in total) one compares the predicted curves with the cumulative number of neutrinos detected.

The average observed temperature of the neutrinos is $3 - 5$ MeV, and agrees well with that estimated before the explosion. A recent calculation of the mean energy, $\langle E_\nu \rangle \sim 3 T_\nu$. The shorter duration of the IMB signal, which has a higher threshold, actually gives some indication that the source is cooling.

7.2 Supernova classification

The classification of supernovae into different types and subtypes is basically an empirical scheme, based on spectral features and light curves. However,

it turns out that this classification also corresponds to important physical differences between them.

Most SNe are discovered shortly after explosion when they are near maximum luminosity. It is therefore important to be able to distinguish the different types from the early spectrum. In Fig. 25 we show a collection of spectra representing the most important types.

The observationally most obvious difference between various SNe is whether or not they have any hydrogen features in their spectra. Type I SNe are defined as those without and Type II as those with $H\alpha$. A closer examination, however, shows that there are important differences within both of these classes. The most important are between the extremely heterogeneous Type I's. These are shown as the top and lower two spectra in Fig. 25. While neither of them have any trace of $H\alpha$, the observational distinction is that the Type Ia's have a strong Si II $\lambda 6150$ line, while the Type Ib's and Ic's lack a strong feature at this wavelength. The distinguishing feature between the Ib's and Ic's is the presence of He I lines in the former, while the Ic's lack any trace of helium.

While the differences between the Ia's and Ib/c's seem marginal, it turns out that they originate from completely different explosion mechanisms. The Ia's are thermonuclear explosions of white dwarfs, completely disrupting the star, while the Ib/c SNe and the Type II's are core collapse explosions of massive stars, leaving a neutron star or black hole.

This distinction can somewhat better be understood from an examination of the spectra at late epochs. Fig. 26 shows a collection of spectra taken 5 months after explosion. Unfortunately, however, no Type Ib is included, but they are qualitatively similar to the Type Ib's. Now the difference between the Ia's and Ib/c SNe become very large. While the Ia spectra are dominated by [Fe II], [Fe III] and [Co III] features, the most prominent features in the Ic spectrum are due to [O I], [Mg I] and [Ca II], with only weak iron lines. Also the Type II's have late spectra where the same lines are strong, although they tend to appear later than for the Ib/c's.

One can now start to appreciate the physical difference between these classes. The presence of substantial amounts of oxygen, magnesium and calcium is characteristic of the processed regions of a massive star, The Type Ia spectra with only weak features of lines from these elements and strong lines of iron are more typical of matter which has undergone complete burning to nuclear statistical equilibrium. The reason that the Type Ib/c's lack hydrogen is most likely because they have lost their hydrogen envelopes, either as a result of mass loss or binary mass exchange. The progenitors are therefore believed to be Wolf-Rayet stars. A more quantitative confirmation of

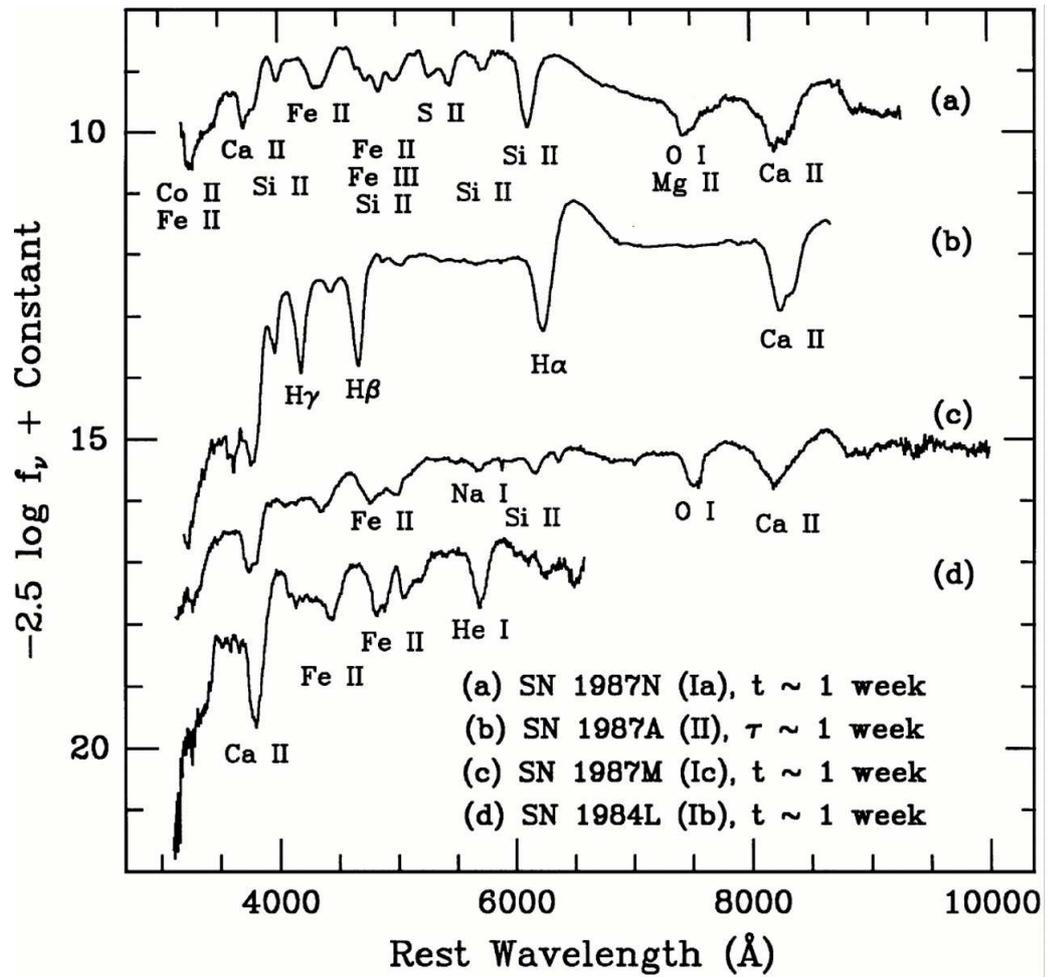


Figure 25: Spectra of different supernova types one week after explosion (Filippenko 1997).

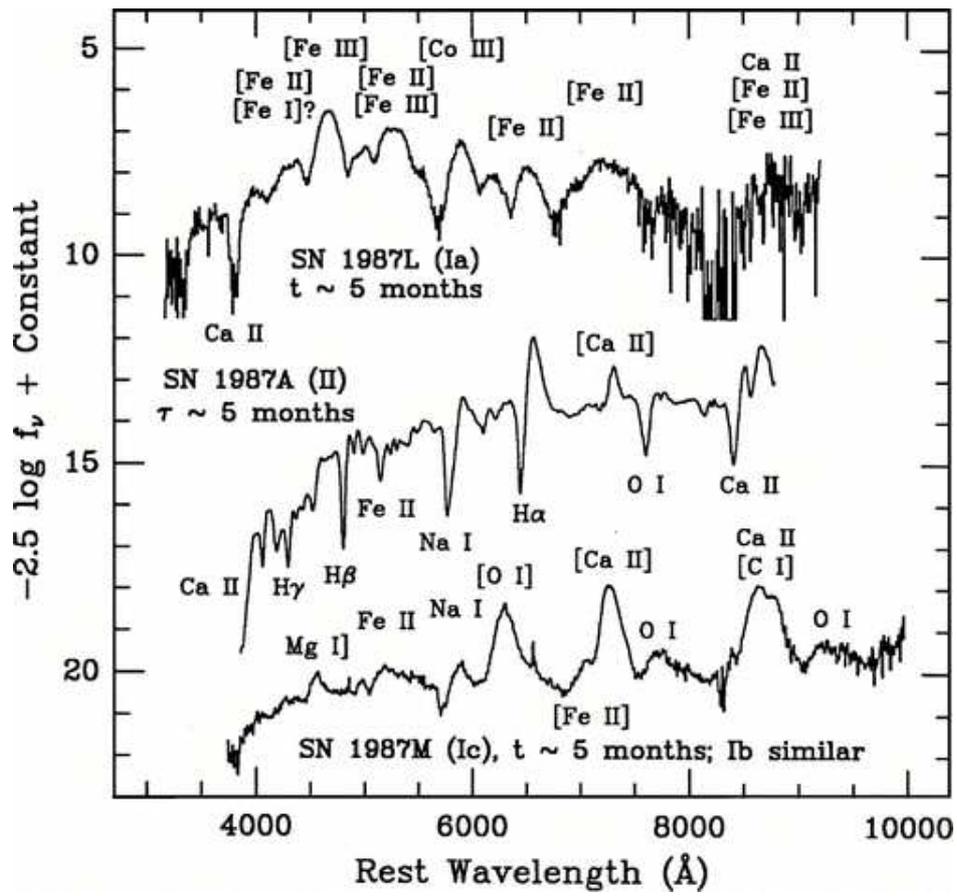


Figure 26: Spectra of different supernova types 5 months after explosion (Filippenko 1997).

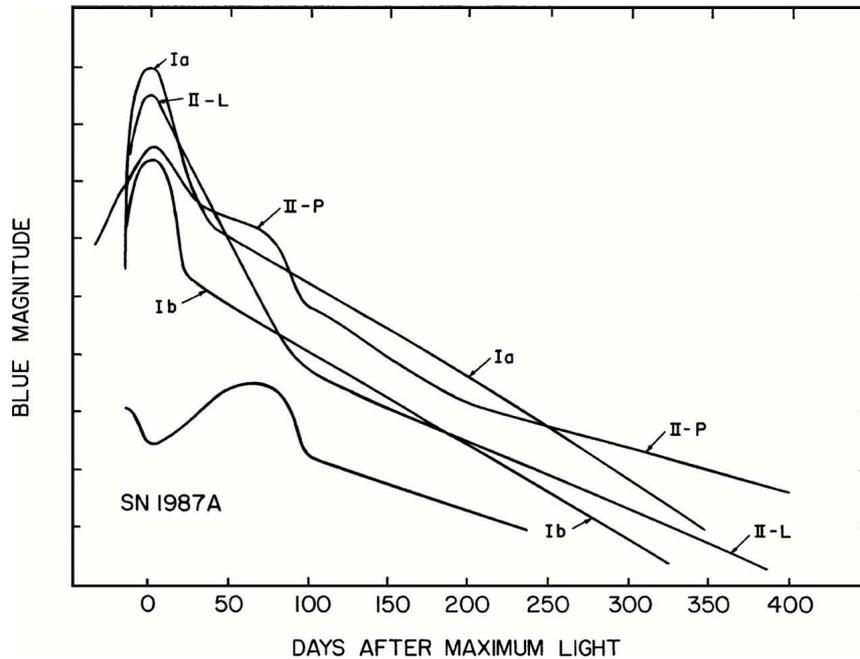


Figure 27: Light curves of different supernova types (Filippenko 1997).

this, requires a much more detailed analysis of their spectra. This, however, completely confirms these conclusions.

While the previous classification has only been discussed from the spectra, there are also important differences with regard to the light curves of the different SN types. In Fig. 27 we show light curves representative of the different types. Here the Type Ia and Type Ib/c curves mainly differ in terms of absolute luminosity. For the first two months they are characterized by a bell shaped peak, occurring 2-3 weeks after explosion. They then have a nearly linear decline in a time - magnitude plot for the rest of the evolution. However, while the Type Ias are highly standardized, there is a large dispersion within the Type Ib/c curves, both in absolute luminosity and in the shape. In particular, the rate of decline after ~ 50 days differ considerably.

While the Type IIs have a fairly large range of light curves, one can distinguish two main types. The IIPs which are the most frequent, are characterized by a fairly fast rise to a peak. After a decline by a magnitude or so, they then stay at nearly constant magnitude for ~ 100 days. This is the reason for the P = plateau. After this there is a drop by a magnitude

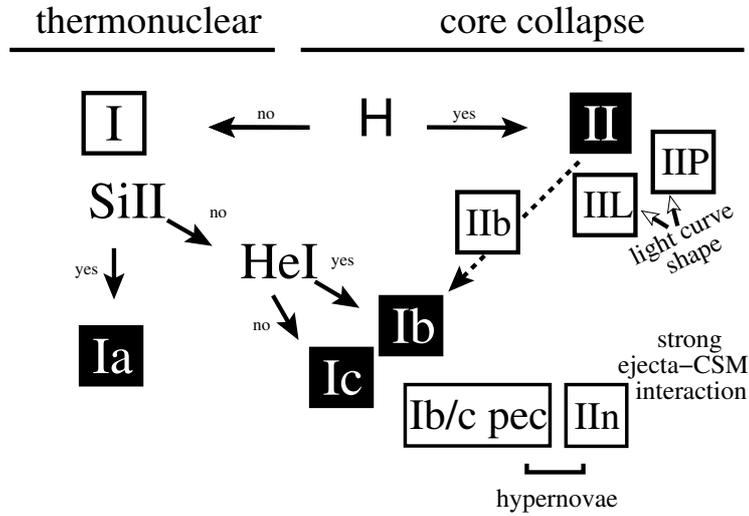


Figure 28: Classification scheme based on spectra and light curves of different kinds of supernovae (Turatto, 2003).

or more, and then they decline with a fairly uniform rate of ~ 1 magnitude per 100 days.

A less frequent class are the Type IILs, where the L stands for linear. The plateau is here lacking and the linear decline sets in shortly after the peak. Usually, the Type IILs are considerably brighter in absolute luminosity than the Type IIPs. Spectroscopically, the Type IILs have already at early epochs a strong $H\alpha$ line in emission, while $H\alpha$ has usually a classical P-Cygni profile in the Type IIPs.

In addition to these main classes, there are several other subtypes, with more or less distinct properties. This is usually connected to interacting with a dense circumstellar medium. In Fig. 28 we summarize the whole SN classification scheme, and we will now discuss the physical interpretation of these characteristics, and the differences between the various types.

7.3 Radioactivity

As we saw in §6.15, the explosive nucleosynthesis in the silicon core resulted in several radioactive isotopes, the most important being ^{56}Ni , ^{57}Ni and ^{44}Ti . All of these have comparatively short half-lives, and the decays of these elements can therefore be directly observed, and are in fact crucial for the observability of the supernova. The decays are characterized by either

the half-life, $t_{1/2}$, or the exponential decay time scale, τ . It is easy to see that $\tau = t_{1/2} / \ln 2$.

^{56}Ni decays on a time scale of $\tau = 8.8$ days by electron capture as



In this process it emits gamma-rays with energies 0.158 – 0.812 MeV (see Fig. 29). The ^{56}Co isotope resulting from this decay is, however, not stable either, but decays by electron capture or by positron decay according to



The first decay occurs in 81% of the cases and the second in the remaining 19%. In terms of energy going into gamma-rays and positrons these numbers are 96.4% and 3.6%, respectively. The strongest gamma-ray lines are at 0.847 MeV and 1.238 MeV. The average positron energy is 0.658 MeV. Similarly, ^{57}Ni decays by electron capture as



with a very short decay time $\tau = 52$ hours. The more interesting decay is

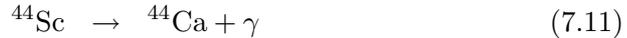


with $\tau = 390$ days.

Finally, ^{44}Ti decays first to ^{44}Sc on a time scale of ~ 89 years.



and then rapidly ($\tau = 5.4$ hours) to



(see Fig. 30).

The result of these radioactive decays are either gamma-rays or positrons. The gamma-rays are scattered by the electrons in the ejecta through Compton scattering. In each scattering they lose roughly half of their energy to the electrons. Because the energy of the gamma-rays are initially in the MeV range, much higher than the binding energies of the bound electrons in the atoms, both free and bound electrons contribute to the scattering.

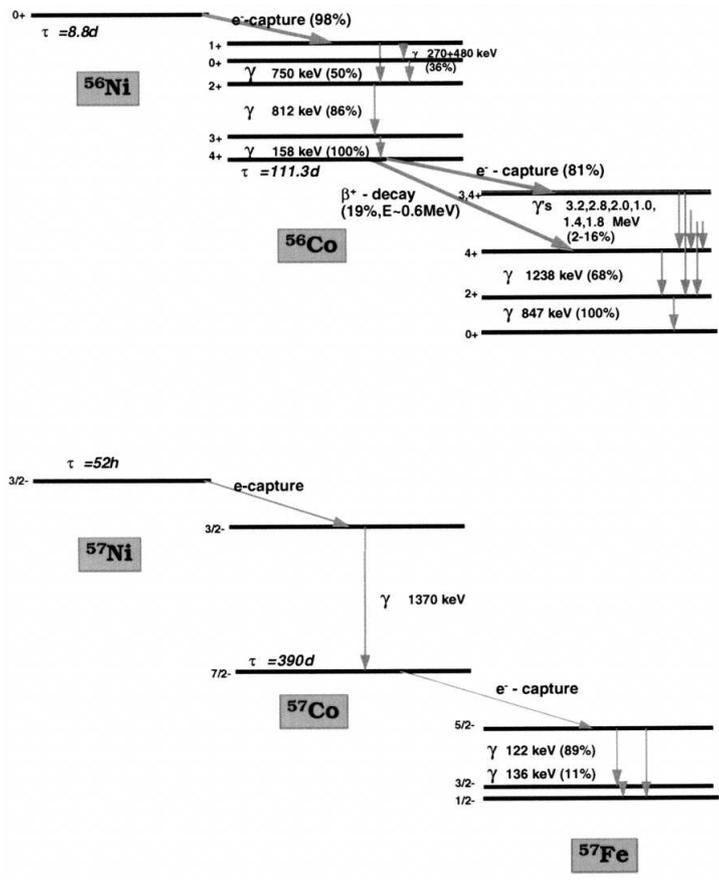


Figure 29: Transitions in the ^{56}Ni and ^{57}Ni decays (Diehl & Timmes 1998).

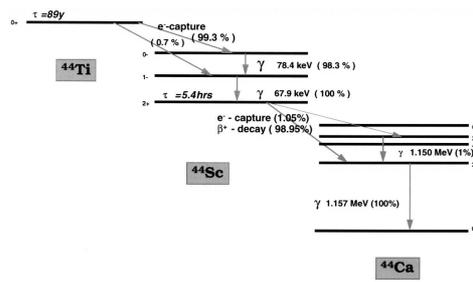


Figure 30: Transitions involved in the ^{44}Ti and ^{44}Sc decays (Diehl & Timmes 1998).

This down-scattering of the gamma-rays continues until the cross section for photoelectric absorption is larger than the Compton cross section, which occurs at an energy of $\sim 10 - 100$ keV, depending on the composition. The most important element for the photoelectric absorption is iron.

The total gamma-ray luminosity from the various decays is given by

$$L_\gamma = 1.27 \times 10^{42} \left(\frac{M(^{56}\text{Ni})}{0.1 M_\odot} \right) e^{-t/111.3^{\text{d}}} + \quad (7.13)$$

$$6.9 \times 10^{38} \left(\frac{M(^{57}\text{Co})}{5. \times 10^{-3} M_\odot} \right) e^{-t/390.^{\text{d}}} + \quad (7.14)$$

$$4.1 \times 10^{36} \left(\frac{M(^{44}\text{Ti})}{10^{-4} M_\odot} \right) e^{-t/89.^{\text{yrs}}} \text{ erg s}^{-1}, \quad (7.15)$$

and the positron input by

$$L_+ = 4.44 \times 10^{40} \left(\frac{M(^{56}\text{Ni})}{0.1 M_\odot} \right) e^{-t/111.3^{\text{d}}} + \quad (7.16)$$

$$1.3 \times 10^{36} \left(\frac{M(^{44}\text{Ti})}{10^{-4} M_\odot} \right) e^{-t/89.^{\text{yrs}}} \text{ erg s}^{-1} \quad (7.17)$$

7.4 Light curves

7.4.1 The diffusion phase of the light curve

After shock breakout the radiation will leak out on a diffusion time scale. We have already estimated this in Eq. (6.38), which we write as

$$t_{diff} = \frac{3R^2 \rho \kappa}{\pi^2 c} \quad (7.18)$$

This should be compared to the expansion time scale $t = R/V$. Taking the opacity to be that of Thompson scattering, $\kappa = 0.4 \text{ cm}^2 \text{ g}^{-1}$, and assuming a uniform density for the envelope we get

$$\frac{t_{diff}}{t_{exp}} = 1.9 \left(\frac{M}{M_\odot} \right) \left(\frac{V}{10^4 \text{ km s}^{-1}} \right) \left(\frac{R}{10^{15} \text{ cm}} \right)^{-2} \quad (7.19)$$

For a typical mass of $10 M_\odot$ we therefore find that not until the supernova has expanded to $R_{peak} \sim 4 \times 10^{15} \text{ cm}$, after $t_{peak} = R/V \sim 40$ days, can the radiation leak out faster than the ejecta expand. This is analogous to the neutrino trapping discussed in §6.11, although in this case it is expansion, rather than collapse.

Before t_{peak} the expansion is nearly adiabatic. Because the ejecta is radiation dominated it behaves as a $\gamma = 4/3$ gas, and the total thermal energy behaves like

$$E_{int} = (\gamma - 1)^{-1} pV = 3K\rho^{4/3}V \propto \rho^{1/3} \propto R^{-1} \quad (7.20)$$

Therefore, if the progenitor has a radius R_0 , the internal energy has decreased by a factor R_{peak}/R_0 once the photons can leak out. A small initial radius therefore means that almost all the internal energy produced by the shock has been lost into adiabatic expansion, i.e., to kinetic energy. If the thermal, shock energy was the only source of energy, supernovae coming from this kind of stars would be very faint. A red supergiant, on the other hand, could be bright for weeks just from the thermal shock energy.

Besides the thermal energy from the shock, there is one more important source for the light curve. As we saw in §7.3, the radioactive isotopes created in the explosion give rise to gamma-rays and positrons as they decay. These are losing their energy in the ejecta, thermalizing their energy into UV and soft X-ray photons, and therefore acts like an additional energy source. In the same way as the thermal energy from the shock, the photons undergo scatterings in the ejecta and only leak out when the diffusion time scale becomes comparable to the expansion time scale. The difference compared to the shock energy is, however, that this source is not affected by adiabatic expansion. The number of radioactive nuclei, of course, remain the same independent of the expansion. Therefore, even if nearly all the internal heat has been lost in the expansion, radioactivity provides a source for the light curve even at late times.

7.4.2 The late light curve and radioactive isotopes

After the peak diffusion plays a steadily decreasing role, and the light curve becomes simpler. This is in particular the case if we consider the bolometric light curve, i.e., the frequency integrated light curve. In this case the emitted luminosity is just the instantaneous gamma-ray and positron energy absorbed by the ejecta.

If we neglect the scattering in space and energy of the gamma-rays and just consider it as an absorption process, which is a reasonable approximation, although not very accurate, we can calculate the bolometric light curve just from the absorbed energy. As an averaged opacity one can for ^{56}Ni and ^{56}Co use $\kappa_\gamma = 0.06Z/A \text{ cm}^2\text{g}^{-1}$, where Z/A is the average charge to mass ratio of the ejecta. The positrons have a considerably smaller mean free path, and they can be considered to be stopped and annihilate on the same

spot as the radioactive decay. Further, since we neglect diffusion we are only considering epochs later than ~ 100 days. At these epochs all ^{56}Ni has decayed into ^{56}Co , and we can therefore neglect the first step in this chain.

For $t \gg \tau(^{56}\text{Ni}) = 8.8$ days we need only consider the ^{56}Co decay. Further, we assume that a fraction $(1 - e^{-\tau_\gamma})$ of the gamma-ray energy is trapped in the ejecta. Here τ_γ is an average optical depth to the gamma-rays. Adding the gamma-ray and positron contributions we get

$$L_{bol} = 1.27 \times 10^{42} \left(\frac{M(^{56}\text{Ni})}{0.1 M_\odot} \right) e^{-t/111.3^d} [(1 - e^{-\tau_\gamma}) + 0.035] \text{ erg s}^{-1}. \quad (7.21)$$

The first term in the square bracket represents the gamma-ray input and the second the positron input. As an estimate of the gamma-ray optical depth we take

$$\tau_\gamma = \kappa_\gamma \rho R = \kappa_\gamma \frac{3}{4\pi} \frac{M}{V^2 t^2} \quad (7.22)$$

$$= 0.38 \frac{Z}{A} \left(\frac{M}{M_\odot} \right) \left(\frac{V}{10^4 \text{ km s}^{-1}} \right)^{-2} \left(\frac{t}{100 \text{ days}} \right)^{-2} \quad (7.23)$$

We therefore see that the gamma-ray trapping is sensitive to both the mass and the expansion velocity. As an example we take SN 1987A, where most of the gamma-rays were trapped in the core. For the mass we therefore take $M \sim 4 M_\odot$ and for the expansion velocity of the core $\sim 2000 \text{ km s}^{-1}$. We then get $\tau_\gamma \sim 40(t/100 \text{ days})^{-2}$. The gamma-rays are therefore in this case trapped up to ~ 600 days. As another extreme case we can take a Type Ia supernova, with $M \sim 1.4 M_\odot$, $Z/A \sim 0.5$ and an expansion velocity $\sim 10,000 \text{ km s}^{-1}$. In this case we get $\tau_\gamma \sim 0.25(t/100 \text{ days})^{-2}$, and the ejecta is therefore transparent already at ~ 50 days, or earlier for higher ejecta velocities.

Eq. (7.21) shows that for $\tau_\gamma \gtrsim 1$ the bolometric light curve follows the radioactive decay time scale closely, $L_{bol} \propto e^{-t/111.3^d}$. For $0.035 \ll \tau_\gamma \ll 1$ the decay is, however, steeper with $L_{bol} \propto e^{-t/111.3^d}/t^2$. This dependence explains the steeper late light curves of the Type Ia, Ib, and Ic supernovae (§7.2).

From Eq. (7.21) we also see that the positrons become important when $\tau_\gamma \sim 0.035$. For slowly expanding ejecta, as for SN 1987A, the positron contribution does not become important before the next abundant radioactive isotope, ^{57}Ni , dominates the ^{56}Ni contribution. For rapidly expanding supernovae, like Type Ib/c supernovae or Type Ia supernovae, the positron contribution, however, becomes dominant for $t \gtrsim 300 - 500$ days. The bolometric luminosity then again follows the radioactive decay.

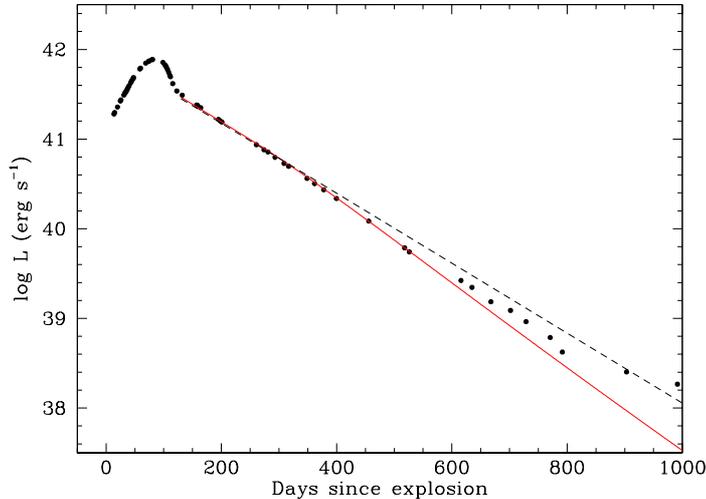


Figure 31: Bolometric light curve from ESO data for SN 1987A during the first 1000 days. The solid line gives the bolometric light curve from Eq. (7.21) with $M(^{56}\text{Ni}) = 0.07 M_{\odot}$ and $\tau_{\gamma} = 30(t/100^d)^{-2}$. The dashed line shows the total radioactive energy, including that escaping the ejecta. (Data from Bouchet et al. 1990).

7.4.3 The bolometric light curve of SN 1987A

As an example of the usefulness of the bolometric light curve we take SN 1987A. In Fig. 31 we show this during the first 1000 days. After the diffusion phase, which ends by day ~ 130 , and up to day ~ 400 , the light curve closely follows the predicted linear relation expected for full trapping, $M_{bol} = -t2.5 \log e/111.3^d + \text{const} = -t/102.5^d + \text{const}$. After day 400 there is an increasing deviation from the full trapping case, well fit by Eq. (7.21) with $\tau_{\gamma} = 30(t/100^d)^{-2}$, showing that some of the gamma-rays now escape the ejecta. Optical depth unity is reached after ~ 550 days. Most important, from the normalization of the curve one can determine the total nickel mass to $M(^{56}\text{Ni}) = 0.07 M_{\odot}$. The error in this mass is not more than 10%.

By day 600 the bolometric light curve starts to deviate from that predicted by the pure ^{56}Ni decay. This is a clear indication that the next most abundant radioactive isotope ^{57}Ni comes into play. Analysis shows that one finds a good agreement with $M(^{57}\text{Ni}) = 3.3 \times 10^{-3} M_{\odot}$ (Fig. 32).

Thanks to its very long decay time scale, 89 years, ^{44}Ti takes over as the dominant source of energy to the ejecta at ~ 1700 days. Although only

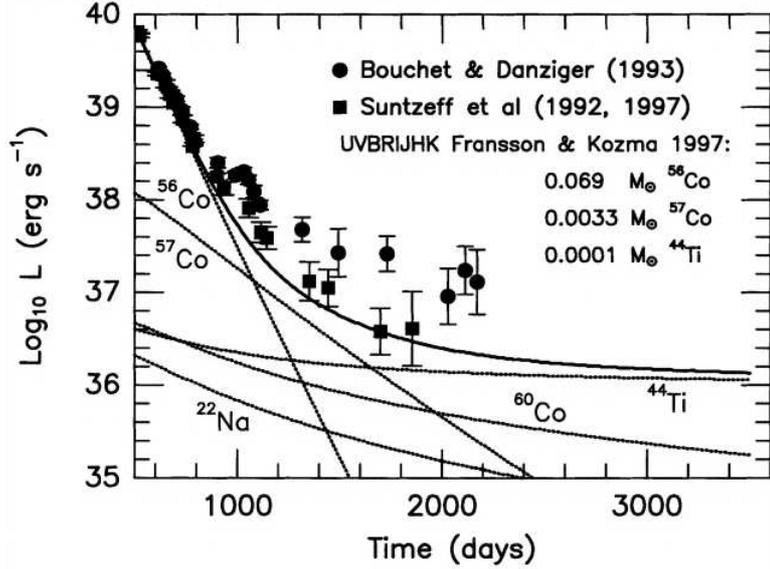


Figure 32: Bolometric light curve for SN 1987A compared to theoretical light curves from different isotopes. (Diehl & Timmes 1998).

$\sim 25\%$ of the energy in the decay is in the form of positrons, these dominate the energy input. The reason is that the positrons are most likely trapped by collisions and even a weak magnetic field, while most of the gamma-rays escape, since $\tau_\gamma \lesssim 0.1$. The trapping in combination with the long decay time scale means that the light curve will be essentially flat after this epoch. Careful modeling by Cecilia Kozma shows that a mass of $(1 - 2) \times 10^{-4} M_\odot$ of ^{44}Ti was formed in the explosion. As we discussed in §7.3, this provides us with a very useful diagnostic of the explosion conditions.

SN 1987A is not the only supernova for which the decays of ^{44}Ti has been observed. For Cas A (age ~ 330 years) COMPTEL on the Compton Gamma Ray Observatory detected the strongest gamma-ray line from the ^{44}Sc at 1.157 MeV (Fig. 33). The strength of the line corresponds to $M(^{44}\text{Ti}) \sim 1.7 \times 10^{-4} M_\odot$, close to that inferred for SN 1987A.

8 Type Ia supernovae

Observationally there are two types of SN explosions, Type I and Type II. The Type I SNe are mainly characterized by the complete lack of hydrogen in their spectra. As we have seen the Type Is can be divided into Type Ia

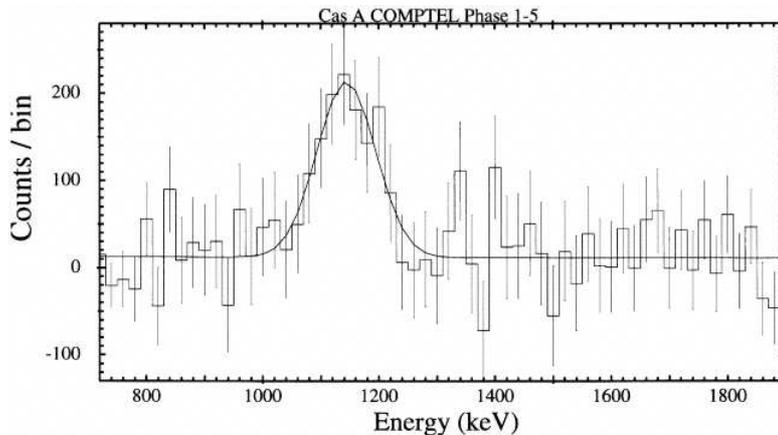


Figure 33: Gamma-ray spectrum of Cas A obtained with COMPTEL on the Compton Gamma Ray Observatory. (Iyudin et al. 1994).

and Type Ib/c. The latter are physically similar to the Type IIs, originating from the collapse of massive stars. The Type Ia SNe have, however, a very different origin, often occurring in elliptical galaxies with a very old stellar population. Observationally they are very similar to each other, both their light curves and spectra. Fig. 34 shows a compilation of three different spectra of Type Ia supernovae and it is clear that they show very similar spectral features.

In addition, the absolute luminosities are also similar within a few tenths of a magnitude. This is particularly true if these are corrected for an observed correlation of the absolute luminosity and decline rate after maximum. This is usually characterized by a quantity known as Δm_{15} , which is the decrease in magnitudes from maximum until 15 days after maximum. It is then found that a nearly linear relation between the absolute magnitude at the peak and Δm_{15} . This is usually known as the Philips relation. Using this the dispersion in absolute magnitude is only ~ 0.1 magnitudes. They have therefore become the most useful 'standard candle' for determining distances in cosmology.

The extreme uniformity of the Type Ia SNe show that they must come from very similar types of stars. These in addition have to be very old, of the order of billions of years, and can therefore not originate from massive stars, like the core collapse SNe. The standard picture is therefore the explosion of a white dwarf, with mass close to the Chandrasekhar mass. If this is in a close binary system mass transfer from the 'normal' star to the white dwarf

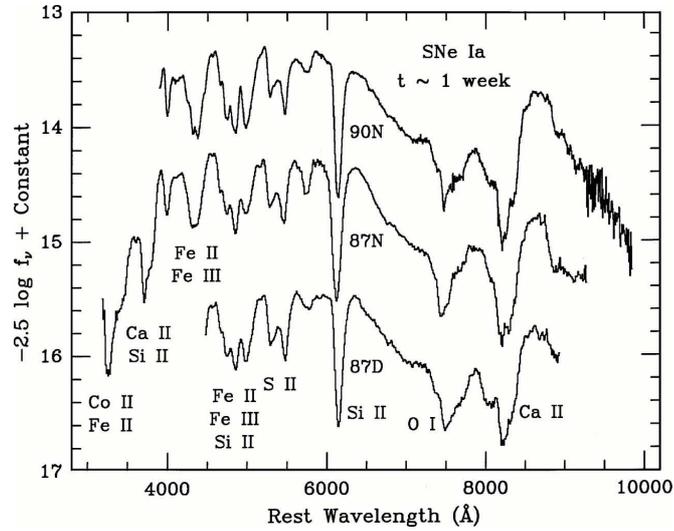


Figure 34: Compilation of three different Type Ia spectra at an age of about one week (Filippenko 1997).

may take place. Because of the angular momentum this will spiral in and form a disc around the white dwarf, and later be accreted onto the white dwarf. The accreted gas will then accumulate and normally will after some time explode in explosive nuclear hydrogen burning. This will give rise to an ordinary nova, occurring frequently every year in our Galaxy. In a fairly restrictive range of mass accretion rates, $10^{-7} - 10^{-6} M_{\odot} \text{ yr}^{-1}$, the mass of the white dwarf will, however, increase steadily. As the mass increases the radius of the WD will decrease, $R \propto M^{-3}$ (Sect. 4), and the density will therefore increase. This will heat up the degenerate core and when the density reaches $1.5 \times 10^9 \text{ g cm}^{-3}$ and the temperature $\sim 8 \times 10^8 \text{ K}$ the nuclear burning will become explosive.

The ignition of the explosion is still not well understood and may occur in two different modes. Either the nuclear burning will propagate from the ignition site subsonically, with velocity less than the sound velocity, or supersonically, with velocity larger than the sound velocity. In the former case, known as a deflagration, the WD will have time to expand as a result of the increased temperature and pressure, resulting in a decreasing density as the deflagration wave propagates to the surface of the star. (Fig. 36) The nuclear burning will at the center go all the way to nuclear statistical equilibrium, mainly $M^{56}\text{Ni}$. As the explosion proceeds and the WD expands the nuclear burning will occur at lower densities and only a partial burning will

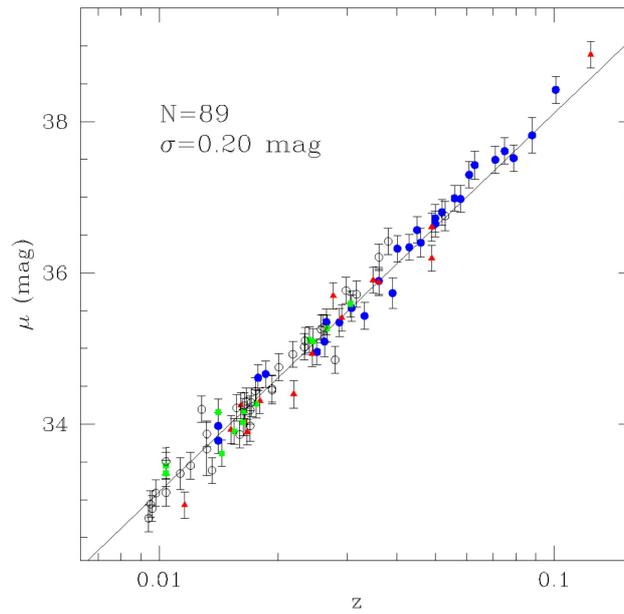


Figure 35: Hubble diagram based on Type Ia supernovae. The distance module μ is the difference in apparent and absolute magnitude, $\mu = m - M$, and z is the redshift. Note the small scatter illustrating the usefulness of the lightcurve as a distance measure. The dispersion is $\sigma = 0.2$ magn. (From Prieto et al 2006).

take place, leading to intermediate mass elements, like Si, S and Ca. In the outer parts of the exploding WD remains of the original carbon and oxygen may be present (Fig. 37).

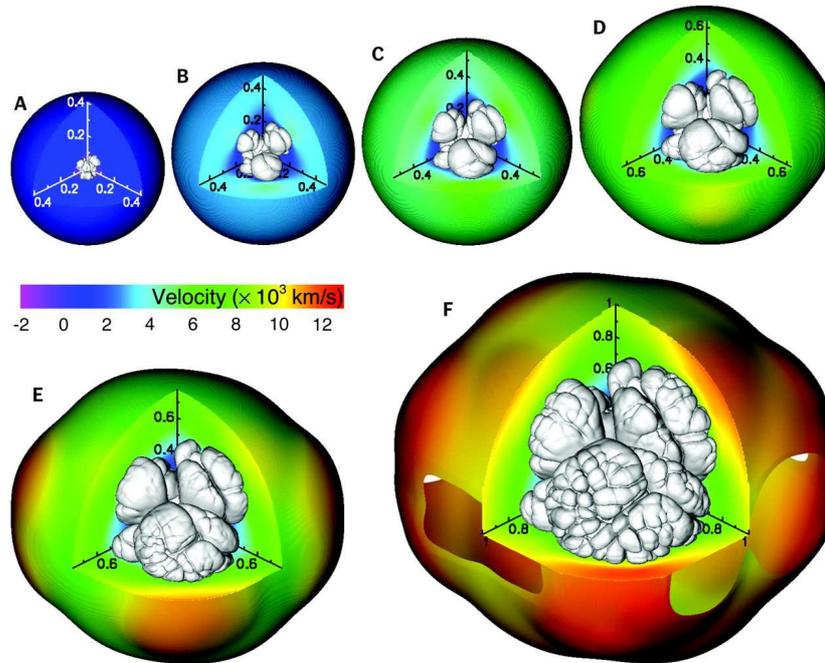


Figure 36: The development of the deflagration flame seen at (A) 1.26, (B) 1.49, (C) 1.57, (D) 1.65, (E) 1.76, and (F) 1.90 s after ignition. The color code shows the velocity. (Gamezo et al. 2003)

In the supersonic case the pressure of the star does not have time to adjust and the explosion takes place at the density of the original WD. This is known as a detonation. The result of this is that the whole WD, consisting originally of $\sim 50\%$ of carbon and $\sim 50\%$ of oxygen, will be burnt into nuclear statistical equilibrium. This will therefore result in a SN with only iron peak elements.

Observationally one finds that the spectra of Type Ia SNe show clear lines of both oxygen and intermediate mass elements at high velocities, close to the surface (see Fig. 37). This strongly argues against a pure detonation. Hydrodynamic simulations of pure deflagrations, however, show that in these a substantial amount of unburnt carbon and oxygen occurs also in the center of the supernova. This is in contradiction to observations of Type Ia SNe

at late epochs, when the central regions dominate the light. Therefore, a popular model is that of a 'delayed detonation', where the burning starts as a deflagration, expanding the WD. After a few seconds the deflagration will turn into a detonation, transforming the unburnt fuel in the center to higher mass elements. Because the expansion during the deflagration phase leads to a lower density the burning will, however, during the detonation now not go all the way to nuclear statistical equilibrium. A substantial mass of intermediate mass elements will therefore be found at especially high velocities, and even unburned material at the surface.

What causes the transition from deflagration to detonation is however, not understood, and much work remains to be done. Also other aspects of the Type Ia explosions are not well understood. In particular the nature of the progenitor systems are unknown. Only in one case, that of Tycho's SN 1572, has a binary companion probably been detected. The mass transfer process and the accretion is also ill-understood. Most important, the influence of different initial metallicities in the progenitor star, and its effect on especially the C/O ratio is also unknown. This will influence the total nuclear energy available, and therefore the total luminosity of the SN. This is especially serious for the use of Type Ia SNe as cosmological standard candles. The fact that they seem to work so well for this purpose is therefore somewhat surprising and needs to be better understood.

9 Pulsars

9.1 What are the pulsars?

Some basic facts: ~ 1000 pulsars detected. Of these ~ 30 millisecond pulsars, most of these in binary systems. The periods are in the range 0.00156 - 8.5 secs. The pulse periods are stable to the same level as atomic clocks, approximately one part in 10^{12} or better. A fact that is important for the formation process of the pulsars is that they in general have very high space velocities, $100 - 1000 \text{ km s}^{-1}$. It is likely that this was either caused by an asymmetric explosion in connection to the supernova collapse and explosion, or because the supernova occurred in a binary system, which was disrupted by the explosion. Currently, the former explanation is the most popular.

Today we know that pulsars are rapidly rotating neutron stars. It is, however, instructive to review some of the arguments in favor of this. Let us first calculate the maximum velocity before break-up by requiring that the centrifugal force should be less than the gravitational force on the surface,

$$\frac{V^2}{r} < \frac{GM}{R^2} \quad (9.1)$$

But $P = 2\pi R/V$, and $M \approx 4\pi R^3 \rho/3$, so

$$\rho > \frac{5.6 \times 10^8}{P^2} \text{ g cm}^{-3}. \quad (9.2)$$

The maximum density of a white dwarf is $\sim 10^{10} \text{ g cm}^{-3}$. The fact that pulsars with as short periods as a few milli-seconds have been observed clearly rules out white dwarfs. Instead, these periods indicate densities close to nuclear, $\gtrsim 10^{14} \text{ g cm}^{-3}$. This is what is expected for a neutron star. Also pulsations of a white dwarf is excluded, based on both the short period and the stability of the pulses.

9.2 Pulsar slowing down (Longair p. 100-104)

Far from the pulsar the magnetic field can be approximated by a dipole field with magnetic moment \mathbf{p}_m . If this dipole field has an angle α with the rotation axis (defined to be the z-axis), the magnetic moment will vary as

$$\mathbf{p} = p_0(\cos \alpha \mathbf{e}_z + \sin \alpha \cos \Omega t \mathbf{e}_x + \sin \alpha \sin \Omega t \mathbf{e}_y) \quad (9.3)$$

The total energy loss is

$$\frac{dE}{dt} = -\frac{2\ddot{\mathbf{p}}^2}{3c^3} \quad (9.4)$$

which is the same as the energy loss from an electric dipole (see e.g., Jackson or Landau-Lifshitz Eq. 71.5). Using Eq. (9.3) for $\ddot{\mathbf{p}}$ we get

$$\frac{dE}{dt} = -\frac{2p_0^2\Omega^4 \sin^2 \alpha}{3c^3} \quad (9.5)$$

Note that there is no energy loss if the dipole is aligned to the rotational axis.

The energy loss by radiation is taken from the loss of rotational, kinetic energy, given by $E_k = I\Omega^2/2$, where I is the moment of inertia. For a uniform sphere this is given by

$$I = \frac{2}{5}MR^2 \quad (9.6)$$

Therefore, setting $dE_k/dt = dE/dt$ we get

$$\frac{d}{dt} \frac{I\Omega^2}{2} = I\Omega \frac{d\Omega}{dt} = -\frac{2p_0^2 \sin^2 \alpha}{3c^3} \Omega^4 \quad (9.7)$$

The dipole field is given by

$$\mathbf{B} = \frac{p_0}{r^3}(2 \cos \theta \mathbf{e}_r + \sin \theta \mathbf{e}_\theta) \quad (9.8)$$

(see e.g., Jackson 1962, Eq. 5.41). Here θ is the polar angle, now from the *magnetic* axis, and r the radial direction. For a current I in a circle with radius a we have $p_0 = \pi a^2 I/c$.

Let us now assume that we can estimate the mass and radius of the neutron star, as $M \approx 1 M_\odot$ and $R \approx 12$ km. We can then, if we can measure Ω and $d\Omega/dt$, from Eq. (9.7) estimate p_0 . Because $B \approx p_0/R^3$ at the surface, we can then estimate the magnetic field at the surface,

$$B \approx \left(\frac{3c^3 I P \dot{P}}{8\pi^2 R_s^6} \right)^{1/2} \quad (9.9)$$

or using $I \approx 1.6 \times 10^{45}$ g cm²

$$B \approx 4 \times 10^{19} (P \dot{P})^{1/2} \text{ G} \quad (9.10)$$

As an example we take the Crab where $P = 0.033$ s (i.e., $\Omega = 190$ Hz) and $\dot{P} = 4.2 \times 10^{-13}$ s/s. With the above numbers we get $I = 1.6 \times 10^{45}$ g cm². The energy loss rate is therefore $dE/dt = I\Omega\dot{\Omega} = 7 \times 10^{38}$ erg s⁻¹. Therefore, $p_0 \sin \alpha \approx 5 \times 10^{38}$ and $B \approx 3 \times 10^{12}$ G, if we set $\sin \alpha \sim 1$.

There are from this two interesting observations. The first is, of course, the extremely high magnetic field at the surface. The second is the fact that the amount of energy lost due to the spin-down is very close to the total amount of energy emitted in the Crab nebula. This is dominated by synchrotron emission in the X-rays, and this shows that the energy needed from this is likely to originate from the pulsar. Fig. 39 shows the central region of the Crab with the pulsar. The blue is X-ray emission from synchrotron emitting electrons, which clearly shows a flow of relativistic particles from the neutron star, as well as a torus like structure. Variations in both in the optical and X-ray structure is seen on time scales of days to months.

9.3 Electrodynamics of the pulsar

The structure of a rotating, strongly magnetized star is an interesting exercise in electrodynamics, and more important, gives some crucial insight in the peculiar conditions around these objects. This is the basis of a now famous analysis by Goldreich & Julian from 1969.

Consider a rotating perfectly conducting sphere with a magnetic field \mathbf{B} . In the interior Ohms law gives

$$\mathbf{J} = \sigma \mathbf{E}' \quad (9.11)$$

where \mathbf{E}' is the electric field in the rotating frame of the conductor. The relation between the rest frame field and that of the conductor is $\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}$. If the conductivity is perfect (infinite) then we must have $\mathbf{E} + \mathbf{v} \times \mathbf{B}/c = 0$, or

$$\mathbf{E} = -\mathbf{v} \times \mathbf{B}/c = -\Omega \mathbf{r} \times \mathbf{B}/c \quad (9.12)$$

With the exterior B-field given by Eq (9.8) the field just inside the surface is

$$\mathbf{B} = \frac{p_0}{R^3} (2 \cos \theta \mathbf{e}_r + \sin \theta \mathbf{e}_\theta) = B_s (\cos \theta \mathbf{e}_r + \frac{1}{2} \sin \theta \mathbf{e}_\theta) \quad (9.13)$$

and

$$\mathbf{E} = \frac{R\Omega B_s \sin \theta}{2} (\frac{1}{2} \sin \theta \mathbf{e}_r - \cos \theta \mathbf{e}_\theta) \quad (9.14)$$

At the surface the θ component of the electric field is continuous, and the outside (vacuum) θ component of this field at the surface is therefore equal

to

$$\mathbf{E}_\theta = -\frac{R\Omega B_s \sin\theta \cos\theta}{2} \mathbf{e}_\theta \quad (9.15)$$

In the vacuum outside the star one can define a potential $\mathbf{E} = \nabla\phi$, with ϕ given by

$$\nabla^2\phi = 0 \quad (9.16)$$

With the boundary condition Eq. (9.15) one finds that

$$\phi = -\frac{B_s\Omega R^5}{6cr^3}(3\cos^2\theta - 1) \quad (9.17)$$

The potential is therefore that of a quadrupole.

From the equations above we finally find that

$$\mathbf{E} \cdot \mathbf{B} = -\frac{B_s^2\Omega R^8}{cr^7} \cos^3\theta \quad (9.18)$$

A direct conclusion of this is that the magnitude of the surface electric field along the magnetic field lines therefore is

$$E \approx -\frac{B_s\Omega R}{c} \approx 2 \times 10^8 \left(\frac{B}{10^{12}\text{G}} \right) P^{-1} \text{ V cm}^{-1} \quad (9.19)$$

The ratio between the electrical force and the gravitational therefore is

$$\frac{F_E}{F_G} = -\frac{eB_s\Omega R/c}{GMm/R^2} \approx 10^9 \quad (9.20)$$

Therefore, even if there was a vacuum outside of the pulsar, the electric force would be so strong that electrons and protons would most likely escape from the surface and flow into the surrounding magnetosphere. The region outside the neutron star will therefore become filled with charged particles which will neutralize the electric field here.

Exercise: Show and motivate the different steps leading up to Eq. (9.20).

9.4 The magnetosphere. Longair p. 107-108

Because the magnetic field is rotating with the neutron star, there will be a radius for which the corotation velocity is equal to the speed of light. This is given by $R_{l.c.} = c/\Omega = 5 \times 10^4 P \text{ km}$. For radii larger than this the field lines would rotate faster than the speed of light. Field lines outside this radius must therefore lag behind rigid rotation, and get a toroidal component, even if this was absent at the surface.

For the case of a dipole field this is given by Eq. (9.8). It is not difficult to show that the magnetic field lines in this case are given by

$$r = K \sin^2 \theta \tag{9.21}$$

where K is the parameter which defines the line. Of special interest is the last field line which is closed within the light cylinder. At the surface of the neutron star the azimuthal angle of this line is given by

$$\sin \theta_{l.c.} = (R\Omega/c)^{1/2}. \tag{9.22}$$

For the parameters of the Crab pulsar $\theta_{l.c.} \sim 5$ deg. For $\theta > \theta_{l.c.}$ the field lines are closed, while for $\theta < \theta_{l.c.}$ they are open.

The electric force will mainly be parallel to the magnetic field lines. Particles escaping on field lines which close will mainly accumulate within this magnetosphere, while those coming from the poles on open field lines will escape to infinity, and therefore in this way transport energy to the surrounding medium.

The emission process of the pulsar is still not well understood. An important ingredient in most models is the enormous potential drop at the surface. This can lead to pair production of electrons and positrons in the vacuum. These will then be accelerated by the electric field along the magnetic field lines and give rise to radiation similar to synchrotron radiation, but with the bending of the field lines as the radius of curvature. This is therefore known as curvature radiation. The gamma-rays produced by this radiation, can then give rise to new pairs in the strong field and a cascade may result.

10 Neutron stars. Longair 15.3.3

The Chandrasekhar mass does not involve the mass of the particle responsible for the degenerate pressure. The maximum mass of a star made of neutrons should therefore be similar. The structure of the star should, however, be much more compact, as can be seen from the mass – radius relation, Eq. (4.8), which shows that the radius should be a factor $m_e/m_n \approx 1800$ smaller. Except for this we do, however, expect a similar relation between mass and radius.

At the next level of approximations there are, however, two effects which change these conclusions. The first is that the equation of state is not that of a pure non-interacting neutron gas. Instead nuclear interacting and other effects become important as the density is close or above nuclear.

The second effect comes from the fact that the radius of a neutron star, ~ 10 km, is of the same order as the Schwarzschild radius, ~ 4 km. General relativistic effects are therefore important for the structure of neutron stars. In this case the hydrostatic equation, Eq. (2.3), is replaced by the GR analogue which is known as the Oppenheimer-Volkoff equation

$$\frac{dP}{dr} = -\frac{G[m(r) + 4\pi r^3 p/c^2](\rho + p/c^2)}{r[r - 2Gm(r)/c^2]} \quad (10.1)$$

The mass conservation equation is the same as before. Compared to the Newtonian case the pressure here adds to the mass as a source to the gravitational force. In addition, the curvature of space changes the r^2 term into $r(r - 2Gm(r)/c^2)$. All these corrections tend to increase the effect of gravity and therefore leads to a smaller mass compared to the Newtonian case. The neutron star analogue to the Chandrasekhar mass for a pure gas of neutrons, the Oppenheimer-Volkoff mass, is therefore $0.71 M_\odot$ corresponding to a radius 9.14 km. In contrast to the white dwarf case the pure neutron star EOS is, however, not a very good approximation and more realistic cases result in considerably higher masses, as we will see below. The uncertainties in the structure and the EOS at especially densities larger than nuclear are unfortunately large. We will now discuss the structure in more detail, starting from the surface.

10.1 Crust

Close to the surface there is an 'atmosphere' where the conditions very fast go from non-degenerate to that of a white dwarf EOS (equation of state). The thickness of this region is only ~ 1 cm.

The crust extends to a density of $\sim 10^{14}$ g cm $^{-3}$, and the thickness of the crust is ~ 1 km, although this may vary by a factor of about two between different models.

In the interior the density quickly increases, and with that the Fermi energy, $E_F = (p_f^2 - m_e^2 c^2)^{1/2}$ where $p_F = (3h^3 n_e / 8\pi)^{1/3}$ (Eq. (3.4)). For densities $\gtrsim 10^6$ g cm $^{-3}$ this energy may become large enough to overcome the mass difference between the proton and neutron, converting protons in the nuclei to neutrons, making the matter increasingly neutron rich, referred to as neutronisation. The matter is therefore changing from ordinary ^{56}Fe to extremely neutron rich nuclei with $A \sim 200$ and $Z/A \sim 0.1$. At this point the neutron starts to drip out of the nuclei, forming a neutron gas together with the degenerate electrons and nuclei. This is usually referred to as the neutron drip and occurs at $\sim 4 \times 10^{11}$ g cm $^{-3}$. Most of the neutrons are, however, contained within the nuclei until a density of $\sim 10^{14}$ g cm $^{-3}$. The nuclei are located in a lattice whose structure changes from that of a individual nuclei to distinctly non-spherical (from 'meatballs, to spaghetti, to lasagna, to o Swiss cheese'). At $\sim 10^{14}$ g cm $^{-3}$ the nuclei begin to break up and form a gas of free neutrons together with degenerate electrons. This occurs at about half nuclear density, or $\sim 10^{14}$ g cm $^{-3}$.

Because pairing of free neutrons give a lower energy state the neutrons will probably form a superfluid.

10.2 Core

Nuclear density corresponds to $\rho_n \sim 2.8 \times 10^{14}$ g cm $^{-3}$.

The core region contains $\sim 99\%$ of the mass of the NS. The density in the core is rather constant, varying by only a factor of ~ 2 . The outer core with density $\lesssim 2\rho_n$ is reasonably well understood, and consists of neutrons with a few percent of protons and electrons, and possibly muons. All are degenerate. Neutrons and protons form a superfluid.

The extent of the inner core depends sensitive on the EOS and the mass of the star. Superfluidity of protons also means that the protons are superconducting. At very high densities, well above ρ_n , the pairing may, however, be suppressed.

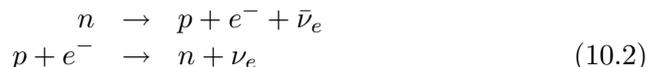
The composition of the inner core is to a large extent open. There are several possibilities discussed. The most conservative is that it has the same composition as the outer core, i.e., neutrons with a few percent of protons, electrons and muons. An often discussed possibility is that a Bose-Einstein condensate of pions may form. A variation of this is a condensate of K-mesons (kaons). Finally, at high enough density the nucleons may merge

into a soup of quarks, up, down and strange, and gluons.

The main effect of these different possibilities for the EOS is that the EOS will be more or less stiff. The creation of pion and kaon condensates, as well as quarks, will in general make the EOS soft. A soft EOS has in turn the consequence that the matter can be compressed more, making the radius of the NS smaller. A measurement of the radius would therefore make it possible to test these different possibilities.

The radius can be determined from e.g. the thermal emission, if such a component can be observed (see below). A related effect is that the cooling of the neutron star is affected by the composition, and therefore the EOS in the inner core.

Shortly after the collapse and bounce the temperature in the core is $\sim 5 \times 10^{11}$ K at 15 s, decreasing to $\sim 5 \times 10^9$ K at 50 s. Most of the cooling then occurs by emission of neutrinos, which in turn depends on the composition. The most effective is the so called URCA process, named after a casino in Rio de Janeiro by Gamow,

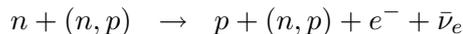


These two process gives no net change of the number of neutrons or protons, but produces two neutrinos, which carry away energy. The fact that this process always results in a loss of energy is the reason for its name. For this process to occur the proton concentration must be sufficiently high, $\gtrsim 1/9$. To see why this is the case we consider the case when the neutrino carries away a negligible amount of momentum. Momentum conservation in the last reaction is then $p_e + p_p = p_n$. But for completely degenerate gases $p \propto n^{1/3}$ for all fermions. Therefore, momentum conservation becomes $p_e + p_p = 2p_p = p_n$. The ratio between protons to total number of nucleons can then be written as

$$\frac{n_p}{n_p + n_n} = \frac{p_p^3}{p_p^3 + p_n^3} = \frac{1}{9} \tag{10.3}$$

The cooling time can be estimated to $\sim 20T_9^{-4}$ s. The high proton ratio requires a high density, 2-3 times that of nuclear density or $\sim 2.7 \times 10^{14}$ g cm $^{-3}$. This in turn requires a soft EOS, and the cooling time is therefore a diagnostic of the EOS.

For lower densities the modified URCA process may still occur



$$p + (n, p) \rightarrow n + (n, p) + e^+ + \nu_e \quad (10.4)$$

where (n, p) is an additional nucleon, neutron or proton, which ensures momentum conservation. This process can occur under more general conditions, but is much slower than the direct URCA process. The cooling therefore takes longer time and the interior temperature will be higher. After ~ 100 years the temperature will be $\sim 3 \times 10^8$ K in the former case, and $\sim 1.2 \times 10^9$ K in the latter. The interior is now isothermal. From 100 years to $\sim 3 \times 10^5$ years the temperature is $\sim 2 \times 10^8$ K in the URCA case and $\sim 6 \times 10^8$ K in the modified URCA case.

The connection between the interior temperature and the surface temperature depends on the heat conduction at the surface. This is in turn dependent on magnetic fields and composition. The typical surface temperatures are in the range 3×10^5 to 10^6 K. Figure 42 shows the expected temperatures for different assumptions about the cooling, together with observations of several NSs. Note the rapid cooling for the direct URCA case at ~ 15 years, while the surface temperature in modified URCA case remains $\gtrsim 3 \times 10^5$ for $10^5 - 10^6$ years.

A determination of the surface temperature can therefore be a test on both which of the URCA processes is operating and indirectly the EOS. It should, however, be pointed out that the determination of the surface temperature is difficult because for most pulsars the thermal emission from the surface is dominated by the non-thermal radiation connected with the pulsar mechanism in the magnetosphere. Further, most of the emission is in the UV and in soft X-rays, which are strongly absorbed by interstellar gas. The observations, however, indicate that cooling occurs mainly by the modified URCA process, and that the direct is suppressed.

Another effect of the different EOS in the core is the maximum mass of the NS. A soft EOS gives in general a lower maximum NS mass, while conversely a stiff EOS results in a high. Fig. 43 shows the radius versus mass for different EOSs. The maximum mass for the different EOSs ranges from $\sim 1.4 M_\odot$ to $\sim 2.7 M_\odot$. A very firm upper bound comes from the requirement that the EOS should be such that the speed of sound is less than the velocity of light. This mass is $\sim 3.0 M_\odot$. A determination of a larger mass for a compact object is therefore extremely strong evidence for the presence of a black hole (see below).

The masses of a large number of NSs in binary systems have been determined with different degree of accuracy. The most accurate are those from NSs in neutron star – neutron star systems, including the famous PSR 1913+16 (the Hulse and Taylor binary pulsar), as well as PSR 1534+12

and PSR J0737-3039. In these cases the masses are $1.4408 \pm 0.0003 M_{\odot}$ and $1.3873 \pm 0.0003 M_{\odot}$ (PSR 1913+16), $1.3332 \pm 0.0010 M_{\odot}$ and $1.3452 \pm 0.0010 M_{\odot}$ (PSR 1534+12), and $1.337 M_{\odot}$ and $1.250 M_{\odot}$ (PSR J0737-3039). In Fig. 44 we show a compilation of masses for most of the observed neutron star binaries. These are determined either from timing of the pulsar frequency in different binaries containing at least one pulsar, or from the motion of the companion star in the case of X-ray binaries. The most accurate of these methods is obviously those systems containing two neutron stars, which is a very clean case of two point masses. As we see, the maximum mass of the most accurately determined are $\sim 1.44 M_{\odot}$. There are, however, several systems which are compatible with considerably higher masses and neither the stiff EOSs (with large masses) or the soft (with low masses) can unfortunately be ruled out yet.

It can be noted that there is a tendency for somewhat higher masses for neutron stars with white dwarf companions.

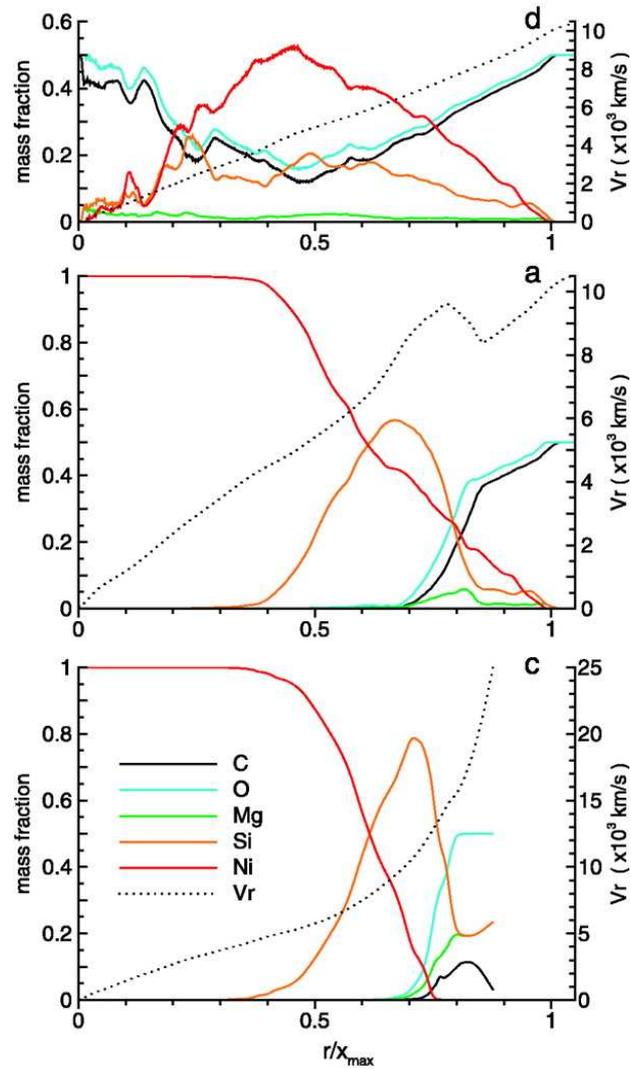


Figure 37: Abundances in one pure deflagration model (top) and two delayed detonation models started at 1.62 s (b) and at 1.51 s (c). Note the absence of oxygen and carbon in the delayed detonation models (Gamezo et al. 2005)

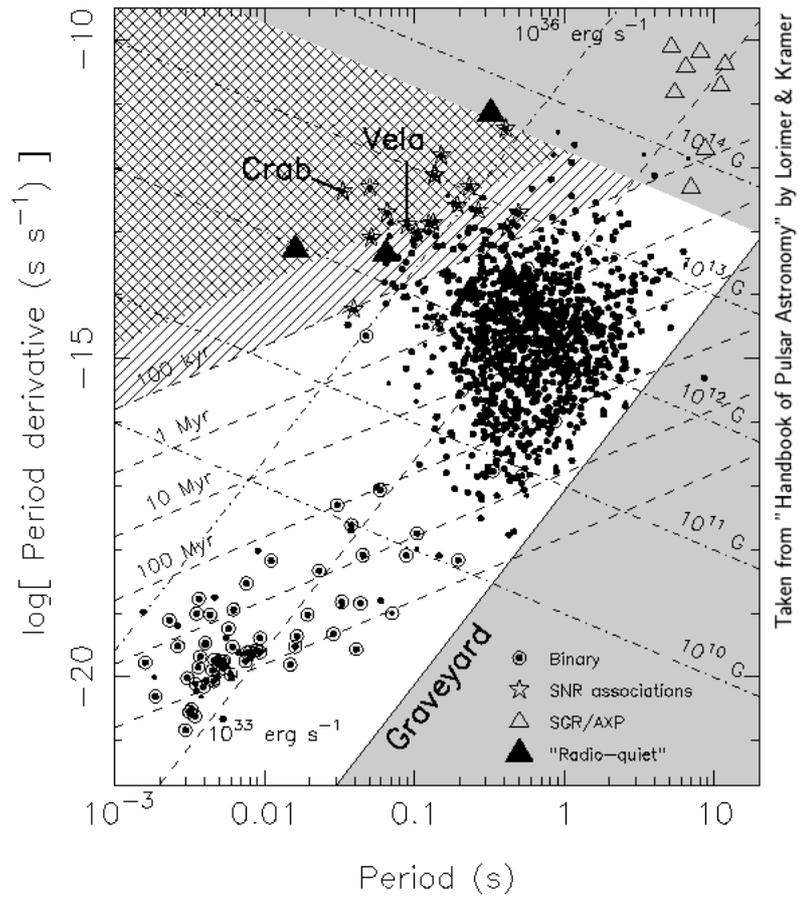


Figure 38: An updated version of the \dot{P} versus P plot. (Lorimer & Kramer)

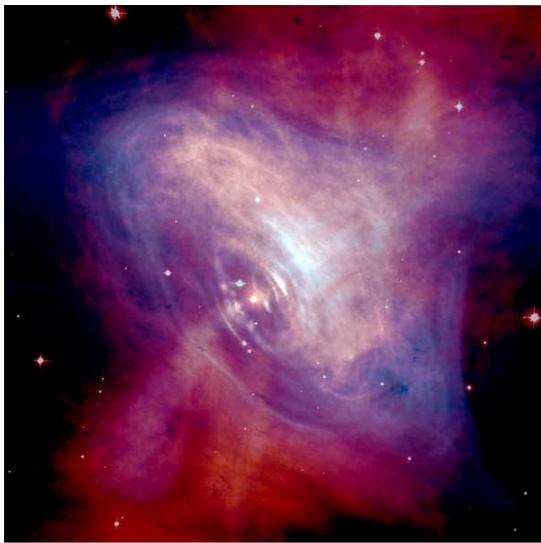


Figure 39: The central region of the Crab nebula as seen by Chandra in X-rays (blue) and HST in the optical (red).

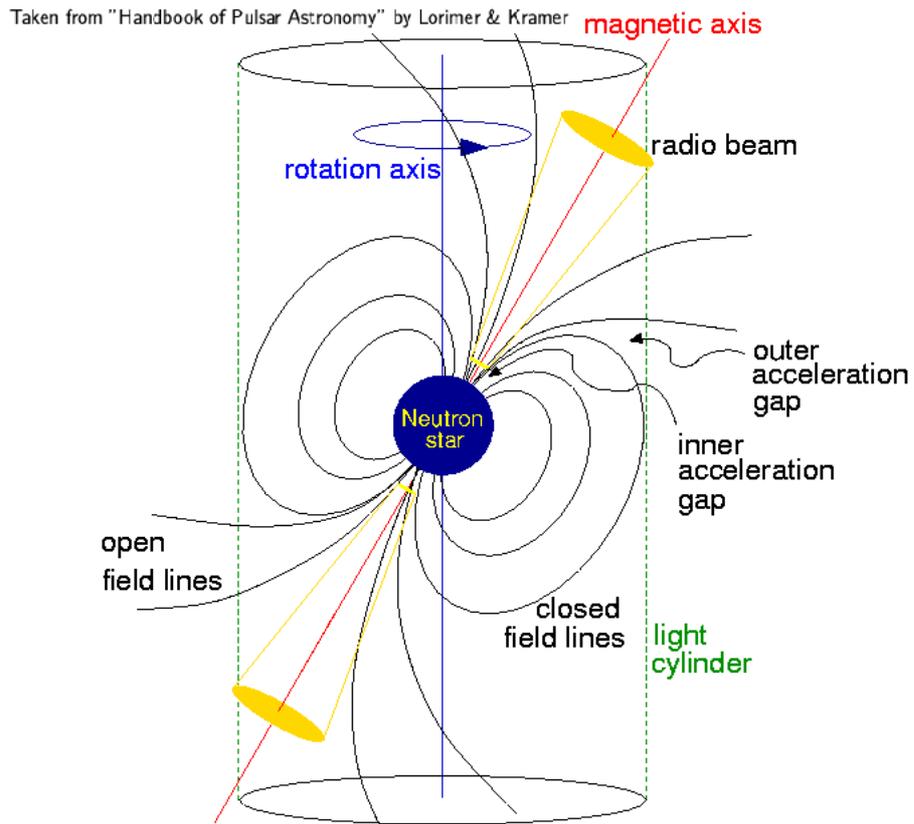


Figure 40: Magnetic field structure around the neutron star. (Lorimer & Kramer)

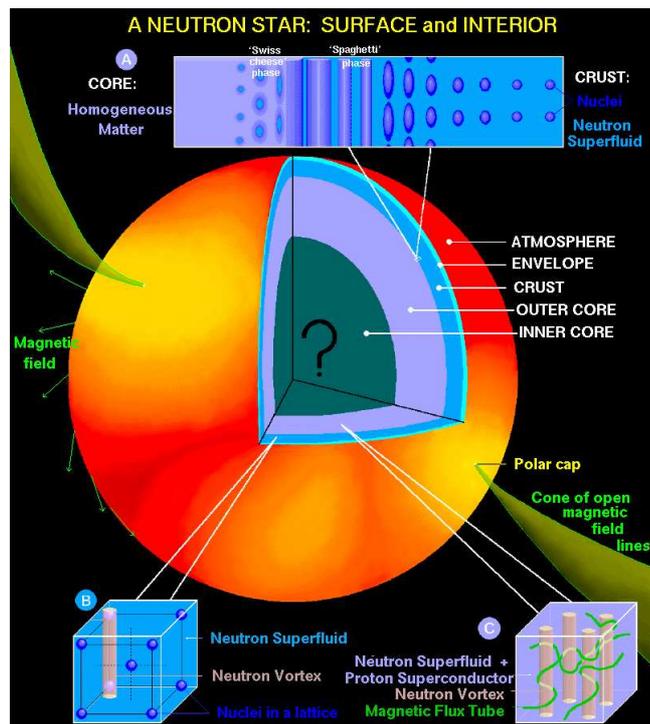


Figure 41: Structure of a neutron star with different regions discussed in the text. (from Lattimer and Prakash)

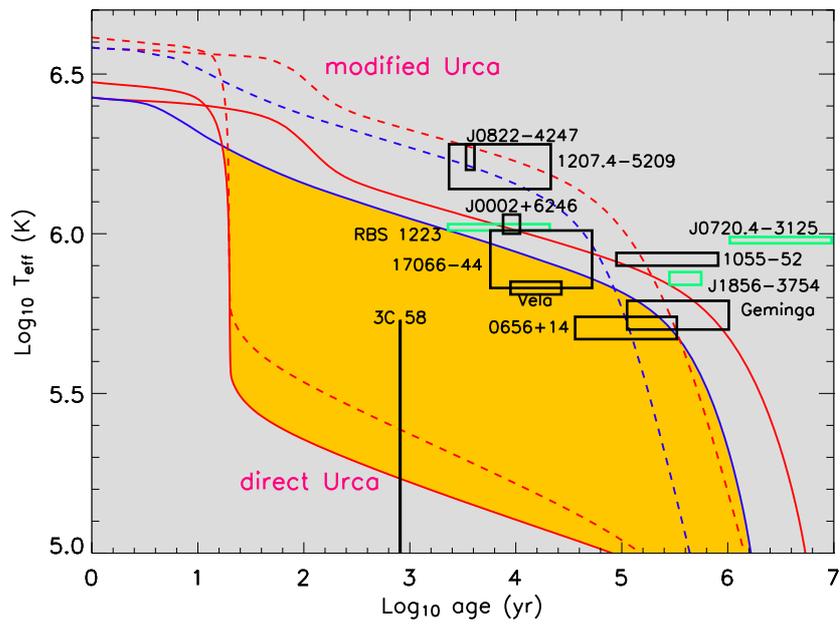


Figure 42: Surface temperature as function of age for different cooling processes. The boxes show observed temperatures and ages (with errors) (from Lattimer and Prakash).

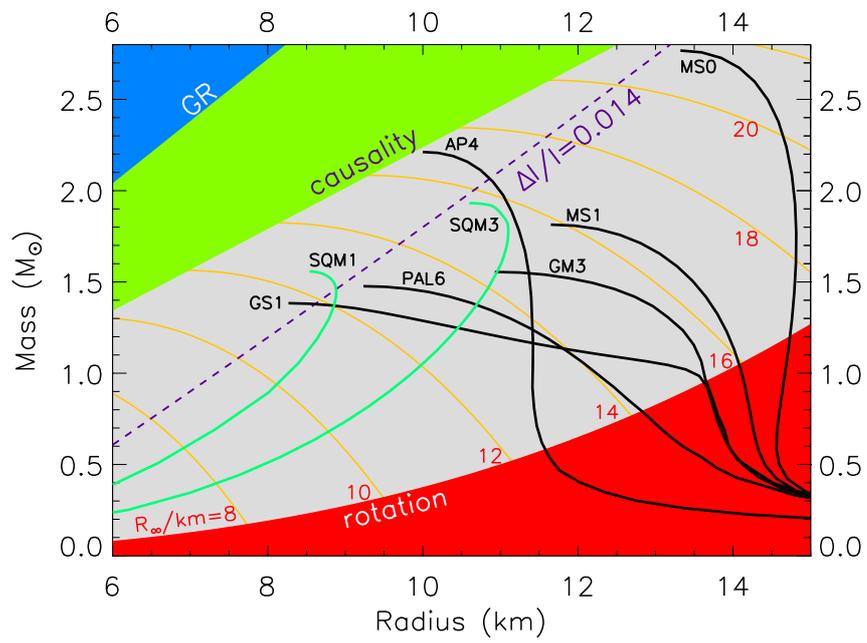


Figure 43: Neutron star mass as function of central density for different EOSs. (from Lattimer and Prakash)

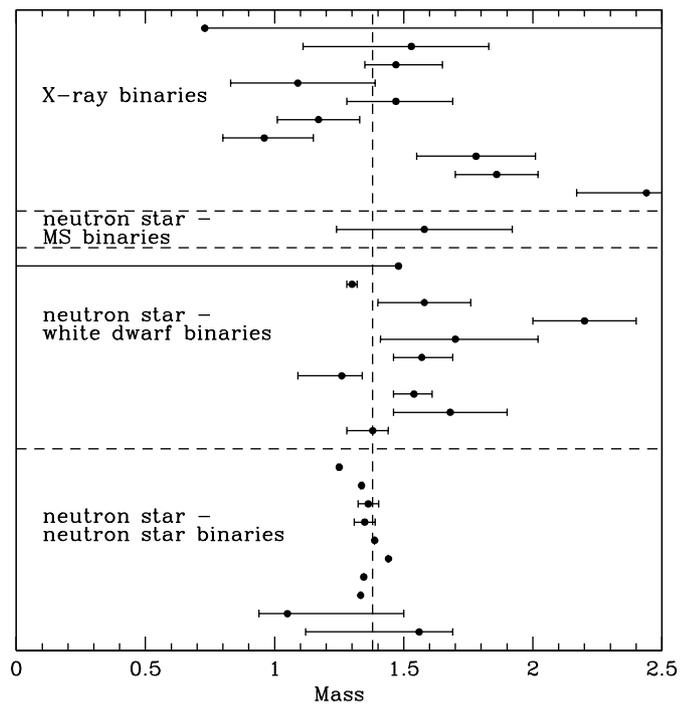


Figure 44: Compilation of neutron star masses determined from different types of neutron star binaries. The vertical dashed line shows a rough average mass of $1.38 M_{\odot}$ (data from Lattimer and Prakash)

10.3 Binary pulsars

The first discovered binary pulsar was PSR 1913+16, discovered by Hulse and Taylor. The system contains two neutron stars, one of which is observed as a pulsar. The pulsar period of this component is 22.7 millisecond and the orbital period 7.75 hours. From the short pulsing and orbital periods it was immediately clear that this system could be used as a test of different effects of General Relativity, which ultimately resulted in the Nobel prize to Hulse and Taylor in 1993. Furthermore, the eccentricity of the orbit is very high, $e = 0.617$, which is important for both the advance of the periastron and the decay by gravitational radiation.

Usually the relativistic effects are described by the so called Post-Newtonian parameters, which are given by

$$\dot{\omega} = 3T_{\odot}^{2/3} \left(\frac{P_b}{2\pi} \right)^{-5/3} \frac{1}{1-e^2} (M_p + M_c)^{2/3}, \quad (10.5)$$

$$\gamma = T_{\odot}^{2/3} \left(\frac{P_b}{2\pi} \right)^{1/3} e \frac{M_c(M_p + 2M_c)}{(M_p + M_c)^{4/3}}, \quad (10.6)$$

$$\dot{P}_b = -\frac{192\pi}{5} T_{\odot}^{5/3} \left(\frac{P_b}{2\pi} \right)^{-5/3} \frac{(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4)}{(1-e^2)^{7/2}} \frac{M_p M_c}{(M_p + M_c)^{1/3}} \quad (10.7)$$

$$r = T_{\odot} M_c, \quad (10.8)$$

$$s = T_{\odot}^{-1/3} \left(\frac{P_b}{2\pi} \right)^{-2/3} x \frac{(M_p + M_c)^{2/3}}{M_c}, \quad (10.9)$$

where P_b is the period and e the eccentricity of the binary orbit. The masses M_p and M_c of pulsar and companion, respectively, are expressed in solar masses (M_{\odot}). The constant $T_{\odot} = GM_{\odot}/c^3 = 4.925490947\mu\text{s}$. The first PN parameter, $\dot{\omega}$, describes the relativistic advance of periastron. According to Eq. (10.5) it gives a direct measurement of the total mass of the system, $(M_p + M_c)$. The parameter γ denotes the amplitude of delays in arrival times caused by the varying effects of the gravitational redshift and time dilation (second order Doppler) as the pulsar moves in its elliptical orbit at varying distances from the companion and with varying speeds. The decay of the orbit due to gravitational wave damping is expressed by the change in orbital period, \dot{P}_b . The other two parameters, r and s , are related to the Shapiro delay caused by the gravitational field of the companion. These parameters are only measurable, depending on timing precision, if the orbit is seen nearly edge-on.

For PSR 1913+16 the gravitational decay parameter is $\dot{P}_b = -2.4184 \times 10^{-12}$ s/s (Fig. 45). Based on the gravitational decay the system should

merge in $\sim 3 \times 10^8$ years. The periastron advance is given by $\dot{\omega} = 4.22$ degrees/yr. The most recent determination of the individual pulsar masses are $1.4414 \pm 0.0002 M_{\odot}$ and $1.3867 \pm 0.0002 M_{\odot}$. Besides PSR 1913+16, the binary PSR B1534+14 has provided a similar test of relativity.

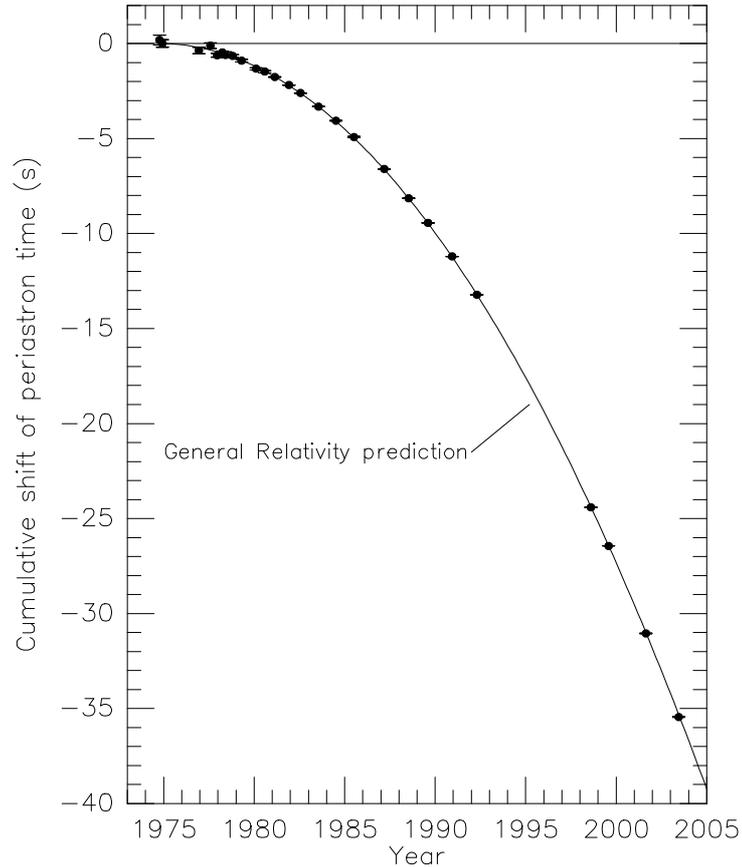


Figure 45: The gravitational radiation orbital decay parameter for PSR 1913+16 together with that predicted from gravitational radiation loss.

The most interesting of the binary pulsars discovered after PSR 1913+16 is the double-pulsar PSR J0737-3039. This is unique in the sense that two pulsars are seen in this system, which gives additional constraints on the parameters of the system. The individual periods are 22.7 ms and 2.773 s. Another unique feature is that it is seen nearly edge-on with $\sin i = 0.9995$, i.e. $i = 87$ degrees. This makes it possible not only to study the orbital parameters, but also to probe the pulsar magnetospheres. The orbital period is 2.4 hours. The periastron advance $\dot{\Omega} = 16.90$ degrees, which is four times

larger than for PSR 1913+16. The edge-on nature of the system has also provided a new observational constraint from the ratio of the masses of the system, $R = M_1/M_2$. The system will be coalescing within 85 million years, much shorter than for PSR 1913+16. This in connection to the low luminosity has increased the frequency of merging neutron stars by an order of magnitude, which is important both for the detection of gravitational waves and for merging neutron stars as a source of the short gamma-ray bursts. In Fig. 46 we show the solution for the two masses as function of the different post-Newtonian parameter, including R . From it is found that the neutron star masses are $1.3381 \pm 0.0007 M_\odot$ and $1.2489 \pm 0.0007 M_\odot$. The ratio of the predicted and observed values of the Shapiro time delay parameter is 0.9987 ± 0.00050 , confirming this effect of the General Theory of Relativity.

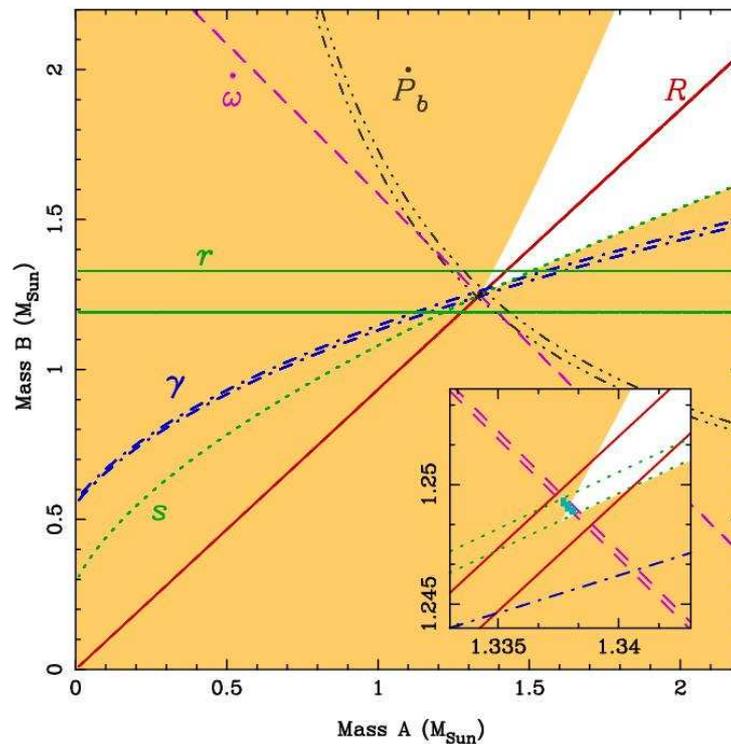


Figure 46: Orbital parameters in the M_1, M_2 plane for PSR J0737-3039. (Kramer et al. 2006)

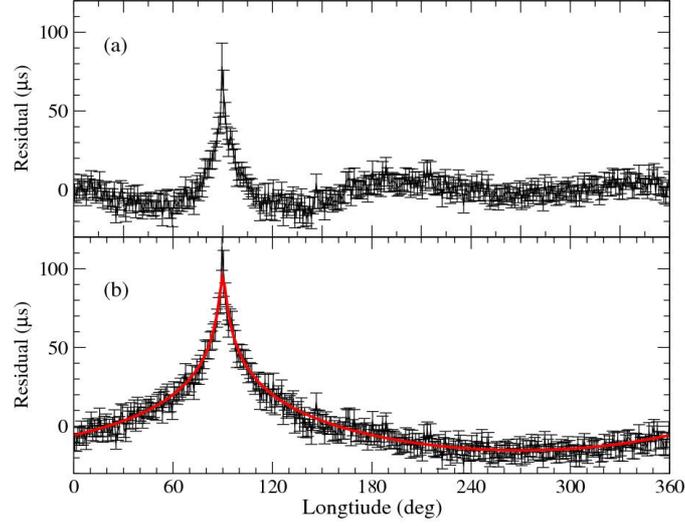


Figure 47: The Shapiro time delay as function of the phase for PSR J0737-3039. The solid lines shows that predicted from General Relativity. (Kramer et al. 2006)

10.4 The binary mass function. (Longair p. 113)

Let a_1 and a_2 be the distance from the CM, i.e.

$$a = a_1 + a_2 \quad (10.10)$$

and

$$M_1 a_1 - M_2 a_2 = 0 \quad (10.11)$$

or

$$a = \frac{(M_1 + M_2)}{M_2} a_1 \quad (10.12)$$

We observe the projected orbital velocity v_1 , given by

$$v_1 = \frac{2\pi}{P} = a_1 \sin i \quad (10.13)$$

where i is the inclination. Keplers law says that

$$\frac{G(M_1 + M_2)}{a^3} = \left(\frac{2\pi}{P}\right)^2 \quad (10.14)$$

Therefore

$$f(M_1, M_2, i) = \frac{(M_2 \sin i)^3}{(M_1 + M_2)} = \frac{Pv_1^3}{2\pi G} \quad (10.15)$$

We can observe the RHS of this equation. As shown by this equation, this, however, only gives the mass function. In addition we need to know the inclination and the mass ratio.

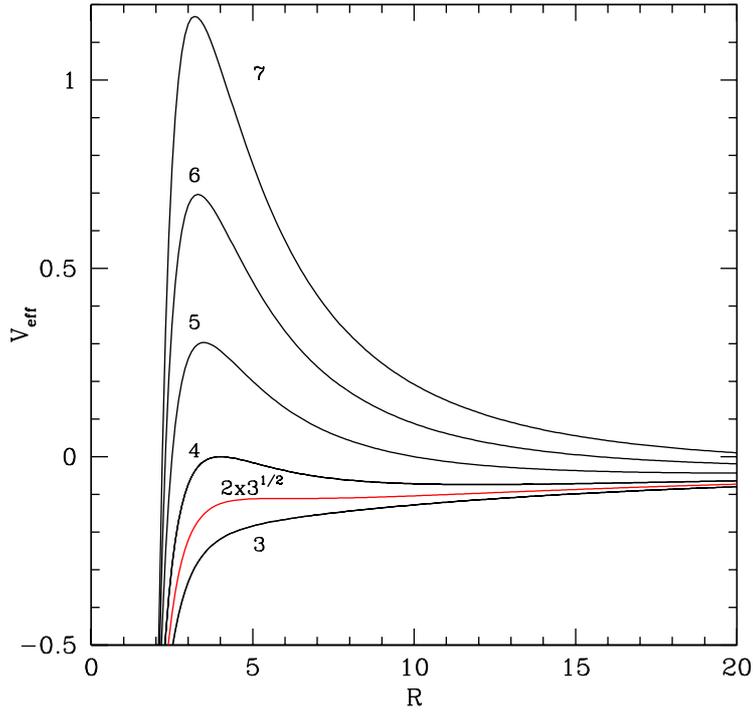


Figure 48: Effective potential for different values of the angular momentum. The $l = 2\sqrt{3}$ curve marks the minimum angular momentum before capture.

11 Hydrodynamics

11.1 The equations of fluid dynamics

Consider a volume, V , surrounded by a surface, S , with a flux of particles flowing in and out of the volume through the surface. . The mass flowing out through a surface element $d\mathbf{S}$ (with normal out from the volume) per

unit time is $\rho \mathbf{v} d\mathbf{S}$. This flow integrated over the whole surface is of course the same as the decrease in mass in the volume, or

$$-\int \frac{\partial \rho}{\partial t} dV = \int \rho \mathbf{v} d\mathbf{S} \quad (11.1)$$

With Gauss theorem the surface integral can be converted to a volume integral,

$$-\int \frac{\partial \rho}{\partial t} dV = \int \nabla \cdot \rho \mathbf{v} dV \quad (11.2)$$

Because this should be true for any volume we must have

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \rho \mathbf{v} \quad (11.3)$$

which is the continuity equation for the mass.

Next let us consider the force acting on the surface through the pressure, $-pd\mathbf{S}$ (note sign). Again using Gauss theorem we have

$$-\int pd\mathbf{S} = -\int \nabla p dV \quad (11.4)$$

The force on dV as it moves around is therefore $-\nabla p dV$, so that

$$\rho \frac{d\mathbf{v}}{dt} = -\nabla p + \rho \nabla \phi \quad (11.5)$$

where ϕ is the gravitational potential. This is the Lagrangian (comoving) derivative, so in Eulerian coordinates we get

$$\rho \frac{\partial \mathbf{v}}{\partial t} + \rho (\mathbf{v} \cdot \nabla) \mathbf{v} = -\nabla p + \rho \nabla \phi \quad (11.6)$$

This is the Euler equation.

Euler's equation can also be written as a conservation law of momentum per unit volume, ρv_i , similar to that of mass. Using the mass and momentum conservation above we have

$$\begin{aligned} \frac{\partial \rho v_i}{\partial t} &= \rho \frac{\partial v_i}{\partial t} + v_i \frac{\partial \rho}{\partial t} = -\rho \sum_j v_j \frac{\partial v_i}{\partial x_j} - \frac{\partial p}{\partial x_i} + \rho \frac{\partial \phi}{\partial x_i} - v_i \sum_j \frac{\partial \rho v_j}{\partial x_j} \\ &= -\sum_j \frac{\partial \rho v_j v_i}{\partial x_j} - \frac{\partial p}{\partial x_i} + \rho \frac{\partial \phi}{\partial x_i} \end{aligned} \quad (11.7)$$

If we write

$$T_{ij} = p \delta_{ij} + \rho v_i v_j \quad (11.8)$$

where δ_{ij} is 0 for $i \neq j$ and 1 for $i = j$, this can be written as

$$\frac{\partial \rho v_i}{\partial t} = - \sum_j \frac{\partial T_{ij}}{\partial x_j} + \rho \frac{\partial \phi}{\partial x_i} \quad (11.9)$$

The tensor T_{ij} gives the momentum flux of the j momentum component in direction i , including both the mass flux and the pressure. It is usually called the energy-momentum tensor.

Finally we consider the energy of the volume. This is the sum of the kinetic and internal energy $\rho v^2/2 + \rho \epsilon$. Here, ϵ is the internal energy per unit mass. Let us consider the evolution of this quantity with time,

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho v^2 + \rho \epsilon \right) = \frac{v^2}{2} \frac{\partial \rho}{\partial t} + \sum_i \rho v_i \frac{\partial v_i}{\partial t} + \epsilon \frac{\partial \rho}{\partial t} + \rho \frac{\partial \epsilon}{\partial t} \quad (11.10)$$

which clearly has the form of a conservation law, and is analogous to the mass conservation law Eq. (11.6).

The connection between internal energy, pressure, volume and heat loss (Tds) is given by

$$d\epsilon = -pdV + Tds = -pdV + Tds \quad (11.11)$$

where V is the specific volume (i.e., volume per unit mass), so $V = 1/\rho$ and s is the entropy, also per unit mass.

Using this together with mass conservation and momentum conservation Eqns. (11.3) and (11.6) and writing the heat loss term as

$$\rho T \left(\frac{\partial s}{\partial t} + \sum_i v_i \frac{\partial s}{\partial x_i} \right) = \rho T \frac{ds}{dt} = \Lambda \quad (11.12)$$

one can now transform Eq. (11.10) to

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho v^2 + \rho \epsilon \right) = - \sum_j \frac{\partial}{\partial x_j} \left[\rho v_j \left(\frac{v^2}{2} + w \right) \right] + v_j \frac{\partial \phi}{\partial x_j} + \Lambda \quad (11.13)$$

where $w = \epsilon + p/\rho = \gamma/(\gamma - 1) p/\rho$ is the heat function. This is equivalent to

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho v^2 + \rho \epsilon \right) = - \nabla \cdot \left[\mathbf{v} \left(\frac{1}{2} \rho v^2 + \rho \epsilon + p \right) \right] + \rho \mathbf{v} \cdot \nabla \phi + \Lambda \quad (11.14)$$

This equation says that the change in total energy density in a volume is equal to the flux of kinetic and internal energy through its surface (the

$\rho \mathbf{v}(v^2/2 + \epsilon)$ term) plus the work done on this volume by the pressure (the $\mathbf{v}p$ term), plus the heat lost by other processes, like conduction or radiation.

Equations (11.3), (11.6) and (11.14) (or Eq. (11.13)) constitute the complete set of hydrodynamic equations.

A case which often is important is that of a stationary, spherically symmetric flow. In this case all $\partial/\partial t$ -terms, as well as angular derivatives are zero. Further, $\nabla \cdot \mathbf{f} = r^{-2} \partial(r^2 f_r)/\partial r$ (since $\partial/\partial \theta = \partial/\partial \phi = 0$), and we get

$$0 = -\frac{1}{r^2} \frac{\partial}{\partial r} [vr^2 (\frac{1}{2} \rho v^2 + \rho \epsilon + p)] + \rho \frac{\partial \phi}{\partial r} + \Lambda \quad (11.15)$$

Finally we note that Eq. (11.6) can be written

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{v} = \frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho + \rho \nabla \cdot \mathbf{v} = \frac{d\rho}{dt} + \rho \nabla \cdot \mathbf{v} = 0 \quad (11.16)$$

For terrestrial applications it is often a good approximation to consider the fluid as *incompressible*, i.e. that the density does not change as the fluid element moves. This means that $d\rho/dt = 0$, or from the equation above that $\nabla \cdot \mathbf{v} = 0$.

11.2 The shock jump conditions

The speed of sound for an adiabatic gas is given by

$$c_s^2 = \frac{\partial p}{\partial \rho} = \gamma \frac{p}{\rho} \quad (11.17)$$

where γ is the adiabatic index. For a non relativistic gas $\gamma = 5/3$ and for a relativistic $\gamma = 4/3$.

Consider a wave propagating with the sound velocity, initially with a sinusoidal form (see Fig. 49). The crest of the wave will have a density slightly higher than the through. Because $c_s^2 \propto p/\rho \propto \rho^{\gamma-1}$ the sound velocity will therefore be higher in the crest than in the through, which means that the crest will propagate faster than the through. The initially sinusoidal wave will therefore gradually become steeper, and at some point the crest will catch up with the through. The wave will then break, as is familiar from the sea. Mathematically this is the same as a discontinuity.

Before the breaking of the wave the motion of the particles in the fluid could be considered as dissipation less, i.e., the viscosity played little role. One the wave brakes the particles collide and the viscosity will lead to dissipation, resulting in a non-adiabatic process.

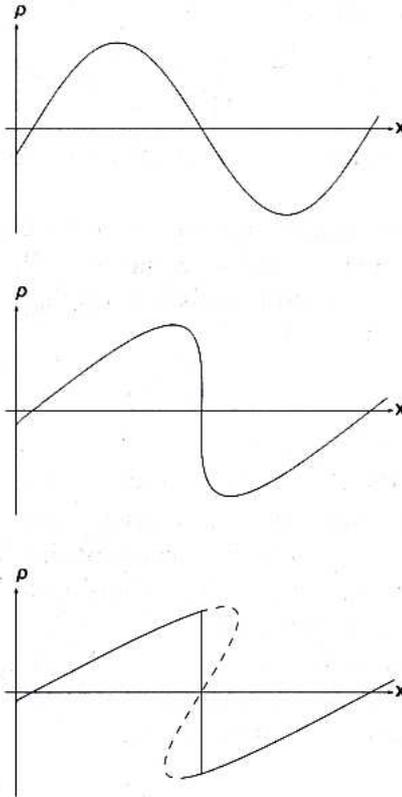


Figure 49: Non-linear development of a wave.

An even more dramatic situation will occur if some object, e.g., an airplane, is moving with a velocity higher than that of the sound. For objects with velocity below that of the sound, sound waves will propagate ahead of it, which will affect the density and pressure of the medium. The medium is therefore 'prepared' for the arrival of the object, and a fairly smooth transition takes place. This is not possible for a supersonic object. Here, the sound waves can not send this kind of early warning to the system, since they constantly lag behind the object. The arrival of the object will therefore only affect the medium once it has already arrived. The transition from the undisturbed medium to the disturbed will therefore take place very suddenly, and will for a thin transition region where the atoms of swept up by the object will collide with those of the undisturbed medium. Here viscosity will play a very important role and as above a dissipative process with a

large change in entropy will take place. The thickness of the transition will be of the order of the mean free path of the particles. For most purposes this shock wave can be considered as a mathematical discontinuity, although for certain processes like non-thermal particle acceleration or non-thermal heating the actual structure is important.

To describe the relation between the conditions ahead and behind the shock wave we use the conservation laws derived in the previous section.

Let us therefore consider a very thin region of the flow with the pre-shock gas on one side. Further, we make a Galilean transformation into the reference system of the shock. *All velocities will therefore from now on refer to that of the shock discontinuity.* If e.g., the shock moves with velocity v_s relative to an observer at rest relative to the gas into which the shock propagates, the velocity of the gas coming in to the shock *in the shock system* will be v_s . The density, velocity and pressure are here denoted by ρ_1, v_1, p_1 , while those behind the shock are denoted by ρ_2, v_2, p_2 . Because the region is very thin the flow can be considered one-dimensional, and time independent.

The mass conservation equation (11.3) then become

$$\frac{d}{dx}\rho v = 0 \quad (11.18)$$

Therefore we have

$$\rho_2 v_2 = \rho_1 v_1 = J \quad (11.19)$$

The momentum equation, Eq. (11.6), becomes

$$\rho v \frac{dv}{dx} + \frac{dp}{dx} = 0 \quad (11.20)$$

From Eq. (11.18) $\rho v = \text{const.}$, so $\rho v^2 + p = \text{constant}$, and

$$\rho_1 v_1^2 + p_1 = \rho_2 v_2^2 + p_2 \quad (11.21)$$

Finally, the energy equation, Eq. (11.14), becomes

$$\frac{d}{dx}[\rho v(\frac{v^2}{2} + w)] = 0 \quad (11.22)$$

Again using $\rho v = \text{const.}$ we get

$$\frac{v_1^2}{2} + w_1 = \frac{v_2^2}{2} + w_2 \quad (11.23)$$

Let us now specialize this to a perfect gas where $\epsilon = (\gamma - 1)^{-1}p/\rho$. The heat function is $w = \epsilon + p/\rho = \gamma/(\gamma - 1) p/\rho$. Therefore Eq. (11.23) becomes

$$\frac{v_1^2}{2} + \frac{\gamma}{(\gamma - 1)} \frac{p_1}{\rho_1} = \frac{v_2^2}{2} + \frac{\gamma}{(\gamma - 1)} \frac{p_2}{\rho_2} \quad (11.24)$$

Let us now further specialize this to the case when the pressure of the gas in front of the shock is much smaller than that in the post shock gas, i.e., $p_1 \ll p_2$. This is known as the strong shock condition. The momentum condition Eq. (11.21) then becomes

$$\rho_1 v_1^2 = \rho_2 v_2^2 + p_2 \quad (11.25)$$

and the energy condition Eq. (11.24)

$$\frac{v_1^2}{2} = \frac{v_2^2}{2} + \frac{\gamma}{(\gamma - 1)} \frac{p_2}{\rho_2} \quad (11.26)$$

Using Eqns. (11.19), (11.25) and (11.26) one then finds that

$$\frac{v_2}{v_1} = \frac{\rho_1}{\rho_2} = \frac{(\gamma - 1)}{(\gamma + 1)} \quad (11.27)$$

One then finds that

$$p_2 = \frac{2}{(\gamma + 1)} \rho_1 v_1^2 \quad (11.28)$$

and using $p = k\rho T/\mu m_p$, where μ is the molecular weight and m_p the atomic mass unit, we get

$$T_2 = \frac{2(\gamma - 1)}{(\gamma + 1)^2} \frac{\mu m_p}{k} v_1^2 \quad (11.29)$$

For $\gamma = 5/3$ we have

$$\frac{v_2}{v_1} = \frac{\rho_1}{\rho_2} = \frac{1}{4} \quad (11.30)$$

One then finds that

$$p_2 = \frac{3}{4} \rho_1 v_1^2 \quad (11.31)$$

and

$$T_2 = \frac{3}{16} \frac{\mu m_p}{k} v_1^2 \quad (11.32)$$

The mean molecular weight is for a fully ionized mix of hydrogen and helium $\mu = (n(\text{H}) + 4n(\text{He})) / (2n(\text{H}) + 3n(\text{He}))$. For $n(\text{He})/n(\text{H}) \approx 0.1$ we get $\mu = 0.61$ and

$$T_2 = 1.38 \times 10^7 \left(\frac{v_1}{1000 \text{ km s}^{-1}} \right)^2 \text{ K.} \quad (11.33)$$

12 Viscosity

The momentum transfer connected to the particle flow and pressure as given by Euler's equation Eq. (11.6), only involves reversible processes. There may, however, also be processes involving irreversible, dissipative processes. This is known as viscous processes, and are familiar from fluids like water, oil etc. Also gases are characterized by a viscosity. For these the viscosity is connected with molecular processes, There may, however, also be macroscopic processes which can have the same effect. An example is turbulence, which is important in e.g., accretion disks.

To describe the viscous momentum transfer we should add a term to the Euler equation. This means an additional term to the energy momentum tensor in Eq. (11.8). This term should obviously depend on the derivative of the velocity. If we have a flow in the x-direction, the momentum transfer connected with the friction should be proportional to the derivative of the velocity in the y-direction,

$$\tau_{xy} = a \frac{\partial v_x}{\partial y} \quad (12.1)$$

For a fluid in circular motion the friction should depend on both $\partial v_x/\partial y$ and $\partial v_y/\partial x$. In addition for a fluid with angular velocity independent of the distance from the center (solid body rotation) the friction should vanish. For this $\mathbf{v} = \boldsymbol{\Omega} \mathbf{r}$, where $\boldsymbol{\Omega}$ is independent of \mathbf{r} . Therefore $v_x = \Omega y$ and $v_y = -\Omega x$, The only combination of the above derivatives which gives zero friction is

$$\tau_{xy} = a \left(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x} \right) = \tau_{yx} \quad (12.2)$$

In addition to these shear stresses there may also be dissipation due to compression in either direction. This should be proportional to $d\rho/dt$ or $\nabla \cdot \mathbf{v}$ (see Eq. (11.16)),

$$\tau_{xx} = b \nabla \cdot \mathbf{v} = \tau_{yy} \quad (12.3)$$

The most general expression generalized to three dimensions is therefore

$$\tau_{ij} = a \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) + b \delta_{ij} \sum_k \frac{\partial v_k}{\partial x_k} \quad (12.4)$$

Usually one writes this as a traceless part and a diagonal

$$\tau_{ij} = \eta \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3} \delta_{ij} \sum_k \frac{\partial v_k}{\partial x_k} \right) + \zeta \delta_{ij} \sum_k \frac{\partial v_k}{\partial x_k} \quad (12.5)$$

Here η is called the shear viscosity and ζ the bulk viscosity. For an incompressible fluid only the former is important. This should be added to the perfect gas expression for the energy-momentum tensor in Eq. (11.8).

$$T_{ij} = p\delta_{ij} + \rho v_i v_j + \tau_{ij} \quad (12.6)$$

12.1 The Navier-Stokes equation

Adding the viscosity tensor to the energy-momentum tensor in Eq.(11.9) we get the equation of motion, including viscosity

$$\frac{\partial \rho v_i}{\partial t} = - \sum_j \frac{\partial T_{ij}}{\partial x_j} + \rho \frac{\partial \phi}{\partial x_i} = - \frac{\partial p}{\partial x_i} - \sum_j \frac{\partial \rho v_i v_j}{\partial x_j} - \sum_j \frac{\partial \tau_{ij}}{\partial x_j} + \rho \frac{\partial \phi}{\partial x_i} \quad (12.7)$$

To simplify the writing we will from now on use the Einsteins sum convention. This says that *whenever an index occurs twice this implicitly means a summation over this index*. E.g., $\sum_i A_i B_i \equiv A_i B_i$ and

$$\nabla \cdot \mathbf{v} = \sum_i \frac{\partial v_i}{\partial x_i} \equiv \frac{\partial v_i}{\partial x_i}$$

The equation above is therefore equivalent to

$$\frac{\partial \rho v_i}{\partial t} = - \frac{\partial T_{ij}}{\partial x_j} + \rho \frac{\partial \phi}{\partial x_i} = - \frac{\partial p}{\partial x_i} - \frac{\partial \rho v_i v_j}{\partial x_j} - \frac{\partial \tau_{ij}}{\partial x_j} + \rho \frac{\partial \phi}{\partial x_i} \quad (12.8)$$

If we assume that the viscosity coefficients η and ζ are independent of temperature and pressure, i.e., of x_i , we get

$$\frac{\partial \tau_{ij}}{\partial x_j} = \eta \frac{\partial^2 v_i}{\partial x_j \partial x_j} + \left(\zeta + \frac{2}{3}\eta\right) \frac{\partial^2 v_j}{\partial x_i \partial x_j} \quad (12.9)$$

(note that $\delta_{ij} \partial / \partial x_j = \partial / \partial x_i$). Using the mass continuity equation we can now write Eq. (12.7) as

$$\rho \frac{\partial v_i}{\partial t} + \rho v_j \frac{\partial v_i}{\partial x_j} = - \frac{\partial p}{\partial x_i} - \eta \frac{\partial^2 v_i}{\partial x_j \partial x_j} - \left(\zeta + \frac{2}{3}\eta\right) \frac{\partial^2 v_j}{\partial x_i \partial x_j} + \rho \frac{\partial \phi}{\partial x_i} \quad (12.10)$$

or in vector notation

$$\rho \frac{\partial \mathbf{v}}{\partial t} + \rho \mathbf{v} \cdot \nabla \mathbf{v} = - \nabla p - \eta \Delta \mathbf{v} - \left(\zeta + \frac{2}{3}\eta\right) \nabla \nabla \cdot \mathbf{v} + \rho \nabla \phi \quad (12.11)$$

where Δ is the Laplace operator, $\Delta = \sum_i \partial^2 / \partial x_i^2$. This is the Navier-Stokes equation for a viscous fluid.

In the case of an incompressible fluid, $\nabla \cdot \mathbf{v} = 0$, the momentum tensor due to viscosity, Eq. (12.5), and the Navier-Stokes equation above take a simpler form,

$$\tau_{ij} = \eta \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) \quad (12.12)$$

and

$$\frac{\partial \rho}{\partial t} + \rho \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla p - \eta \Delta \mathbf{v} + \rho \nabla \phi \quad (12.13)$$

12.2 The energy equation for a viscous fluid

For a fluid with no viscosity the energy conservation equation Eq. (11.14) is

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho v^2 + \rho \epsilon \right) = -\nabla \cdot \left[\mathbf{v} \left(\frac{1}{2} \rho v^2 + \rho \epsilon + p \right) + \rho \mathbf{v} \cdot \nabla \phi + \Lambda \right] \quad (12.14)$$

which says that the change of total energy in a volume (LHS) is equal to the energy flux across its boundary (RHS).

Also the work done by the viscous force generates an energy flux. To see this we note that $\tau_{ij} dS_j$ is the viscous force in the direction i on the surface j . The work done per unit time on the volume in direction i is therefore $v_i \tau_{ij} dS_j$. (Because we are here considering the individual contributions for each i and j , there is here no summation implied.) The total change in energy is therefore

$$\frac{dE}{dt} = \int v_i \tau_{ij} dS_i \quad (12.15)$$

where as usual a sum over i and j is implied, since we now consider the total contribution from all directions and on all sides. Applying Gauss theorem to the surface integral we get

$$\frac{dE}{dt} = \int \frac{\partial}{\partial x_j} (v_i \tau_{ij}) dV \quad (12.16)$$

Therefore, we should add this term to Eq. (12.14) giving in component form

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho v^2 + \rho \epsilon \right) = -\frac{\partial}{\partial x_i} \left[v_i \left(\frac{1}{2} \rho v^2 + \rho w \right) + v_j \tau_{ij} \right] + \rho v_i \frac{\partial \phi}{\partial x_i} + \Lambda \quad (12.17)$$

where $w = \epsilon + p/\rho$ is the heat function. This is the full energy equation including viscosity.

To more clearly see the effect of the viscous heating we calculate the total energy dissipation $\rho T ds/dt$ per volume and unit time as the fluid element

moves. Transforming from the comoving, Lagrangian system to a fixed Eulerian we get

$$\rho T \frac{ds}{dt} = \rho T \left[\frac{\partial s}{\partial t} + v_i \frac{\partial s}{\partial x_i} \right] \quad (12.18)$$

To rewrite this equation we use the law of thermodynamics, $d\epsilon = -pdV + Tds = p/\rho^2 d\rho + Tds$ to eliminate $\partial s/\partial t$. In the same way we use $dw = TdS + dp/\rho$ to eliminate $\partial s/\partial x_i$ to get

$$\rho T \frac{ds}{dt} = \rho T \left[\frac{\partial \epsilon}{\partial t} - \frac{p}{\rho^2} \frac{\partial \rho}{\partial t} + v_i \frac{\partial w}{\partial x_i} - \frac{v_i}{\rho} \frac{\partial p}{\partial x_i} \right] \quad (12.19)$$

We now use the energy conservation equation Eq. (12.17) to calculate the term $\rho(\partial\epsilon/\partial t + v_i\partial w/\partial x_i)$. We neglect for the moment the gravitational potential. Using mass conservation we find from Eq. (12.17)

$$\rho \left(\frac{\partial \epsilon}{\partial t} + v_i \frac{\partial w}{\partial x_i} \right) = -\rho v_i \frac{\partial v_i}{\partial t} - \frac{p}{\rho} \frac{\partial \rho v_i}{\partial x_i} - \rho v_i v_j \frac{\partial v_j}{\partial x_i} + \frac{\partial v_i \tau_{ij}}{\partial x_j} \quad (12.20)$$

We now use the Navier-Stokes equation

$$\rho \frac{dv_i}{dt} = \rho \frac{\partial v_i}{\partial t} + \rho v_j \frac{\partial v_i}{\partial x_j} = -\frac{\partial p}{\partial x_i} - \frac{\partial \tau_{ij}}{\partial x_j} + \rho \frac{\partial \phi}{\partial x_i} \quad (12.21)$$

multiplied by v_i to eliminate the $\rho v_i \partial v_i / \partial t$ term in Eq. (12.20), resulting in

$$\rho \left(\frac{\partial \epsilon}{\partial t} + v_i \frac{\partial w}{\partial x_i} \right) = v_i \frac{\partial p}{\partial x_i} - \frac{p}{\rho} \frac{\partial \rho v_i}{\partial x_i} + \tau_{ij} \frac{\partial v_i}{\partial x_j} \quad (12.22)$$

Using this in Eq. (12.19) we finally get

$$\rho T \frac{ds}{dt} = \tau_{ij} \frac{\partial v_i}{\partial x_j} \quad (12.23)$$

for the energy dissipation due to the viscous forces per volume and unit time. If there are radiation losses this adds the term Λ (see Eq. (12.17)). In addition, heat conduction or other diffusive energy losses may also contribute. In particular, heat conduction is often important, either through Coulomb collisions or through radiation (see Eq. 2.10). The energy flux is for this proportional to the temperature gradient,

$$F_i = -\kappa \frac{\partial T}{\partial x_i} \quad (12.24)$$

Note that the flux is opposite to the temperature gradient. The energy loss per volume is then (using again Gauss theorem)

$$\frac{\partial F_i}{\partial x_i} = -\frac{\partial}{\partial x_i} \left(\kappa \frac{\partial T}{\partial x_i} \right) \quad (12.25)$$

which should be added to Eq. (12.23). In general the conductivity κ is a function of temperature.

The Navier-Stokes equation, Eq. (12.21), and Eq. (12.23) for the viscous energy dissipation will be useful especially in connection to the hydrodynamics of accretion disks.

In the case of an incompressible fluid Eq. (12.23) takes a simple form. Using Eq. (12.12) in this we get

$$\rho T \frac{ds}{dt} = \eta \frac{\partial v_i}{\partial x_j} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) \quad (12.26)$$

which can also be written

$$\rho T \frac{ds}{dt} = \frac{\eta}{2} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right)^2 \quad (12.27)$$

13 Spherical accretion

As an application of the hydrodynamical equations we consider the adiabatic accretion of gas from a constant density medium at rest onto a central gravitating body with mass M . We therefore start with the time independent version of Eq. 11.14 with $\phi = GM/r$

$$\frac{1}{2}v^2 + (\epsilon + p/\rho) - \frac{GM}{r} = \text{constant} \quad (13.1)$$

or

$$\frac{1}{2}v^2 + \frac{\gamma}{(\gamma-1)} \frac{p}{\rho} - \frac{GM}{r} = \text{constant} \quad (13.2)$$

But $c_s^2 = \gamma p/\rho$, so

$$\frac{1}{2}v^2 + \frac{c_s^2}{(\gamma-1)} - \frac{GM}{r} = \text{constant} \quad (13.3)$$

The value of the constant depends on the problem we consider. In our case with accretion from a stationary medium with $v \approx 0$ as $r \rightarrow \infty$ we have $\text{constant} = c_s^2(\infty)/(\gamma-1)$, so

$$\frac{1}{2}v^2 + \frac{c_s^2}{(\gamma-1)} - \frac{GM}{r} = \frac{c_s^2(\infty)}{(\gamma-1)} \quad (13.4)$$

In addition we have the mass conservation which says that

$$\dot{M} = 4\pi r^2 v \rho \quad (13.5)$$

where \dot{M} is the accretion rate.

The relation above contains both $c_s(r)$ and $v(r)$. In principle we can use the relation $c_s^2 = \gamma p / \rho = \gamma K \rho^{\gamma-1}$ and Eq. (13.5) to eliminate c_s^2 in Eq. (13.4). This, however, only gives v as a function of r and \dot{M} . The latter should, however, not be a free parameter, but should be determined from the values of the density and sound velocity at infinity.

To proceed further we go back to the momentum equation which for a stationary, spherically symmetric flow gives

$$v \frac{dv}{dr} = -\frac{1}{\rho} \frac{dp}{dr} - \frac{GM}{r^2} \quad (13.6)$$

But $dp/dr = dp/d\rho \, d\rho/dr = c_s^2 d\rho/dr$. Eq. (13.5) shows that $d \ln \rho / dr = -d \ln v r^2$. Therefore

$$v \frac{dv}{dr} = c_s^2 \frac{d \ln v r^2}{dr} - \frac{GM}{r^2} \quad (13.7)$$

or

$$\left(1 - \frac{c_s^2}{v^2}\right) \frac{dv^2}{dr} = -\frac{2GM}{r^2} + \frac{4c_s^2}{r} \quad (13.8)$$

This equation is interesting because it has an obvious singularity for the RHS at

$$r_c = \frac{GM}{2c_s^2}. \quad (13.9)$$

This means that also the LHS must be zero for $r = r_c$. There are two possibilities for this. Either $dv/dr = 0$ at r_c , or $v = c_s$ at r_c .

For our outer boundary condition the $dv/dr = 0$ possibility corresponds to solutions which accelerate to a maximum velocity and then decreases (see Fig. 50). This means that the density and pressure will increase rapidly for small r . For this to be possible the accretion solution must match some kind of atmosphere. The flow will everywhere be subsonic.

The $v = c_s$ at r_c solution means that the flow becomes supersonic at r_c , which is the reason for sometimes naming it the sonic radius. This is, however, only correct for this type of solutions.

Using Eq. (13.9) in Eq. (13.4) with $v = c_s$ at $r = r_c$ we get

$$\frac{1}{2} c_s^2 + \frac{c_s^2}{(\gamma - 1)} - 2c_s^2 = \frac{c_s^2(\infty)}{(\gamma - 1)} \quad (13.10)$$

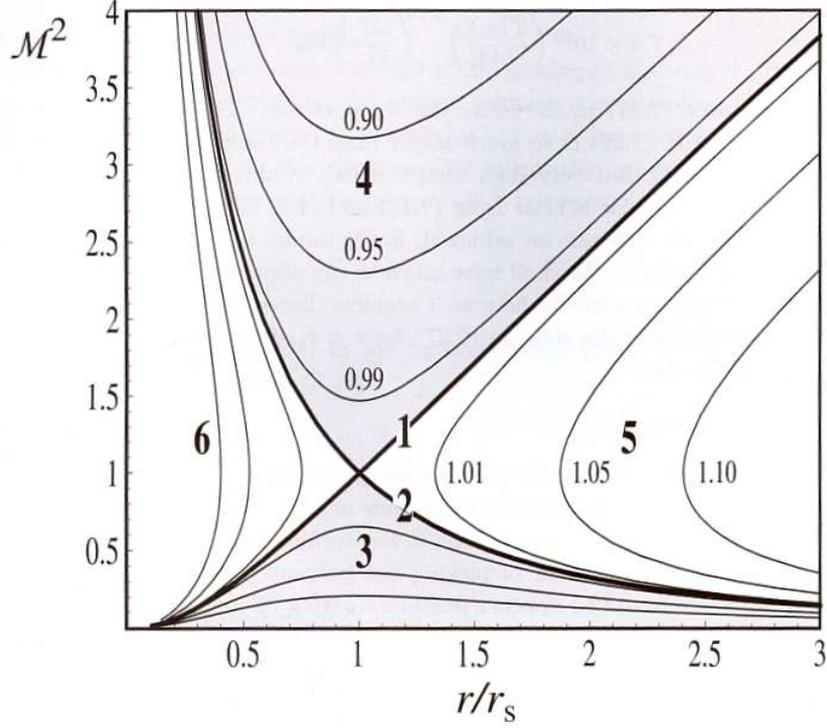


Figure 50: Different solutions corresponding to different boundary conditions of the flow equations. Solution number 2 corresponds to the accretion onto a compact object. Solution number 1 corresponds to a wind, with an outer boundary condition with negligible pressure.

or

$$c_s(r_c) = c_s(\infty) \left(\frac{2}{5 - 3\gamma} \right)^{1/2} \quad (13.11)$$

We can now use this in Eq. (13.5) evaluated at r_c

$$\dot{M} = 4\pi r_c^2 c_s(r_c) \rho(r_c) = \frac{\pi G^2 M^2 \rho(r_c)}{c_s(r_c)^3} \quad (13.12)$$

For an adiabatic flow $\rho(r_c) = \rho(\infty) (c_s(r_c)/c_s(\infty))^{2/(\gamma-1)}$, so

$$\dot{M} = \frac{\pi G^2 M^2 \rho(\infty)}{c_s(\infty)^3} \left(\frac{2}{5 - 3\gamma} \right)^{(5-3\gamma)/2(\gamma-1)} \quad (13.13)$$

This gives finally the accretion rate as function of the density and sound velocity at infinity. This is usually referred to as the Bondi accretion rate. The sound velocity is given by

$$c_s = \left(\gamma \frac{p}{\rho}\right)^{1/2} = \left(\gamma \frac{qk\mu T}{m_p}\right)^{1/2} = \left(\gamma \frac{kT}{\mu m_p}\right)^{1/2} = 12 \left(\frac{T}{10^4 \text{K}}\right)^{1/2} \text{ km s}^{-1} \quad (13.14)$$

for $\gamma = 1$. For the ISM we have $c_s \approx 10 \text{ km s}^{-1}$ and $\rho \approx 1.6 \times 10^{-24} \text{ g cm}^{-3}$ (i.e., one atom per cm^{-3}). This gives for $\gamma \approx 1.4$

$$\dot{M} = 10^{11} \left(\frac{M}{1 M_\odot}\right)^2 \frac{\rho(\infty)}{1.6 \times 10^{-24} \text{ g cm}^{-3}} \left(\frac{c_s(\infty)}{10 \text{ km s}^{-1}}\right)^3 \text{ g s}^{-1} \quad (13.15)$$

This is a very small accretion rate ($\sim 10^{-15} M_\odot \text{ yr}^{-1}$) and accretion from the ISM is usually unimportant.

For $r \ll r_c$ the gravitational attraction dominates over the pressure term in Eq. (13.4) and the gas is in free fall,

$$v = \left(\frac{2GM}{r}\right)^{1/2} \quad (13.16)$$

The density is then

$$\rho = \frac{\dot{M}}{4\pi(2GM)^{1/2}r^{3/2}} \quad (13.17)$$

We also remark that Eq. (13.3) also applies for an outflow, like the solar wind, but now with a different constant and a solution which starts subsonic at small radii, and ends with a supersonic flow at large radii.

14 Accretion efficiency

Consider first radial accretion onto a central object. Inside the accretion radius it is a good approximation to consider the velocity to be close to free fall,

$$\frac{1}{2}mv^2 \approx \frac{GmM}{R} \quad (14.1)$$

Assume that the flow ends at a hard surface, e.g., a white dwarf or neutron star. The kinetic energy will then be converted to radiation, and the luminosity will be given by

$$L = \frac{G\dot{m}M}{R} = \frac{1}{2}\dot{m}c^2 \frac{R_g}{R} \quad (14.2)$$

where $r_g = 2GM/c^2$ is the Schwarzschild radius.

We can now define an accretion efficiency by

$$L = \epsilon \dot{m} c^2, \quad (14.3)$$

so that

$$\epsilon = \frac{1}{2} \dot{m} \frac{R_g}{R} \quad (14.4)$$

For a white dwarf $\epsilon = 4 \times 10^{-4}$ while for a neutron star $\epsilon = 0.2$. For a black hole the *maximum* energy which can be released is the binding energy of the last stable orbit. For a Schwarzschild hole this corresponds to $\epsilon = 0.06$, while for a maximally rotating Kerr hole $\epsilon = 0.42$. These numbers should be compared to that of nuclear burning which corresponds to $\epsilon \approx 0.01$. Note, however, that the accretion efficiencies for a black hole depends on the energy release mechanism, because of the absence of a hard surface, and the numbers just given are only upper limits, while there is no lower limit. In fact, some models for black hole accretion have extremely low accretion efficiencies.

15 The Eddington luminosity

The accretion luminosity results in a radiation pressure acting on the accretion flow. An important upper limit to the accretion luminosity of a steady flow can be obtained by balancing the gravitational force to the radiation force acting on the gas. For a completely ionized gas, which is often a good approximation at the high temperatures in the inflowing gas, electron scattering is the most important source of opacity. The Coulomb force between the protons and electrons then transfers this momentum to the protons, which are responsible for most of the gravitational force.

Consider a spherical shell with radius r and thickness dr . The absorbed momentum of the electrons in this shell is then

$$F_s = \frac{\text{absorbed momentum in } dr}{dt} = \frac{\Delta E}{cdt} = \frac{d\tau L dA}{4\pi r^2 c} \quad (15.1)$$

where $d\tau$ is the optical depth and ΔE is the absorbed energy in a time dt . Now, $d\tau = \sigma_T n_e dr$, where σ_T is the Thompson cross section for electron scattering. Therefore,

$$\frac{F_s}{4\pi r^2 c} = L n_e \sigma_T dr dA \quad (15.2)$$

The gravitational force on the protons in the shell is

$$F_g = \frac{GMm_p n_p dr dA}{r^2} \quad (15.3)$$

We then find that for a luminosity

$$L_{Edd} = \frac{4\pi GMm_p c}{\sigma_T} \quad (15.4)$$

the radiation force is equal to the gravitational force. Steady accretion is therefore not possible for higher luminosities. This upper limit is known as the Eddington luminosity and is given by

$$L_{Edd} = 1.3 \times 10^{38} \frac{M}{M_O} \text{ erg s}^{-1} \quad (15.5)$$

In terms of accretion rate this corresponds to

$$\dot{m} = 1.4 \times 10^{18} \frac{M}{M_O} \left(\frac{\epsilon}{0.1}\right)^{-1} \text{ g/s} = 2.3 \times 10^{-8} \frac{M}{M_O} \left(\frac{\epsilon}{0.1}\right)^{-1} M_\odot \text{ yr}^{-1} \quad (15.6)$$

Note that this assumes that the flow is spherical, the opacity is dominated by electron scattering and that the flow is steady. Each of these assumptions can be violated. E.g., the inflow may be in the form of a disk, while most of the radiation may be in the polar direction. Nevertheless, the Eddington luminosity constitutes an important measure for the accretion luminosity of a steady source. Unsteady sources, such as supernovae or gamma-ray bursts, obviously violates this limit by orders of magnitude.

16 Accretion disks

16.1 Vertical structure

Let us first consider the vertical structure of the disk. The hydrostatic equation in the vertical direction gives

$$\frac{\partial p}{\partial z} = \rho F_z = -\frac{GM\rho \sin \theta}{r^2} \quad (16.1)$$

where θ is the azimuthal angle from the disk. Because $z = r \sin \theta$

$$\frac{\partial p}{\partial z} = -\frac{GM\rho z}{r^3} . \quad (16.2)$$

Assuming now that the disk is approximately homogeneous in the vertical direction we get

$$\frac{p}{H} \approx \frac{GM\rho H}{r^3} \quad (16.3)$$

where H is the thickness of the disk. If the motion of the gas in the disk is approximately Keplerian

$$\frac{v_\phi^2}{r} = \frac{GM}{r^2} . \quad (16.4)$$

With Eq. (16.3) this becomes

$$\frac{p}{\rho} \approx v_\phi^2 \frac{H^2}{r^2} \quad (16.5)$$

The sound velocity c_s is approximately given by

$$\frac{p}{\rho} = c_s^2 \quad (16.6)$$

Using this we can write Eq. (16.5) as

$$\frac{H}{r} \approx \frac{c_s}{v_\phi} = \frac{1}{M} \quad (16.7)$$

where M is the Mach number in the disk. We therefore find that a disk where the gas motion is supersonic is also a thin disk, and conversely.

16.2 Radial structure

We now consider the radial structure of a thin disk. If we integrate over the vertical direction, mass conservation in the radial direction becomes

$$\dot{m} = 2\pi r v_r \int \rho dz . \quad (16.8)$$

Defining the surface density Σ by

$$\Sigma = \int \rho dz \quad (16.9)$$

this equation becomes

$$\dot{m} = 2\pi r v_r \Sigma = \text{constant} . \quad (16.10)$$

To determine the surface density we therefore need to know the radial velocity of the gas as function of the radius.

Because of the viscosity in the disk there will be a shear force at each radius given by fA where f is the shear force per unit area. The total torque at r is then $G = Afr$.

To evaluate the shear force we use the expression for f derived earlier. If the flow is in the x direction and the velocity gradient in the y direction the shear stress is given by Eq. (12.12)

$$f_{xy} = \eta \left(\frac{\partial v_x}{\partial v_y} + \frac{\partial v_y}{\partial v_x} \right) \quad (16.11)$$

Using now polar coordinates given by

$$x = r \cos \phi \quad y = r \sin \phi \quad (16.12)$$

one obtains for the velocity vectors

$$v_x = v_r \cos \phi - v_\phi \sin \phi \quad v_y = v_r \sin \phi + v_\phi \cos \phi . \quad (16.13)$$

Further,

$$\frac{\partial}{\partial x} = \cos \phi \frac{\partial}{\partial r} - \frac{\sin \phi}{r} \frac{\partial}{\partial \phi} \quad \frac{\partial}{\partial y} = \sin \phi \frac{\partial}{\partial r} + \frac{\cos \phi}{r} \frac{\partial}{\partial \phi} \quad (16.14)$$

Collecting these results we get

$$f_{xy} = \eta \left(-\frac{v_\phi}{r} + \frac{\partial v_\phi}{\partial r} \right) = \eta r \frac{\partial}{\partial r} \left(\frac{v_\phi}{r} \right) = \eta r \frac{\partial \Omega}{\partial r} \quad (16.15)$$

Because $f_{xy} = f_{r\phi}$ we therefore get for the total torque

$$G = f_{r\phi} r A = \int \eta r^2 \frac{\partial \Omega}{\partial r} 2\pi r dz = 2\pi r^3 \nu \frac{\partial \Omega}{\partial r} \int \rho dz , \quad (16.16)$$

or

$$G = 2\pi r^3 \nu \frac{\partial \Omega}{\partial r} \Sigma \quad (16.17)$$

Consider now the net torque on a thin annular ring with thickness dr

$$\frac{dG}{dr} dr = G(r + dr) - G(r) \quad (16.18)$$

The angular momentum flow through r is

$$\dot{L}(r) = \dot{m} v_\phi(r) r \quad (16.19)$$

But $G = \dot{L}(r)$ and the change of angular momentum flow over dr is then

$$\frac{d\dot{L}(r)}{dr}dr = \frac{dG}{dr}dr \quad (16.20)$$

Therefore,

$$\frac{dG}{dr} = \dot{m} \frac{d(v_\phi r)}{dr} . \quad (16.21)$$

This can be integrated to

$$G = \dot{m}v_\phi r + \text{constant} \quad (16.22)$$

and using Eqns. (16.4) and (16.10) we get

$$G = 2\pi r^3 v_r \Sigma \Omega + C \quad (16.23)$$

But from Eq. (16.17) we have

$$G = 2\pi r^3 \nu \frac{\partial \Omega}{\partial r} \Sigma , \quad (16.24)$$

so

$$\nu \Sigma \frac{\partial \Omega}{\partial r} = v_r \Sigma \Omega + \frac{C}{2\pi r^3} \quad (16.25)$$

The value of the constant depends on the inner boundary condition. In the simple case that we have a thin layer where the disk slows down to the value of that of the accreting star

$$\frac{\partial \Omega}{\partial r} = 0 \quad \text{at } R_s \quad (16.26)$$

and

$$C = -2\pi r^3 v_r \Sigma \Omega \quad (16.27)$$

At the surface

$$\frac{v_\phi^2}{R_s} = R_s \Omega^2 = \frac{GM_s}{R_s^2} \quad (16.28)$$

and

$$\dot{m} = 2\pi R_s v_r \Sigma \quad (16.29)$$

so

$$C = -\dot{m}(GM_s R_s)^{1/2} . \quad (16.30)$$

Using

$$\Omega = \left(\frac{GM_s}{r^3} \right)^{1/2} \quad (16.31)$$

and

$$\dot{m} = 2\pi r v_r \Sigma \quad (16.32)$$

with the above results in Eq. (16.25) we finally get

$$\nu \Sigma = \frac{\dot{m}}{3\pi} \left[1 - \left(\frac{R_s}{r} \right)^{1/2} \right] \quad (16.33)$$

More general boundary conditions are discussed in FKR. The important point about Eq. (16.33) is that the surface density depends on the precise nature of the viscosity and radius. We will discuss this in Sect. .

16.3 Energy loss from the disk

We now want to estimate the total energy loss rate of the disk as function of the radius. For this we start at the expression for the energy loss due to viscosity, Eq. (12.27)

$$\frac{dE}{dt} = \frac{\eta}{2} \sum_{i,j} \left(\frac{\partial v_i}{\partial v_j} + \frac{\partial v_j}{\partial v_i} \right)^2 \quad (16.34)$$

In the same way as above we can calculate this in cylindrical coordinates and find

$$\frac{\partial v_x}{\partial v_y} + \frac{\partial v_y}{\partial v_x} = -\frac{v_\phi}{r} + \frac{\partial v_\phi}{\partial r} = r \frac{\partial \Omega}{\partial r} \quad (16.35)$$

Therefore,

$$\frac{dE}{dt} = \eta \left(r \frac{\partial \Omega}{\partial r} \right)^2 \quad (16.36)$$

Using the relation between the kinematic ν and dynamic η viscosity,

$$\eta = \nu \rho \quad (16.37)$$

we get

$$\int \frac{dE}{dt} dz = \nu \Sigma \left(r \frac{\partial \Omega}{\partial r} \right)^2 \quad (16.38)$$

Using now Eq. (16.33) for $\nu \Sigma$ we finally get the simple result

$$\int \frac{dE}{dt} dz = \frac{3GM_s \dot{m}}{4\pi r^3} \left[1 - \left(\frac{R_s}{r} \right)^{1/2} \right] \quad (16.39)$$

This is the energy generation per surface area. The total energy generation in an annulus dr is therefore

$$\frac{dL}{dr}dr = \int \frac{dE}{dt} dz 2\pi r dr = \frac{3GM_s \dot{m}}{2r^2} \left[1 - \left(\frac{R_s}{r} \right)^{1/2} \right] dr \quad (16.40)$$

and the total luminosity from the disk is easily obtained as

$$L = \frac{GM_s \dot{m}}{2r_s} \quad (16.41)$$

where r_s is the inner radius of the disk. Therefore, the energy radiated is one half of the binding energy. The rest is in the kinetic energy of the gas.

If we assume that the disk radiates locally as a blackbody we have

$$2 \times 2\pi r \sigma T^4 = \frac{3GM_s \dot{m}}{2r^2} \quad r \gg R_s \quad (16.42)$$

where the factor of two comes from the two sides of the disk. The disk temperature is therefore given by

$$T = \left(\frac{3GM_s \dot{m}}{8\pi\sigma r^3} \right)^{1/4} \quad (16.43)$$

Knowing the temperature and again assuming locally blackbody radiation we can calculate the spectrum from the disk. The integrated intensity is then

$$I_\nu = 2\pi \int_{R_s}^{\infty} B_\nu[T(r)] r dr \quad (16.44)$$

Using now

$$B_\nu(T) = \frac{2h\nu^3}{c^2(e^{h\nu/kT} - 1)} \quad (16.45)$$

and

$$T \propto r^{-3/4} \quad y = \frac{h\nu}{kT} \propto \nu r^{3/4} \quad dy = \nu^{4/3} y^{-1/3} dr \quad (16.46)$$

we get

$$I_\nu \propto \nu^3 \int_{R_s}^{\infty} B_\nu(T(r)) T dT \propto \nu^{1/3} \int_{x(R_s)}^{\infty} \frac{x^{5/3}}{(e^x - 1) dx} \quad (16.47)$$

The last integral is just a number and we therefore get for the spectrum

$$I_\nu \propto \nu^{1/3} \quad (16.48)$$

This is the canonical disk spectrum. Note that this is a power law and not a blackbody. The cutoff at low and high frequencies are caused by the inner and outer disk temperatures.

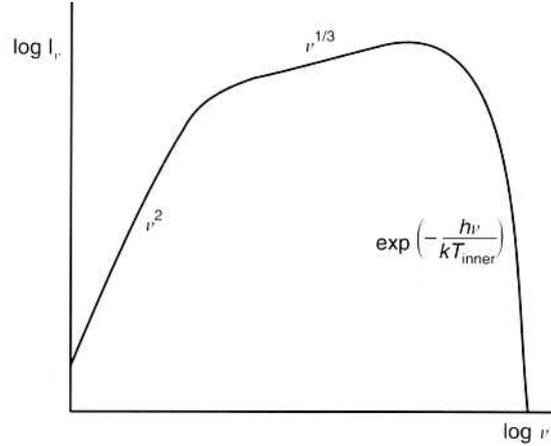


Figure 51: The spectrum from an optically thick standard disk. (From ?)

16.4 Viscosity

The surface density as well as the radial velocity depend on the viscosity (see Eqns. (16.10) and (16.33)). To describe the structure of the disk, and in particular the emission from it without making the assumption of blackbody emission as well as describing the pressure we need to relate the viscosity to the local properties of the disk.

In an ordinary gas the molecular viscosity is given by $\nu = 1/3 v_{thermal} \lambda_{mfp}$, where $v_{thermal}$ is the thermal velocity and λ_{mfp} is the mean free path of the particles. The mean free path is determined by Coulomb collisions and given by

$$\lambda_{mfp} \approx \frac{1}{n\sigma} \quad (16.49)$$

where σ is the cross section for collisions between the particles and n the particle density. The cross section for Coulomb collisions can be estimated from the distance the Coulomb energy is equal to the thermal energy, $e^2 Z^2 / r_C \approx 3/2 kT$. The cross section is therefore $\sigma = \pi r_C^2 \approx 4\pi Z^4 e^4 / 9(kT)^2$. The mean free path is therefore given by

$$\lambda \approx \frac{9(kT)^2 \ln \Lambda}{4\pi Z^4 e^4 n} \quad (16.50)$$

Here Z the charge of the particles and we have included the Coulomb logarithm, Λ , which takes the effect of distant collisions into account, $\ln \Lambda \approx 10 - 30$ (see e.g., Spitzer, Physics of Fully Ionized Gases).

To estimate the importance of viscosity we can take the ratio between the inertial term $\rho \mathbf{v} \cdot \nabla \mathbf{v}$ and the viscosity term $\eta \Delta \mathbf{v} = \rho \nu \Delta \mathbf{v}$ in the Navier-Stokes equation Eq. (12.11). If L is the typical length scale of the variations in velocity and V the velocity we get

$$R \approx \frac{VL}{\nu} \quad (16.51)$$

R is known as the Reynolds number. Experimentally it is known that flows with large Reynolds number tend to become turbulent, while the molecular viscosity is important for $R \sim 1$.

If typical values for the temperature, velocity and dimensions of an accretion disk is inserted in the expressions above one finds that $R \sim 10^{12}$. It is therefore unlikely that molecular viscosity is important, and the disk is instead turbulent. This turbulence will instead give rise to a macroscopic viscosity as the different turbulence elements interact with each other. This is not well understood in the context of accretion disks and historically one has instead used a simplified expression, based on dimensional analysis and simple arguments. In particular, it is now understood that a weak magnetic field is needed for turbulence to develop (see below).

As one of the simplest models one may argue that

$$\nu = 1/3 v_t \lambda_t \approx \alpha v_s H, \quad (16.52)$$

where one approximates the turbulent velocity with that of sound $v_t \approx v_s$ and the scale of the turbulence with the thickness of the disk, H . Because ordinary turbulence is severely damped as it becomes supersonic and shocks develop, and the scale of the turbulence is unlikely to exceed the thickness of the disk, the scaling parameter $\alpha < 1$. This simple prescription, introduced by Shakura and Sunyaev in 1973, is known as the α -disk model.

In the 1990's there has been some very interesting development in the understanding of the disk physics, and the effects of magnetic instabilities as a source of turbulence and viscosity. Balbus and Hawley, in particular, rediscovered an instability, first discussed by Chandrasekhar, known as the magnetorotational instability (MRI). They found that the presence of even a very weak magnetic field in a differentially rotating disk will lead to rapidly growing MHD instabilities, which act like a source of turbulence (Fig. 52). In contrast, a non-magnetic disk is stabilized against instabilities by the Coriolis force. The growth of the instability is very fast, of the order of the rotation period of the disk. The energy driving the instability is taken from the differential rotation and not from the magnetic field itself. Using MHD

simulations it is now possible to make considerably more realistic models of accretion disks in different contexts. Nevertheless, the simple models based on the above scaling show many of the most important features.

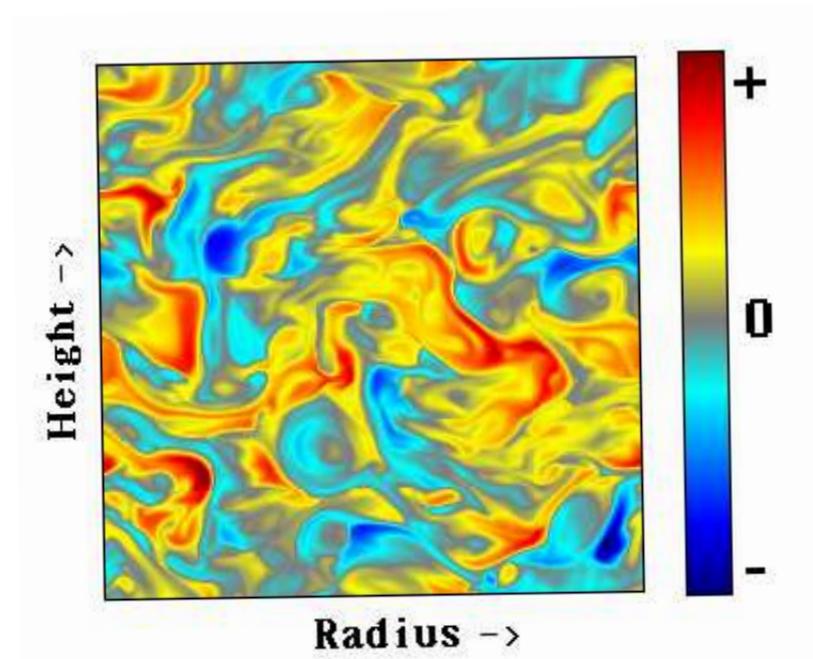


Figure 52: Cross section through a magnetized disk, showing the development of turbulence. The colors denote different values of the angular momentum. (From J. Hawley)

16.5 α -disks

In Sect. 16.3 we assumed that the disk was optically thick and locally radiated like a blackbody. For a more realistic model we need to know the density and vertical structure of the disk. The disk equations in the previous section then need to be supplemented by an equation of state and an expression for the opacity. For a non-degenerate gas we have

$$P = \frac{k\rho T}{\mu m_p} + \frac{1}{3}acT^4 \quad (16.53)$$

where the first term is the gas pressure and the second the radiation pressure.

Further, the energy losses in the previous section were assumed to be described by an optically thick blackbody. A more realistic model is obtained if we solve for the temperature using the diffusion equation with a realistic opacity. From Eq. (2.10) we have

$$F = -\frac{4acT^3}{3\kappa\rho} \frac{dT}{dr} \quad (16.54)$$

The optical depth is given by

$$\tau = \int \kappa\rho dz \approx \kappa\rho H = \kappa\Sigma \quad (16.55)$$

The flux should be equal to the vertically integrated viscous dissipation given by

$$\int_0^H \frac{dE}{dt} dz = \frac{1}{2} \frac{3GM_s\dot{M}}{4\pi r^3} \left[1 - \left(\frac{R_s}{r} \right)^{1/2} \right] \quad (16.56)$$

where the factor 1/2 accounts for the fact that only one half of the total flux emerges on each side of the disk. The flux is given by Eq. (16.54), which we can approximate with

$$F = -\frac{ac}{3\kappa\rho} \frac{dT^4}{dr} \approx \frac{ac}{3\kappa\rho} \frac{T^4}{H} = \frac{acT^4}{3\tau} \quad (16.57)$$

In this equation the temperature should be interpreted as the temperature at the center of the disk, $z = 0$. Combining these equations we get

$$\frac{acT^4}{3\tau} = \frac{3GM_s\dot{M}}{8\pi r^3} \left[1 - \left(\frac{R_s}{r} \right)^{1/2} \right] \quad (16.58)$$

To proceed further we need to calculate the optical depth, $\tau = \kappa\Sigma$. This requires that we know the density and temperature, as well as the surface density (or equivalently the thickness). Because Σ depends on the viscosity, we must have a model for the viscosity, and in this section we study the famous α -disk prescription from Eq. (16.52).

The form of the opacity depends on the temperature. At high temperature, $\gtrsim 10^8$ K electron scattering dominates. Below this Kramer's opacity, which is an approximation to the free-free and free-bound (i.e., photoelectric absorption) opacity is the dominant.

To illustrate the properties of an α -disk we show the solution for the case of an optically thick, gas pressure dominated disk where the opacity is dominated by Kramer's opacity, is given by

$$\kappa = 5 \times 10^{24} \rho T^{-7/2} \text{ cm}^2\text{g}^{-1}. \quad (16.59)$$

For this case the solution of the disk equations is given by

$$\begin{aligned}
\Sigma &= 5.2\alpha^{-4/5}\dot{M}_{16}^{7/10}m^{1/4}r_{10}^{-3/4}f^{14/5} \text{ g cm}^{-2} \\
H &= 1.7 \times 10^8\alpha^{-1/10}\dot{M}_{16}^{3/20}m^{-3/8}r_{10}^{9/8}f^{3/5} \text{ cm} \\
\rho &= 3.1 \times 10^{-8}\alpha^{-7/10}\dot{M}_{16}^{11/20}m^{5/8}r_{10}^{-15/8}f^{11/5} \text{ g cm}^{-3} \\
T_c &= 1.4 \times 10^4\alpha^{-1/5}\dot{M}_{16}^{3/10}m^{1/4}r_{10}^{-3/4}f^{6/5} \text{ K} \\
\tau &= 1.9 \times 10^2\alpha^{-4/5}\dot{M}_{16}^{1/5}f^{4/5} \\
\nu &= 1.8 \times 10^{14}\alpha^{4/5}\dot{M}_{16}^{3/10}m^{-1/4}r_{10}^{3/4}f^{6/5} \text{ cm}^2 \text{ s}^{-1} \\
v_r &= 2.7 \times 10^4\alpha^{4/5}\dot{M}_{16}^{3/10}m^{-1/4}r_{10}^{-1/4}f^{-14/5} \text{ cm s}^{-1} \quad (16.60)
\end{aligned}$$

where $f = (1 - (R_s/r)^{1/2})^{1/4}$. The parameters are given by $m = M/M_\odot$, $\dot{M}_{16} = \dot{M}/10^{16} \text{ g s}^{-1}$, and $r_{10} = r/10^{10} \text{ cm}$.

From Eq. (16.60) we note several things:

1. The density, thickness of the disk and temperature are only weakly dependent on α , while the optical depth and radial velocity are more sensitive to this parameter.
2. For typical values of \dot{M} and m the disk is indeed thin as we have assumed, with $H/r \propto \alpha^{-1/10}\dot{M}^{3/20}r^{1/8}$
3. The radial velocity ($\sim 0.3 \text{ km s}^{-1}$ for $m = \dot{M}_{16} = r_{10} = 1$) is much less than the azimuthal ($v_\phi = 1150m^{1/2}r_{10}^{-1/2} \text{ km s}^{-1}$). This justifies the Keplerian approximation for v_ϕ .

Finally a word of caution: The scalings in Eq. (16.60) depend on the assumption that $\nu = \alpha c_s H$. The above scalings are therefore likely to change if we find a more realistic model for the viscosity.

The solution in Eq. (16.60) was based on the assumption that gas pressure dominates radiation pressure and that Kramer's opacity is more important than electron scattering. The ratio of radiation pressure to gas pressure is

$$\frac{p_{rad}}{p_{gas}} = 2.8 \times 10^{-3}\alpha^{1/10}\dot{M}_{16}^{7/10}r_{10}^{-3/8}f^{7/5} \quad (16.61)$$

and one finds that Kramer's' opacity dominates electron scattering for

$$r \gtrsim 2.5 \times 10^7 \dot{M}_{16}^{2/3} m^{1/3} f^{8/3} \text{ cm} \quad (16.62)$$

In Fig.53 we show the different regimes in the $r - \dot{M}$ plane, assuming that $\alpha = 1$ for the $p_{rad}/p_{gas} = 1$ boundary, and that $M = 1 M_\odot$.

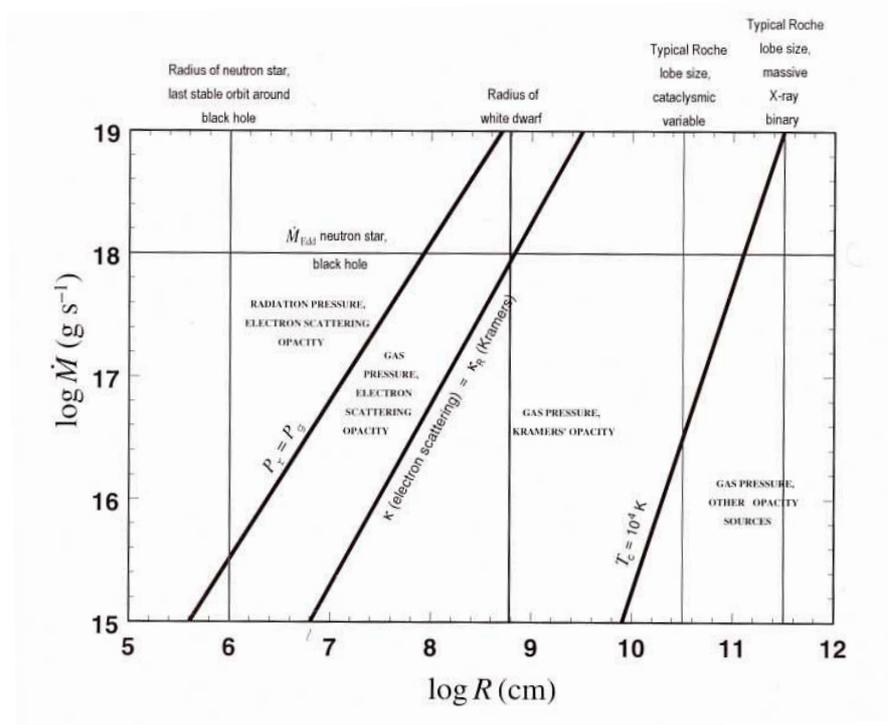


Figure 53: Different cases for the thin disk solution for $M = 1 M_{\odot}$.

16.6 Thick disks. Skip this section!

The thin disk solution we have discussed is not obviously present in all accreting sources. The condition for this to be the case is that

$$\frac{H}{r} \approx \frac{c_s}{v_{\phi}} \ll 1 \quad (16.63)$$

Therefore, if the sound velocity becomes large the thickness of the disk grows. Because $c_s^2 \approx p/\rho = kT/m_p\mu$ this will occur if the temperature is large. This may occur if e.g. the radiative efficiency is low, so that the binding energy gained by the accretion is not radiate away. This could in turn depend on if either the density is low, since for optically thin disks the cooling is in most cases proportional to the density squared. In addition, if the coupling between electrons and protons is weak the result may be also be a high temperature for the ions. The reason is that most of the binding energy is in the protons. Most of the energy losses are, however,

by the electrons through processes such as free-free radiation or Compton cooling. Therefore, if the coupling between these two components is weak this will result in a two temperature plasma, with ions at a considerably higher temperature than the electron.

The temperature of the protons may be close to the virial temperature,

$$kT_p \sim \frac{GMm_p}{r} = \frac{2GMm_pc^2}{2c^2r} = m_pc^2 \frac{r_g}{2r} = 5 \times 10^{12} \frac{r_g}{r} \text{ K} \quad (16.64)$$

The time scale of energy transfer between protons and electrons is in the non-relativistic limit

$$t_{\text{ep}} \approx \frac{(2\pi)^{1/2}}{2n_e\sigma_Tc \ln \Lambda} \frac{m_p}{m_e} (\theta_e + \theta_p)^{3/2} \approx 5.8 \times 10^{15} (\theta_e + \theta_p)^{3/2} n_e^{-1} \text{ s} \quad (16.65)$$

where $\theta_e = kT_e/m_e c^2$ and $\theta_p = kT_p/m_p c^2$.

To estimate the density we consider the spherical accretion model inside the sonic radius where the velocity is close to free-fall, $v \approx (2GM/r^3)^{1/2}$. If we parameterize the accretion rate in terms of the Eddington luminosity as $\epsilon \dot{M}_{Edd} c^2 = L_{Edd}$, or

$$\dot{M}_{Edd} = \frac{4\pi GMm_p}{\epsilon\sigma_Tc} = 1.4 \times 10^{18} \left(\frac{\epsilon}{0.1}\right)^{-1} \left(\frac{M}{M_\odot}\right) \text{ g s}^{-1} \quad (16.66)$$

we can write the density as

$$\begin{aligned} n_e &\approx \frac{\dot{M}}{4\pi m_p r^2 v} = \frac{1}{4\epsilon} \left(\frac{r_g}{r}\right)^{3/2} \left(\frac{\dot{M}}{\dot{M}_{Edd}}\right) \frac{1}{r_g \sigma_T} \\ &= 1.2 \times 10^{19} \left(\frac{\epsilon}{0.1}\right)^{-1} \left(\frac{r_g}{r}\right)^{3/2} \left(\frac{\dot{M}}{\dot{M}_{Edd}}\right) \left(\frac{M}{M_\odot}\right)^{-1} \text{ cm}^{-3} \end{aligned} \quad (16.67)$$

The accretion time scale for spherical accretion is

$$t_{\text{acc}} = \frac{r}{v_{\text{ff}}} = \frac{r_g}{c} \left(\frac{r}{r_g}\right)^{3/2} = 10^{-5} \left(\frac{M}{M_\odot}\right) \left(\frac{r}{r_g}\right)^{3/2} \text{ s} \quad (16.68)$$

Comparing these time scales we get

$$\frac{t_{\text{acc}}}{t_{\text{ep}}} = 2 \times 10^{-2} \left(\frac{\epsilon}{0.1}\right)^{-1} \left(\frac{\dot{M}}{\dot{M}_{Edd}}\right) (\theta_e + \theta_p)^{-3/2} \quad (16.69)$$

independent of mass and radius. Because $\theta_p = kT_p/m_p c^2 = 1/2(r_g/r)$ we find that in the neighborhood of the gravitational radius the time scale

for electron-ion equilibrium may be longer than the accretion time scale, especially for sub-Eddington accretion rates. Therefore, a situation where $T_p \gg T_e$ may occur. An important caveat is, however, that collisionless processes, involving plasma instabilities, may couple the electrons and ions more efficiently than Coulomb collisions. The fact that $T_e \ll T_p$ means that we have to calculate the two temperatures separately and further that we have to take into account the advection (flow term) of the energy explicitly.

The above discussion is one particular case of a more general situation. The energy equation, including viscosity and radiation losses is given by Eq. (12.23) and Eq. (12.25).

$$\rho T \frac{ds}{dt} = \tau_{ij} \frac{\partial v_i}{\partial x_j} - \frac{\partial}{\partial x_i} \left(\kappa \frac{\partial T}{\partial x_i} \right) \quad (16.70)$$

which schematically, and in a more general form, can be written

$$\rho T \frac{ds}{dt} \equiv q_{adv} = q_+ - q_- \quad (16.71)$$

where q_{adv} is the term describing the heat advected by the flow, q_+ the different heating gains by e.g. viscosity and q_- the energy losses by radiation or conduction. Depending on the relative size of the three terms one can then differentiate between three cases:

1. $q_{adv} \ll q_+ \approx q_-$. This is the standard disk model, where the local heating and cooling balance.
2. $q_{adv} \approx q_+$. In this case the advective term dominates the energy losses, and balances heating. This is known as the *advection-dominated* case.
3. $q_+ \ll q_{adv} \approx -q_-$. Heating due to viscosity is unimportant and the adiabatic heating is balanced by radiation losses. This is the case for the spherical Bondi solution.

The second case, the advection-dominated case, is of special importance for accretion onto black holes. In this case most of the internal energy may be carried along the flow and finally accrete onto the black hole. Once behind the horizon this energy will never be observed to an outside observer. There is therefore a possibility that of a very inefficient accretion process in terms of luminosity. This is only possible for a black hole, because a neutron star or white dwarf has a hard surface where the binding energy will be released in either a shock or a boundary layer.

17 Magnetospheric accretion

In the case of a central neutron star the accretion flow will be strongly affected by the magnetic field. For simplicity consider a spherical accretion inside the sonic radius. The gas will there be in approximately free fall, with velocity

$$V_{ff} \approx \left(\frac{2GM}{r} \right)^{1/2}. \quad (17.1)$$

The ram pressure (momentum flux) is given by ρV^2 , and the density by $\rho = \dot{M}/4\pi r^2 v_{ff}$, if the flow is spherically symmetric. Here \dot{M} is the accretion rate. We therefore get

$$p_{ram} = \frac{\dot{M}}{4\pi r^2} \left(\frac{2GM}{r} \right)^{1/2}. \quad (17.2)$$

As the accretion flow approaches the central object the magnetic field will become more and more important. As a first approximation one may assume that the magnetic field of the neutron star is that of a dipole,

$$B = \left(\frac{R_s}{r} \right)^3 B_s \quad (17.3)$$

where B_s is the magnetic field at the surface of the neutron star. The magnetic pressure is therefore

$$p_{magn} = \frac{B^2}{8\pi} = \left(\frac{R_s}{r} \right)^6 B_s^2. \quad (17.4)$$

Assuming now that the gas is in almost free fall to the surface, one then finds from Eqns. (17.4) and (17.2) that inside a radius

$$r_M = \left(\frac{R_s^{12} B_s^4}{8GM\dot{M}^2} \right)^{1/7} \quad (17.5)$$

the magnetic field dominates the dynamics. The radius r_M is called the *Alfvén radius*. For $R_s = 10$ km, $B_s = 10^{12}$ G, $M_s = 1.4 M_\odot$ and $\dot{M} = 10^{-8} M_\odot \text{ yr}^{-1}$ the Alfvén radius is at ~ 3000 km.

Using Eq. (14.2) we can write this in terms of the luminosity of the source as

$$r_M = \left(\frac{GR_s^{10} B_s^4 M}{8L^2} \right)^{1/7} \quad (17.6)$$

or

$$r_M = 2.9 \times 10^3 \left(\frac{R_s}{10 \text{ km}} \right)^{10/7} \left(\frac{B_s}{10^{12} \text{ G}} \right)^{4/7} \left(\frac{M}{M_\odot} \right)^{1/7} \left(\frac{L}{10^{37} \text{ erg s}^{-1}} \right)^{-2/7} \text{ km.} \quad (17.7)$$

Therefore, for a neutron star the magnetic field is dominant far before the accretion reaches the surface of the neutron star.

If we instead insert typical values for an accreting normal white dwarf, $R \sim 10^4$ km, $B_0 \sim 10^4$ G, and $L \sim 10^{34}$ erg s⁻¹ we get $r_M \sim 2.1 \times 10^4$ km. This is of the same order as the radius of the star, and depending on the precise numbers the accretion may or may not reach the surface of the star. In particular, there is a class of accreting white dwarfs with very strong magnetic fields, the AM Hercules objects or polars, where the magnetic field is $\sim 10^7$ G and therefore dominates the flow at large radii from the star. In these cases the Alfvén radius may even be comparable to the separation to the companion star, and the accretion occurs directly from the inner Lagrange point, without the formation of a disk. In the opposite limit, in accreting dwarf novae the fields are weak enough for the disk to reach the surface of the white dwarf, and the energy release occurs in a boundary layer at the equator.

The fact that the magnetic field dominates for accreting neutron stars, and is likely to have a dipolar form, has important consequences for the accretion. Even if at large radii the accretion is in a disk as a result of e.g., Roche lobe overflow from a companion star in a binary system, the gas flow inside the Alfvén radius will be channeled along the magnetic field lines. Instead of ending up on the equator in a boundary layer, where most of the dissipation will occur, it will fall down on the magnetic poles of the star (Fig. 55).

Assuming that the magnetic field is dipolar we can estimate the angle, β , from the magnetic pole over which the accretion will take place (see Fig. 54). For simplicity we assume that the magnetic rotation axis is perpendicular to the disk. This is easily generalized to arbitrary inclinations. Eq. (9.21) gives the form of the magnetic field lines, $r = K \sin^2 \theta$. At the equator we have $r \approx r_M$, giving $K = r_M$. On the surface of the star we therefore have

$$\sin^2 \theta = \frac{R_s}{r_M}. \quad (17.8)$$

Because $\theta \ll 1$, the fraction of the surface covered by the accretion column is therefore

$$f = 2 \frac{\pi \theta^2}{4\pi} = \frac{R_s}{2r_M}. \quad (17.9)$$

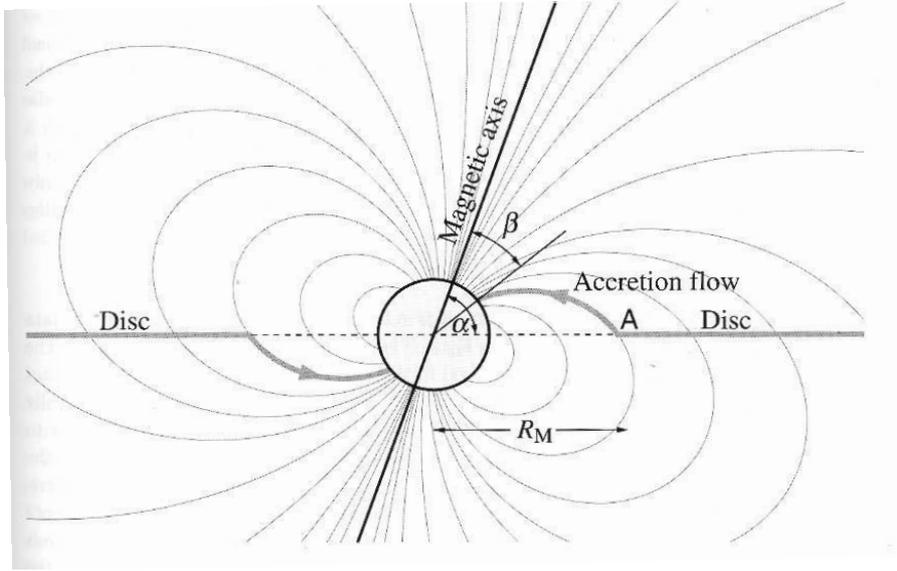


Figure 54: Schematic illustration of the transition from disk accretion to polar accretion at the magnetosphere (Frank, King and Reine).

where the factor of 2 comes from the two poles. As we have seen, $r_M \gtrsim 10^3$ km, and therefore $f \lesssim 10^{-2}$.

The radiation therefore comes mainly from an accretion column at each pole, covering only a small fraction of the surface. Therefore, it is likely that the emission will be pulsed with the frequency of the rotating neutron star, as is indeed the case in most binary neutron stars. Because the flow is highly supersonic above the surface a strong shock will form, where most of the kinetic energy is thermalized. Behind the shock the flow will gradually lose energy in radiation and slow down until it reaches the surface. Because the strong magnetic field the flow may be either in a hollow or filled cylinder, and the radiation will be strongly beamed and also polarized.

If we *assume* that the protons and electrons are thermalized by the shock (see below) the temperature behind the shock is

$$\frac{3}{2}k(T_e + T_p) = \frac{GMm_p}{R_s} \quad (17.10)$$

Assuming now that $T_e \ll T_p$, because the energy transfer to the electrons is inefficient, we find

$$T_p = \frac{GMm_p}{3kR_s} \approx 1.5 \times 10^{12} \text{K} \quad (17.11)$$

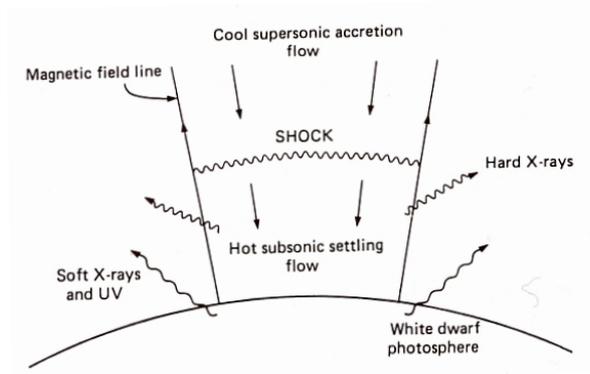


Figure 55: Schematic picture of the accretion column in an accreting X-ray binary where the magnetic field channels the accretion on to the poles. (Frank, King and Reine).

for $R_s = 10$ km and $M_s = 1.4 M_\odot$.

Because the protons radiate very inefficiently the energy has to be transferred to the electrons by some process. Ordinary Coulomb scattering is relatively inefficient,

It is, however, far from obvious that the electrons and protons are thermalized in the shock, and even less that the protons have time to transfer their energy to the electrons.

18 Spin-up by accretion

The magnetic field of the neutron star co-rotates approximately out to the light cylinder, $r_{lc} = c/\Omega = Pc/2\pi = 4.8 \times 10^4 P$ km. This is in general larger than the Alfvén radius, r_M . If the rotational frequency, Ω , of the neutron star is less than the Keplerian frequency of the disk, $\Omega_K = (GM/r_M^3)^{1/2}$, and the magnetic fields couples to the disk, there will be a torque from the disk on the neutron star. This will lead to a spin-up of the rotation of the neutron star.

The angular momentum flux inwards is given by Eq. (16.19), $\dot{L}(r) = \dot{m}v_\phi(r)r$. This is transferred to the neutron star by the coupling of the magnetic field at the magnetosphere. The change in angular momentum of the neutron star is therefore

$$I\dot{\Omega} = \dot{m}v_\phi(r_M)r_M = \dot{m}\Omega_K(r_M)r_M^2 = \frac{LR_s r_M^{1/2}}{(GM)^{1/2}} \quad (18.1)$$

Using Eq. (17.6) for the radius of the magnetosphere we get

$$I\dot{\Omega} = \frac{L^{6/7} R_s^{12/7} B_s^{2/7}}{2^{3/14} G^{3/7} M^{1/4}} \quad (18.2)$$

For a quantitative estimate we take $I \approx 2MR_s^2/5$, which is valid for a uniform star, and use $\dot{\Omega} = 2\pi\dot{P}/P^2$. Using the canonical values $R_s = 10$ km, $M = 1 M_\odot$ and $B_s = 10^{12}$ G we find

$$\log\left(\frac{-\dot{P}}{P}\right) = -4.4 + \log P + \frac{6}{7} \log L_{37} . \quad (18.3)$$

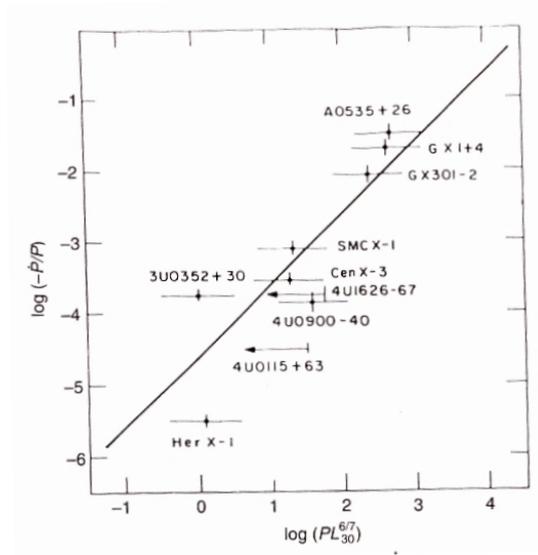


Figure 56: Spin-up rates for a number of X-ray binaries compared with Eq. (18.3) (Rappaport and Joss (1983)).

In Fig. 56 we compare this with observations of a number of X-ray binaries. In many sources the X-ray luminosity varies by a large factor. Eq. (18.3) then predicts that there should be a relationship between the instantaneous luminosity and spin-up rate. In Fig. 57 we show this for the X-ray binary GRO J1744-28. In both cases we see a very good agreement with this comparatively simple picture, giving strong support to this general accretion picture.

This spin-up can only continue as long as the Keplerian frequency at the magnetosphere is larger than that of the neutron star. We can therefore

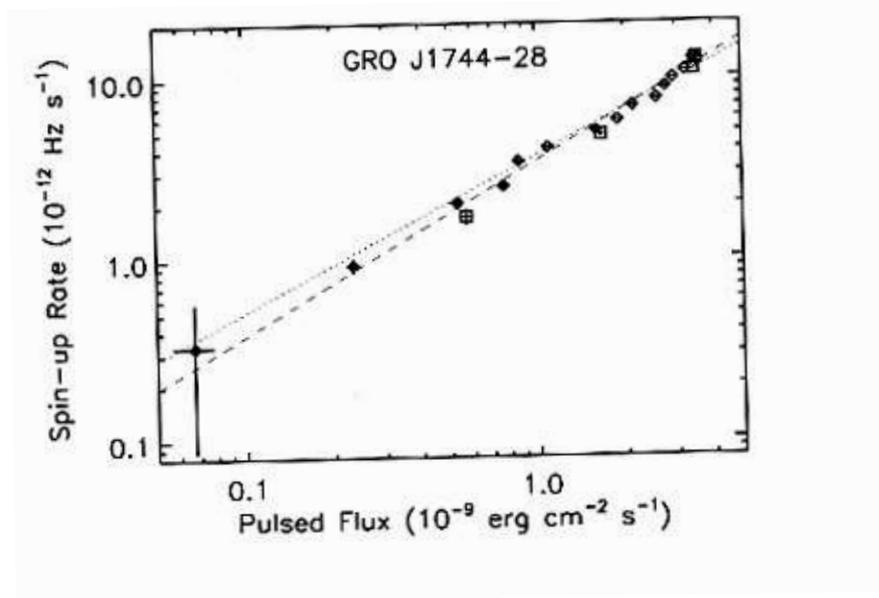


Figure 57: Spin-up rates as function of the X-ray luminosity for GRO J1744-28, compared with Eq. (18.3) (??).

calculate the minimum period of a neutron star spun-up by accretion by setting $\Omega_s = 2\pi/P_s = \Omega_K = (GM/r_M^3)^{1/2}$, or

$$P_s = 2\pi \left(\frac{r_M^3}{GM} \right)^{1/2} \quad (18.4)$$

Again using Eq. (17.6) we find

$$P_{min} = 2^{5/14} \pi \left(\frac{R_s^{15} B_s^6}{G^2 M^2 L^3} \right)^{1/7} \quad (18.5)$$

or

$$P_{min} = 1.0 \left(\frac{B_s}{10^{12} \text{ G}} \right)^{6/7} \left(\frac{L}{10^{38} \text{ erg s}^{-1}} \right)^{-3/7} = 2.7 \times 10^{-3} \left(\frac{B_s}{10^9 \text{ G}} \right)^{6/7} \left(\frac{L}{10^{38} \text{ erg s}^{-1}} \right)^{-3/7} \text{ s.} \quad (18.6)$$

The first estimate refers to a young neutron star with a strong magnetic field. The observed milli-second pulsars all have very weak fields, $B_s \lesssim 10^9$ G, meaning that the minimum period can be of the order of milli-seconds, as observed.

If we combine this expression with the relation between the relation between the spin-down for pulsars and the magnetic field, Eq. (9.10), $B \approx 4 \times 10^{19} (P\dot{P})^{1/2}$ G, we can get a relation between the spin-down of the pulsar and the period for a pulsar in a binary system,

$$\dot{P} < 1 \times 10^{-15} P^{4/3} \text{ s/s.} \quad (18.7)$$

We have here assumed that the luminosity of the binary is close to the Eddington luminosity, 1.3×10^{38} erg s⁻¹. Note here that the spin-up occurs while the neutron star is an accreting X-ray binary, while the spin-down of the pulsar takes place after this has stopped, and the neutron star has become a radio pulsar. If we compare this relation with Fig. 40 we note that there is indeed an absence of millisecond pulsars with \dot{P} above this relation. The second thing to note from this figure is that all pulsars with $P \lesssim 0.01$ s are in binary systems. This is, of course, consistent with the expectation if the spin-up is caused by accretion. We therefore have a consistent picture for the formation of these milli-second pulsars.

19 Characteristics of X-ray Binaries

There are about 130 known high mass X-ray binaries (HMXBs). About half of them show pulsations, most of them with periods 10 – 300 s, but extending from 0.069 s to 20 min. The companion stars are high mass stars with masses $\gtrsim 10 M_{\odot}$. The luminosities are very high, $\gtrsim 10^5 L_{\odot}$, more characteristic of stars of mass $\gtrsim 20 M_{\odot}$. This indicates that they must have lost a large fraction of the mass as a result of binary evolution, either to the companion by Roche lobe overflow, or by a stellar wind (Fig. 58). These massive stars have very strong winds with mass loss rates of $\sim 10^{-6} M_{\odot} \text{ yr}^{-1}$ and wind velocities of 2000 – 3000 km s⁻¹. Well-known examples are Cen X-3 and Cyg X-1 and LMC X-1. The latter two do not show any pulsations, have very large masses for the unseen component, and are therefore believed to have black holes.

A large fraction of the HMXBs have a Be-star companion (i.e., B star with emission lines) show transient X-ray outburst lasting several days. These are thought to be the result off accretion onto the neutron star as it passes periastron in a very elliptic orbit (i.e., closest to the companion). A well-known example is A0535+26.

The low mass X-ray binaries (LMXBs) belong to an old population of stars, older than a billion years, and therefore must have low mass companion stars. The LMXBs are seldom X-ray pulsars. The main reason is that

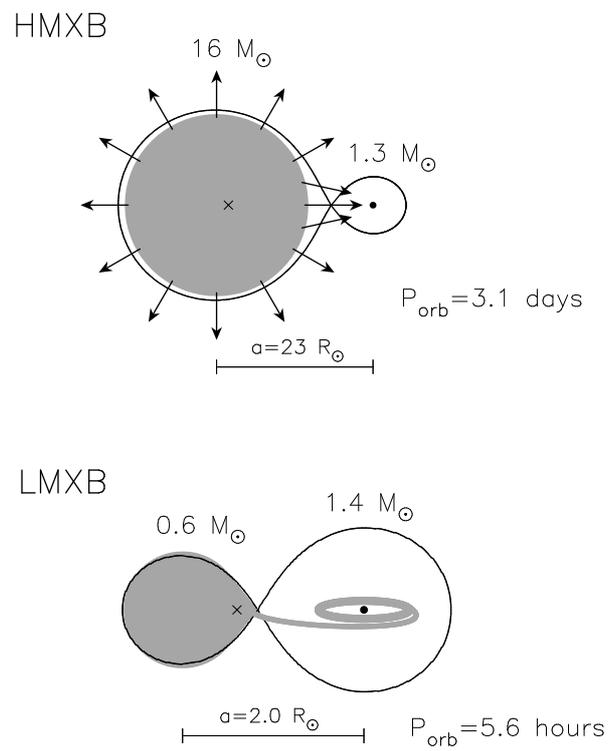


Figure 58: Schematic illustrations of high and low mass X-ray binaries. (From Tauris & van den Heuvel 2006)

Table 3: The two main classes of strong Galactic X-ray sources (From Tauris & van den Heuvel 2006)

	HMXB	LMXB
X-ray spectra:	$kT \geq 15$ keV (hard)	$kT \leq 10$ keV (soft)
Type of time variability:	regular X-ray pulsations no X-ray bursts	only a very few pulsars often X-ray bursts
Accretion process:	wind (or atmos. RLO)	Roche-lobe overflow
Timescale of accretion:	10^5 yr	$10^7 - 10^9$ yr
Accreting compact star:	high \vec{B} -field NS (or BH)	low \vec{B} -field NS (or BH)
Spatial distribution:	Galactic plane	Galactic center and spread around the plane
Stellar population:	young, age $< 10^7$ yr	old, age $> 10^9$ yr
Companion stars:	luminous, $L_{\text{opt}}/L_x > 1$ early-type O(B)-stars $> 10 M_{\odot}$ (Pop. I)	faint, $L_{\text{opt}}/L_x \ll 0.1$ blue optical counterparts $\leq 1 M_{\odot}$ (Pop. I and II)

the magnetic fields are in general weak, $10^9 - 10^{10}$ G compared to HMXBs. Many of them show X-ray bursts, while HMXBs do not. The orbital periods range from 11 min to 17 days. The optical spectra are dominated by the accretion disk, because of the comparatively low luminosity of the companion star (Fig. 58). Many show kiloHertz quasi-periodic oscillations (QPOs) presumably arising in the accretion disks. These have proved extremely valuable for probing both the physics of the accretion disk and in particular relativistic effects near the compact object.

More than ten LMXBs show the characteristics of a black hole. These are all so called soft X-ray transients where the X-ray emission increase dramatically to $\sim 10^{38}$ erg s $^{-1}$ during a few weeks. Also the optical flux increases by 6 – 10 magnitudes, and they therefore resemble ordinary novae, although the origin of the emission is different. The best known example is A0620-00.

In Table 3 we summarize the main characteristics of the HMXBs and LMXBs.

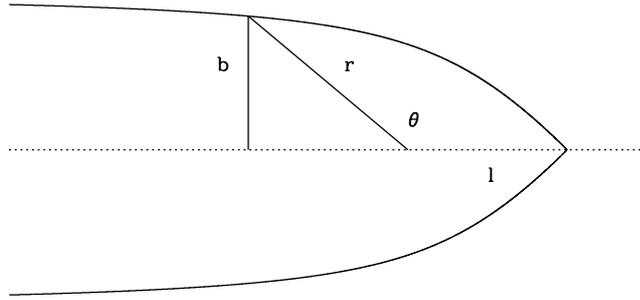


Figure 59: Geometry of the accretion flow onto a gravitating object, discussed in the text.

20 Wind accretion

While most low mass X-ray binaries lose mass by Roche lobe overflow, the high mass X-ray binaries have very strong stellar winds. If the companion star is not filling its Roche lobe the wind will dominate the accretion rate to the compact object.

The total velocity of the compact object relative to the wind is the sum of the orbital and wind velocities,

$$v_t = v_w^2 + v_{orbit}^2 . \quad (20.1)$$

The equations of motion for a particle with impact parameter b are

$$\frac{d^2 r}{dt^2} - r \frac{d\theta}{dt} = -\frac{GM_X}{r^2} \quad (20.2)$$

and conservation of angular momentum

$$r^2 \frac{d\theta}{dt} = bv_\infty \equiv h \quad (20.3)$$

Substituting $u = 1/r$ we get

$$\frac{d^2u}{d\theta^2} + u = \frac{GM_X}{h^2} \quad (20.4)$$

This has the solution

$$u = A \cos \theta + B \sin \theta + \frac{GM_X}{h^2} \quad (20.5)$$

where the constants are fixed by the boundary conditions. As $\theta = \pi$ we must have $r \rightarrow \infty$, i.e., $u \rightarrow 0$. Therefore, $A = GM/h^2$. Further, at infinity we must have $b = r \sin \theta$ (see Fig. 59). Therefore, $B = 1/b$, so that

$$u = \frac{GM_X}{h^2}(1 + \cos \theta) + \frac{1}{b} \sin \theta \quad (20.6)$$

The flow will reach $b = 0$ at $\theta = 0$, which occurs at a distance l behind the X-ray source, given by

$$l = \frac{h^2}{2GM_X} \quad (20.7)$$

Because we have a fluid, the particles will collide on this line, $\theta = 0$, and by symmetry lose their angular velocity. The radial velocity should, however, not be seriously affected and close to v_∞ .

For capture by the compact object we now require that

$$\frac{GM_x}{l_c} > \frac{v_\infty^2}{2} \quad (20.8)$$

Using Eq. (20.7) l_c corresponds to a critical impact parameter

$$b_c = \frac{2GM_X}{v_\infty^2} \quad (20.9)$$

Using Eq. (20.1) we can generalize this to

$$R_c = \frac{2GM_X}{v_w^2 + v_{orbit}^2} . \quad (20.10)$$

This is known as the Hoyle-Lyttleton radius.

If we know ask what the accretion rate from the wind is, this is given by the fraction of the wind which is inside R_c , or

$$\dot{m} = \frac{\pi R_c^2}{4\pi R_p^2} \dot{M}_p = \frac{R_c^2}{4R_p^2} \dot{M}_p, \quad (20.11)$$

where \dot{M}_p is the mass loss rate of the companion star. For an accretion efficiency ϵ the luminosity then becomes

$$L_X = \epsilon \dot{m} c^2 = \epsilon \frac{R_c^2}{4R_p^2} \dot{M}_p c^2 = \frac{\epsilon \dot{M}_p c^2}{4} \left(\frac{2GM_X}{R_p(v_w^2 + v_{orbit}^2)} \right)^2. \quad (20.12)$$

If $v_w \gg v_{orbit}$ we get

$$L_X \approx \epsilon \dot{M}_p c^2 \left(\frac{GM_X}{R_p} \right)^2 v_w^{-4}. \quad (20.13)$$

We here note the sensitivity to both the wind velocity and the orbital separation between the companion and compact object.

21 Cyclotron radiation

The strong magnetic field in the accretion column in X-ray binaries can have important effects of the observed X-rays. In particular, cyclotron processes may leave some imprint on the spectrum. The cyclotron frequency is given by $\Omega = eB/cm_e$. The observed energy of the first harmonic is therefore

$$E = 11.6 \left(\frac{B}{10^{12} \text{ G}} \right) (1+z)^{-1} \text{ keV} \quad (21.1)$$

where the last factor accounts for the gravitational redshift.

The first source for which a cyclotron feature was observed was Her X-1 in a balloon flight in 1978. There are now ~ 14 X-ray binaries for which at least one absorption line has been identified. In some case higher harmonics have also been identified, the best example is X0115+63. X0115+63 is one of the best-studied X-ray transients. The source shows pulsation at 3.6 s, and is orbiting an O9e companion with a period of 24.3 days. In this source four harmonics have been identified at 12.74, 24.16, 35.74, and 49.5 keV (Fig. 60). In a weak magnetic field the spacing between the harmonics is just $E_N = nE_0$, where E_0 is the energy of the first harmonic. In a strong field relativistic effects change this to

$$E_N = m_e c^2 \left\{ \left[1 + 2n \frac{B}{B_{crit}} \sin^2 \theta \right]^{1/2} - 1 \right\} / \sin^2 \theta \quad (21.2)$$

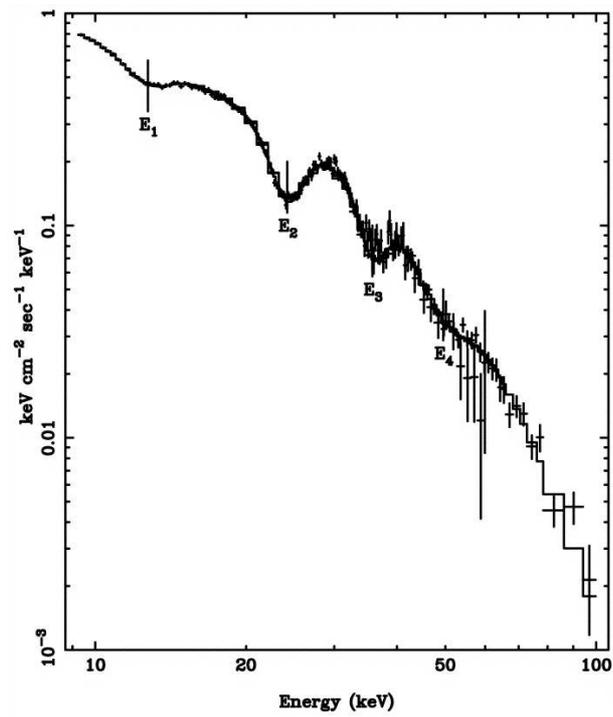


Figure 60: Hard X-ray spectrum of X0115+63 observed with Beppo-SAX, showing the first four cyclotron harmonics. binaries. (From Santangelo et al. 1999)

Here m_e is the electron mass, θ is the angle between the photon propagation angle and the magnetic field, and $B_{crit} = 4.414 \times 10^{13}$ G is the critical field (see, e.g., Wang, Wasserman, & Lamb 1993). The typical magnetic field depend on the angle to the line of sight, but is in general $10^{12} - 10^{13}$ G, close to those found for radio pulsars.

22 Non-thermal emission

Many objects in high energy astrophysics are characterized by non-thermal spectra, typically power laws. Examples are the radio emission from supernova remnants, radio emission from active galaxies, and the diffuse radio emission from the Galaxy. We therefore assume that the energy spectrum of the emitting electrons can be approximated by

$$\frac{dn}{dE} \propto \frac{dn}{d\gamma} \propto E^{-p} \propto \gamma^{-p} \quad (22.1)$$

where γ is the Lorentz factor of the electrons. Assuming that the electrons radiate in a narrow band around

$$\nu_c = \frac{eB}{2\pi m_e c} \gamma^2 = 4.2 \times 10^6 \gamma^2 B(G) \text{ Hz} \quad (22.2)$$

one may obtain the emitted spectrum from the expression

$$j(\nu) \propto B^2 \int \gamma^2 \delta(\nu - \gamma^2 \nu_B) \gamma^{-p} d\gamma \propto B^{(p+1)/2} \nu^{-(p-1)/2} \quad (22.3)$$

where

$$\nu_B = 4.2 \times 10^6 B(G) \text{ Hz} \quad (22.4)$$

The observed spectral index α is therefore related to the electron spectral index by $\alpha = (p - 1)/2$. The observed spectra vary between different objects, but are in general between $\alpha \approx 0.5$ and $\alpha \approx 1$, corresponding to $p = 2 - 3$. In some cases the spectra may flatter or steeper than this.

ISM $B \sim 3 \times 10^{-6}$ G I up to 5 GHz

One may now ask two fundamental questions with regard to these electrons. 1. What is the origin of the NT electrons (and cosmic rays in general)? 2. Why do we observe different values of p in different sources and at different frequencies?

The first question is one of the most important and to a large extent unanswered questions in astrophysics. The second is better understood and we will therefore discuss this first.

23 Energy losses of relativistic electrons

The electrons lose energy on a time scale τ given by

$$\tau = \frac{E}{dE/dt} \quad (23.1)$$

where dE/dt gives the energy loss rate. For ionizations this is given by

$$\frac{dE}{dt} = 7.6 \times 10^{-9} n_{HI} (3 \ln \gamma + 19.8) \text{ eV s}^{-1} \quad (23.2)$$

where n_{HI} is the hydrogen density. Therefore, for this process

$$\tau \propto E \quad (23.3)$$

For bremsstrahlung the energy loss is given by

$$\frac{dE}{dt} = 4n_p Z^2 \alpha r_e^2 c g E . \quad (23.4)$$

For a fully ionized plasma the Gaunt factor is

$$g = \ln \gamma + 0.36 \quad (23.5)$$

so that

$$\frac{dE}{dt} = 7 \times 10^{-17} n_p (\ln \gamma + 0.36) E (eV) \text{ eV s}^{-1} \quad (23.6)$$

For a neutral plasma

$$g = \ln(183/Z^{1/3}) - \frac{1}{18} \quad (23.7)$$

and the energy loss is now given by

$$\frac{dE}{dt} = 3.7 \times 10^{-16} n_{HI} E (eV) \text{ eV s}^{-1}. \quad (23.8)$$

Therefore, as for the ionization losses, the bremsstrahlung time scale is

$$\tau \text{ independent of } E \quad (23.9)$$

Synchrotron losses are given by

$$\frac{dE}{dt} = 2\sigma_T c \gamma^2 U_B \sin^2 \theta \quad (23.10)$$

Averaging over all pitch angles θ one finds

$$\frac{1}{2} \int_0^\pi \sin^2 \theta d \cos \theta = \frac{2}{3} \quad (23.11)$$

and therefore the averaged energy loss rate

$$\frac{dE}{dt} = \frac{4}{3} \sigma_T c \gamma^2 U_B = 6.6 \times 10^{-4} \gamma^2 B^2 \text{ eV s}^{-1} \quad (23.12)$$

For inverse Compton losses the same expression with the magnetic field energy density replaced by the radiation density applies,

$$\frac{dE}{dt} = \frac{4}{3}\sigma_T c \gamma^2 U_{rad} \quad (23.13)$$

The ratio between the two time scales is therefore given by the wellknown relation

$$\frac{dE}{dt}_{IC} / \frac{dE}{dt}_{synch} = \frac{U_{rad}}{U_B} \quad (23.14)$$

For both these processes

$$\tau \propto E^{-1} \quad (23.15)$$

Finally, we have the adiabatic energy loss rate

$$\frac{dE}{dt} = -p \frac{dV}{dt} \quad (23.16)$$

Using

$$\rho = \frac{1}{V} \quad (23.17)$$

and the continuity equation

$$\frac{d\rho}{dt} = -\rho \nabla \cdot \mathbf{v} \quad (23.18)$$

we can write

$$\frac{dV}{dt} = -V^2 \frac{d\rho}{dt} = -V \nabla \cdot \mathbf{v} \quad (23.19)$$

In general

$$E = \frac{pV}{(\gamma - 1)} \quad (23.20)$$

so

$$\frac{dE}{dt} = -p \frac{dV}{dt} = (\gamma - 1) E \nabla \cdot \mathbf{v} \quad (23.21)$$

The rate of energy loss due to adiabatic expansion therefore depends on the velocity field of the expansion (or contraction).

For the special, but interesting, case of a uniformly expanding sphere

$$v(r) = v_0 \frac{r}{r_0} \quad (23.22)$$

The divergence of this is easy to calculate since all angular terms are zero, and we get

$$\nabla \cdot \mathbf{v} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \mathbf{v}) = \frac{v_0}{r_0} \frac{1}{r^2} \frac{\partial}{\partial r} (r^3) = 3 \frac{v_0}{r_0} = 3 \frac{v}{r} \quad (23.23)$$

Using this in Eq. (23.21) we find

$$\frac{dE}{dt} = 3(\gamma - 1)E\frac{v}{r} = 3(\gamma - 1)\frac{E}{t} = \frac{E}{t}. \quad (23.24)$$

Therefore, not unexpectedly

$$\tau \propto t \quad (23.25)$$

24 The equation for the diffusion losses

Let us now consider the flow of particles in and out of an energy - space volume element, $x, x + dx$ and $E, E + dE$. Let ϕ_x be the flux of particles in the x -direction while ϕ_E is the flux of particles in the E -direction. The change in the number of particles in the volume $dx dE$ is then

$$\begin{aligned} \frac{dn(E, x, t)}{dt} dE dx &= [\phi_x(E, x, t) - \phi_x(E, x + dx, t)] dE + \\ &[\phi_E(E, x, t) - \phi_E(E + dE, x, t)] dx + \\ &Q(E, x, t) dE dx \end{aligned} \quad (24.1)$$

or

$$\frac{dn(E, x, t)}{dt} = -\frac{\partial \phi_x(E, x, t)}{\partial x} - \frac{\partial \phi_E(E, x, t)}{\partial E} + Q(E, x, t) \quad (24.2)$$

where $Q(E, x, t)$ is the source term for creation or destruction of particles in the volume.

Assuming now that the spatial flux can be described as a diffusion process, we can write the flux in space as a gradient with a diffusion coefficient D , which in general is a function of both energy and position. Therefore,

$$\phi_x(E, x, t) = -D \frac{\partial n(E, x, t)}{\partial x} \quad (24.3)$$

Inserting this in the equation above we find

$$\frac{dn(E, x, t)}{dt} = D \frac{\partial^2 n(E, x, t)}{\partial x^2} - \frac{\partial \phi_E(E, x, t)}{\partial E} + Q(E, x, t) \quad (24.4)$$

which can be generalized to three spatial dimensions as

$$\frac{dn(E, x, t)}{dt} = D \nabla^2 n(E, x, t) - \frac{\partial \phi_E(E, x, t)}{\partial E} + Q(E, x, t) \quad (24.5)$$

The flux of particles in energy is just the slowing down rate ('velocity' in energy space) times the particle density

$$\phi_E(E, x, t) = n(E) \frac{dE}{dt} \equiv -b(E)n(E) \quad (24.6)$$

and the diffusion equation now takes the final form

$$\frac{dn(E, x, t)}{dt} = D\nabla^2 n(E, x, t) + \frac{\partial}{\partial E}[b(E)n(E, x, t)] + Q(E, x, t) \quad (24.7)$$

In this equation $b(E)$ includes all energy loss (or gain) processes.

24.1 Example

Consider first the case when we have a uniform, infinite medium with a uniform distribution of sources acting over an infinite time. Eq. (24.7) then takes the simple form

$$\frac{d}{dE}[b(E)n(E)] = -Q(E) \quad (24.8)$$

which can be integrated to

$$\int d[b(E)n(E)] = - \int Q(E)dE \quad (24.9)$$

Now, if $n(\infty) \rightarrow 0$ we find

$$b(E)n(E) = \int_E^\infty Q(E)dE \quad (24.10)$$

Assuming that we inject a power law spectrum for the electrons $Q(E) = SE^{-p}$ the integral is easily evaluated to

$$n(E) = \frac{\kappa E^{-(p-1)}}{(p-1)b(E)} \quad (24.11)$$

If we know write the energy loss term as

$$b(E) = A_1(\ln E/m_e c^2 + 19.8) + A_2 E + A_3 E^2 \quad (24.12)$$

where the first term is the ionization loss, the second bremsstrahlung and adiabatic losses, and the third the sum of the synchrotron and inverse Compton losses.

From Eqns (24.11) and (24.12) we then conclude that ionization losses flatten the electron spectrum by one power, adiabatic and bremsstrahlung losses leaves the slope unaffected, and synchrotron and inverse Compton steepens the spectrum by one power. These are general and very useful rules to remember.

24.2 The minimum energy argument

Many radio galaxies and supernova remnants show synchrotron power law spectra. From this one infers that the non-thermal electrons also have power law spectra

$$\frac{dn}{dE} = \kappa E^{-p} . \quad (24.13)$$

The relation between the emitted frequency ν_{max} and the Lorentz factor of the electrons γ is

$$\nu_{max} = 0.29\nu_c = 0.29 \cdot 4.2 \times 10^6 \gamma^2 B(G) = 1.2 \times 10^6 \gamma^2 B(G) \equiv CE^2 B \quad (24.14)$$

where

$$\alpha = (p - 1)/2 \quad (24.15)$$

The total synchrotron luminosity is then given by

$$L_\nu = A(\alpha)V\kappa B^{1+\alpha}\nu^{-\alpha} \quad (24.16)$$

with $\alpha = (p - 1)/2$. The total energy of the source is then

$$W = V(\epsilon_e + \epsilon_p + \frac{B^2}{8\pi}) \quad (24.17)$$

where ϵ_e and ϵ_p are the energy density in non-thermal electrons and protons, respectively, and the last term is the energy density in magnetic fields. For convenience we write

$$\epsilon_e + \epsilon_p \equiv \eta\epsilon_e \quad (24.18)$$

Using Eq. (24.13) we find

$$W = V(\epsilon_e\eta + \frac{B^2}{8\pi}) = V(\int_{E_{min}}^{E_{max}} \kappa E^{-p} E dE + \frac{B^2}{8\pi}) \quad (24.19)$$

where E_{min} and E_{max} are the minimum and maximum energies, respectively.

We then find for the electron energy

$$W_e = V \int_{E_{min}}^{E_{max}} \kappa E^{-p+1} dE = V \frac{\kappa}{(p-2)} (E_{min}^{-p+2} - E_{max}^{-p+2}) \quad (24.20)$$

Using Eq. (24.14) this becomes

$$W_e = V \frac{\kappa(CB)^{(p-2)/2}}{(p-2)} [\nu_{min}^{-(p-2)/2} - \nu_{max}^{-(p-2)/2}] \quad (24.21)$$

If we use Eq. (24.16) to eliminate the normalization of the total number of electrons we get

$$W_e = \frac{V(CB)^{(p-2)/2}}{(p-2)} \frac{L_\nu}{A(\alpha)VB^{1+\alpha}\nu^{-\alpha}} [\nu_{min}^{-(p-2)/2} - \nu_{max}^{-(p-2)/2}] \quad (24.22)$$

We can now write the total energy in particles as

$$W_{part} = \eta G(\alpha, \nu_{min}, \nu_{max}) L_\nu B^{-3/2} \nu^\alpha \quad (24.23)$$

and the total energy

$$W_{tot} = \eta GL_\nu B^{-3/2} \nu^\alpha + V \frac{B^2}{8\pi} \quad (24.24)$$

For a given luminosity the total energy is therefore a function only of the magnetic field. The two terms clearly define a minimum of the total energy and we can therefore ask what the magnetic field which corresponds to this is. This is easily found to be given by

$$B_{min} = \left[\frac{6\pi\eta GL_\nu \nu^\alpha}{V} \right]^{2/7} \quad (24.25)$$

This corresponds to an energy density in magnetic field and particles given by

$$W_{minB} = V \frac{B^2}{8\pi} = \frac{3}{4W_{min\ part}} \quad (24.26)$$

and a total energy

$$W_{min\ tot} = \frac{7}{4} \frac{V^{3/7}}{8\pi} [6\pi\eta GL_\nu \nu^\alpha]^{4/7} \quad (24.27)$$

The importance of this is that one can from this expression calculate a strict value of the minimum total energy in terms of observable parameters, i.e., the luminosity and emitting volume. It is therefore often used as a first estimate of the energy of different kinds of sources. There is, however, a number of points which should be remembered in this type of analysis.

- There is no obvious physical reason why the magnetic field and particle energies should adjust to this minimum energy requirement!
- For $B = B_{min}$ we have $W_B \approx W_{part}$. The minimum energy field is therefore often called the equipartition field

- $W_{min\ tot}$ is a strict lower limit to the total energy.
- $W_{min\ tot}$ is sensitive to the non-thermal proton/electron ratio η . This can be large, but is not directly observable.
- W_{part} depends on the minimum energy (for $\alpha > 0.5$, or $p > 2$). ν_{min} is difficult to determine from observations.
- W depends on the emitting volume. The filling factor may of the particle may be low!

We finally list some practical formulae. For this we assume a typical spectral index $\alpha = 0.75$, i.e. $p = 2.5$ and a minimum energy equal to the minimum observed frequency $\nu_{min} = \nu$.

$$B_{min} = 1.8 \times 10^4 \left[\frac{\eta L_\nu}{V} \right]^{2/7} \nu^{1/7} \text{ G} \quad (24.28)$$

$$W_{min\ tot} = 3 \times 10^{13} V^{3/7} (\eta L_\nu)^{4/7} \nu^{2/7} \text{ erg} \quad (24.29)$$

As an example we take the radio lobes of the radio galaxy Cyg A. These have a size of $R \approx 50 \text{ kpc} = 1.5 \times 10^{23} \text{ cm}$. The total luminosity is $L_\nu \approx 8 \times 10^{35} \text{ erg/s/Hz}$ at 178 MHz. We then get $W_{min\ tot} = 2 \times 10^{59} \eta^{4/7} \text{ erg}$. This corresponds to a rest mass energy equal to $\sim 3 \times 10^5 M_\odot c^2$, which is of course a minimum.

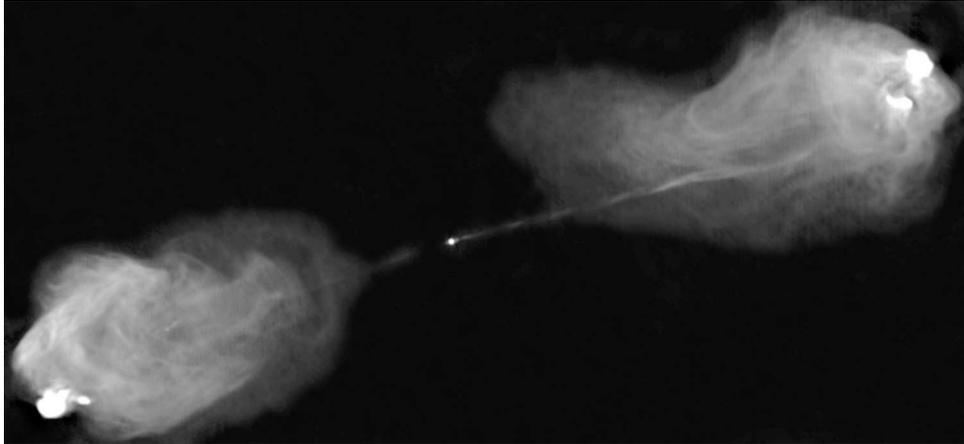


Figure 61: Radio image at 6 cm (5 GHz) with VLA.

25 The expansion of a point explosion

The sudden energy release into a medium is of general interest in many contexts. Historically this was first considered by Sedov and Taylor in the USSR and UK, respectively, in connection to the interpretation of nuclear explosions on Earth. Using the analytic solution for this problem Taylor could estimate the total energy release in the first nuclear tests, which was classified material in the US.

In astrophysics this has applications for supernova remnants and other explosive events, and variations on this have been developed in several other contexts. A strict analysis based on the time dependent hydrodynamical equations is fairly complicated and can be found in e.g., Landau and Lifshitz, Fluid Dynamics. A simple derivation can, however, be obtained if a few approximations are made.

We therefore consider the injection of an energy E_{tot} into a medium of constant density ρ . This energy will create a large overpressure and a shock wave propagating out from the centre. This shock will sweep up the surrounding medium in a hot shell. To proceed we now make the main simplification: Instead of calculating the detailed interior structure we consider all swept up mass to be contained in the shell, which is assumed to be very thin and we neglect the structure in this. Further, we neglect the mass of the exploding object compared to that swept up by the shock wave, and also the external pressure p_{ext} compared to that of the hot bubble. We also consider the explosion to be adiabatic so that the total energy is constant. We will consider radiative losses later.

To see why the thin shell approximation is reasonable we make a simple estimate the thickness of the shell. If Δr is the thickness we have

$$M = \frac{4\pi}{3}r^3\rho = 4\pi r^2\rho_2\Delta r \quad (25.1)$$

where ρ_2 is the density behind the shock. For a strong shock we have $\rho_2 = 4\rho$ and therefore

$$\Delta r = \frac{r}{12} \quad (25.2)$$

We therefore find that the shell indeed is thin.

With this assumption the equation of motion for the shell becomes

$$\frac{d}{dt}(Mv) = 4\pi r^2(p - p_{ext}) \approx 4\pi r^2p \quad (25.3)$$

where M is the mass, v the velocity, and p the interior pressure working on the shell.

We now consider the velocity and density at the shock. We assume that this is strong (consistent with our neglect of p_{ext}). Let v'_1 be the pre-shock velocity of the incoming gas in the shock frame. Because the outside medium is at rest this is also the velocity of the shock wave V_s as seen by observer at rest (with a minus sign).

$$v'_1 = V_s \quad (25.4)$$

The velocity in the shock frame of the gas behind the shock is $v'_2/4$, so

$$v'_2 = \frac{V_s}{4} \quad (25.5)$$

In the rest frame the velocity of the gas behind the shock is given by

$$v = -v' + V_s = \frac{3}{4}V_s = \frac{3}{4}\frac{dr}{dt} \quad (25.6)$$

With Eq. (25.1) and Eq. (25.3) the equation of motion, Eq. (25.3), becomes

$$\frac{1}{4}\rho\frac{d}{dt}\left(r^3\frac{dr}{dt}\right) = r^2p \quad (25.7)$$

We now want to relate the pressure to the conserved quantity E_{tot} . This consists of the kinetic energy of the shell and the thermal energy of the hot bubble interior of the shell

$$E_{tot} = E_{therm} + E_{kin} = \epsilon V + \frac{1}{2}Mv^2 = \text{constant} \quad (25.8)$$

where ϵ is the internal energy, given by

$$\epsilon = \frac{p}{(\gamma - 1)} = \frac{3}{2}p \quad (25.9)$$

where we have assumed an adiabatic gas with adiabatic index $\gamma = 5/3$. Therefore

$$E_{tot} = \frac{3}{2}pV + \frac{1}{2}Mv^2 \quad (25.10)$$

If we now replace p with Eq. (25.7) and v from Eq. (25.6) we get a differential equation for the radius of the shell

$$E_{tot} = \frac{3}{8}\frac{\rho}{r^2}\frac{d}{dt}\left(r^3\frac{dr}{dt}\right)\frac{4\pi}{3}r^3 + \frac{1}{2}\rho\frac{4\pi}{3}r^3\left(\frac{3}{4}\frac{dr}{dt}\right)^2 \quad (25.11)$$

Although this looks fairly complicated a simple power law solution can be found. If we try

$$r = At^\alpha \quad (25.12)$$

the above equation becomes

$$E_{tot} = A^5 \frac{\pi}{2} \rho t^{5\alpha-2} [\alpha(4\alpha - 1) + \frac{3}{4}\alpha^2] \quad (25.13)$$

Because the right hand side should be constant this means that

$$\alpha = \frac{2}{5} \quad (25.14)$$

and

$$E_{tot} = A^5 \rho \frac{9\pi}{50} \quad (25.15)$$

Our final solution for the radius of the shell is therefore

$$r = At^\alpha = \left(\frac{50}{9\pi}\right)^{1/5} \left(\frac{E_{tot}}{\rho}\right)^{1/5} t^{2/5} \approx 1.12 \left(\frac{E_{tot}}{\rho}\right)^{1/5} t^{2/5} \quad (25.16)$$

This is the famous Sedov-Taylor solution.

Before proceeding we make a few remarks. Our solution above shows that the radius (and velocity) of the shell only depends on the ratio E_{tot}/ρ . This should not be surprising since these two numbers are the only characteristic quantities of the problem, since we have assumed that the mass of the exploding object and the external pressure can be neglected. This in turn is the reason why we could find the simple power law solution above. In fact, except for numerical factors the solution for the radius could have been written down on purely dimensional grounds (Do this!). Such solutions are known as similarity solutions, because the properties of the solution only depend in this scaling factor. For this type of solutions the density and pressure in the interior as they are found from the exact hydrodynamical calculation have e.g., the same form, independent of time and only change in the scaling of the radius and values at the shock front (see Landau and Lifshitz, Fluid Dynamics). We also remark that the approximate solution in Eq. (25.16) only differs by a small factor from the exact solution. Instead of the factor 1.12 the exact solution has 1.152. This approximate solution is therefore surprisingly accurate, although it does not give any information about e.g., the density inside the shock.

26 Collisionally ionized plasmas

ALthough non-thermal emission is characteristic for most object in high energy astrophysics, most of them also emit thermal radiation. In many objects this actually dominates the non-thermal. In addition, this component

has the advantage of having a very large information content in the form of spectral lines from a wide range of ionization states, and thus temperatures. Further, detailed velocity information can be derived from the line profiles. They therefore offer a very useful complement to the non-thermal emission. We will therefore in this section consider some of the properties of plasmas which are collisionally ionized. This includes especially the emission from very hot gas, such as encountered in the solar corona, supernova remnants and clusters of galaxies.

We first consider the ionization balance. If an electron has an energy larger than the ionization potential, χ_i , of an ion X_i this may result in a higher ionization stage X_{i+1}



For this to occur the thermal energy has to fulfill

$$\frac{1}{2}m_e v^2 \approx kT_e \approx \chi_i \quad (26.2)$$

This corresponds to

$$T_e \approx 1.16 \times 10^4 \chi_i (\text{eV}) \text{ K} \quad (26.3)$$

Because most ionizations result from the tail of the Maxwellian distribution an element usually becomes ionized at a temperature well below that in Eq. (26.3).

The number of ionizations per unit time is given by

$$\frac{dn_i}{dt} = n_e n_i C_i(T_e) \quad (26.4)$$

where $C_i(T_e)$ is the collisional ionization rate, which is a Maxwellian average of the cross section for ionizations.

The process balancing the ionizations are in most low density plasmas two-body recombination



where an ion captures an electron with emission of a photon. In plasmas of very high density also three-body recombination may become important.

The number of recombinations are given by

$$\frac{dn_i}{dt} = n_e n_{i+1} \alpha_{i+1}(T_e) \quad (26.6)$$

where $\alpha_{i+1}(T_e)$ is the recombination coefficient, which is again a thermal average of the recombination cross section.

In a steady state plasma the above processes balance and the state of ionization is given by a solution to the equations

$$n_i C_i(T_e) = n_{i+1} \alpha_{i+1}(T_e) \quad (26.7)$$

for all ionization stages i of the element. This is supplemented by a number conservation equation

$$\sum_i n_i = n_{tot} \quad (26.8)$$

where n_{tot} is the total number of ions of this element. Note that the ionization balance is independent of the electron density, because both collisional ionizations and recombinations are proportional to this.

The temperature dependence of the ionization and recombination rates are given approximately by

$$C_i(T_e) \propto T_e^{-1/2} e^{-kT_e/\chi_i} \quad \alpha_{i+1}(T_e) \propto T_e^{-0.7} \quad (26.9)$$

Therefore in the case of collisional ionizations the ionization balance is very sensitive to the temperature. In Fig. (61) we show the ionization balance of iron as function of temperature. The presence of a specific ion in

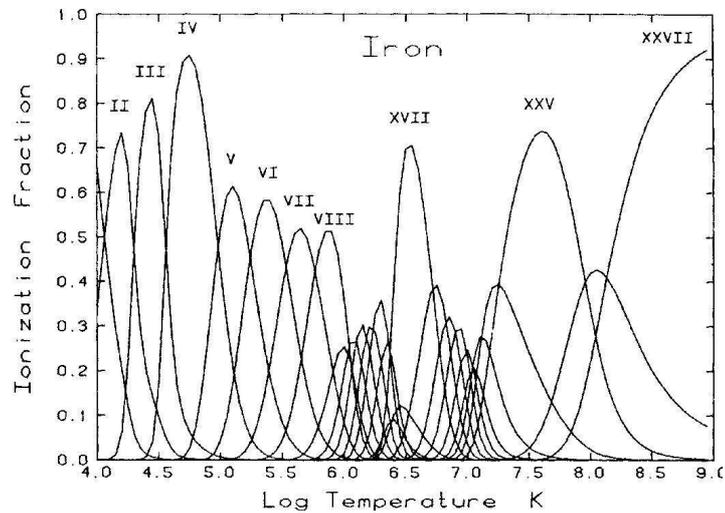


Figure 62: Ionization balance of iron as function of temperature. (Shull 1979)

e.g. the spectrum of a collisionally ionized plasma is therefore a very good temperature indicator.

An example is the coronal line from Fe XIV seen in the optical spectrum the solar corona at solar eclipses at a wavelength of 5303 Å. This ion is only present at a temperature of $\sim 2 \times 10^6$ K, which is therefore a good estimate of the temperature of the solar corona. What is giving rise to this high temperature is not yet completely clear.

In supernova remnants the temperature of the gas is determined by the speed of the shock wave. The radius of the remnant is given by

$$r = 1.1 \left(\frac{E_{tot}}{\rho} \right)^{1/5} t^{2/5} = 1.6 \times 10^{19} E_{51}^{0.2} n^{-0.4} \left(\frac{t}{1000 \text{ years}} \right)^{0.4} \text{ cm} \quad (26.10)$$

The velocity of the shock is therefore

$$V_s = \frac{dr}{dt} = \frac{2}{5} \frac{r}{t} = 2.0 \times 10^3 E_{51}^0.2 n^{-0.4} \left(\frac{t}{1000 \text{ years}} \right)^{-0.6} \text{ km s}^{-1} \quad (26.11)$$

From Eq. (11.33) we have for a strong shock

$$T = 1.38 \times 10^7 \left(\frac{V_s}{1000 \text{ km s}^{-1}} \right)^2 = 5.6 \times 10^7 E_{51}^{0.40} n^{-0.8} \left(\frac{t}{1000 \text{ years}} \right)^{-1.2} \text{ K} \quad (26.12)$$

27 Emission processes

The emission from a hot thermal plasma has several components. The continuum emission originates from free-free emission (bremsstrahlung), bound-free emission (recombination emission) and two-photon emission. The latter is a result of the forbidden continuum emission originating from the $2s - 1s$ transition in hydrogenic atoms and the $2^3S - 1^1S$ and $2^3P - 1^1S$ transitions in helium like ions.

The free-free emission has a form

$$\frac{dE}{d\nu dt dV} = n_e n_i \frac{C}{T_e^{0.5}} e^{-h\nu/kT_e} \quad (27.1)$$

while the bound-free emission

$$\frac{dE}{d\nu dt dV} = n_e n_i \frac{C}{T_e^{1.5}} e^{-(h\nu - \chi_i)/kT_e} \quad (27.2)$$

for $h\nu > \chi_i$.

In addition to the continuum processes there are discrete transitions where bound electrons are excited by collision with free, thermal electrons. These are then either de-excited by radiative transitions, giving rise to line emission, or by collisional de-excitations. The rate for collisional excitations from a level i to level j can be written as

$$\frac{dE_{ij}}{dt} = n_e n_i h\nu_{ij} \frac{\Omega_{ij}}{g_i T_e^{0.5}} e^{-h\nu_{ij}/kT_e} \quad (27.3)$$

where Ω_{ij} is the collision strength and g_i the statistical weight of the lower level. The collision strength is a thermal average over the cross section for collisional excitation, and is in general only slowly varying with temperature.

Because all these emission processes are triggered by thermal electrons and proportional to the number of ions present the total emission can be written as

$$\frac{dE}{dt dV} = n_e n_i \Lambda(T_e) \quad (27.4)$$

where the factor $\Lambda(T_e)$ mainly depends on the temperature. This is, however, only true if collisional de-excitation processes as well as radiative heating and ionization processes can be ignored. In particular, collisional ionization has to be much more important than photoionization. One is usually referring to this as the coronal approximation, because it closely applies to that of the solar corona, as well as clusters of galaxies. In Fig. (62) we show this cooling curve for a plasma of 'cosmic' composition, i.e., to a chemical composition close to that of the sun or interstellar medium in the Galaxy.

A good and very useful approximation to $\Lambda(T_e)$ is given by

$$\Lambda(T_e) = 8.0 \times 10^{-23} (Z/Z_\odot) \left(\frac{T_e}{10^6 K} \right)^{-0.90} + 2.3 \times 10^{-24} \left(\frac{T_e}{10^6 K} \right)^{0.50} \text{ erg cm}^3 \text{ s}^{-1} \quad (27.5)$$

where Z is the metallicity and Z_\odot that of the sun. In this expression the first term comes from collisional excitation and radiative recombination, important for $T_e \lesssim 2 \times 10^7$ K. The last term is from free-free emission, which dominates at high temperatures and has a $\Lambda(T_e) \propto T_e^{0.5}$ dependence.

28 Supernova remnants

28.1 The adiabatic stage

After the SN explosion the ejecta expands with a velocity of $V_s \sim (1 - 2) \times 10^4$ km s⁻¹ into the CSM of the progenitor star (see Sect. 6.7). This

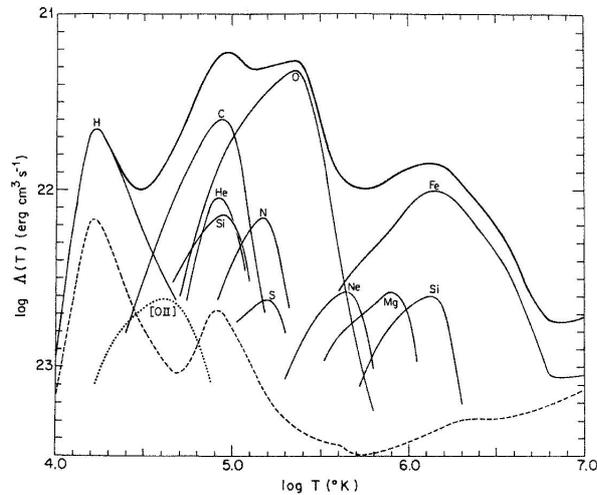


Figure 63: Cooling function for a plasma of solar composition. The contribution from different elements are also shown. (Shull 1979??)

results in a strong shock wave, and as we will discuss later, a reverse shock into the SN ejecta. The interaction gives rise to both radio, X-ray and optical emission. Because the temperature behind the outer shock is $1.4 \times 10^9 (V_s/10^4 \text{ km s}^{-1})^2 \text{ K}$ the cooling time of this is long and the shock will in general be adiabatic. The reverse shock is, however, at least during the first years, usually radiative. Because the mass of the swept-up CSM is smaller than the ejecta mass the Sedov-Taylor solution in Sect. 25 does not apply. Instead the expansion is close to free expansion. The exact rate depends on the density gradient of the ejecta which determines how the reverse shock will propagate inwards in mass.

The structure of the CSM can be highly complex with remains of previous evolutionary stages in the form of fast and slow winds and their interactions. While the fast blue supergiant wind during the main sequence stage creates a low density bubble around the SN. Most SNe explode as red supergiants. These have slow winds with a high mass loss rate, which expand out into the bubble. The interaction of the SN ejecta with this wind will therefore dominate the evolution during the first decades or centuries after the explosion. Because the circumstellar density is high in this case the interaction can be very strong, with strong emission both in radio and in X-rays. For some SNe (the Type II_n) even the optical light curve may be

dominated by circumstellar interaction.

For the Type Ib/c SNe the situation may be even more complex. After the red supergiant stage, with a slow, dense wind, the star evolves back to the blue, as it loses the hydrogen envelope and becomes a Wolf-Rayet star. The wind now has roughly the same mass loss rate ($\sim 10^{-5} M_{\odot} \text{ yr}^{-1}$) but the wind velocity increases to $1000 - 3000 \text{ km s}^{-1}$. This fast wind will therefore sweep up the red supergiant wind into a very dense shell. Between this dense shell and the freely expanding wind there will be a region of shocked wind from the Wolf-Rayet star. Fig. 63 shows the evolution of this complex structure from the main sequence to the Wolf-Rayet stage.

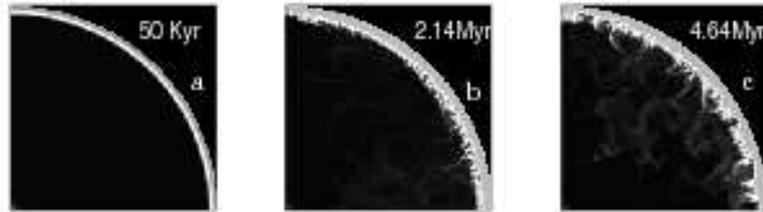


Fig. 1.4. Time-sequence of images of the formation of the CSM around a $35 M_{\odot}$ O-Star during the main sequence.

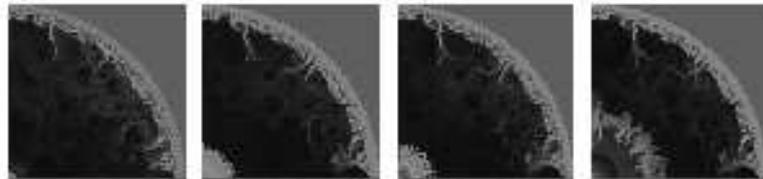


Fig. 1.5. The first two density images from left show the formation of the inner RSG shell, which is unstable to thin-shell perturbations. The next two display the onset of the WR wind and its collision with the RSG shell, causing it to fragment and the RSG material to be mixed in with the rest of the nebula.

Figure 64: Structure of CSM of a $35 M_{\odot}$ star from the main sequence until it explodes as a Wolf-Rayet star. (Dwarkadas 2004)

From this discussion it is clear that the interaction of the SN will be very complex during the first few hundred years, as the shock sweeps up the CSM. During this time the SN ejecta sweeps up mainly the CSM, and the mass of the swept up gas is comparable to the ejecta mass. The conditions for the Sedov-Taylor solution to apply are therefore not satisfied either for

the uniform density or for the ejecta mass to be neglected. Only after at least 1000 years can this be expected to be a reasonable approximation. Even then the non-uniform nature of the ISM means that the evolution will deviate from the Sedov-Taylor solution.

An object in this early stage of evolution is the youngest SN remnant in the Galaxy, Cas A, which exploded around 1670. Fig. 64 shows a 24μ Spitzer (red), optical HST (yellow) and Chandra X-ray image (green and blue) of Cas A. The high energy X-rays are blue and trace especially the continuum radiation from non-thermal synchrotron emission close to the outer shock. The green is dominated by line emission and coincides well with the region between the reverse shock and the forward shock of optical emission from cool gas, as well as dust emission seen in the IR. It is likely that the SN ejecta are interacting with CS gas from the progenitor star, possibly a Type IIn SN.

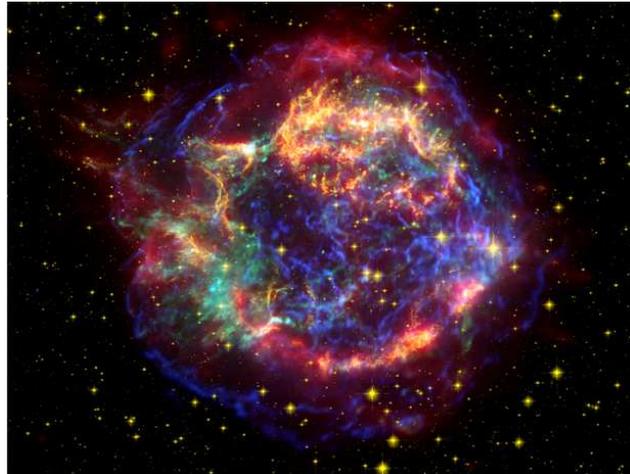


Figure 65: Composite of a 24μ Spitzer (red), optical HST (yellow) and Chandra X-ray image (green and blue) of Cas A.

28.2 The radiative stage

As the SN remnant expands it sweeps up more gas and gradually slows down. Because $T_s \propto V_s^2$ the temperature decreases and cooling becomes more and more important. Because the density in the shell at the shock is highest it will cool fastest. The interior may, however, stay hot and exert a pressure on the cool shell.

If we, however, neglect this pressure the equation of motion for the shell is

$$\frac{d}{dt}(Mv) = 4\pi r^2(p - p_{ext}) \approx 0 \quad (28.1)$$

saying that the momentum of the shell is constant. The mass, however, increases and the shell therefore slows down. Because Mv is constant and the shell mass

$$M = \frac{4\pi}{3}r^3\rho \quad (28.2)$$

we find that

$$r^3v = \text{constant} . \quad (28.3)$$

Therefore,

$$r = At^{1/4} , \quad (28.4)$$

which is often referred to as the snowplow solution. Because we have neglected the interior pressure of the still hot gas, the actual solution will be somewhere between the Sedov-Taylor and snowplow solutions.

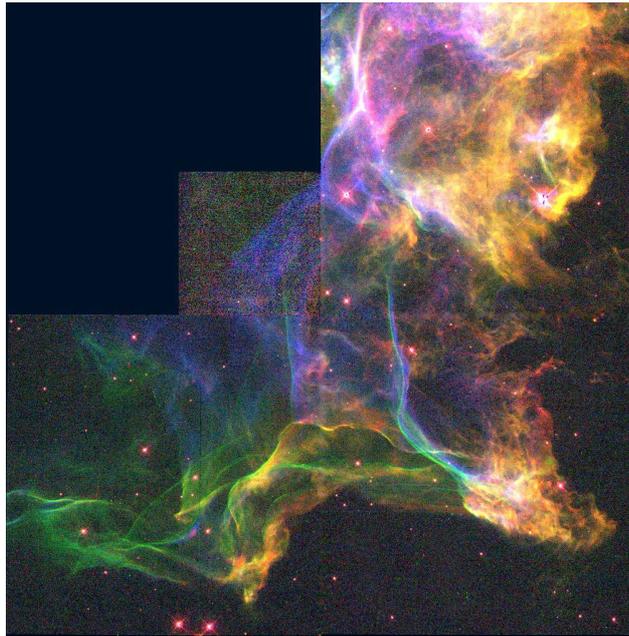


Figure 66: Part of the Cygnus Loop with HST. Note the thin filaments which are shocked interstellar gas which have cooled from $\sim 10^6$ K to $\sim (1-2) \times 10^4$ K.

Using the expression for the shock velocity and the cooling function one can estimate the cooling time of the gas. One then finds that cooling sets in at $\sim 10^4$ years. This has the character of a thermal instability. Because $\Lambda(T_e) \propto T_e^{-0.90}$ for $T_e \lesssim 10^7$ K (see Eq. (27.5)), the cooling accelerates as the temperature drops. The gas therefore cools from $\sim 10^6$ K to $\sim (1-2) \times 10^4$ K very rapidly.

During the adiabatic stage most of the emission is in X-rays. In the radiative phase most of the emission from the shell now emerges as optical and UV emission. Fig. 65 shows one of the most well-known SNRs in this category, the Cygnus Loop. The total size of the remnant is ~ 20 pc and the expansion velocity ~ 115 km s $^{-1}$. The thin filaments seen are likely to be a result of the thermal instability and the effects of magnetic fields.

29 Hot gas in clusters of galaxies

Clusters of galaxies contain 100–1000 galaxies within a radius of 1.5–3 Mpc, with a very large fraction elliptical galaxies. From the velocity dispersion 500 – 1500 km s $^{-1}$, one infers total masses of $10^{14} - 3 \times 10^{15} M_\odot$. This corresponds to a mass to light ratio of 200 – 300 M_\odot / L_\odot , and already Zwicky concluded that this was most likely a result of dark matter dominating the gravitational potential. Clusters of galaxies were very early identified as an important class of X-ray emitting objects, and it was also clear that this came from hot, diffuse gas in the cluster. The typical X-ray luminosities are $10^{43} - 10^{45}$ erg s $^{-1}$. Observations of this gas is important since it gives information about the cluster formation process, the degree of relaxation in the cluster and it can be used as a tracer of the gravitational potential of the cluster. In Fig. 66 we show a sample of X-ray images of clusters with different degrees of relaxation.

The hot gas in clusters of galaxies show many similarities with SN remnants. The gas is dominated by collisional ionization and most of the emission is coming out as X-rays. The typical temperature of the gas is $\sim 10^8$ K and the emission is therefore dominated by free-free radiation. Because $\Lambda(T_e) \propto T_e^{0.5}$ for this process the total luminosity is

$$L_X \approx V n_e^2 \Lambda \propto n_e^2 R^3 T_e^{1/2} \quad (29.1)$$

where $V \approx 4\pi R^3/3$ is the volume of the cluster and n_e the electron density.

The heating of the gas is provided by the motion of the galaxies, as well as the initial heating in connection to the formation of the cluster. The gas is therefore expected to be in virial equilibrium and the temperature

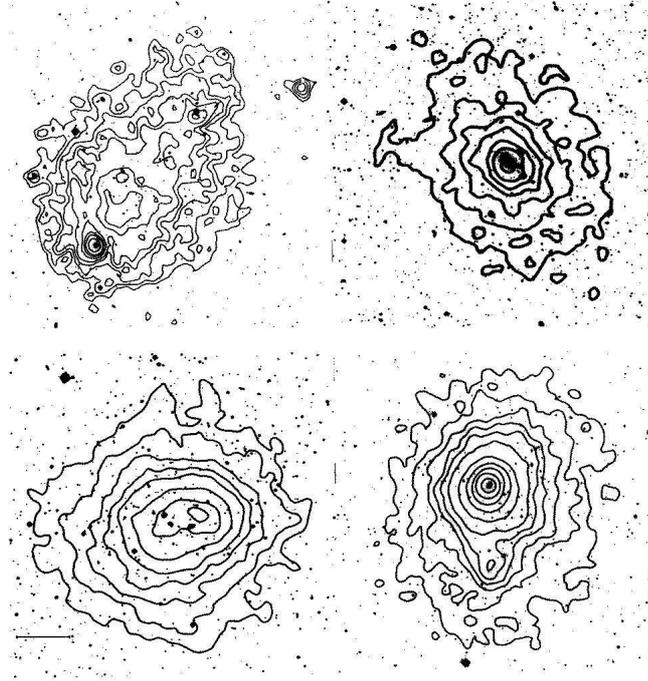


Figure 67: X-ray contours of four clusters with different morphology. The upper left is highly irregular, indicating a non-relaxed state, while the others are more relaxed (Ref?).

therefore

$$kT_e \approx \sigma^2 \approx \frac{GM_{tot}}{R} \quad (29.2)$$

where σ is the velocity dispersion in the cluster, as obtained from measurements of the dispersion of the spectra of the cluster members. Using this we find that

$$n_e \approx \frac{M_{gas}}{V} \propto \frac{M_{tot}}{R^3} \propto \frac{T_e}{R^2} \quad (29.3)$$

If we use this in Eq. (29.1) we obtain

$$L_X \propto n_e^2 (T_e/n_e)^{3/2} T_e^{1/2} = n_e^{1/2} T_e^2 \quad (29.4)$$

Based on this simple scaling we therefore expect a relation between the cluster luminosity and the density and temperature of the gas. The density can be estimated from the luminosity and temperature, using Eq. (29.1), while the temperature can be obtained from the X-ray spectrum. In Fig.

67 we show the observed temperature against the X-ray luminosity together with the above relation.

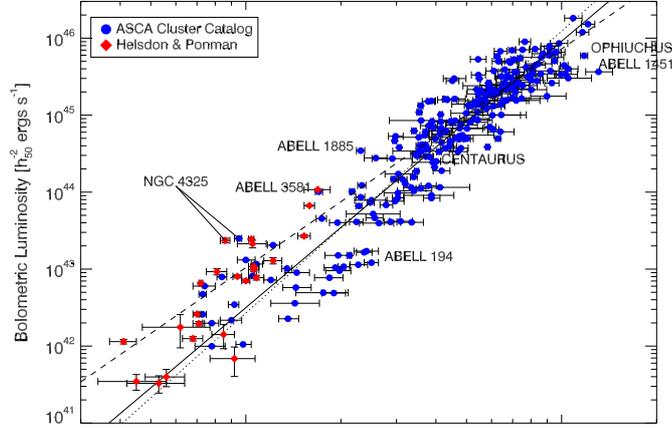


Figure 68: The X-ray luminosity as function of the temperature of the gas for a sample of clusters (Ref?).

A more detailed analysis fits the surface brightness $S(R)$, i.e., the emissivity integrated along the line of sight

$$S(R) = 2 \int_0^\infty n_e(l)^2 \Lambda[T(l)] dl = 2 \int_R^\infty \frac{n_e(r)^2 r \Lambda[T(r)]}{(r^2 - R^2)^{1/2}} dr \quad (29.5)$$

As a simple parameterization of the density profile we can use

$$n_e(r) = n_0 \left[1 + (r/R_c)^2 \right]^{-3\beta/2} \quad (29.6)$$

where R_c is the core radius, which is a fitting parameter. β gives the radial slope of the distribution. Typical values for the core density is $\sim 2.5 \times 10^{-3} \text{ cm}^{-3}$, $\beta = 0.7$ and $R_c = 0.15 - 0.4 \text{ Mpc}$. Using

$$\Lambda[T] \approx 2.4 \times 10^{-27} T^{1/2} \quad \text{for } T_e > 2 \times 10^7 \text{ K} \quad (29.7)$$

for free-free emission one can then from the surface brightness obtain the parameters n_0 , β and R_c if $T(r)$ is known. The latter can be obtained from the spectrum, since $dE/d\nu dt dV \propto n_e n_i T_e^{-0.5} e^{-h\nu/kT_e}$ (Eq. (27.1)). In this way we can determine the density profile of the cluster gas.

This is very useful since we can now relate this to the mass distribution in the cluster. If the cluster is relaxed we expect it to be in hydrostatic equilibrium. If we assume that the cluster is spherically symmetric we then have

$$\frac{dp(r)}{dr} = -\frac{GM(r)\rho(r)}{r^2} \quad (29.8)$$

where $p(r)$ is the pressure and $M(r)$ is the mass inside radius r . For the pressure we have the usual relation

$$p = \frac{k}{m_p \mu} \rho T \quad (29.9)$$

where μ is the mean molecular weight. If we differentiate this we obtain

$$\frac{dp}{dr} = \frac{k}{m_p \mu} \left(T \frac{d\rho}{dr} + \rho \frac{dT}{dr} \right), \quad (29.10)$$

which can be replaced for the left hand side in Eq. (29.8) to obtain

$$M(r) = -\frac{kTr}{Gm_p\mu} \left(\frac{d \ln \rho}{d \ln r} + \frac{d \ln T}{d \ln r} \right) \quad (29.11)$$

Because we can determine $T(r)$ and $n_e(r)$ from the spectrum and surface brightness mass distribution we can now derive the mass profile $M(r)$. Note that this is the total gravitational mass, not only the mass in the gas. It is therefore a very useful probe of the dark matter distribution in the cluster.

The analysis of the cluster Abell 478 illustrates this procedure. In Fig. 68 we show the X-ray image of the cluster, where the contours show the surface brightness levels. In Fig. 69 the hardness ratio, i.e. the flux ratio between 2.0-5.0 keV and 0.2-2.0 keV, is shown. This ratio is useful since it provides a measure of the temperature of the gas. These observations are used to obtain the surface brightness, electron density, temperature and pressure as function of the radius from the center shown in Fig. 70. This can then be used in Eq. (29.11) to obtain the mass profile in Fig. 71. The total mass of the cluster is $\sim 10^{15} M_\odot$, typical for a massive cluster. The mass is however much

The cooling time of the cluster gas is

$$t_{cool} \approx \frac{3kT_e}{n_e \Lambda(T_e)} \quad (29.12)$$

For free-free emission this gives

$$t_{cool} \approx 8.5 \times 10^{10} \left(\frac{n}{10^{-3} \text{cm}^{-3}} \right)^{-1} \left(\frac{T}{10^8 \text{K}} \right)^{1/2} \text{ years} \quad (29.13)$$

For high densities, such as those encountered in the core, the cooling time may be much less than the Hubble time. Unless the gas is heated by some process it is therefore expected to cool in the core. As with the radiative SN

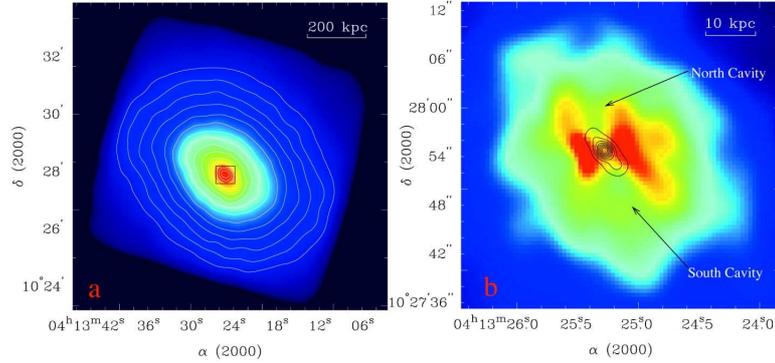


Figure 69: Chandra 0.5-5 keV X-ray image showing the surface brightness of the cluster Abell 478 from Chandra? The right panel shows the central region (Sun et al 2003).

remnants this will lead to a thermal instability until the gas temperature is $\sim 10^4$ K, or even less. The pressure will at the same time decrease and a flow of gas towards the center may result. This is known as a cooling cluster. In the central region the cool gas may give rise to optical line emission from the gas, and to star formation in center.

This is indeed seen in some clusters, like the Perseus cluster. The frequency is, however, much lower than expected which has been a problem. Different explanations have been proposed to solve this problem, in particular heating by a central AGN or heat conduction may possibly prevent the instability. This is, however, not yet solved.

29.1 The Sunyaev-Zeldovich effect

$$\frac{\Delta\nu}{\nu} = \frac{kT_e}{m_e c^2} \quad (29.14)$$

$$y = \int \frac{kT_e}{m_e c^2} n_e \sigma_T dl \quad (29.15)$$

30 The Fermi mechanism

Let us first consider the effect of an elastic scattering of a particle on a moving object. Suppose also that this scattering is elastic. An example is the scattering of an electron against a magnetic irregularity in the interstellar medium.

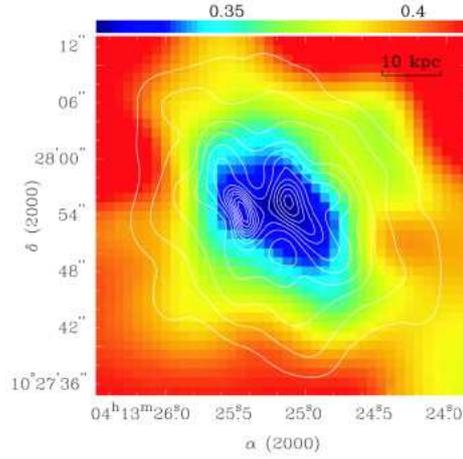


Figure 70: The hardness ratio giving a measure of the temperature of the gas. (Sun et al 2003).

To see the result of this we first consider the transformation of the momentum and energy from the rest system K to that of the moving object, K'. The momentum in the frame K' is then given by

$$p'_{x1} = \gamma(p_{x1} + \frac{V}{c^2}E_1) \quad (30.1)$$

where p_{x1} is the momentum before the collision along the x-axis. In terms of the total momentum $p_{x1} = p_1 \cos \theta$, where θ is the angle of the velocity with the x-axis. Similarly, the energy in frame K' is

$$E'_{x1} = \gamma(Vp_{x1} + E_1) \quad (30.2)$$

Suppose now that the scattering is elastic in the K' frame, i.e., $E'_{x2} = E'_{x1}$ and $p'_{x2} = -p'_{x1}$. In the rest frame the energy after the scattering will therefore be

$$E_2 = \gamma(-Vp'_{x2} + E'_2) = \gamma(Vp'_{x1} + E'_1) \quad (30.3)$$

Inserting the relation Eq. (30) between the energy and momentum in K' and K we get

$$E_2 = \gamma(V\gamma(p_{x1} + \frac{V}{c^2}E_1) + \gamma(Vp_{x1} + E_1)) \quad (30.4)$$

or

$$E_2 = \gamma^2 E_1 [1 + 2\frac{Vp_{x1}}{E_1} + \frac{V^2}{c^2}] \quad (30.5)$$

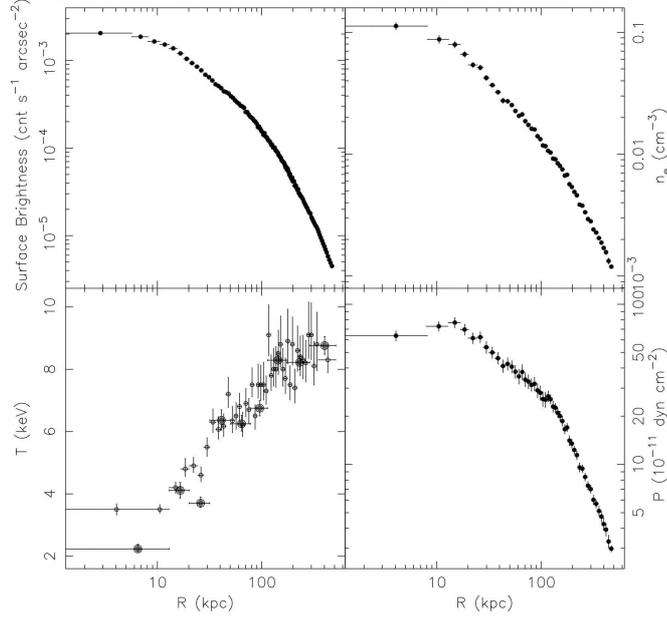


Figure 71: The surface brightness, electron density, temperature and pressure as function of the radius from the center(Sun et al 2003).

Suppose the particle velocity has an angle θ against the scatterer ('wall'). Because $p_{x1} = p_1 \cos \theta$ and $p_1 = E_1 v/c$ we get

$$E_2 = \gamma^2 E_1 \left[1 + 2 \frac{Vv \cos \theta}{c^2} + \frac{V^2}{c^2} \right] \quad (30.6)$$

Note here that v is the velocity of the particle and V that of the scatterer.

Let us now assume that the scatterer has a velocity much less than that of light, while the particle velocity may be close to this. In this case $\gamma \approx 1$ and we can neglect the last term in the bracket,

$$E_2 \approx E_1 \left[1 + 2 \frac{Vv \cos \theta}{c^2} + O(V/c)^2 \right] \quad (30.7)$$

The change in energy in one scattering is therefore

$$\frac{\Delta E}{E} \approx \frac{2Vv \cos \theta}{c^2} \quad (30.8)$$

$$F = n(v + V \cos \theta) \approx n(c + V \cos \theta) \quad (30.9)$$

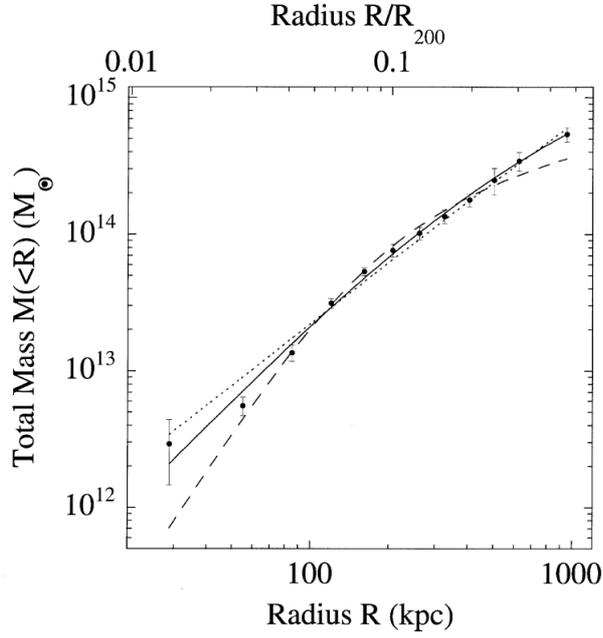


Figure 72: The derived mass profile of Abell 478. The solid line gives the mass profile predicted by the CDM model of Navarro, Frenk and White (Ref ?).

$$F\Delta E = n(c + V \cos \theta)E_1 \frac{2V \cos \theta}{c} \quad (30.10)$$

$$\Delta E_{net} = n[(c + V \cos \theta)E_1 2V \cos \theta - (c - V \cos \theta)E_1 2V \cos \theta] \quad (30.11)$$

$$\Delta E_{net} = 4n(V \cos \theta)^2 \quad (30.12)$$

$$\Delta E_{net} = \frac{\int_0^{\pi/2} 4\frac{n}{c}(V \cos \theta)^2 d \cos \theta}{\int_0^{\pi/2} d \cos \theta} = \frac{4}{3}nc \left(\frac{V}{c}\right)^2 \quad (30.13)$$

$$E' = \gamma(Vp_x + E) \quad (30.14)$$

$$E' \approx \frac{V}{c} \cos \theta E + E \quad (30.15)$$

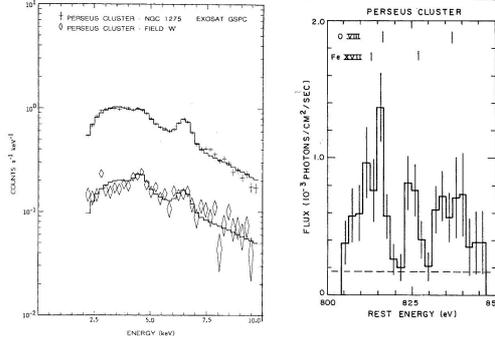


Figure 73: Part of the Cygnus Loop with HST. Note the thin filaments which are shocked interstellar gas cooling to $\sim (1 - 2) \times 10^4$ K.

$$\Delta E = \frac{V}{c} \cos \theta E \quad (30.16)$$

$$E(k) = \left(1 + \frac{\Delta E}{E}\right) E(k-1) = \dots = \left(1 + \frac{V_s}{c}\right)^k E_0 \quad (30.17)$$

$$N(k) = N_0 P^k \quad (30.18)$$

$$\ln[N(E)/N_0] = k \ln P = \frac{\ln P \ln[E/E_0]}{\ln(1 + \frac{V_s}{c})} \quad (30.19)$$

$$\frac{N(E)}{N_0} = \left[\frac{E}{E_0}\right]^{\frac{\ln P}{\ln(1 + \frac{V_s}{c})}} \quad (30.20)$$

$$\frac{dN}{dE} = \left[\frac{E}{E_0}\right]^{\frac{\ln P}{\ln(1 + \frac{V_s}{c})} - 1} \quad (30.21)$$

$$dn = \cos \theta d\Omega = 2 \cos \theta d \cos \theta \quad (30.22)$$

$$\Delta E = \frac{V}{c} \cos \theta E \quad (30.23)$$

$$\left\langle \frac{\Delta E}{E} \right\rangle = 2 \int_0^{\pi/2} \frac{V}{c} \cos \theta \cos \theta d \cos \theta = \frac{2V}{3c} \quad (30.24)$$

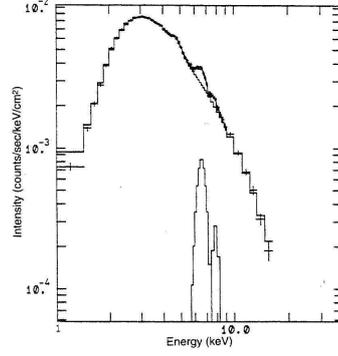


Figure 74: Part of the Cygnus Loop with HST. Note the thin filaments which are shocked interstellar gas cooling to $\sim (1 - 2) \times 10^4$ K.

$$\left\langle \frac{\Delta E}{E} \right\rangle = \frac{4V}{3c} = V_s \quad (30.25)$$

$$\frac{dn}{dt} = Nc \frac{1}{2} \int_0^{\pi/2} \cos \theta d \cos \theta = \frac{Nc}{4} \quad (30.26)$$

$$\frac{dn}{dt} = NV = \frac{NV_s}{4} \quad (30.27)$$

$$P_{esc} = \frac{V_s}{c} \quad (30.28)$$

$$\frac{dN}{dE} = CE^{\frac{\ln P}{\ln(1 + \frac{V_s}{c})} - 1} \quad (30.29)$$

$$\frac{\ln P}{\ln(1 + \frac{V_s}{c})} - 1 = \frac{\ln(1 - \frac{V_s}{c})}{\ln(1 + \frac{V_s}{c})} - 1 \approx -1 - 1 = -2 \quad (30.30)$$

$$\frac{dN}{dE} = CE^{-2} \quad (30.31)$$

$$\frac{dE}{dt} = -p \frac{dV}{dt} = (\gamma - 1)E \nabla \cdot \mathbf{v} \quad (30.32)$$

$$v(r) = v_0 \frac{r}{r_0} \quad (30.33)$$

$$\nabla \cdot \mathbf{v} = \frac{1}{r^2} \frac{\partial}{\partial r}(r^2 \mathbf{v}) = \frac{v_0}{r_0} \frac{1}{r^2} \frac{\partial}{\partial r}(r^3) = 3 \frac{v_0}{r_0} = 3 \frac{v}{r} \quad (30.34)$$

$$\frac{dE}{dt} = 3(\gamma - 1)E \frac{v}{r} = 3(\gamma - 1) \frac{E}{t} = \frac{E}{t} \quad (30.35)$$

$$T = 5.6 \times 10^7 E_{51}^0 \cdot 4n^{-0.8} \left(\frac{t}{1000 \text{ years}} \right)^{-1.2} K \quad (30.36)$$

$$t_{cool} \approx \frac{3kT_e}{n_e \Lambda(T_e)} \quad (30.37)$$

$$\Lambda(T_e) = 8.0 \times 10^{-23} \left(\frac{T_e}{10^6 K} \right)^{-0.90} \quad (30.38)$$

$$n_2 = 4n_1 \equiv 4n \quad (30.39)$$

$$t_{cool} \approx 4.0 \times 10^4 n^{-1} \left(\frac{T}{10^6 K} \right)^{1.9} \text{ years} \quad (30.40)$$

$$t \approx t_{cool} \Rightarrow t_{cool} \approx 2 \times 10^4 E_{51}^{0.23} n^{-0.77} \text{ years} \quad (30.41)$$

31 Gamma-ray Bursts

31.1 Historical overview

Gamma-ray bursts (GRB's) were discovered in 1967 as a result of the Cold War. At that time the results were classified, and it was not until 1973 the first results were published. This immediately inspired a large number of more or less exotic theories, and in connection to the Texas conference on Relativistic Astrophysics in 1974 Ruderman could summarize more theories than discovered bursts at that time. Among these were more or less exotic candidates, like mini-black holes, white holes, comets falling down on neutron stars, etc. It is, however, worth noting that supernovae were already in 1974 proposed as a candidate by Colgate.

With the Compton Gamma-ray Observatory in 1991 it was found that the distribution on the sky was highly isotropic, indicating a cosmological origin, although an extended halo population could not be completely ruled out (Fig. 74). CGRO also found that the bursts could roughly be divided

2704 BATSE Gamma-Ray Bursts

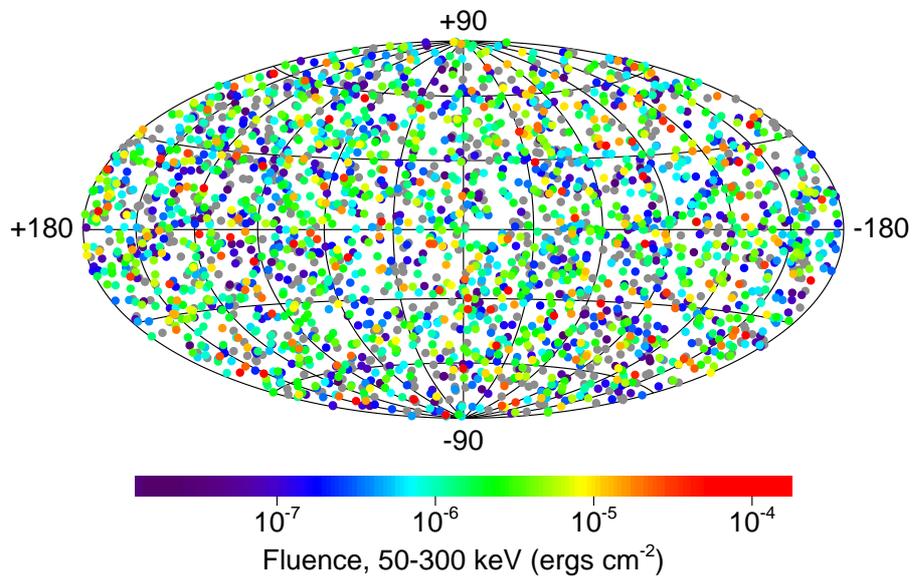


Figure 75: GRB distribution on the sky for bursts observed with BATSE

into short ($\Delta t \lesssim 2$ s) and long ($\Delta t \gtrsim 2$ s) bursts, although the distribution of these formed a continuum.

A problem with CGRO was the fact that the localization could only be done within $\sim 3^\circ$. An optical identification, which requires a position within arc minutes, was therefore impossible.

A breakthrough came with the Dutch-Italian satellite Beppo-Sax which had both a gamma-ray trigger and an X-ray telescope which could localize the X-ray emission in connection to the burst within a few arc minutes. This allowed in 1997 the first optical identification of an afterglow from GRB970228. Shortly afterwards, GRB970508 was identified, and, moreover, the spectrum showed absorption lines from intervening galaxies up to $z = 0.835$, once and for all demonstrating that the GRBs are at cosmological distances.

In 1998 the GRB 980425 was found to coincide with the Type Ic SN 1998bw in ESO 184-G82 at $z=0.0085$. This SN was highly unusual from several points of view. The radio emission was the strongest seen among all SNe (see Fig. ??). Also its optical luminosity was an order of magnitude higher than the typical Type Ic luminosity, indicating $M(^{56}\text{Ni}) \sim 0.7 M_\odot$, which is larger than e.g., SN 1987A by a factor of ten. Finally the spectrum indicated an expansion velocity of $\gtrsim 60,000$ km s $^{-1}$, which was probably only a lower limit. Modeling of the radio observations showed that these could be well fitted with a synchrotron-self absorption spectrum of a source expanding with a Lorentz factor $\Gamma \sim 2$. This gave rise to the notion hypernovae. Although many in especially the SN community saw this as the confirming evidence for the SN – GRB connection, which was proposed already 1974 by Colgate and others, this GRB was extremely weak compared to typical GRBs.

More evidence, however, came from the afterglow light curves which in several cases showed a clear bump in the light curves, which was interpreted as a SN signature. Complete confirmation came with the identification of a SN 1998bw-like spectrum in the afterglow of the GRB 030329.

31.2 Summary of observations

31.2.1 Prompt phase

In Fig. 75 we show a sample of burst profiles detected by BATSE. It is obvious that both the length of the burst and its light curve shape differ greatly. Some, like Triggers 1406 and 2571, have sharply rising bursts and then a smooth decay, while others like Trigger 1606 have highly irregular

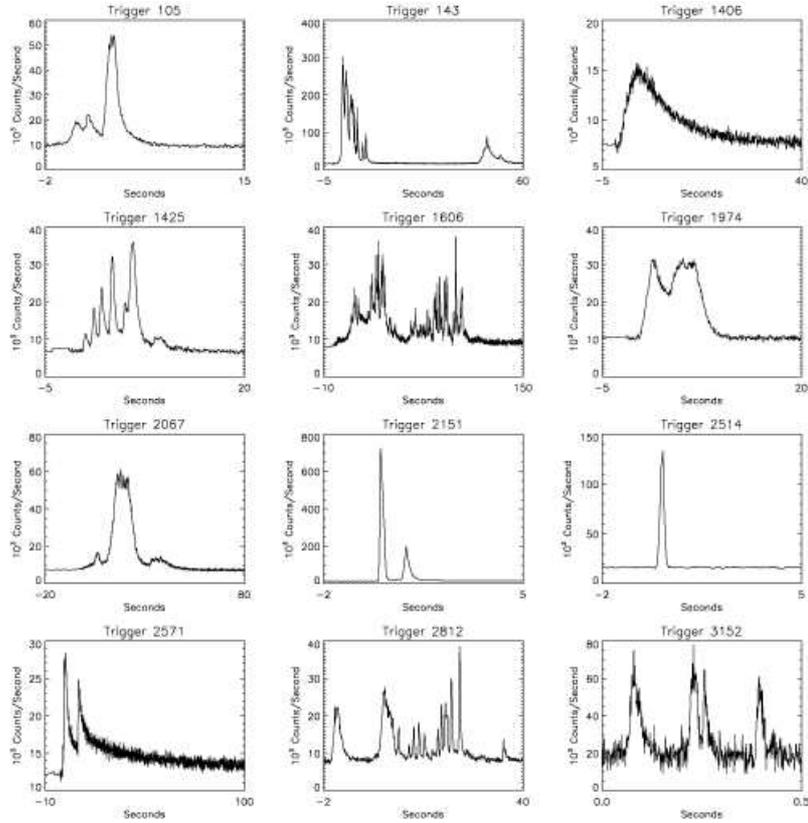


Figure 76: Examples of light curves observed with BATSE.

profiles.

The duration of the bursts vary from 10^{-2} s to 10^3 s. As is seen from Fig. 75, many long bursts show considerable substructure, with peaks with a duration of the order of milliseconds or even less. The distribution of the durations show a clearly bimodal structure, with one peak at ~ 0.2 s and one at ~ 30 s (Fig. 76). Because of this, one usually divides the burst into short $\lesssim 2$ s and long $\gtrsim 2$ s. The short account for $\sim 25\%$, but there may be selection effects which may increase the true fraction. This bimodal distribution has led some people to the suggestion that this represents two different physical mechanisms for the bursts. We are coming back to this later.

The spectra of the prompt emission can be described as two power laws with a break between, $dN(E)/dE \propto E^{-\alpha}$ where α is the photon number

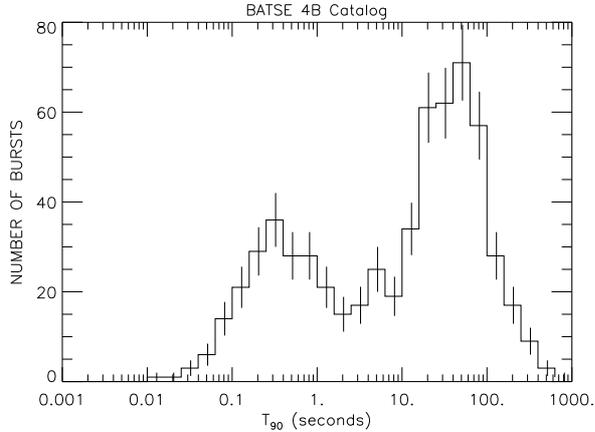


Figure 77: Distribution of durations of gamma-ray bursts detected by BATSE. The duration is defined as the time, T_{90} , between when 5% and 95% of the total number of counts are measured.

index. At energies less than the peak energy, E_p , $\alpha \approx 1 \pm 1$, while above E_p the bursts have a wide range of $\alpha \approx 1-4$. The fact that the $EF(E)$ spectrum shows a peak in the gamma-ray range, implies that for most GRBs most of the energy of the burst is really coming out as gamma-rays. Recently, many bursts have, however, been discovered which have their peak energy in the X-rays.

The range in E_p is very large ranging from MeVs down to tens of keVs. Unfortunately, both at high and low energies selection effects makes this highly uncertain.

There is a correlation between the peak energy and the duration of the burst, so that short burst in general have harder spectra than long bursts. This can be seen in the hardness ratio defined as the ratio of the fluency (time integrated flux) in the 100-300 keV channel divided by that in the 50 – 100 keV channel of the BATSE instrument (Fig. 77). Clearly, the short bursts have considerably harder spectra. The transition between the two groups is, however, smooth.

31.2.2 Afterglow phase

GRB 970508 was the first GRB to show an afterglow emission in the radio as well as in the optical. This allowed first of all an identification of the object at these wavelength, and secondly a very valuable diagnostic of the global properties of the GRB. This includes such parameters as the total

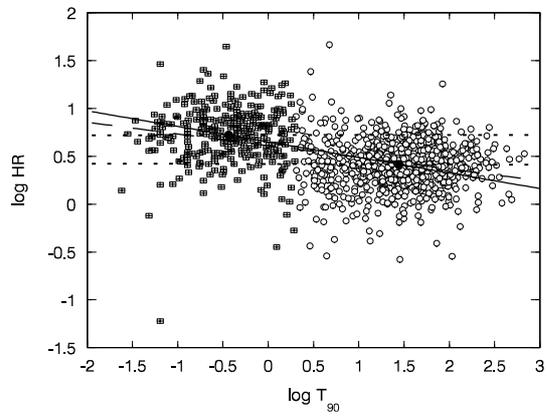


Fig. 1

Figure 78: Hardness – duration correlation of BATSE bursts. The HR is defined as the ratio of the fluency in the 100-300 keV channel divided by that in the 50 – 100 keV channel (from Qin et al. 1999)

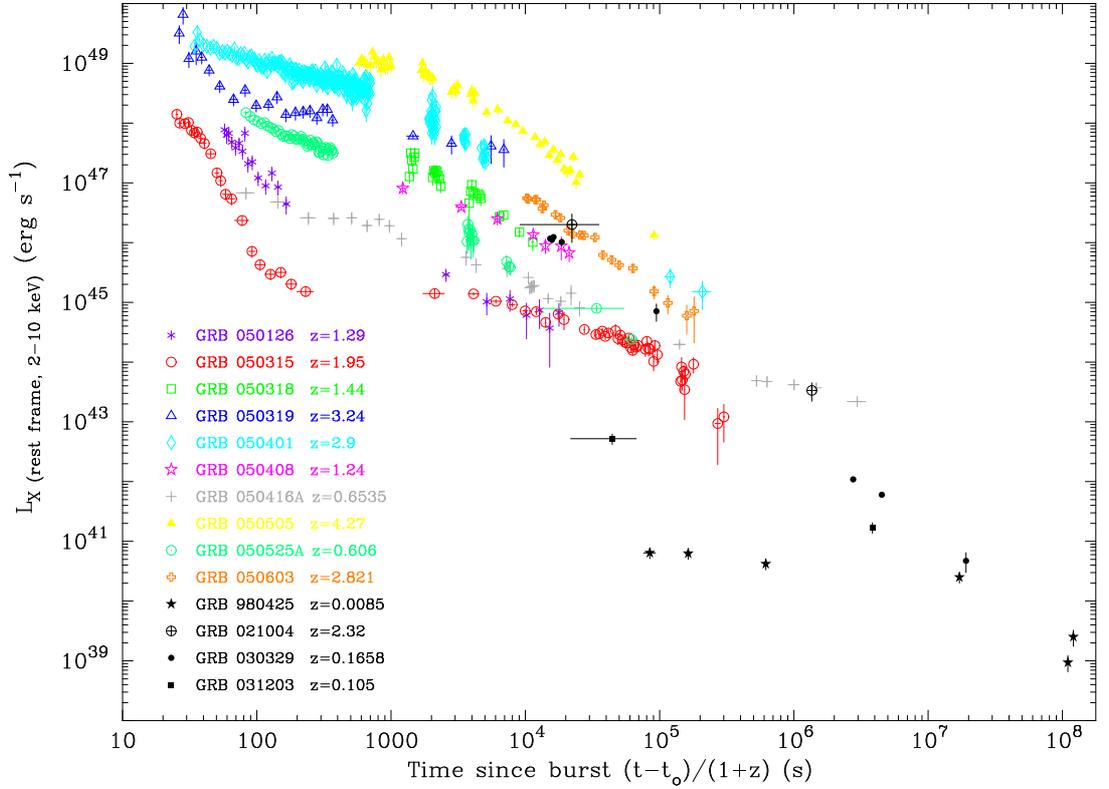


Figure 79: Sample of X-ray light curves observed with SWIFT. (Nousek et al. 2005)

energy, the Lorentz factor of the expanding matter, magnetic fields and particle energies, as well as information about the environment of the GRB. In particular, this has allowed a direct identification of the type of object which is responsible for the GRB. As examples of afterglow light curves in the X-ray range we show in Fig. 78 a sample of recently observed GRBs with SWIFT.

The afterglow phase is a direct relativistic version of the Sedov solution for a supernova remnant. In the same way as for these the dynamics reflects the properties of the radiative emission and spectrum. Before discussing this, we, however, consider a few very basic constraints which can be derived from observations of the prompt phase.

31.3 The necessity of relativistic expansion

First assume that the source is non-relativistic. The fluency is then given by $F = L\Delta t/4\pi D^2$, where Δt is the time scale of the burst, L the luminosity and D the luminosity distance to the GRB. If we assume that the source has a radius R , the energy density is $L/4\pi R^2 c$. As an estimate we take $R \sim \Delta t$. Further, we assume that a fraction f_p of the photons have energies above the pair creation threshold, $\sim 2m_e c^2$, the density of energetic photons is $n_\gamma = f_p L/4\pi R^2 m_e c^3$. The optical depth to pair production is therefore

$$\tau_{\gamma\gamma} = \sigma_T n_\gamma R = \frac{\sigma_T f_p L}{4\pi R m_e c^3} = \frac{\sigma_T f_p F D^2}{R \Delta t m_e c^3} = \frac{\sigma_T f_p F D^2}{(\Delta t c)^2 m_e c^2} \quad (31.1)$$

As a typical value for the fluency we take $F \sim 10^{-6} \text{ erg cm}^{-2}$ and $D \sim 2000 \text{ Mpc}$, corresponding to a total energy of $5 \times 10^{50} \text{ ergs}$. If we take $\Delta t \sim 0.01 \text{ s}$ we get

$$\tau_{\gamma\gamma} = 3 \times 10^{14} f_p \frac{F}{10^{-6} \text{ erg cm}^{-2}} \left(\frac{\Delta t}{0.01 \text{ s}} \right)^{-2} \quad (31.2)$$

Therefore for any reasonable values of f_p the source would be extremely optically thick to pair production and would therefore show a thermal spectrum, contrary to the observations.

This paradox is solved if the source is expanding relativistically with a large Lorentz factor, Γ . This has several consequences which help in the right direction.

First, if the source is moving towards us with a velocity v , the observed time interval, dt_{obs} between two photons emitted in an interval dt_{em} will be smaller by a factor $2\Gamma^2$. To see this we consider a photon emitted from the shell at a radius r_1 from the origin at a time $t_{1\ em}$ in the GRB frame, and at an angle θ . The time when it will arrive to the observer is therefore $t_{1\ obs} = t_{1\ em} + (D - r_1 \cos \theta)/c$. Now, let a second photon be emitted at a time $t_{1\ em} + dt_{em}$. The radius will now be $r_2 = r_1 + v dt_{em}$, and the time when it will be observed is therefore $t_{2\ obs} = t_{2\ em} + (D - r_2 \cos \theta)/c = [D - (r_1 + v dt_{em}) \cos \theta]/c$. The time interval it will be received in is therefore

$$dt_{obs} = dt_{em} - \beta dt_{em} \cos \theta = dt_{em}(1 - \beta \cos \theta) \quad (31.3)$$

where $\beta = v/c$.

Because $v \approx c$ it is more useful to write this in terms of the Lorentz factor. For this we note that $\Gamma^2 = 1/(1 - \beta^2) = 1/[(1 + \beta)(1 - \beta)] \approx 1/[2(1 - \beta)]$.

If we now assume that $\cos \theta \approx 1$, we can write Eq. (31.3) as

$$dt_{obs} = \frac{dt_{em}}{2\Gamma^2} \quad (31.4)$$

If the source is expanding with constant velocity the true size is therefore not ct_{obs} but $2\Gamma^2 ct_{obs}$.

Note that Eq. (31.4) is not the result of a Lorentz transformation, but is only a result of the fast expansion and the finite velocity of light.

Secondly, for a relativistically expanding source the radiation we receive will be blue shifted by a the Doppler effect which gives a factor Γ higher frequency. Therefore, the number of photons above the pair production threshold will decrease by a factor $\Gamma^{2\alpha}$.

Putting everything together, one gains a factor of $\Gamma^{2(1+\alpha)}$ from the relativistic motion. With $\alpha \sim 2$ this becomes $\sim \Gamma^6$. The Lorentz factors needed to have $\tau_{\gamma\gamma} \ll 1$ are therefore in the range $\Gamma \sim 100 - 1000$.

Further evidence of relativistic expansion comes from radio observations of interstellar scintillations in GRB light curves. An example of this is shown in Fig. 79 for GRB 970508. During the first ~ 50 days the radio flux showed large excursions, which later decreased, consistent with that expected for an expanding source. From the size of the plasma fluctuations the angular extent of the radio emission could be estimated, and one found that the source must have had a size of $\gtrsim 10^{17}$ cm, showing that the expansion was close to the velocity of light.

The fact that we need a highly relativistic expansion means that the mass involved in this must be very small, since

$$E \sim \Gamma M c^2 \quad (31.5)$$

which means that

$$M \approx 5 \times 10^{-6} \left(\frac{\Gamma}{10^3} \right)^{-1} \left(\frac{E}{10^{52} \text{ergs}} \right) M_{\odot} \quad (31.6)$$

This means that the explosion has to have a very small fraction of baryons to photons.

31.4 General scenario for the prompt and afterglow emission.

Most models for the prompt emission, as well as the afterglow, do not specify the way the explosion takes place. The only assumption is that a very

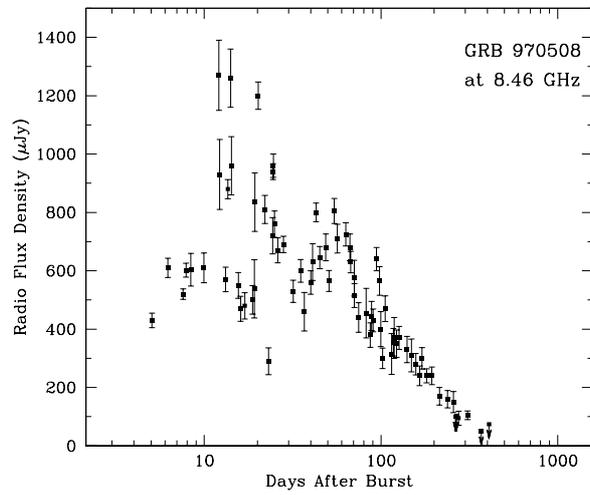


Figure 80: Radio light curve for GRB970508 at 8.46 GHz. Note the rapid fluctuations due to interstellar scintillations in the light curve during the first ~ 50 days (Frail 2003).

large amount of energy is released in either a spherical explosion or, as we will discuss below, in a narrow conical jet. This is both the strength and weakness of the model. On the one hand it is free of assumptions about this early not well understood phase. On the other hand, the model gives very little information about this crucial stage.

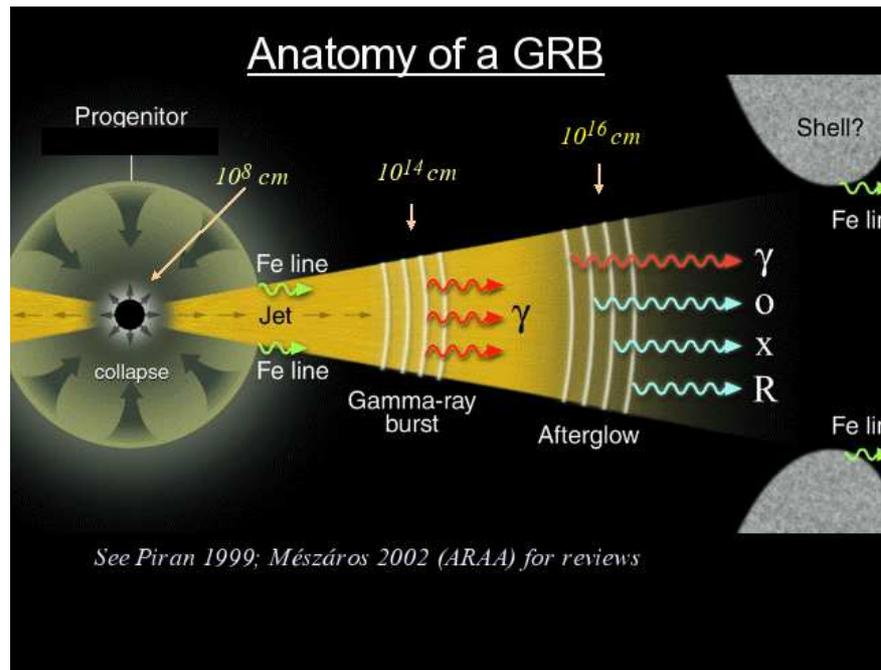


Figure 81: Schematic representations of the different stages in the evolution of a GRB.

In its simplest version one releases a large amount of energy in a medium and let this expand. Depending on the structure of the medium, like radial density variation, asphericity and rotation, etc, the explosion may either be spherical or, more realistically, confined to a jet. In the currently popular models the latter is usually assumed.

The prompt emission may be produced in two quite different alternatives (see Fig. 81). In one type of models it is a result of the interaction of the blast wave with the external medium, in a similar way to what happens in a supernova remnant. The properties of the emission is therefore sensitive to the details of the circumstellar medium, such as clumping and the presence of a stellar wind from the progenitor. This model is usually called the external shock model.

In the other type of models the prompt emission is caused by the fact that it is likely that the 'central engine' does not have a steady energy and momentum output as function of time. Instead, the Lorentz factor may e.g., vary with time depending on the details of the energy production mechanism. This variation will lead to a situation where you may have matter with a higher Lorentz factor ejected after that of lower. The more energetic ejection will therefore at some point catch up with the less energetic, and a shock wave will form in the outflow itself. The energy release is in this type of models therefore internal to the outflow and the models are therefore referred to as the internal shock model.

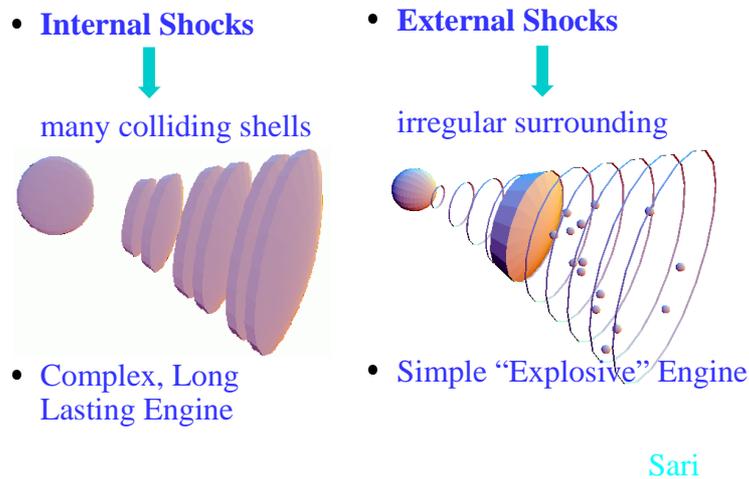


Figure 82: Schematic representations of the internal and external shock model for the prompt emission. (Sari).

An important clue to the the cause of the prompt phase is the rapid variations in intensity on a time scale of milli-seconds seen in most bursts. There are basically two possibilities to create this.

In the external shock model the variations are caused by encounters of the relativistic blast wave by a large number of clumps in the circumstellar medium. A problem for this model is, however, that the shock emission

resulting from this does not react fast enough to the clumping but instead smooths the variation with time. The internal shock model has no problems in this respect, since the variations in the radiative flux will directly reflect the variation in the outflow. For this reason it is the currently favored model for the prompt emission. After the central engine has been switched off and these internal variations have had time to set up a smooth outflow, one expects the internal energy release to switch off. At that point the interaction with the external medium takes over, and the GRB has entered the afterglow phase. The details of the model, in particular the reason for the variable Lorentz factor, are, however not at all clear.

The emission in the afterglow phase is better understood. In principle it is just a relativistic version of the blast wave for a supernova remnant. In the same way as this the kinetic energy of the outflow is converted into thermal energy behind the shock. One can then formulate shock conditions similar to the non-relativistic, which gives the relation between the conditions in front and behind the shock (for details see the Appendix). For an adiabatic expansion of the blast wave one can then find a similarity solution for the dynamics, similar to the Sedov solution, which describes the expansion of the blast wave in the surrounding medium. Compared to the Sedov solution this is complicated first by the relativistic effects from the Lorentz transformations and secondly by the difference in the observed time evolution and the time evolution in the frame of the GRB, as we discussed in the previous section. Without any details the result is

$$R \approx \left(\frac{9E}{2\pi n_1 m_u c} \right)^{1/8} t_{obs}^{1/4} \quad (31.7)$$

or

$$R \approx 4 \times 10^{17} \left(\frac{E}{10^{52} \text{ ergs}} \right)^{1/8} \left(\frac{n_1}{1 \text{ cm}^{-3}} \right)^{-1/8} \left(\frac{t_{obs}}{\text{days}} \right)^{1/4} \text{ cm}. \quad (31.8)$$

The Lorentz factor of the shocked gas behind the blast wave, with parameters considered to be typical for a GRB, is given by

$$\Gamma_2 \approx 4.4 \left(\frac{E}{10^{52} \text{ ergs}} \right)^{1/8} \left(\frac{n_1}{1 \text{ cm}^{-3}} \right)^{-1/8} \left(\frac{t_{obs}}{\text{days}} \right)^{-3/8} \quad (31.9)$$

The energy, as well as the density, can vary by large factors, and can in particular be considerably higher respectively lower. It is therefore conceivable that larger Lorentz factors than the indicated are at hand in some cases. Note, however, the fairly strong dependence on the observer time. Lorentz

factors of 100–1000 are therefore likely to be present during the first minutes and hours after the burst, as is needed from the discussion earlier.

The derivation of these relations are given in the Appendix for those interested.

31.5 Afterglow spectra

For synchrotron radiation an electron in a magnetic field, B , with energy $\gamma m_e c^2$ radiates a total power $P = 4/3 \sigma_T c B^2 \gamma^2 / 8\pi$, at a frequency

$$\nu_0 = \frac{eB\gamma^2}{2\pi m_e c} \equiv \nu_B \gamma^2 B, \quad (31.10)$$

where $\nu_B = 4.2 \times 10^6$ Hz. The spectral distribution can be approximated by $P(x) \approx x^{1/3} \exp(-x)$ where $x = \nu/\nu_0$ (e.g., Rybicki & Lightman). This is the frequency in the comoving frame. The frequency in the observer frame is $\nu_{obs} = \Gamma_2 \nu_0$. Both observationally and theoretically there are strong reasons for assuming that the non-thermal electron spectrum is given by a power law, $dn(\gamma)/d\gamma \propto \gamma^{-p}$, where $p \sim 2$.

The spectrum of the afterglow is characterized by a number of power law segments, separated by several breaks. The frequencies of these breaks correspond to the minimum energy of the electron distribution, $\nu_{min} = \nu_B \gamma_{min}^2 B$, the energy where synchrotron cooling becomes important, $\nu_c = \nu_B \gamma_c^2$, and the frequency where synchrotron self-absorption becomes important ν_{SSA} .

Let us for a moment ignore the synchrotron self-absorption, and also assume that $\nu_{min} < \nu_c$. Below ν_{min} the radiation will be dominated by electrons close to γ_{min} . The spectrum from these will be the same as a mono-energetic spectrum with $\nu \ll \nu_0 = \nu_{min}$. i.e., $F_\nu \propto P(\nu/\nu_{min}) \propto (\nu/\nu_{min})^{1/3}$. Between ν_{min} and ν_c the spectrum will be $F_\nu \propto (\nu/\nu_{min})^{-(p-1)/2}$, the usual synchrotron spectrum. Finally, for $\nu > \nu_c$ cooling is important, and $F_\nu \propto (\nu/\nu_c)^{-p/2}$.

Let us now include synchrotron self-absorption. We then have to consider two cases. Let us first assume that $\nu_{min} < \nu_{SSA}$. For $\nu_{min} < \nu < \nu_{SSA}$ the spectrum will then be $F_\nu \propto (\nu/\nu_c)^{5/2}$, characteristic of an optically thick source. For $\nu < \nu_{min}$ the spectrum will, however, be somewhat flatter $F_\nu \propto (\nu/\nu_{min})^2$. The reason for this can be understood if we write the spectrum in the optically thick Rayleigh-Wien limit as $F_\nu = 2\nu^2 E_{mean}/c^2$, where E_{mean} is the mean energy of the radiating particles. For a thermal distribution $E_{mean} \sim kT$, while for a non-thermal $E_{mean} \propto \gamma_{mean} \propto (\nu_{mean}/B)^{1/2}$. In this case $F_\nu \propto \nu^2 \gamma_{mean} \propto \nu^{5/2}/B^{1/2}$. However, if $\nu < \nu_{min}$, then the electrons at γ_{min} are doing most of the absorption and emission, so that

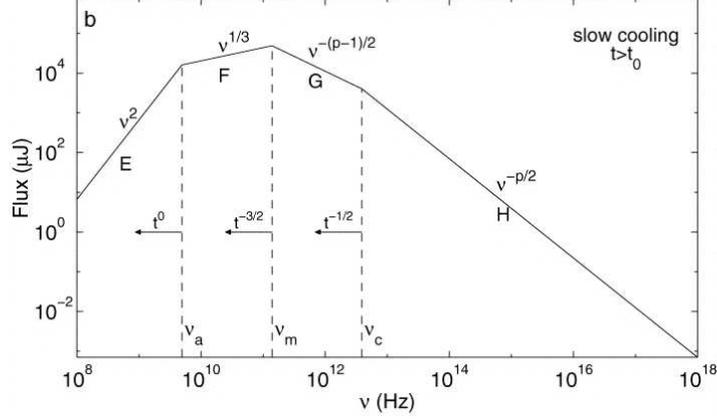


Figure 83: Schematic afterglow spectrum (Sari et al 1998).

$\gamma_{mean} = \gamma_{min}$ and $F_\nu \propto \nu^2 \gamma_{min} \propto (\nu_{min}/B)^{1/2} \nu^2$. In Fig. 82 we summarize the different sections of the spectrum.

To demonstrate this with real observations, we show in Fig. 83 a fit to the broad band spectrum of GRB970508, which is one of the best examples of this kind of fit. From the fit, in combination with the time evolution, values of $n_1, E, \varepsilon_e, \varepsilon_B, \gamma_{min}, p$ can be derived. The value of p is obtained directly from the spectral slope and typical values are $p \approx 2.1$, which is in good accordance with theoretical expectations. The other parameters are more uncertain and model dependent.

From the fit, one can determine the frequencies of the spectral breaks ν_{SSA}, ν_{min} and ν_c , as well as the flux at the peak, $F_\nu(\nu_{min})$. The value of p is obtained directly from the spectral slope, and typical values are $p \approx 2.2$, which is in good accordance with theoretical expectations. The remaining parameters characterizing the blast wave $n_1, E, \varepsilon_e, \varepsilon_B$ can then be derived. Assuming a constant external density Panaitescu & Kumar get $n_1 = 0.1 - 30 \text{ cm}^{-3}$, $E \sim (1 - 5) \times 10^{50} \text{ ergs}$, $\varepsilon_e \approx 0.1$, $\varepsilon_B \approx 10^{-4} - 0.1$. The energy is corrected for beaming, as will be discussed below. These numbers should be taken with caution, since they depend on uncertain observations, as well as questionable assumptions.

Note that we have until now not made any assumptions about the time evolution of the remnant. This, however, enters in the time evolution of the frequencies of the spectral breaks, and therefore depend on the density profile of the environment and whether the blast wave is adiabatic or not.

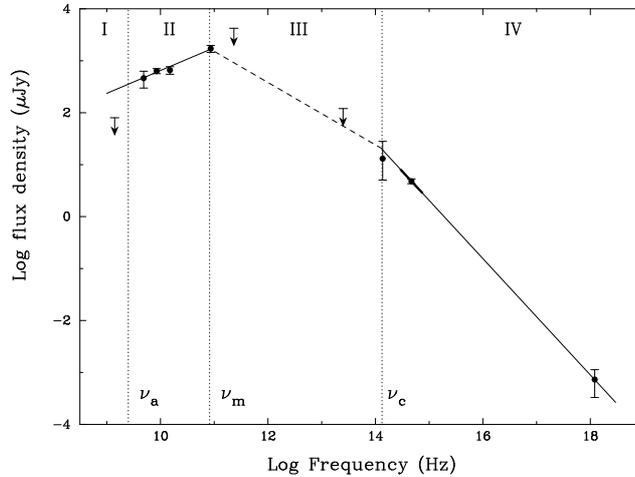


Figure 84: Broad band spectrum of the burst GRB970508 12.1 days after the burst, together with a model fit. (Galama et al. 1998)

31.6 Jet steepening

The fact that the energy, assuming isotropic emission, is so enormous has led to the suggestion that the relativistic outflow occurs in two narrow jets, in analogy with e.g., jets from compact radio galaxies. This is also motivated from hydrodynamical models for the GRB, as will be discussed later. Because of relativistic aberration, the radiation emitted at an angle θ' relative to its velocity in the rest frame, will be seen to be emitted at an angle θ given by

$$\cos \theta = \frac{\cos \theta' + \beta}{1 + \beta \cos \theta'} \quad (31.11)$$

(e.g., Rybicki & Lightman §4.1.3, or the similar effect for synchrotron emission in the electron and observer frames). Consider a light ray emitted at $\theta' = 90^\circ$. Then $\cos \theta = \beta$. Using $1 - \beta \approx 1/2\Gamma^2$, and $\cos \theta \approx 1 - \theta^2/2$, we find $\theta \approx 1/\Gamma$ for $\Gamma \gg 1$.

Therefore, for relativistic velocities the radiation is seen only within an angle $\theta \sim 1/\Gamma$. As long as the jet opening angle is larger than this, there is no difference between a spherical shell and a jet. However, as the jet is slowing down there will be a point when this condition is no longer true. Because only part of the emitting cone will now be filled this leads to a steepening of the light curve of the afterglow. Therefore, if one can determine when this occurs one can from the afterglow model estimate the value of $\theta \sim 1/\Gamma$ (see Eq. (31.9)), if one from the spectral modeling has determined the other

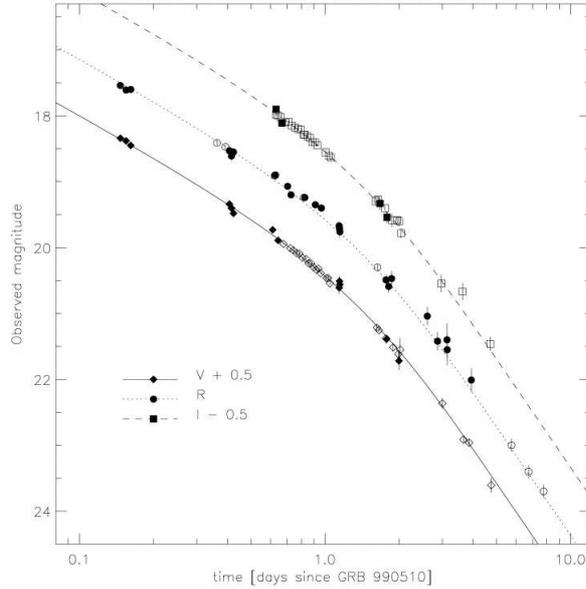


Figure 85: Optical light curves of GRB 990510 in the V, R and I bands (Harrison et al. 1999).

parameters. Knowing θ , one can now determine the correct total energy as $E = 2\theta^2 E_{iso}/4\pi$ (for a two-sided jet), where E_{iso} is the energy assuming an isotropic shell.

In Fig. 84 we show the optical light curves of GRB 990510. At ~ 1 day there is a clear steepening in the light curves of all colors. This has now been done for a number of afterglows, and in Fig. 85 we show a distribution of isotropic and corrected total energies. Typical jet angles are $\sim 10^\circ$. It is here seen that while the isotropic energies are in the range $10^{52} - 10^{54}$ ergs, the beam-corrected energies are $\sim 5 \times 10^{50}$ ergs, with a small dispersion. This has led to the suggestion that one can use GRBs as standard candles in the same way as TypeIa supernovae. Because of many systematic effects this is in my view optimistic.

32 GRB Progenitors

Up to now we have made no assumptions about the nature of the exploding object, but only assumed an instantaneous injection of a large amount of energy with a large $E/M_0 c^2$. To explain the large energies involved $\sim 10^{51} -$

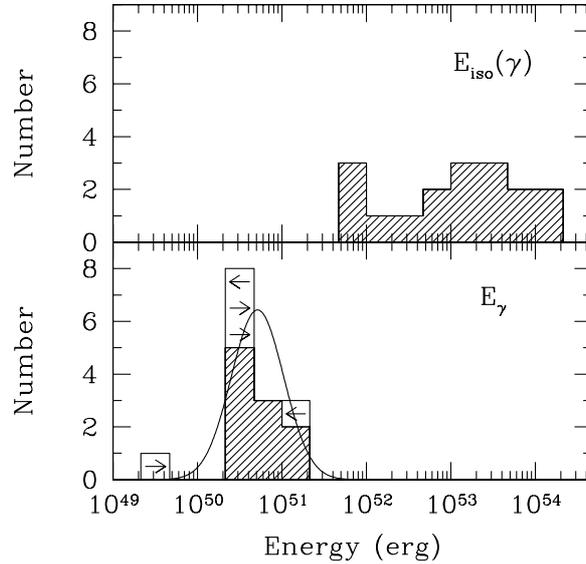


Figure 86: Distribution of isotropic end corrected energies (Frail et al. 2001).

10^{52} ergs it is, however, obvious that the formation of some kind of compact object is involved. This can either be a neutron star or a black hole. There are then two main classes of scenarios, which have quite different progenitors. The physics involved in the generation of the energy may, however, be fairly similar. We will now discuss these one by one.

32.1 The supernova - GRB connection

Supernovae have from the theoretical point of view for a long time been proposed as an origin for GRBs. When the first afterglows were identified, it was also noted that these in most cases were in the central regions of star forming galaxies, typical of massive stars. Direct evidence for a connection between these was, however, lacking. This changed when in April 1998 the error box of GRB 980425 was found to coincide with the supernova SN 1998bw in the galaxy ESO 184-G82 with a very low redshift, 2550 km s^{-1} or $z=0.0085$. The supernova which was a Type Ic SN, was very remarkable from several points of view. The radio emission from the supernova was found to be more luminous than any other radio SN, and was well fitted by a synchrotron self-absorption spectrum. From modeling of the radio emission the expansion velocity of the emitting material was found to have

a Lorentz factor of $\Gamma \sim 2$. Also the optical spectrum showed very broad, smooth features indicating an expansion velocity of at least $60,000 \text{ km s}^{-1}$. The luminosity of the SN was nearly a factor of ten larger than a typical Type Ic SN, and close to that of Type Ia's. The light curve indicated a total ^{56}Ni mass of $\sim 0.5 M_{\odot}$, much higher than that in e.g., SN 1987A. The gamma-ray luminosity was, however, about four orders of magnitude less than a typical GRB, $\sim 5 \times 10^{47}$ ergs. This has led to some doubt about the GRB-SN connection in this case. The coincidence of the SN and GRB as well as the remarkable properties of the SN, makes it in my view, however, completely clear that the GRB and the SN really originated from the same object.

In addition to this direct evidence there has for a number of GRBs been seen evidence for a bump in the light curve of the afterglow. While the early evolution in most cases follows a power law, there has been several examples where a red bump has been seen in the light curve at ~ 20 days (Fig. 86). The luminosity of these bumps as well as the shape and color are roughly consistent with that of SN 1998bw, indicating that it really is the emission from the SN which is seen.

Besides SN 1998bw, the most direct evidence for the SN/GRB connection came from GRB 030329. This was by GRB standards an extremely nearby GRB with $z = 0.168$. As was immediately recognized by several groups, this was a unique opportunity of getting high S/N spectra of the afterglow during the first months. While the first spectra showed basically a power law spectrum with $F_{\nu} \propto \nu^{-1.2}$, there was after ~ 8 days a clear excess emission above a power law fit (Fig. 87). This component became increasingly stronger, and when the power law spectrum seen during the first days was subtracted it was found that this coincided almost perfectly with that of SN 1998bw. The supernova consequently got the designation SN 2003dh. In addition to this there has been several other GRBs where there is strong evidence for an underlying supernova. With GRB980425, GRB030329 and these other cases, the SN/GRB connection is now firmly established,

Note, however, that the optically identified GRBs all belong to the long GRBs. Because there is some evidence from the distribution of the durations that there may be two different populations of progenitors, it is fair to say that the SN-GRB connection is only established for the long bursts. The short could have a different class of progenitors.

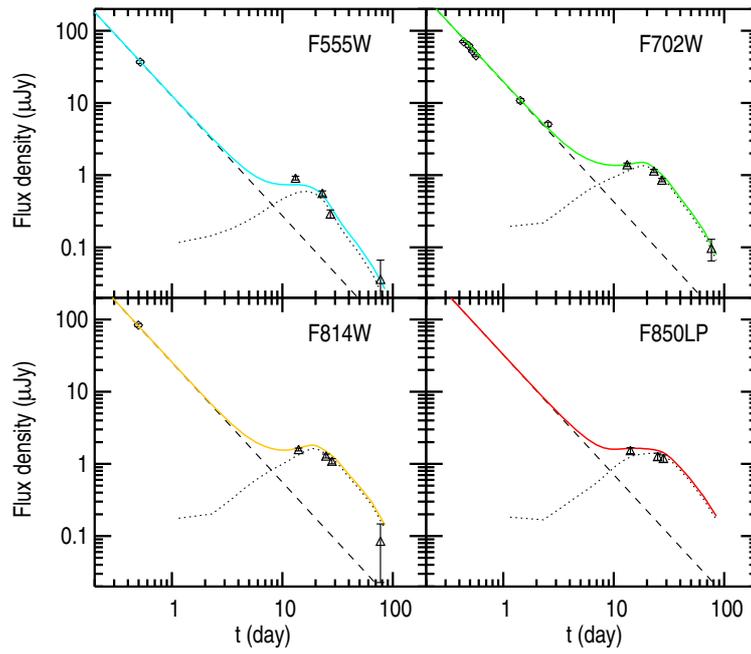


Figure 87: Light curves of the afterglow of GRB011121 obtained with HST (triangles) and ground based telescopes (diamonds). Note the bump in the light curve at 10-30 days, consistent with that from a of SN 1998bw, dimmed by $\sim 55\%$ (Bloom et al 2002).

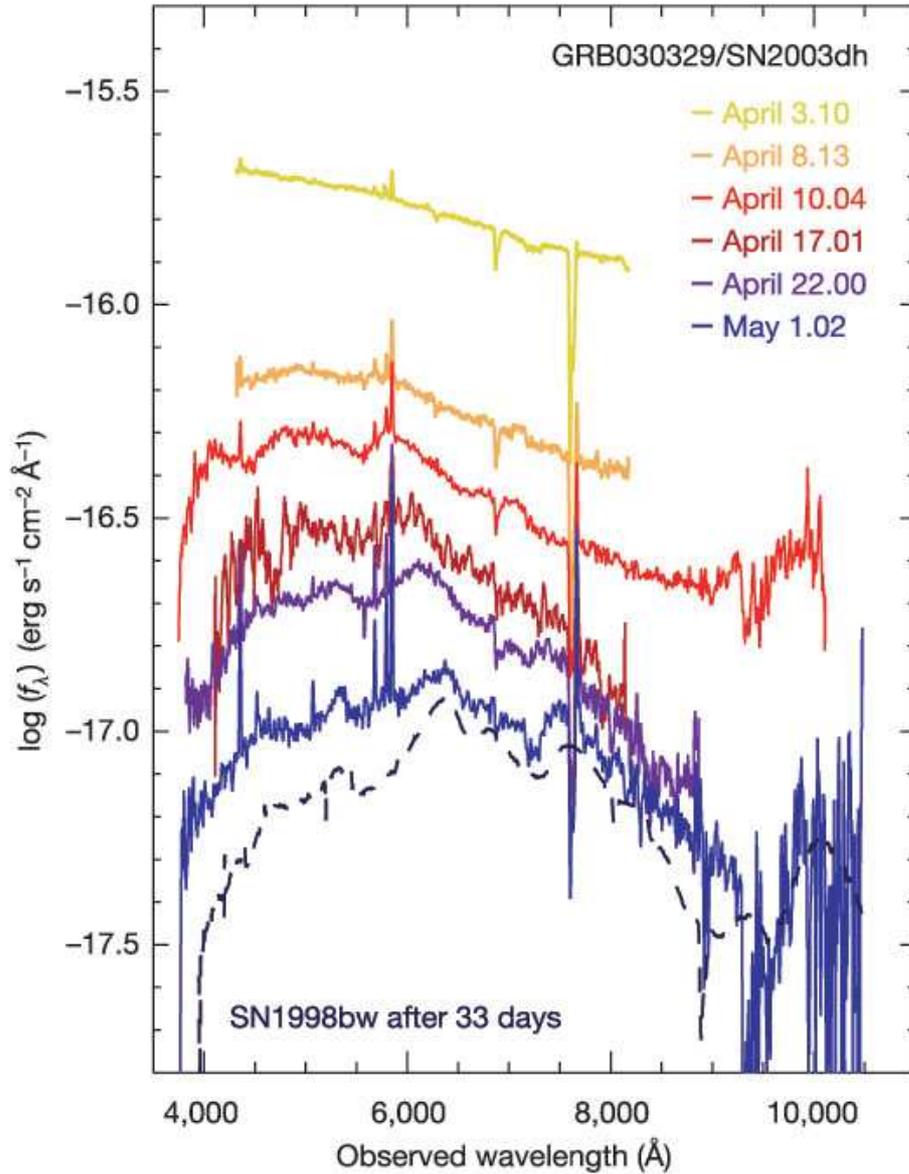


Figure 88: Spectral sequence of GRB030329/SN2003dh with VLT (Hjorth et al 2003). Note the power law spectrum on April 3 and the gradually stronger supernova contribution. The dashed line shows the spectrum of SN 1998bw at an age of 33 days.

32.2 The GRB environment

32.3 Collapsars

The collapsar model for GRBs is based on the partial failure to produce an explosion from models of core collapse SNe. However, to get a gamma-ray burst several special properties of the collapsing star are likely to be needed. This is also indicated by the fact that only a very small fraction of all SNe produce GRBs. From the beaming angle, corresponding to a solid angle of $\Omega \sim 0.03$, we observe only every $\lesssim 200$ of all GRBs. More detailed estimates of this factor vary between 75–500. The total GRB rate (including the ones with beaming away from us) is estimated to be $\sim 33 \text{ Gpc}^{-3} \text{ year}^{-1}$. The typical rate should then be one GRB per $\sim 3 \times 10^5$ years for a typical galaxy. Therefore, only a fraction of one per $\sim 3 \times 10^3$ SNe will become a GRB.

In the standard GRB scenario the main ingredients is a rapidly rotating stellar core, and a low mass or absent stellar envelope. The former is needed to produce a jet, while the latter is needed to get the jet out of the star.

The main parameters of the collapsing core are the specific angular momentum, $j = J/M$. During the first seconds a centrifugally supported disk forms with interior to $R \approx j^2/GM$. For reasonable values of j this is $\sim 100 - 200$ km. Because the centrifugal support is much smaller in the polar direction the matter in this direction continues to accrete onto the black hole, until this region is nearly empty (Fig. 88). The density contrast between the equatorial disk, where the density is $\sim 10^9 \text{ g cm}^{-3}$, and the polar direction will therefore be very large.

Most of the energy losses from the disk will, because of the high temperature, $\sim 10^{10}$ K, be in the form of neutrinos. Because the disk dominates the neutrino luminosity, the neutrino radiation field will be highly anisotropic. Neutrino annihilation, $\nu + \bar{\nu} \rightarrow e^- + e^+$, above the disk may then produce electron-positron pairs in this region. These can then give rise to a jet perpendicular to the disk. The details of this mechanism are, however, uncertain.

Another suggestion uses some kind of electromagnetic extraction of the energy from the disk in a similar manner as a pulsar. A magnetic field anchored in the disk and treading the black hole horizon can tap the black hole on rotational energy by the so called Blandford-Znajek mechanism. The total amount of rotational energy is in principle enormous, $\sim 10^{54}$ ergs, but again, this mechanisms is not worked out in sufficient detail for a proper evaluation of its merits.

In some way or another a large amount of energy is likely to be deposited

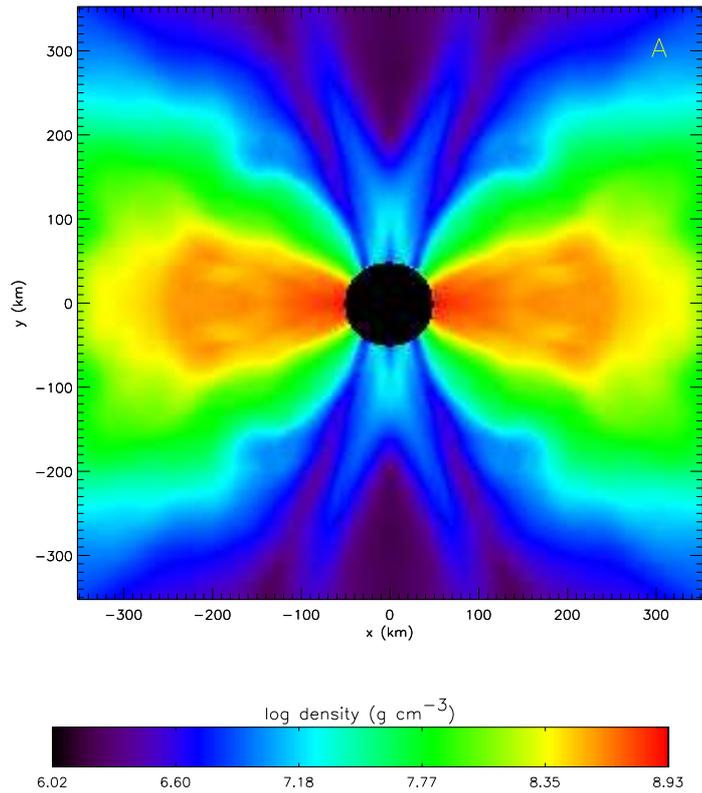


Figure 89: Density in the center ~ 7 s after collapse. Inside of ~ 200 km the centrifugally supported torus can be seen. In the polar direction the density is very low because the lack of centrifugal support has emptied this region. (MacFadyen & Woosley 1999)

in the polar directions above the disk on a time scale of the order of 1-100 s. This is the starting point of the two jets along the rotational axis of the star.

As the jet is launched from the center, it propagates outwards through the star. Fig. 89 shows a simulation of this from the inner region up to the point when it interacts with the circumstellar medium in the form of a wind from the progenitor. The radius of the Wolf-Rayet progenitor is in this model 8×10^{10} cm and the He-core mass $15 M_{\odot}$.

While inside the star, the narrow jet will be preceded by a cocoon, consisting of shocked material from the stellar core and envelope, as well as the shocked jet. This cocoon propagates through the star with a sub-relativistic velocity, $\sim (5 - 10) \times 10^4$ km s⁻¹, although the jet itself is relativistic with $\Gamma \sim 10$. As it penetrates through the surface of the star the cocoon spreads in angle and also accelerates down the steep density at the surface. This results in a Lorentz factor of $\sim 5 - 10$ for the cocoon and an angular extent of $\sim 30^\circ$.

In Fig. 90 we show the density and Lorentz factor at the final epoch of the model above. Although only ~ 20 at the time of the jet break-out, the final Lorentz factor in the jet reaches $\gtrsim 100$ as the internal energy is converted to kinetic energy by the adiabatic expansion. A most important thing to note in the figure is the highly variable Lorentz factor in the jet. As the faster material will catch up with the slower, internal shocks in the jet will form, explaining the initial burst. This can explain the prompt burst as we discussed earlier.

The cocoon mentioned above is interesting because although it only contains a minor fraction of the total energy, it has a factor of 5–10 larger angular extent than the jet itself. The solid angle, and thus the probability of observing it, is therefore a factor of 25–100 larger. The lower Lorentz factor means that the radiation from the cocoon should be considerably softer. It has been proposed that this may explain the so called X-ray flashes (XRFs), which has most of their energy in the X-ray rather than gamma-ray domain.

The collapsar model is a very likely candidate to explain the long bursts. The duration of the burst is set roughly by the time scale of the launch of the jet. It is, however, difficult to see that this can be much shorter than seconds, and the model has therefor problems explaining the short bursts. In principle, it is, however, possible also to get very short bursts from the interaction of the jet and the head of the cocoon.

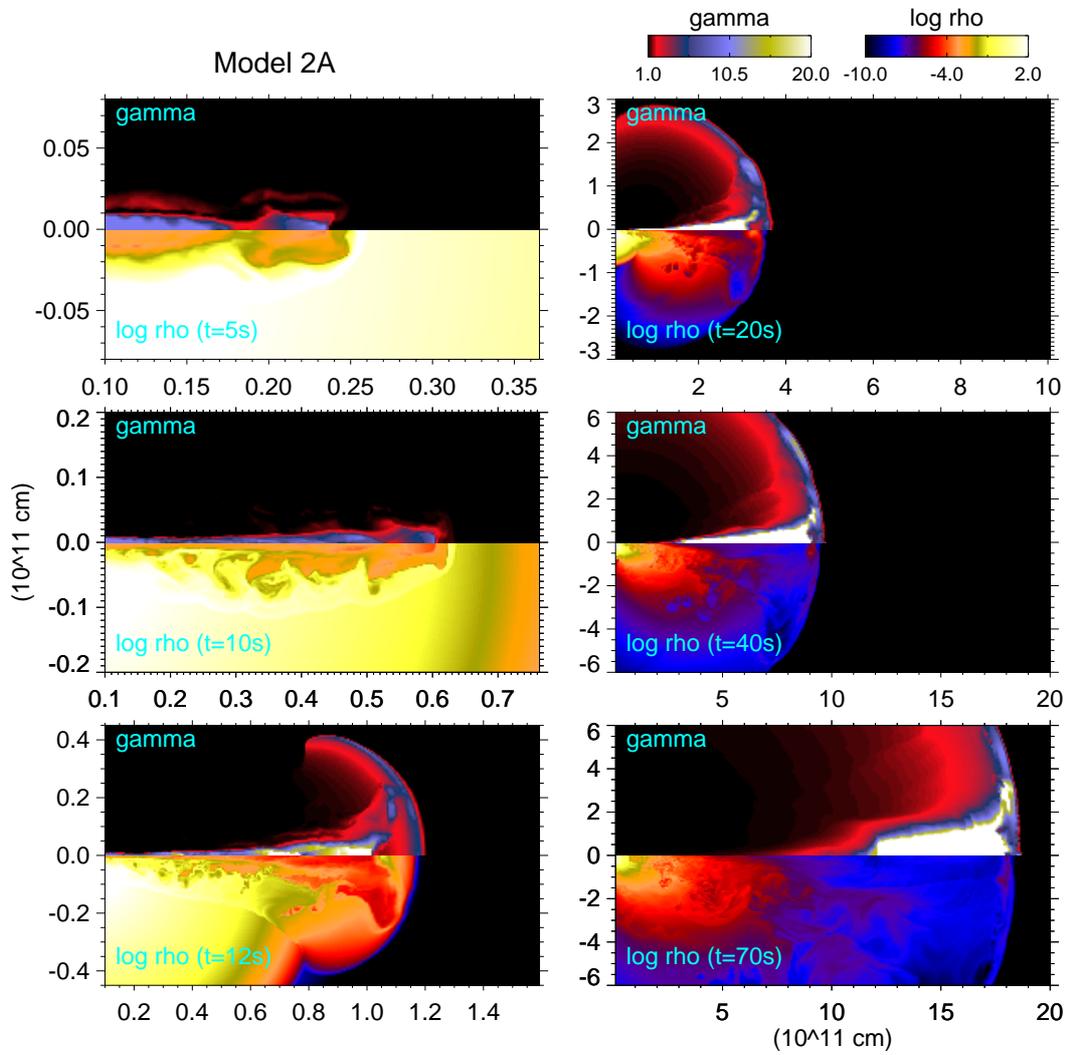


Figure 90: Simulation of jet propagation and break-out. The different panels show the Lorentz factor and density at six epochs, 5, 10, 12, 20, 40, and 70 s. (from Zhang, Woosley, Heger 2003)

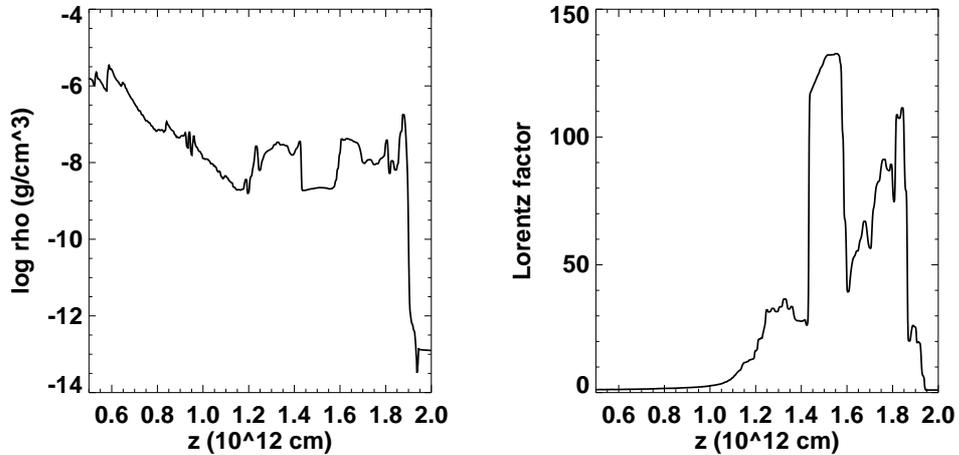


Figure 91: Density and Lorentz factor in the jet in Fig. 89 at 70 s. Note the highly variable density and especially Lorentz factor (Zhang, and Woosley 2002).

32.4 Neutron star mergers

We know that binary neutron stars exist, as the famous case of the Hulse-Taylor pulsar PSR1913+16 shows. This system will decay by gravitational radiation on a time scale of $\sim 10^8$ years. Because both stars have a mass close to $1.4 M_{\odot}$, the result will most likely be a black hole, unless a very large fraction of the mass is expelled. The energy release in connection to this may be very large, comparable to that in an ordinary core collapse supernova. The time scale will be of the order of milliseconds. This has led to the suggestion that merging neutron stars may have something to do with GRBs, and was for a long time the most popular GRB model. As we have seen, there is now compelling evidence that the long GRBs are connected to supernovae. This evidence does, however, not apply to the short GRBs. In particular, this model has some properties which can easier explain the short time scales connected with this class of GRBs.

As the neutron stars spiral in they will lose more and more of the orbital energy by gravitational radiation. The final merger will occur on a time scale of the order of milliseconds. Because of the large angular momentum the tidal forces will distort and tear apart the stars, and a flattened, disk like configuration will form (Fig. 91). While most of the mass results in a black hole of mass $\sim 2.5 M_{\odot}$, a substantial fraction, $\sim 0.1 - 0.2 M_{\odot}$, will

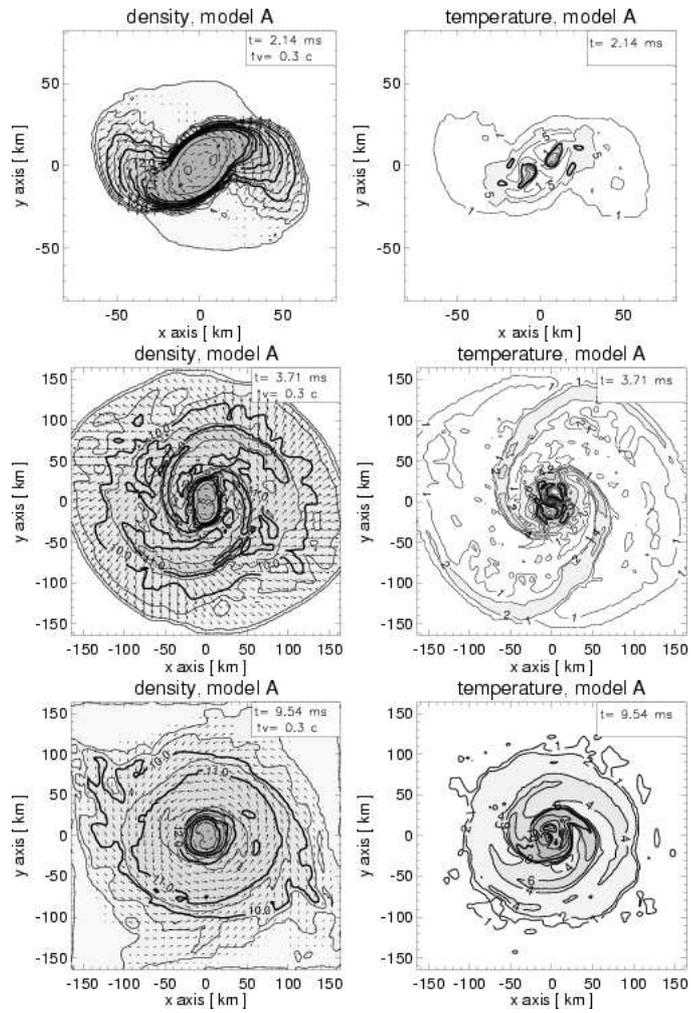


Figure 92: Evolution of the neutron star binary at different epochs after the start of the simulation. The contours show the density and temperature, while the arrows show the velocity field. (Ruffert & Janka 2001).

stay in the form of an extremely hot accretion disk. The temperature of this will be $\sim 10^{10}$ K, and it will therefore lose most of its internal energy as neutrinos. The accretion rate will be $\sim 1 M_{\odot} s^{-1}$, so the disk will have a life time of $\sim 0.1 - 0.2$ s. The total energy in the neutrinos will be of the same order as for a core collapse SN, $\sim 10^{53}$ ergs, more than enough to feed a GRB with a moderate amount of beaming. The fundamental problem is just how to convert the neutrino energy into photons.

Similar to the collapsars, there are two main mechanisms which have been proposed for this, neutrino annihilation or electromagnetic extraction. The neutrinos produce electron-positron pairs, which give rise to a jet perpendicular to the disk. This then convert its kinetic energy into heat by internal shocks, giving rise to a gamma-ray burst. Because the neutron star binary is not expected to have any circumstellar medium, the afterglow is expected to be the result of interaction of the outflow with the interstellar gas, having a constant density.

The main problem with this mechanism is that detailed simulations show that the efficiency of the neutrino pair annihilation is relatively inefficient. The energy converted into pairs is $\sim 5 \times 10^{49}$ ergs, which may be too low. This is especially true since it is difficult to obtain the narrow beaming suggested by the afterglow observations. Although highly uncertain, the MHD extraction of energy may be the most promising, but also most complex to calculate.

Summarizing this model, it has the virtue of being based on events which we know will take place, and that the total energy available is sufficient. The drawbacks is the difficulty of converting this energy to photons. In addition, the frequency of these mergers is highly uncertain, although estimates give a rate of one merger per $\sim 10^6$ years for a typical L_{\star} galaxy.

Recently, there has been observations of some short GRBs with SWIFT which have given support to this progenitor scenario. GRB 050509B, GRB 050709 and GRB 050724 were the first short GRBs to have an X-ray afterglow and therefore allow a precise localization of the GRB. For two of these optical afterglows were also found. The redshift of the galaxies were all comparatively low, 0.16 - 0.25. In two of the cases the host galaxies were ellipticals, while in the other it was a star forming dwarf galaxy. Compared to the long bursts, the gamma-ray luminosity is down by 2-3 orders of magnitude.

No indication of a supernova was found in the optical afterglows and the modeling of the spectrum and light curves of the afterglows indicate very low density environments. Both these results, as well as the location in two of the cases in elliptical galaxies, are consistent with what would be

expected for a merger of two compact objects. While the neutron –neutron star merger is the most favored, also a black hole – neutron star merger is a possibility.

It is likely that SWIFT will discover several more of the short bursts, and that we will therefore get considerably more information about this class of GRBs in the near future.

33 Active Galactic Nuclei

33.1 Classification

Seyfert 1

Seyfert 2

Quasars

Blazars

Host galaxies

Redshift distribution

33.2 Regions

reverberation

33.3 Emission lines

Photoionization

Fluorescence

Narrow and broad emission line regions

Overview of region FKR Fig 7.2

33.4 Superluminal expansion

VLBI observations of radio sources can reveal structures down to milli arc-sec scales. A remarkable discovery with this technique was the fact that some compact radio sources in radio galaxies exhibited apparent expansion velocities larger than the velocity of light, i.e., superluminal velocities.

The existence of superluminal motion indicates that one has physical velocities close to that of light. In fact, superluminal motion were predicted by M. Rees long before they were discovered observationally.

A simple example of superluminal expansion is that of a jet moving out with a speed $v = \beta c$ at an angle θ to the line of sight.

Let a first flash of light be emitted at $t = 0$ when the jet is launched. The jet then moves with constant velocity and emits a second flash at $t = \Delta t$. It has then moved a distance $v\Delta t$ from the origin. This corresponds to a transverse distance $\Delta L = v\Delta t \sin \theta$ as seen by the observer. If the distance from the observer to the source is D the first pulse is received at time D/c . The second pulse is emitted *later by* Δt but has a shorter distance by $v\Delta t \cos \theta$ to travel to the observer, who receives it at $t = \Delta t + (D - v\Delta t \cos \theta)/c$. The observer therefore finds that they arrive with a time interval of $\Delta t_{obs} = \Delta t(1 - v \cos \theta/c)$. This can be much smaller than Δt if the second factor is close to unity. The transverse velocity the observer measures is then

$$\frac{v_{app}}{c} = \frac{\Delta L}{\Delta t_{obs}} = \frac{\beta c \Delta t \sin \theta}{\Delta t(1 - \beta \cos \theta)} = \frac{\beta \sin \theta}{(1 - \beta \cos \theta)} \quad (33.1)$$

Consider now the case when the angle is close to the line of sight and the velocity is close to that of light. In this case $\sin \theta \approx \theta$, $\cos \theta \approx 1 - \theta^2/2$ and $\beta \approx 1 - 1/2\gamma^2$, so that

$$v_{app} \approx \frac{2\theta}{1/\gamma^2 + \theta^2} c \quad (33.2)$$

Therefore, it is only for $\theta \sim 1$ that $v_{app} < c$!

Note also that it does not have to be a physical motion involved, as was the case here. It could just as well have been a pulse of light from the center, which gave rise to some excitation at a distance $c/\Delta t$ from the source. Or it can be a mirror, reflecting the light. The apparent velocity would then be $v_{app} \approx 2c/\theta$, which can be very large for small angles.

33.5 The broadline region

Both quasars and Seyfert 1 galaxies show a large number of broad emission lines from both neutral and highly ionized species on top of a power law continuum. The velocities extend up to $\sim 10^4$ km s⁻¹, indicating motion in a deep gravitational potential.

- ionization parameter
-
- reverberation

33.6 Jets

Here we recapitulate the most important relations between the relevant quantities in the frame of the propagating shock and the observer frame. For a detailed derivation see e.g., Rybicki & Lightman Chap. 4. The relation between the emitted and observed frequency is

$$\nu_{obs} = \frac{\nu_{em}}{\Gamma(1 - \mu_{em}\beta)} \quad (33.3)$$

where μ_{em} is the angle between the jet direction and the line of sight measured in the emitting system.

The relation between the angle of an emitted photon in the frame of the shock and in the observer frame is

$$\mu_{obs} = \frac{\mu_{em} + \beta}{1 + \mu_{em}\beta} \quad (33.4)$$

This illustrates the aberration of the emitted light. In particular, an isotropic emission will be beamed in the direction of the jet propagation. Using $\beta\gamma \approx 1/2\gamma^2$ one finds that a photon emitted at right angle in the emitting frame ($\mu_{em} = 0$) has an angle $\theta_{obs} \approx 1/\gamma$ for $\gamma \gg 1$. This is the beaming effect.

Because of the beaming and Doppler effect the relativistic velocity has an important effect on the intensity of the emission from the jet, depending on the direction of the jet. To see this we note that the quantity I_ν/ν^3 is a Lorentz invariant (RL Chap 4.9). If we define

$$D = \Gamma(1 - \mu_{em}\beta) \quad (33.5)$$

we therefore have

$$I_{\nu, obs}(\nu_{obs}) = \left(\frac{\nu_{obs}}{\nu_{em}}\right)^3 I_{\nu_{em}, em}(\nu_{em}) = D^{-3} I_{\nu, em}(\nu_{em}) \quad (33.6)$$

In the case that the radiation is produced by synchrotron emission $I_{\nu, em}(\nu_{em}) = C\nu_{em}^{-\alpha}$. If we want to see the effect of the Doppler boosting we should compare the observed intensity at a give frequency with that emitted at the same emitted frequency. Therefore,

$$I_{\nu, obs}(\nu_{obs}) = CD^{-3}\nu_{em}^{-\alpha} = CD^{-3-\alpha}\nu_{obs}^{-\alpha} \quad (33.7)$$

- Energetics

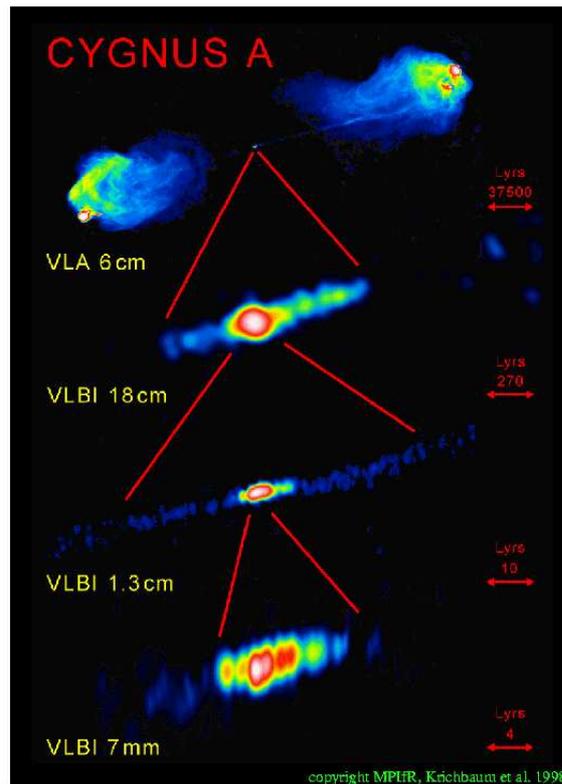


Figure 93: The radio galaxy Cygnus A on different spatial scales. The distance between the two radio lobes is ~ 100 kpc. (VLA, VLBI, Krichbaum et al).

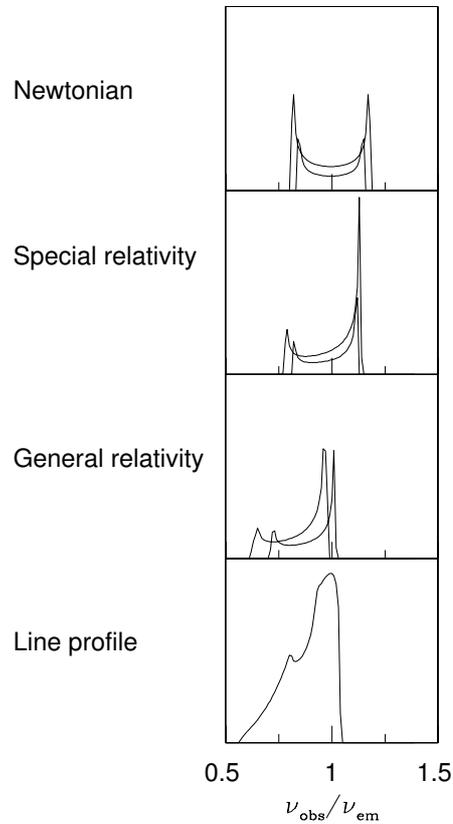


Figure 94: The Fe $K\alpha$ line profile showing the addition of different relativistic effects to the Newtonian line profile. (Fabian 2006).

- Doppler boosting
- collimation
- synchrotron-Compton

33.7 Evidence for black holes in AGNs

33.7.1 Fe $K\alpha$ lines from accretion disks

One of the most interesting indications of the effects of strong gravity around a black hole can be seen in the Fe $K\alpha$ line profile in several AGNs. This line at 6.4–6.9 keV (depending on the state of ionization) arises as a result of photoionization of one of the inner K-shell electrons and the subsequent flu-

orescence as the outer electrons fills the hole after the K-shell electron. The interesting aspect of this line is that because of the high ionization potential iron is not completely ionized even in the inner parts of the accretion disk the $K\alpha$ line may act as a probe of the relativistic effects close to the black hole. These result in a very characteristic, asymmetric line profile, showing the different effects of strong gravity.

The first observations of an asymmetric $K\alpha$ line was by ASCA of the Seyfert galaxy MCG-6-30-15, shown in the left panel of Fig. 94. This has now been observed for a large number of AGNs and also for several Galactic stellar mass black holes.

It is instructive to see the different relativistic contributions to the line profile, shown in Fig. 93. The upper panel of this figure shows the double peaked line profile for a Newtonian disk, where the two peaks originate from the region rotating towards and away from the observer. The second panel shows the effect of including the transverse Doppler effect extending the the profile to lower energies and the relativistic beaming increasing the blue peak and decreasing the red. The gravitational redshift gives an overall shift of the profile (third panel). The fourth panel shows the combination of these effects and the contribution from different radii of the disk.

While there is good qualitative agreement with the expected profile this can be taken one step further. The extension of the line to the 'red' (low energy) is a measure of the depth of the potential and velocity of the last stable orbit. Because the LSO extends all the way to the horizon for a Kerr hole this can have a larger redshift than a non-rotating black hole, where the LSO is at $3R_s$. In the right panel of Fig. 93 we show the line profile of a Schwarzschild hole compared to a maximally rotating Kerr hole. When compared to the observations the red extent of the line indicates that a Schwarzschild hole is not sufficient to explain this. Instead the authors argue that a Kerr hole rotating close to the maximum rate is necessary. This is one of the few ways one can directly study the close environment of a black hole.

33.7.2 The Galactic Center

The strongest evidence for a massive black hole in the center of a galaxy comes from our own Galactic Center. It has long been known that there is interesting activity in the central region of the Galaxy, especially in the radio. This radio source is known as Sgr A*.

Because of the large reddening observations can only be made in the near-IR where extinction is much lower than in the optical. In Fig. 95 a K-band

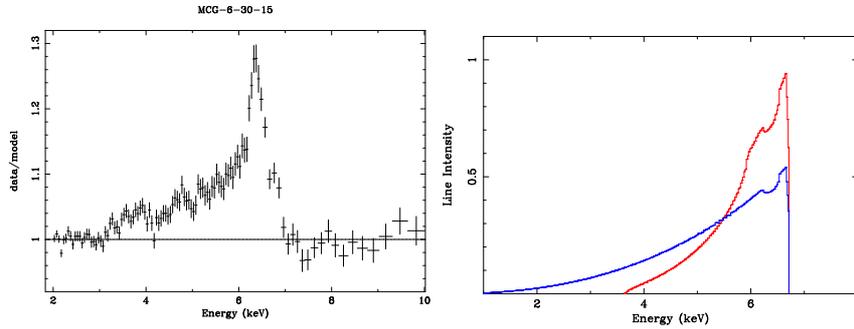


Figure 95: The left panel shows the Fe K α line profile for the Seyfert 1 galaxy MCG-6-30-15. Note the asymmetric line profile with almost no emission on the 'blue' side of the line, while the 'red' extends to ~ 3 keV. The right panel shows the line profile from a non-rotating Schwarzschild black hole (red) and a maximally rotating Kerr hole (blue). Note the larger extension to the red for the Kerr case, in agreement with the observed. (Fabian 2006).

image at 2μ taken by NACO on the VLT is shown. The most spectacular observation of the GC is the cluster of stars near Sgr A* seen in this picture. From measurements using adaptive optics on 8-10 m class telescopes two different groups have been able to measure the proper motions of several stars close to the GC. In particular, they have been able to map the orbit of one star, called S2, which has a very eccentric orbit with $e = 0.87$. The period of the orbit is 15.2 years and at periastron (i.e., closest to Sgr A*) it is only ~ 17 light-hours, or 124 AU, from Sgr A*. From the orbital elements it is straightforward to solve for the mass inside the orbit and they find a total mass of $(3.6 \pm 0.6) \times 10^6 M_{\odot}$. This corresponds to a stellar density of $\sim 10^{17} M_{\odot} pc^{-3}$. Because the luminosity in any wavelength band is very low ($\lesssim 10^{35} \text{ erg s}^{-1}$) this enormous mass can only correspond to a black hole.

A recent discovery of great interest is that monitoring of Sgr A* at both radio, IR and X-ray wavelengths shows a great amount of activity. Compared to any AGNs the luminosity is, however, extremely low, only $10^{-10} - 10^{-9}$ times the Edington luminosity. The emission in the near-IR, as well as radio, show a high degree of polarization, indicating that the emission is synchrotron. The light curve in the near-IR (Fig. 96) shows a quasi periodic activity on a time scale of ~ 20 minutes. This shows that it must originate within ~ 10 Schwarzschild radii from the BH. All these observations indicate that the flares originate as a result of a very low luminosity accretion flow immediately outside the black hole.

bulge- BH mass

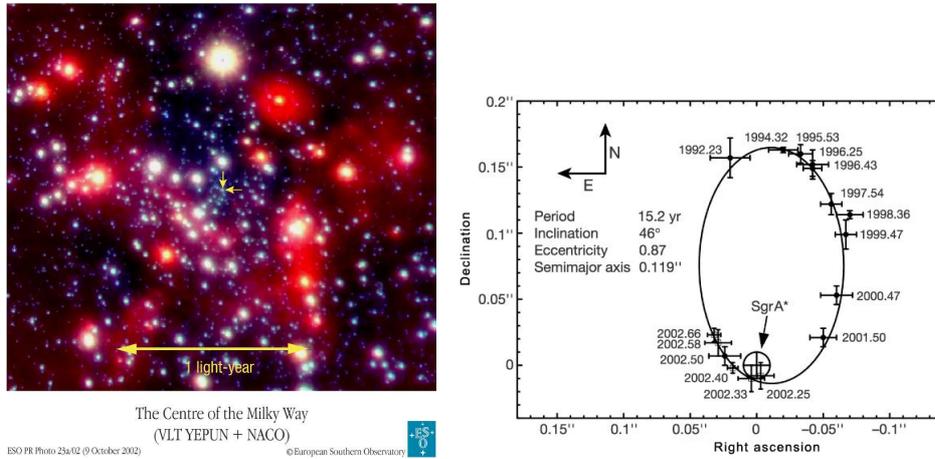


Figure 96: The Galactic Center at 2μ . The arrows point to star S2. The right panel shows the orbit of the star with pericenter closest to the black hole, star S2. (VLT/NACO), Schödel et al. 2002.

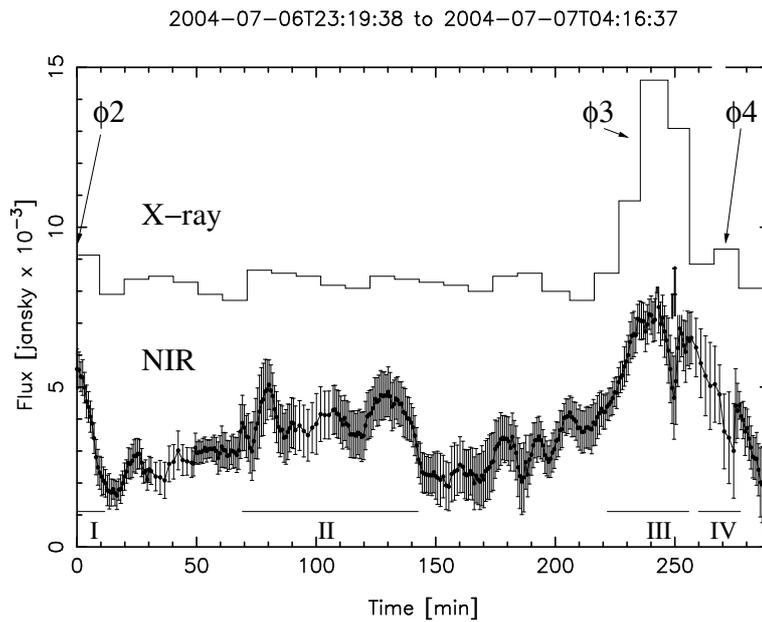


Figure 97: Near-IR and X-ray light curves from the position of the Galactic Center, Sgr A*. (Eckart et al. 2006).

Masers
 Blandford-Znajek
 compact radio sources
 extended radio sources

34 Pair production

The interaction of two photons can lead to pair production provided $E_1 E_2 \gtrsim m_e c^2$. A low energy photon may therefore give rise to a pair if it interacts with a photon with high enough energy.

The typical $\gamma + \gamma \rightarrow e^+ + e^-$ cross section is $\sim 0.2\sigma_T$ around $E_1 = 2m_e c^2/E_2$. The optical depth for a photon of energy E in a region with radius R containing a density, n_γ , of softer photons is therefore

$$\tau(E_1) = R \int_{E_{th}} \sigma_{\gamma\gamma} \frac{dn_\gamma}{dE} dE \approx 0.2\sigma_T \frac{dn_\gamma}{dE}(E_{th}) E_{th} \quad (34.1)$$

where $\frac{dn_\gamma}{dE}(E_{th}) E_{th} \approx n_\gamma$ approximately is the number density of photons above the threshold. But, $n_\gamma \approx L(E > (E_{th}))/4\pi c R^2 E_{th}$, so

$$\tau(E_1) \approx 0.2 \frac{\sigma_T L(E > E_{th})}{4\pi c R E_{th}} \quad (34.2)$$

As a estimate we take $E_{th} \sim m_e c^2$ so we get

$$\tau(E_1) \approx 0.2 \frac{\sigma_T L(E > E_{th})}{4\pi c R m_e c^2} \equiv \frac{0.2}{4\pi} l \quad (34.3)$$

where l is the compactness parameter

$$l \equiv \frac{\sigma_T L(E > E_{th})}{R m_e c^3} \quad (34.4)$$

Therefore, if the compactness parameter is large that means that the plasma will be optically thick to pair production. This will increase the density of electrons (and positrons), which can scatter the photons by Compton scattering and change the observed spectrum dramatically.

A Equations of relativistic hydrodynamics

This Appendix is NOT included in the course. Just for the really interested!! See Weinberg Chap. 2.10 and Landau & Lifshitz, Fluid Dynamics 1985 for more details

When we discuss the dynamics and radiation from the GRB there are three reference frames which are of interest. The rest frame of the exploding star, the comoving frame of the expanding gas and the reference frame of the observer.

An example is the time interval of a process as measured in the comoving frame, dt_c , and that of the GRB, which are related as

$$dt_{GRB} = \gamma dt_c \quad (\text{A.1})$$

In general when we want to relate measurements in different reference systems it is convenient to see these as transformations between different four-vectors. An example of such a four-vector is $x^\mu = (x, y, z, t)$. If the coordinates of an event in a system, K, moving with velocity v relative to another system K' along the x-axis is $x^\mu = (x, y, z, t)$, then the coordinates in the system K' are given by

$$x'^\mu = \sum_{\nu=1}^4 \Lambda^\mu{}_\nu x^\nu \quad (\text{A.2})$$

where $\nu = 1, 2, 3$ denote the space components and $\nu = 4$ the time component. The matrix $\Lambda^\mu{}_\nu$ is called the Lorentz boost, and is for the case of motion only along the x-axis given by

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & 0 & 0 & \gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \gamma\beta & 0 & 0 & \gamma \end{pmatrix} \quad (\text{A.3})$$

Here $\beta = v/c$ and $\gamma = 1/\sqrt{1 - (v/c)^2}$.

The energy density and pressure transform as components of the of the energy-momentum tensor, $T^{\mu\nu}$, which in the rest frame is given by

$$T^{\mu\nu} = \begin{pmatrix} p & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & \epsilon \end{pmatrix} \quad (\text{A.4})$$

To transform this into a system moving along the x -axis, we apply two Lorentz boosts,

$$T'^{\kappa\lambda} = \sum_{\mu=1}^4 \sum_{\nu=1}^4 \Lambda^{\kappa}_{\mu} \Lambda^{\lambda}_{\nu} T^{\mu\nu} . \quad (\text{A.5})$$

We therefore obtain

$$T'^{\mu\nu} = \begin{pmatrix} \gamma^2(\epsilon\beta^2 + p) & 0 & 0 & \gamma^2\beta(\epsilon + p) \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ \gamma^2\beta(\epsilon + p) & 0 & 0 & \gamma^2(\epsilon + \beta^2 p) \end{pmatrix} \quad (\text{A.6})$$

Exercise:

Show Eq. (A.6)!

The hydrodynamic equations are given by the divergence of the energy momentum tensor,

$$\sum_{\nu=1}^4 \frac{\partial T^{\mu\nu}}{\partial x^{\nu}} = 0 \quad (\text{A.7})$$

In spherical symmetry we get

$$\frac{\partial}{\partial t} \gamma^2(\epsilon + \beta^2 p) + \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \gamma^2 \beta(\epsilon + p) = 0 \quad (\text{A.8})$$

$$\frac{\partial}{\partial t} \gamma^2 \beta(\epsilon + p) + \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \gamma^2 \beta^2(\epsilon + p) + \frac{\partial p}{\partial r} = 0 \quad (\text{A.9})$$

In addition we have the conservation of the particle number. The particle number in the rest frame is n . This is the time component of the particle current four-vector. To transform this into an arbitrary frame we apply again a Lorentz boost $J'^{\mu} = \sum_{\nu=1}^4 \Lambda^{\mu}_{\nu} J^{\nu}$, or

$$J'^{\mu} = (\gamma\beta n, 0, 0, \gamma n) \quad (\text{A.10})$$

Taking the four-divergence in spherical geometry we get

$$\frac{\partial}{\partial t} \gamma n + \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \gamma \beta n = 0 \quad (\text{A.11})$$

B Relativistic shocks

Denote the postshock medium by 2 and the preshock medium by 1. As in the non-relativistic case, the relativistic shock conditions are obtained by integrating Eqns. (A.8), (A.9) and (A.11) over an infinitesimal radial distance across the shock. All the time derivative terms then go to zero, while the radial derivatives result in the jump conditions

$$\gamma_1 \beta_1 n_1 = \gamma_2 \beta_2 n_2 \quad (\text{B.1})$$

$$\gamma_1^2 \beta_1 (\epsilon_1 + p_1) = \gamma_2^2 \beta_2 (\epsilon_2 + p_2) \quad (\text{B.2})$$

$$\gamma_1^2 \beta_1^2 (\epsilon_1 + p_1) + p_1 = \gamma_2^2 \beta_2^2 (\epsilon_2 + p_2) + p_2 \quad (\text{B.3})$$

The energy densities and pressure refer to the proper values before and after the shock, respectively. All velocities are relative to the reference frame of the shock.

From the second and third equation one can solve for β_1 and β_2 . After some algebra, most easily done by setting $\beta_1 = \tanh \phi_1$ etc., one obtains

$$\beta_1 = \left[\frac{(p_2 - p_1)(\epsilon_2 + p_1)}{(\epsilon_2 - \epsilon_1)(\epsilon_1 + p_2)} \right]^{1/2} \quad (\text{B.4})$$

$$\beta_2 = \left[\frac{(p_2 - p_1)(\epsilon_1 + p_2)}{(\epsilon_2 - \epsilon_1)(\epsilon_2 + p_1)} \right]^{1/2} \quad (\text{B.5})$$

The relative velocity of the preshock and postshock gas is given by the law of addition of velocities $\beta = (\beta_1 - \beta_2)/(1 - \beta_1 \beta_2)$, so

$$\beta = \left[\frac{(p_2 - p_1)(\epsilon_2 - \epsilon_1)}{(\epsilon_1 + p_2)(\epsilon_2 + p_1)} \right]^{1/2} \quad (\text{B.6})$$

Because the medium into which the shock is propagating is at rest, the Lorentz factor of the shock as seen by an observer is $\Gamma_s = \gamma_1$, and $\Gamma_2 = 1/\sqrt{1 - \beta^2}$ that of the shocked gas as seen by the observer in the medium at rest. It is easy to show that

$$\Gamma_2 = \left[\frac{(\epsilon_1 + p_2)(\epsilon_2 + p_1)}{(\epsilon_1 + p_1)(\epsilon_2 + p_2)} \right]^{1/2} \quad (\text{B.7})$$

Instead of the energy behind the shock it is usually more convenient to consider this, or Γ_s , as the main parameter of the shock.

Let us now consider some special cases. First, assume that the shock is propagating into a cold medium with density $\rho_1 = n_1 m_u$ and negligible

pressure, so that $\epsilon_1 = n_1 m_u c^2$ and $p_1 = 0$. Further, we assume that the gas behind the shock is relativistic, so that $p_2 \approx \epsilon_2/3$. We then obtain from Eq. (B.7)

$$\Gamma_2 = \frac{1}{2} \left(\frac{\epsilon_2}{n_1 m_u c^2} + 3 \right)^{1/2}. \quad (\text{B.8})$$

The velocities of the preshock and postshock gas in the shock frame are Eq. (B.5)

$$\beta_1 \approx 1 - n_1 m_u c^2 / \epsilon_2 \quad (\text{B.9})$$

$$\beta_2 \approx \frac{1}{3} (1 + 2n_1 m_u c^2 / \epsilon_2) \quad (\text{B.10})$$

showing that the gas is flowing in with nearly the speed of light, while the postshock gas flows away with $\beta_2 \sim 1/3$.

Using Eq. (B.10) it is easy to show that therefore

$$\Gamma_s = \gamma_1 \approx \left(\frac{\epsilon_2}{2n_1 m_u c^2} \right)^{1/2}. \quad (\text{B.11})$$

Therefore, the shock Lorentz factor is a factor of $2^{1/2}$ larger than that of the postshock gas as seen by the observer at rest,

$$\Gamma_s \approx 2^{1/2} \Gamma_2 \quad (\text{B.12})$$

Using Eq. (B.10) in the first of the jump conditions in Eq. (B.3) we obtain

$$\frac{n_2}{n_1} \approx 2 \left(\frac{\epsilon_2}{n_1 m_u c^2} \right)^{1/2} \approx 4\Gamma_2 \quad (\text{B.13})$$

In contrast to a non-relativistic shock we can therefore get arbitrarily large compressions behind the shock. This relation, together with Eq. (B.11), also shows that most of the incoming kinetic energy of the particles in the shock frame, $\sim \Gamma_s m_u c^2$, is converted to thermal, internal energy.

Equations (B.8), (B.12), and (B.13) form the basic relations describing the properties of the shock wave as function of its Lorentz factor and preshock density, or alternatively the energy density behind the shock and the preshock density.

The shape and flux of the synchrotron emission is determined by the magnetic field and density of relativistic particles behind the shock. Lacking a fundamental theory, one is usually assuming that the energy densities of these scale with the thermal energy density behind the shock. Therefore

$$\frac{B^2}{8\pi} = \epsilon_B \epsilon_2 \approx 4\epsilon_B \Gamma_2^2 n_1 m_u c^2 \quad (\text{B.14})$$

and

$$u_{rel} = \varepsilon_e \varepsilon_2 \approx 4\varepsilon_e \Gamma_2^2 n_1 m_u c^2 \quad (\text{B.15})$$

For a power law spectrum $n_{rel} = C\gamma^{-p}$ the ratio of the energy density and number density is given by

$$\frac{u_{rel}}{n_{rel}} = \frac{(p-1)}{(p-2)} m_e c^2 \gamma_{min} \quad (\text{B.16})$$

where γ_{min} is the minimum energy, and we assume that $p > 2$, so that we can omit the upper limit to γ .

To determine γ_{min} we divide Eq. (B.15) by Eq. (B.13) to obtain

$$\frac{u_{rel}}{n_2} = \varepsilon_e m_p c^2 \Gamma_2 . \quad (\text{B.17})$$

If we now assume that $n_{rel} \approx n_2$ we get

$$\gamma_{min} = \varepsilon_e \frac{m_p (p-2)}{m_e (p-1)} \Gamma_2 . \quad (\text{B.18})$$

Note that this is based on the assumptions that the relativistic particle density scales as the postshock energy density, and that number of relativistic particles also scale with the postshock particle density. Although, a likely situation, this has to be justified by detailed simulations of collisionless shocks. There is currently substantial work going on in this area.

C Relativistic blast waves

Immediately after the explosion the hot ejecta is basically a ball of hot photons, with a very small baryon load M_0 (Eq. 31.6). It therefore expands with $\Gamma \propto R$, until $\Gamma = \Gamma_0 \approx E/M_0 c^2$, after which it expands with constant Γ until it has swept up an energy comparable to the initial. The shock condition Eq. (B.8) shows that in the rest frame of the shocked fluid the energy is $E \approx \Gamma m c^2$, and therefore in the rest frame of the observer $E_{obs} \approx \Gamma^2 m c^2$. The energy swept up is therefore comparable to the initial thermal energy when the mass of the swept up ejecta is $m \approx M_0/\Gamma_0$. At this point the ejecta will start to slow down.

In the same way as the Sedov solution plays a central role for the dynamics of the interaction of the SN with its environment, one can find relativistic generalizations of these, which describes the slowing down of the blast wave, and the conversion of the kinetic energy to thermal energy behind the blast wave.

From Eq. (B.11) and Eq. (B.12), we find that the energy behind the shock in the comoving frame is

$$\epsilon_2 = 4\Gamma_2^2 m_u c^2 n_1 \quad (\text{C.1})$$

Because the post-shock gas is relativistic, $p_2 = \epsilon_2/3$. To transform ϵ_2 and p_2 to the observer frame we use

$$\epsilon_{obs} = \Gamma_2^2(\epsilon_c + \beta^2 p_c) = \Gamma_2^2 \epsilon_c (1 + \frac{1}{3}\beta^2) = (4\Gamma_2^2 - 1) \frac{\epsilon_c}{3} \quad (\text{C.2})$$

With $\epsilon_c = \epsilon_2$ from Eq. (C.1), we get for large Γ_2 's

$$\epsilon_{obs} \approx \frac{16}{3} \Gamma_2^4 m_u c^2 n_1 \quad (\text{C.3})$$

The total energy in the observer frame is therefore

$$E_{obs} \approx \frac{64\pi}{3} \Gamma_2^4 m_u c^2 n_1 R^2 \Delta R_{obs} \quad (\text{C.4})$$

where ΔR_{obs} is the shock thickness as seen in the observer frame. To estimate ΔR_{obs} we use the conservation of mass. For constant external density the total mass swept up by the shock is $4\pi R^3 m_u n_1/3$. This should be equal to the mass in the shell, $4\pi R^2 \Delta R_{obs} n_{2, obs} m_u$. The density behind the shock in the observer frame is $n_{2, obs} = \Gamma_2 n_{2, c}$. The comoving density is given in terms of the pre-shock density by Eq. (B.13). Therefore,

$$n_{2, obs} = 4 \Gamma_2^2 n_1 \quad (\text{C.5})$$

and

$$\Delta R_{obs} \approx \frac{R}{12\Gamma_2^2} \quad (\text{C.6})$$

In the comoving frame of the shocked gas $\Delta R/R = 1/(12\Gamma_2)$. The shell is therefore even in the comoving frame extremely thin, due to the large compression behind the shock.

Inserting this in Eq. (C.4) we get

$$\Gamma_2 \approx \left(\frac{9E}{16\pi n_1 m_u c^2} \right)^{1/2} R^{-3/2} \quad (\text{C.7})$$

This is the Blandford – McKee solution, and describes together with Eqns. (B.12), (C.1), and (C.5) the physical conditions of the blast wave as function of its radius.

To transform this into a relation of time as measured by the observer we use Eq. (31.4), $dt_{obs} = dt_{em}/2\Gamma_2^2$, and that $dR = c\beta dt_{em} \approx c dt_{em}$, so that

$$t_{obs} = \int \frac{dt_{em}}{2\Gamma_2^2} = \int_0^R \frac{dR'}{2c\Gamma_2(R')^2}. \quad (\text{C.8})$$

Therefore, if $\Gamma_s \propto R^{-\alpha}$, we get

$$R = 2(1 + 2\alpha)\Gamma_2(R)^2 ct_{obs} \quad (\text{C.9})$$

so that with $\alpha = 3/2$ we get $R = 8\Gamma_2^2 ct_{obs}$. Using this in Eq. (C.7) we get

$$\Gamma_2 \approx 0.4 \left(\frac{E}{n_1 m_u c^5} \right)^{1/8} t_{obs}^{-3/8} \quad (\text{C.10})$$

or with parameters considered to be typical for a GRB

$$\Gamma_2 \approx 4.4 \left(\frac{E}{10^{52} \text{ ergs}} \right)^{1/8} \left(\frac{n_1}{1 \text{ cm}^{-3}} \right)^{-1/8} \left(\frac{t_{obs}}{\text{days}} \right)^{-3/8} \quad (\text{C.11})$$

and

$$R \approx \left(\frac{9E}{2\pi n_1 m_u c} \right)^{1/8} t_{obs}^{1/4} \quad (\text{C.12})$$

or

$$R \approx 4 \times 10^{17} \left(\frac{E}{10^{52} \text{ ergs}} \right)^{1/8} \left(\frac{n_1}{1 \text{ cm}^{-3}} \right)^{-1/8} \left(\frac{t_{obs}}{\text{days}} \right)^{1/4} \quad (\text{C.13})$$

The derivation above, as well as the total energy used here, assume a spherical expansion. As we will see, it is, however, likely that the outflow is in the form of two narrow jets. As long as the jet angle is much larger than Γ^{-1} this is, however, of minor importance, except for the fact that the energy going into these equations is the equivalent *isotropic* energy, E_{iso} . The real total energy is only $\Omega E_{iso}/2\pi$, where Ω is the solid angle of each of the two jets.