SUPERNOVAE AND GAMMA-RAY BURSTS

Claes Fransson

April 21, 2008

Contents

1 Pre-supernova evolution of massive stars

1.1 The equation of state

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

$$P = knT \tag{1.1}$$

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

$$P = \frac{k}{m_\mathrm{u}\mu}\rho T \tag{1.2}$$

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

$$n = \frac{8\pi}{h^3}\int_0^\infty f(p)p^2 dp \tag{1.3}$$

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 $p < p_\mathrm{F}$ and $f(p) = 0$ for $p > p_\mathrm{F}$, allowing us to solve for p_F

$$p_\mathrm{F} = \left(\frac{3h^3 n_\mathrm{e}}{8\pi}\right)^{1/3} \tag{1.4}$$

This can be written in terms of the Fermi energy using $E_\mathrm{F} = \sqrt{p_\mathrm{F}^2 c^2 + m_\mathrm{e}^2 c^4}$. For a non-relativistic gas $E_\mathrm{F} = p_\mathrm{F}^2/2m_\mathrm{e}$, while for a relativistic gas $E_\mathrm{F} = p_\mathrm{F}c$.

The pressure, P, is given by

$$P = \frac{1}{3}\int_0^\infty v(p)p\frac{dn(p)}{dp}dp \tag{1.5}$$

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

$$P = \frac{8\pi}{3h^3}\int_0^{p_\mathrm{F}} v(p)p^3 dp\ . \tag{1.6}$$

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

$$P = \frac{8\pi}{3h^3}\int_0^{p_\mathrm{F}} \frac{p^4}{\sqrt{m^2 + p^2/c^2}}dp\ . \tag{1.7}$$

This can be integrated to give

$$P = \frac{\pi m^4 c^5}{3h^3} f(x) \tag{1.8}$$

where

$$f(x) = x(2x^2 - 3)(x^2 + 1)^{1/2} + 3 \sinh^{-1} x \tag{1.9}$$

and $x = p_{\mathrm{F}}/mc$.

For simplicity we consider the non-relativistic and ultra-relativistic limits separately. The transition occurs when $p_{\mathrm{F}} \approx m_{\mathrm{e}} c$. Using $\rho = \mu_{\mathrm{e}} m_{\mathrm{p}} n_{\mathrm{e}}$ this occurs at

$$\rho_r = 9.7 \times 10^5 \mu_{\mathrm{e}} \text{ g cm}^{-3} \tag{1.10}$$

For a non-relativistic gas $p \ll mc$ and Eq. (1.7) shows that

$$P = \frac{8\pi}{15 m h^3} p_{\mathrm{F}}^5 \ . \tag{1.11}$$

With p_{F} from Eq. (1.4) we finally get

$$P = \frac{1}{20} \left(\frac{3}{\pi} \right)^{2/3} \frac{h^2}{m} n^{5/3} \ , \tag{1.12}$$

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

$$P = 1.00 \times 10^{13} \mu_{\mathrm{e}}^{-5/3} \rho^{5/3} \ . \tag{1.13}$$

In the opposite limit of an ultra-relativistic gas we obtain in the same way from Eq. (1.7)

$$P = \frac{2\pi c}{3h^3} p_{\mathrm{F}}^4 \ . \tag{1.14}$$

and

$$P = \frac{1}{8} \left(\frac{3}{\pi} \right)^{1/3} hc\, n^{4/3} \ , \tag{1.15}$$

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

$$P = 1.24 \times 10^{15} \mu_{\mathrm{e}}^{-4/3} \rho^{4/3} \ . \tag{1.16}$$

For a non-relativistic gas the boundary between degeneracy and perfect gas equations of state is obtained by setting the non-degenerate pressure (Eq. (1.2)) equal to the degenerate, given by Eq. (1.12)

$$\frac{T}{\rho^{2/3}} = 1.2 \times 10^5 \frac{\mu}{\mu_\mathrm{e}^{5/3}} \tag{1.17}$$

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

$$\frac{T}{\rho^{1/3}} = 1.5 \times 10^7 \frac{\mu}{\mu_\mathrm{e}^{4/3}} \ . \tag{1.18}$$

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

$$P = P_\mathrm{e} + P_\mathrm{ion} \tag{1.19}$$

where P_e is given by either Eq. (1.12) or Eq. (1.15) and P_ion by Eq. (1.1). 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 For a non-relativistic gas the boundary between degeneracy and perfect gas equations of state is give by

$$\frac{RT\rho}{\mu} = \frac{aT^4}{3} \tag{1.20}$$

or

$$\frac{T}{\rho^{1/3}} = 3.2 \times 10^7 \mu^{-1/3} \tag{1.21}$$

In Fig. 1 we show the different regions defined by Eqns. (1.10), and (1.17) – (1.21) in the $\rho - T$ plane.

1.2 Evolution of the core

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. In this section we will discuss a simplified

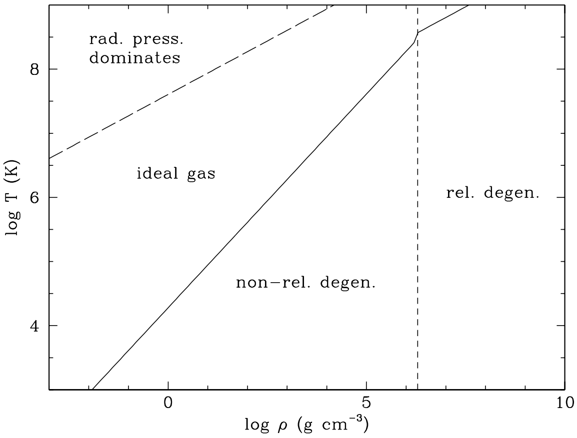

Figure 1: Regions of different equations of state in the $\rho - T$ plane.

model, which in spite of its simplicity, gives a good representation of the main features.

The pressure of a partially degenerate gas can be approximated by the interpolation formula

$$P \approx P_{\rm e} = \frac{k}{m_{\rm p}\mu_{\rm e}}\rho T + K_\gamma \left(\frac{\rho}{\mu_{\rm e}}\right)^\gamma \tag{1.22}$$

where the first term is that of an ideal gas pressure and the second that of the degenerate electrons. Here $\mu_{\rm e}$ is the mean molecular weight of the electrons. We have here neglected the ions because of the much more numerous electrons in the advanced stages of the evolution. The adiabatic index γ varies from 5/3 in the non-relativistic case to 4/3 in the relativistic degenerate case. For these limiting cases $K_{4/3} = 1.24 \times 10^{15}$ and $K_{5/3} = 1.00 \times 10^{13}$ in cgs units (see Eqns. 1.13 and 1.16).

The equation of hydrostatic equilibrium gives

$$\frac{dP}{dr} = -\frac{Gm(r)\rho}{r^2} \tag{1.23}$$

or approximately

$$P_{\rm c} \approx \frac{GM_{\rm c}\rho_{\rm c}}{r} \tag{1.24}$$

Now $\rho_{\rm c} \approx 3M_{\rm c}/(4\pi R_{\rm c}^3)$, so

$$P_{\rm c} \approx fGM_{\rm c}^{2/3}\rho_{\rm c}^{4/3} \tag{1.25}$$

where f is of order unity. For a polytrope of order 3 ($\gamma = 4/3$) it can be shown that $f = 0.364$ (see Arnett App. C).

Therefore we get

$$\frac{k}{m_\mathrm{p}\mu_\mathrm{e}} T_\mathrm{c} \approx f G M_\mathrm{c}^{2/3} \rho_\mathrm{c}^{1/3} - K_\gamma \rho_\mathrm{c}^{\gamma-1} \mu_\mathrm{e}^{-\gamma} \tag{1.26}$$

For low density and $\gamma \sim 5/3$ we can neglet the degeneracy pressure and get

$$k T_\mathrm{c} \approx f m_\mathrm{p} \mu_\mathrm{e} G M_\mathrm{c}^{2/3} \rho_\mathrm{c}^{1/3} \tag{1.27}$$

In the opposite limit of high density, the electron gas becomes relativistic and $\gamma \to 4/3$. For low T we neglect the ideal gas pressure term and obtain

$$M_\mathrm{c} \approx \left(\frac{K_{4/3}}{fG}\right)^{3/2} \mu_\mathrm{e}^{-2} \tag{1.28}$$

independent of T_c. This is the Chandrasekhar mass

$$M_\mathrm{Ch} = 5.85 \mu_\mathrm{e}^{-2}\ \mathrm{M}_\odot\ . \tag{1.29}$$

In terms of this we can write Eq. (1.26) as

$$\frac{k}{m_\mathrm{p}} T_\mathrm{c} \approx \left(\frac{\rho_\mathrm{c}}{\mu_\mathrm{e}}\right)^{1/3} [K_{4/3} \left(\frac{M_\mathrm{c}}{M_\mathrm{Ch}}\right)^{2/3} - K_\gamma \left(\frac{\rho_\mathrm{c}}{m u_\mathrm{e}}\right)^{\gamma-4/3}] \tag{1.30}$$

For $M_\mathrm{c} < M_\mathrm{Ch}$ we have the non-relativistic case ($\gamma = 5/3$), and get

$$\frac{k}{m_\mathrm{p}} T_\mathrm{c} \approx K_{4/3} \left(\frac{\rho_\mathrm{c}}{\mu_\mathrm{e}}\right)^{1/3} \left(\frac{M_\mathrm{c}}{M_\mathrm{Ch}}\right)^{2/3} - K_{5/3} \left(\frac{\rho_\mathrm{c}}{\mu_\mathrm{e}}\right)^{2/3} \tag{1.31}$$

The temperature therefore first increases as ρ_c increases (Eq. 1.27) and reaches a maximum for

$$\frac{\rho_\mathrm{c,max}}{\mu_\mathrm{e}} = \left(\frac{K_{4/3}}{2K_{5/3}}\right)^3 \left(\frac{M_\mathrm{c}}{M_\mathrm{Ch}}\right)^2 \tag{1.32}$$

at which

$$T_\mathrm{c,max} = \frac{m_\mathrm{p} K_{4/3}^2}{4k K_{5/3}} \left(\frac{M_\mathrm{c}}{M_\mathrm{Ch}}\right)^{4/3} \approx 4.7 \times 10^8 \left(\frac{M_\mathrm{c}}{M_\mathrm{Ch}}\right)^{4/3} \mathrm{K}. \tag{1.33}$$

For $\rho_\mathrm{c} > \rho_\mathrm{c,max}$ and $M_\mathrm{c} < M_\mathrm{Ch}$ the temperature decreases. This is therefore the maximum temperature a non-relativistic partially degenerate core with mass less than the Chandrasekhar mass can obtain.

For $M_{\rm c} > M_{\rm Ch}$ the gas becomes degenerate and relativistic, and we have $\gamma = 4/3 + \epsilon$, where ϵ is small. Eq. (1.30) then shows that the temperature rises monotonically as $T_{\rm c} \propto \rho_{\rm c}^{1/3}$. We can therefore separate two qualitatively different behaviors, depending on the value of $M_{\rm c}$. Core masses below $M_{\rm Ch}$ never reach a temperature larger than $T_{\rm c,max}$, while those above can increase their temperature as the core contracts and ignite new fuels.

In Fig. 2 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 reaches a maximum temperature and evolves into the degenerate regime. In all three cases the cores do not reach a temperature high enough for carbon burning to start. The 15 $M_\odot$ model on the other hand continue to evolve in the partially degenerate regime and the temperature increases monotonically, and passes through all the nuclear burning stages.

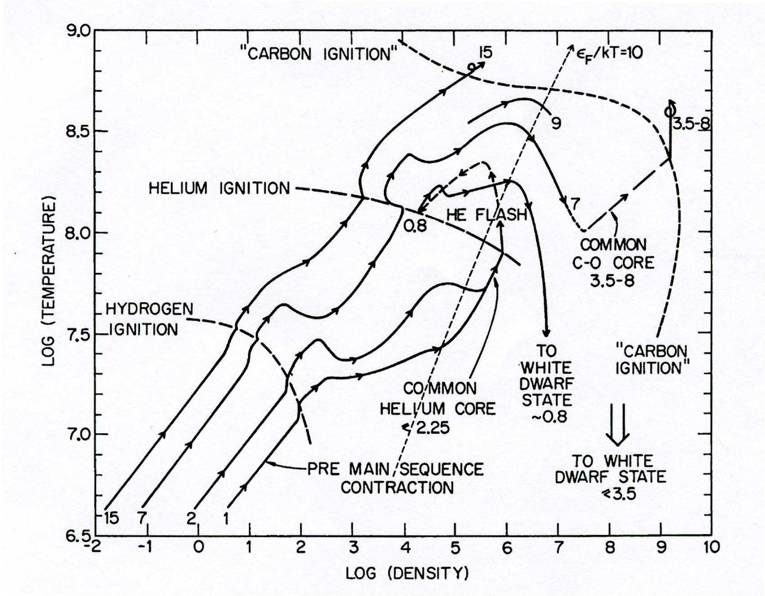

Figure 2: Evolution of the central density and temperature for stars of different masses between 1 $M_\odot$ and 15 $M_\odot$. (Iben 1974)

In Fig. 3 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. Also in this case the $T_{\rm c} \propto \rho_{\rm c}^{1/3}$ scaling is a good approximation to the evolution, although occasional excursions are seen in connection to the ignition of new burning fuels. Because the entropy of a radiation dominated gas in units of k is $s = 4am_{\rm p}T^3/(3k\rho)$ this shows that the evolution of the core is at approximately constant radiation entropy. Typically, $s \sim 1$. The concept of constant entropy will later be used frequently for describing the collapse of the core.

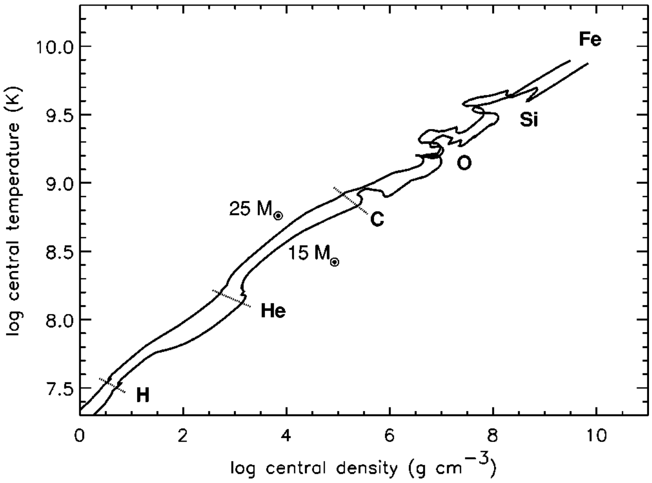

Figure 3: Evolution of the central density and temperature for a 15 $M_\odot$ and 25 $M_\odot$ star. (WHW02)

Exercises:

1. Do a 'back of the envelope' derivation of the Chandrasekhar mass and the mass – radius relation for a white dwarf.

2. **Skip!** Derive the Chandrasekhar mass strictly using the theory of polytropes. You have probably already done this in the undergraduate course, but it is importance enough for a repetition!

1.3 Convection (You may skip this!)

Convection is, besides mass loss, the most uncertain area in stellar evolution. Not only is the actual energy transport rate (i.e., the mixing length) uncertain, but also the very criterion to use for deciding whether the energy transport takes place by convection or radiative diffusion occurs. We therefore repeat some of the discussion on basic stellar convection theory, where we follow KW.

We start, as usual, by considering a gas blob at a radius r with density $\rho(r)$, temperature $T(r)$ and pressure $p(r)$, equal to the surrounding thermodynamic quantities. We then displace this to a radius $r + \Delta r$. At this point the surrounding values of these parameters are $\rho(r) + \Delta\rho$, $T(r) + \Delta T$ and $p(r) + \Delta p$. The difference in density between the blob and the surrounding density is therefore

$$D\rho = \left[\left(\frac{d\rho}{dr}\right)_b - \left(\frac{d\rho}{dr}\right)_s\right] \Delta r \tag{1.34}$$

For stability we require that $D\rho > 0$, or

$$\left[\left(\frac{d\rho}{dr}\right)_b - \left(\frac{d\rho}{dr}\right)_s\right] > 0 \tag{1.35}$$

The problem with applying Eq. (1.35) is that the density-derivatives do not appear in the basic equations of stellar structure, and are therefore difficult to relate to the basic physical parameters at that point. Using the equation of state $\rho = \rho(p, T, \mu)$ we therefore translate the derivate into

$$\frac{d\rho}{\rho} = \alpha\frac{dp}{p} - \delta\frac{dT}{T} + \phi\frac{d\mu}{\mu} \tag{1.36}$$

where

$$\alpha \equiv \frac{\partial \ln \rho}{\partial \ln p}|_{T,\mu}; \delta \equiv -\frac{\partial \ln \rho}{\partial \ln T}|_{p,\mu}; \phi \equiv \frac{\partial \ln \rho}{\partial \ln \mu}|_{p,T} \tag{1.37}$$

Now we assume that the pressure in the blob quickly adjusts itself to the surroundings, i.e., $Dp = 0$. We also assume that the composition in the blob does not change as it rises. The *surrounding* matter may, however, have a finite composition gradient. Using Eq. (1.36) in Eq. (1.35) we therefore get

$$-\left(\frac{\delta}{T}\frac{dT}{dr}\right)_b + \left(\frac{\delta}{T}\frac{dT}{dr}\right)_s - \left(\frac{\phi}{\mu}\frac{d\mu}{dr}\right)_s > 0 \tag{1.38}$$

We now multiply this with the pressure scale height,

$$h = -\frac{dr}{d\ln p} = \frac{p}{g\rho}, \tag{1.39}$$

where we have used the equation of hydrostatic equilibrium. Obviously, $h > 0$, and we can translate Eq. (1.38) into

$$\left(\frac{d\ln T}{d\ln p}\right)_s < \left(\frac{d\ln T}{d\ln p}\right)_b + \frac{\phi}{\delta}\left(\frac{d\ln \mu}{d\ln p}\right)_s \tag{1.40}$$

This is the *Ledoux criterion* for convective stability. If we ignore the last term in Eq. (1.40) we get the original *Schwarzschild criterion.* It is obvious that the difference is only important in regions where the chemical composition varies radially.

It is usual to write Eq. (1.40) in a more compact form using the gradients

$$\nabla \equiv \left(\frac{d\ln T}{d\ln p}\right)_s ; \nabla_b \equiv \left(\frac{d\ln T}{d\ln p}\right)_b ; \nabla_\mu \equiv \left(\frac{d\ln \mu}{d\ln p}\right)_s \tag{1.41}$$

so that Eq. (1.40) becomes

$$\nabla < \nabla_b + \frac{\phi}{\delta}\nabla_\mu \tag{1.42}$$

Therefore, the surrounding temperature gradient should be less than that of the blob, with a correction for the chemical gradient.

To make this more concrete, we test a region for stability assuming that the energy transport is by radiation. Then, using the equation of radiative diffusion, we get for the surrounding gas

$$\nabla = \nabla_{rad} = \frac{3}{16\pi acG}\frac{\kappa Lp}{mT^4} \tag{1.43}$$

where κ is the Rosseland mean opacity, $L(r)$ the local luminosity and $m(r)$ the mass from the center. If we, further, assume that the blob behaves adiabatically, as is likely, we have $\nabla_b = \nabla_{ad}$ and therefore

$$\nabla_{rad} < \nabla_{ad} + \frac{\phi}{\delta}\nabla_\mu \tag{1.44}$$

The mixing of the chemical composition by convection can be treated as a diffusion process according to

$$\frac{\partial Y_i}{\partial t} = \frac{\partial}{\partial M}\left[(4\pi r^2\rho)^2 D\frac{\partial Y_i}{\partial M}\right] . \tag{1.45}$$

where Y_i is the abundance of ion i. The diffusion coefficient D is in the mixing length formalism given by

$$D = \frac{1}{3}V_{\mathrm{c}}l \tag{1.46}$$

where l is the mixing length, usually taken as αh, with h being the pressure scale height, and $\alpha \approx 1$. The convective velocity is

$$V_{\mathrm{c}} = \frac{1}{2}\left(\frac{GM}{\rho r^2}\Delta\nabla\rho\right)^{1/2} l \tag{1.47}$$

In this expression $\Delta\nabla\rho$ is the *excess* gradient above the adiabatic gradient. (Consider the meaning of this! see e.g., KW)

The treatment of convection has several problems and uncertainties, particularly at the interfaces between different zones. In these regions the composition gradient is usually important and it makes a big difference if one is using the Schwarzschild or Ledoux criterion for the convection. Usually regions of this kind, which are stable by the Schwarzschild criterion but unstable by the Ledoux criterion, are referred to as *semi-convective.* The determination of the convective diffusion coefficient is in this case highly uncertain, and only a few numerical hydro simulations with limited applicability have been done to test these prescriptions.

A related problem is that of convective overshooting. This may occur if the convective elements at the boundary between a convective and radiative zone has enough inertia to penetrate the boundary between these zones, and therefore mix the material from the two zones. As with the semi-convection this process is poorly understood and is usually treated in a parametrised way. Unfortunately, the results are fairly sensitive to this treatment.

1.4 Hydrogen burning

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

$$\begin{aligned} {}^1\mathrm{H} + {}^1\mathrm{H} &\rightarrow {}^2\mathrm{H} + e^+ + \nu_e \\ {}^2\mathrm{H} + {}^1\mathrm{H} &\rightarrow {}^3\mathrm{He} + \gamma \\ {}^3\mathrm{He} + {}^3\mathrm{He} &\rightarrow {}^4\mathrm{He} + 2\,{}^1\mathrm{H} \end{aligned}$$

Summarizing

$$4\,{}^1\mathrm{H} \rightarrow {}^4\mathrm{He} + 2e^+ + 2\nu_e + 2\gamma$$

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

$$\begin{aligned} {}^3\mathrm{He} + {}^4\mathrm{He} &\rightarrow {}^7\mathrm{Be} + \gamma \\ {}^7\mathrm{Be} + e^- &\rightarrow {}^7\mathrm{Li} + \nu_e \\ {}^7\mathrm{Li} + {}^1\mathrm{H} &\rightarrow 2\,{}^4\mathrm{He} \end{aligned}$$

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

$$\begin{aligned}
{}^{7}\mathrm{Be} + {}^{1}\mathrm{H} &\rightarrow {}^{8}\mathrm{B} + \gamma \\
{}^{8}\mathrm{B} &\rightarrow {}^{8}\mathrm{Be} + e^{+} + \nu_e \\
{}^{8}\mathrm{Be} &\rightarrow 2{}^{4}\mathrm{He}
\end{aligned}$$

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

For stars with high mass the temperature will be high enough for the CNO cycle to dominate. The most important chain in this cycle is the CN cycle

$$\begin{aligned}
{}^{12}\mathrm{C} + p &\rightarrow {}^{13}\mathrm{N} + \gamma + 1.95 \text{ MeV} \\
{}^{13}\mathrm{N} &\rightarrow {}^{13}\mathrm{C} + \nu_e + e^{+} + 2.22 \text{ MeV} \\
{}^{13}\mathrm{C} + p &\rightarrow {}^{14}\mathrm{N} + \gamma + 7.54 \text{ MeV} \\
{}^{14}\mathrm{N} + p &\rightarrow {}^{15}\mathrm{O} + \gamma + 7.35 \text{ MeV} \\
{}^{15}\mathrm{O} &\rightarrow {}^{15}\mathrm{N} + \nu_e + e^{+} + 2.75 \text{ MeV} \\
{}^{15}\mathrm{N} + p &\rightarrow {}^{12}\mathrm{C} + {}^{4}\mathrm{He} + 4.96 \text{ MeV}
\end{aligned}$$

In this cycle ^{12}C only acts as a catalyst, and the end resut is the same as for the pp-chain, conerting hydrogen to helium.

The temperature where the CNO cycle takes over from the pp-cycle is $\sim 1.7 \times 10^7$ K. This corresponds to a mass of ~ 1 $\mathrm{M}_\odot$. In addition to the cycle above there are two more minor branches which may be important in the late stages. We will discuss the CNO cycle in more detail in 1.13.

1.5 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

$$\begin{aligned}
{}^{4}\mathrm{He} + {}^{4}\mathrm{He} &\rightarrow {}^{8}\mathrm{Be} \\
{}^{8}\mathrm{Be} + {}^{4}\mathrm{He} &\rightarrow {}^{12}\mathrm{C} + \gamma
\end{aligned}$$

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,

$$\begin{aligned} ^{12}\mathrm{C} + {}^{4}\mathrm{He} &\rightarrow {}^{16}\mathrm{O} + \gamma \\ ^{16}\mathrm{O} + {}^{4}\mathrm{He} &\rightarrow {}^{20}\mathrm{Ne} + \gamma \end{aligned}$$

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.

1.6 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).

1.6.1 Carbon burning

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

$$\begin{aligned} ^{12}\mathrm{C} + {}^{12}\mathrm{C} \rightarrow {}^{24}\mathrm{Mg}^{\star} &\rightarrow {}^{24}\mathrm{Mg} + \mathrm{n} - 2.6 \text{ MeV} && (1.48) \\ &\rightarrow {}^{20}\mathrm{Ne} + \alpha + 4.6 \text{ MeV } (\sim 50\%) && (1.49) \\ &\rightarrow {}^{23}\mathrm{Na} + \mathrm{p} + 2.2 \text{ MeV } (\sim 50\%) && (1.50) \end{aligned}$$

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

1.6.2 Neon burning

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

$$^{20}\mathrm{Ne} + \gamma \rightarrow {}^{16}\mathrm{O} + \alpha \qquad (1.51)$$

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

$$2\ {}^{20}\mathrm{Ne} \rightarrow {}^{16}\mathrm{O} + {}^{24}\mathrm{Mg} + 4.6 \text{ MeV} \qquad (1.52)$$

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.

1.6.3 Oxygen burning

Oxygen can burn either as photo-disintegration, $^{16}\mathrm{O}(\gamma, \alpha)^{12}\mathrm{C}$, 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

$$\begin{aligned}
^{16}\mathrm{O} + {}^{16}\mathrm{O} \rightarrow {}^{32}\mathrm{S}^{\star} &\rightarrow {}^{31}\mathrm{S} + \mathrm{n} - 1.5 \text{ MeV } (5\%) && (1.53)\\
&\rightarrow {}^{31}\mathrm{P} + \mathrm{p} + 7.7 \text{ MeV } (56\%) && (1.54)\\
&\rightarrow {}^{30}\mathrm{P} + \mathrm{d} - 2.4 \text{ MeV } (5\%) && (1.55)\\
&\rightarrow {}^{28}\mathrm{Si} + \alpha + 9.6 \text{ MeV } (34\%) && (1.56)
\end{aligned}$$

Although ^{31}P is the main excit channel, these are quickly destroyed by the reaction $^{31}\mathrm{P}(\alpha, \gamma)^{28}\mathrm{Si}$. The main products of oxygen burning are therefore ^{28}Si and ^{32}S. At $\sim 2 \times 10^9$ K the reaction rate depends on the temperature as $q \propto T^{33}$.

1.6.4 Silicon burning

Because of the high Coulomb barrier for $^{28}\mathrm{Si} + {}^{28}\mathrm{Si}$, 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-disintegrates to lighter nuclei, etc. In summary,

$$^{28}\mathrm{Si}(\gamma, \alpha)^{24}\mathrm{Mg}(\gamma, \alpha)^{20}\mathrm{Ne}(\gamma, \alpha)^{16}\mathrm{O}(\gamma, \alpha)^{12}\mathrm{C}(\gamma, 2\alpha)\alpha \ . \qquad (1.57)$$

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.

$$^{28}\mathrm{Si}(\alpha, \gamma)^{32}\mathrm{S}(\alpha, \gamma)^{36}\mathrm{Ar}(\alpha, \gamma)^{40}\mathrm{Ca}(\alpha, \gamma)^{44}\mathrm{Ti}(\alpha, \gamma)^{48}\mathrm{Cr}(\alpha, \gamma)^{52}\mathrm{Fe}(\alpha, \gamma)^{56}\mathrm{Ni} \qquad (1.58)$$

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 half of that of oxygen burning.

1.7 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

$$n_i + \gamma \leftrightarrow n_{i+1} + e^-, \tag{1.59}$$

is

$$\frac{n_{i+1} n_{\mathrm{e}}}{n_i} = \frac{G_{i+1} g_{\mathrm{e}}}{G_i} \frac{(2\pi m_{\mathrm{e}} kT)^{3/2}}{h^3} e^{-\chi_i/kT} \tag{1.60}$$

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. For a more general form one should replace m_{e} by $m_{\mathrm{e}} m_{i+1}/(m_{\mathrm{e}} + m_{i+1})$.

Exercises: Skip this!

Derive the Saha equation.

In exact analogy with this, one can relate the equilibrium densities of two different isotopes, resulting from the photodissociation or inverseley the capture of one neutron, in photodissociation balance

$$n_{Z,A} + \gamma \leftrightarrow n_{Z,A-1} + n \tag{1.61}$$

by

$$\frac{n_{Z,A-1} n_{\mathrm{n}}}{n_{Z,A}} = \frac{2G_{Z,A-1}}{G_{Z,A}} \frac{(2\pi m_{Z,A-1} m_{\mathrm{n}} kT)^{3/2}}{h^3 m_{Z,A}^{3/2}} e^{-Q_{\mathrm{n}}/kT} \tag{1.62}$$

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

$$\frac{n_{Z,A-1} n_\mathrm{n}}{n_{Z,A}} = \frac{2G_{Z,A-1}}{G_{Z,A}} \left(\frac{A-1}{A}\right)^{3/2} \theta\ e^{-Q_\mathrm{n}/kT} \tag{1.63}$$

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

Similarly, we can remove one proton from the nucleus we produced in reaction (1.61) by photodissociation according to

$$n_{Z,A-1} + \gamma \leftrightarrow n_{Z-1,A-2} + p \tag{1.64}$$

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

$$\frac{n_{Z-1,A-2} n_\mathrm{p}}{n_{Z,A-1}} = \frac{2G_{Z-1,A-2}}{G_{Z,A-1}} \left(\frac{A-2}{A-1}\right)^{3/2} \theta\ e^{-Q_\mathrm{p}/kT} \tag{1.65}$$

where $Q_\mathrm{p} = (m_{Z-1,A-2} + m_\mathrm{p} - m_{Z,A-1})c^2$.

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

$$n_{Z,A} = G_{Z,A} \frac{A^{3/2} n_\mathrm{p}^Z n_\mathrm{n}^{A-Z}}{2^A} \theta^{1-A}\ e^{Q_{Z,A}/kT} \tag{1.66}$$

where now $Q_{Z,A} = (Zm_\mathrm{p} + (A-Z)M_\mathrm{n} - m_{Z,A})c^2$ is the total binding energy of the nucleus.

Exercise: Do this step!

This needs to be complemented with an equation relating the total number of neutrons and protons. To supply this we note that, in the absence of beta-decays, the total number of neutrons and protons is conserved, as well as the number of electrons.

The total number of electrons per volume is

$$n_\mathrm{e} = \sum_i Z_i n(Z_i, A_i) \tag{1.67}$$

and the total mass density

$$\rho = m_\mathrm{u} \sum_i A_i n(Z_i, A_i), \tag{1.68}$$

where $m_\mathrm{u} = 1/N_\mathrm{A}$ is the atomic mass unit. If we now define X_i as the abundance of the isotope i by mass, we have

$$X_i = \frac{n_i A_i m_\mathrm{u}}{\rho} \ . \tag{1.69}$$

Eq. (1.68) can then be written as

$$\sum_i X_i = 1 \tag{1.70}$$

and Eq. (1.67) as

$$Y_\mathrm{e} = \sum_i Z_i \frac{X_i}{A_i} \tag{1.71}$$

where we have defined the molar electron fraction of the electrons, by $Y_\mathrm{e} \equiv n_\mathrm{e}/(\rho/m_\mathrm{u}) = N_\mathrm{A} n_\mathrm{e}/\rho$. Instead of Y_e one often uses the neutron excess

$$\eta \equiv \frac{n_\mathrm{n} - n_\mathrm{p}}{n_\mathrm{n} + n_\mathrm{p}} \approx 1 - 2Y_\mathrm{e} \tag{1.72}$$

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. (1.66).

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. 4 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 ^{4}He, as photo-disintegration of ^{56}Fe becomes important.

To illustrate the last point further, we show in Fig. 5 the regions in temperature and density where ^{4}He, ^{54}Fe, and ^{56}Ni, respectively, dominate. As temperature increases, photo-disintegration becomes more and more important, causing ^{4}He to dominate.

1.8 Neutrino cooling

In the advanced burning stages cooling by neutrinos play an increasingly important role. At temperatures of the order of $m_\mathrm{e}c^2/k \sim 5 \times 10^9$ K electron

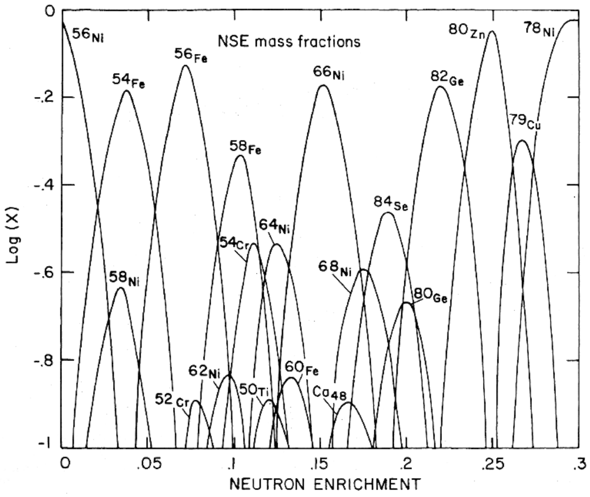

Figure 4: Abundances of iron peak elements at 3.5×10^9 K as function of the neutron excess. From left to right the isotopes are ^{56}Ni, ^{54}Fe, ^{56}Fe, ^{58}Fe. (Hartman, Woosley and el Eid 1985)

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,

$$e^- + e^+ \rightarrow \nu + \bar{\nu} \tag{1.73}$$

The typical neutrino cross section is

$$\sigma \approx 10^{-44} \left(\frac{E}{m_e c^2} \right)^2 \ \mathrm{cm}^{-2}. \tag{1.74}$$

The electron-positron annihilation cross section is of the order of the Thompson cross section, $\sigma_T = 0.665 \times 10^{-24}\mathrm{cm}^2$ (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 is given by

$$\epsilon_\nu = n_- n_+ \int \sigma v E dE \tag{1.75}$$

where v is the relative velocity of the electron and positron and E the energy in the center of mass system. The electron and positron densities can be

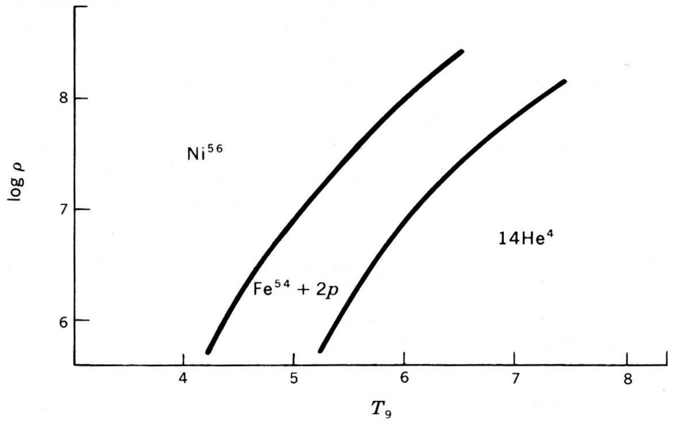

Figure 5: Abundances of ^{4}He, ^{54}Fe, and ^{56}Ni as function of the temperature.

obtained from their thermal equilibrium densities as

$$n_{-,+} \approx \pm\frac{n_\mathrm{e}^0}{2} + \sqrt{(\frac{n_\mathrm{e}^0}{2})^2 + n_1^2}, \tag{1.76}$$

where n_e^0 is the electron density in the absence of pair production, i.e., $n_\mathrm{e}^0 = n_- - n_+$ and

$$n_1 \approx 1.5 \times 10^{29} T_9^{3/2} \exp(-5.93/T_9)\ \mathrm{cm}^{-3}\ . \tag{1.77}$$

Exercise:
Show this!

As a rough approximation to the energy loss rate one finds at $T < 10^9$ K

$$\epsilon_\nu \approx 4.9 \times 10^{18} T_9^3 \exp^{-11.86/T_9}\ \ \mathrm{erg\ cm}^{-3} \tag{1.78}$$

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

$$\epsilon_\nu \approx 4.5 \times 10^{15} T_9^9\ \mathrm{erg\ cm}^{-3} \tag{1.79}$$

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* is important in many circumstances. Although quantitatively complicated, the basic ideas may be understood from simpler arguments. While a single photon in a vacuum can not produces a neutrino – anti-neutrino pair because of momentum conservation, the situation is different in a dense medium. In that case the propagation of a photon is governed by the dispersion relation $\omega^2 = k^2c^2 + \omega_{pl}^2$, where $\omega_{pl} = (4\pi n_{\rm e} e^2/m_{\rm e})^{1/2} = 5.6 \times 10^4 n_{\rm e}^{1/2}$ Hz is the plasma frequency.

This dispersion relation shows that the photon behaves in the plasma as if it has an excess energy in relation to its momentum, kc, compared to when propagating in a vacuum (cf., $E^2 = p^2c^2 + m^2c^4$). This extra energy can be used to create a neutrino–anti-neutrino pair. Plasma neutrino cooling is especially important at very high densities.

A further processes which may be important is the *photo-neutrino* process

$$\gamma + e^- \rightarrow e^- + \nu + \bar{\nu}, \tag{1.80}$$

which is similar to Compton scattering, but where the photon produces a neutrino – anti-neutrino pair. This is important at low temperatures and low densities.

In Fig. 6 we show the dominant neutrino cooling processes in the $\rho - T$ plane for a gas dominated by carbon. The situation is, however, not very different from that of other heavy element compositions.

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.

1.9 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 $\mathrm{M}_\odot$ and a 25 $\mathrm{M}_\odot$ ZAMS star, including mass loss (from WHW02). These models include mass loss (see below), explaining the low masses for the later stages 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

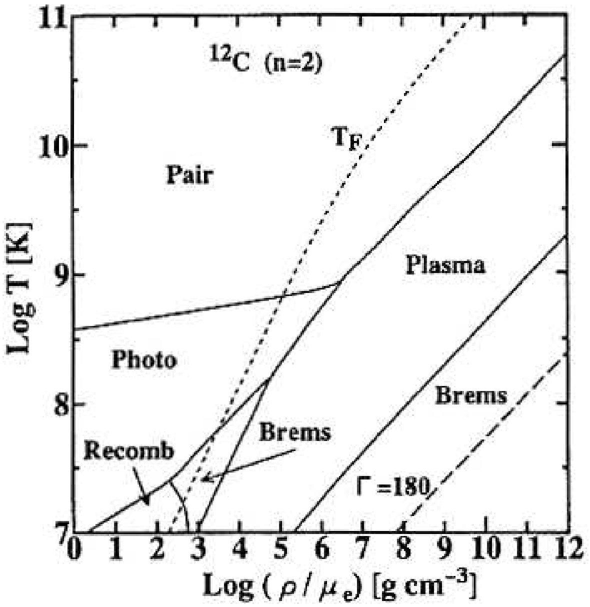

Figure 6: Dominant neutrino cooling processes in the $\rho - T$ plane for a carbon dominated gas (Itoh 1996).

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 by orders of magnitude, as Fig. 7 clearly shows. The instantaneous escape of these from the center means that the temperature will be set by a balance of neutrino cooling and nuclear heating. Because the nuclear heating increases with a very high power, $q \propto T^{30}$ or faster, while neutrino cooling is less sensitive, $\epsilon_\nu \propto T^9$ or so, the heating can balance the neutrino cooling as long as there remains some unburned fuel in the core. The much steeper temperature dependence of the nuclear heating compared to the neutrino cooling also means that the nuclear burning will occur at a nearly unique temperature for the different burning stages, independent of the mass of the star. This is clearly seen from Fig. 8, which shows the evolution of a core contracting as $\rho = 10^6 T_9^3$ g cm^{-3}, typical of a massive core.

We can from this figure estimate the typical time scale for the burning

Table 1: **Burning stages for a** 15 M$_\odot$ **star (WHW02)**

Fuel	Ashes	T	ρ	M	L	R	τ
		10^8 K	g cm^{-3}	M$_\odot$	10^3 L$_\odot$	R$_\odot$	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

Table 2: **Same as above for a** 25 M$_\odot$ **star**

Fuel	Ashes	T	ρ	M	L	R	τ
		10^8 K	g cm^{-3}	M$_\odot$	10^3 L$_\odot$	R$_\odot$	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

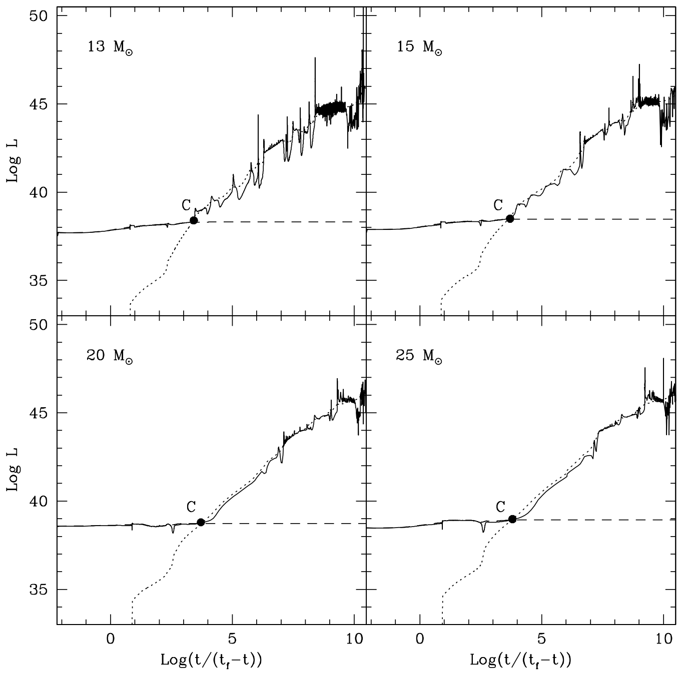

Figure 7: Neutrino luminosity of stars with mass $13 - 25$ $M_\odot$ compared to the nuclear energy generation and the photon luminosity (Limongi et al. 2002).

stages. The energy produced from each gram is q_i. From the figure we can estimate the energy loss rate, ϵ_i, at the temperature where balance between heating and cooling occurs. If the fractional abundance is Y_i the time scale is then

$$\tau_i \approx \frac{Y_i q_i}{\epsilon_i} \tag{1.81}$$

In Table 3 we give the approximate values of q_i and Y_i, as well as for ϵ_i for the different burning stages from Fig. 8. The final column then gives the estimated life times from Eq. (1.81). Comparing with Tables 1 and 2 we find a good agreement with the more realistic models. It is clear that because of the steep dependence of the neutrino cooling with temperature, the life time decreases extremely fast for the last burning stages.

After carbon burning the diffusion time for the photons is much longer than the duration of these stages, and the core evolves independently from the envelope. Unless some kind of shell flash or similar occurs, the envelope is essentially decoupled from the core.

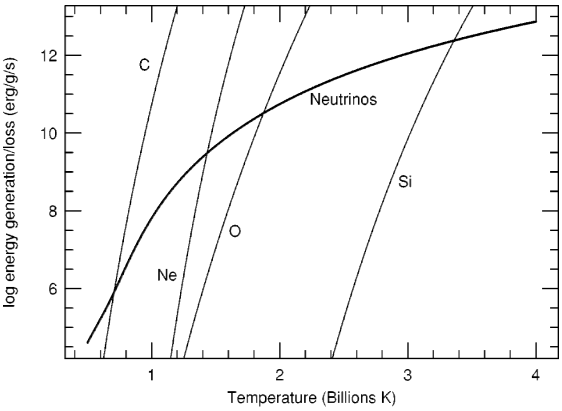

Figure 8: Nuclear heating rate and neutrino cooling rate for a core contracting as $\rho = 10^6 T_9^3$ g cm^{-3} (WHW02).

Table 3: **Estimates of the duration of the advanced burning stages**

Burning stage	$q_i/10^{17}$ erg g^{-1}	Y_i	T_i 10^9 K	$\rho/10^6$ g cm^{-3}	ϵ_i erg g^{-1} s^{-1}	τ_i s
C	4.0	0.2	0.7	0.3	7.1×10^5	3.6×10^3 yrs
Ne	1.1	0.7	1.4	2.7	3.0×10^9	0.81 yrs
O	5.0	0.2	1.9	6.9	3.5×10^{10}	30 days
Si	1.9	0.5	3.4	39.3	2.4×10^{12}	0.5 days

1.10 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$_\odot$ yr^{-1} and the wind velocities are $1,000 - 3,000$ 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 $\sim 10^4$ yrs 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} $\mathrm{M}_\odot$ yr^{-1} and the wind velocities are $10 - 50$ km s^{-1}. In the superwind phase mass loss rates as high as $10^{-4} - 10^{-3}$ $\mathrm{M}_\odot$ yr^{-1} may occur. The duration of this phase must obviously be only of the order of a few times 10^4 yrs.

A useful formula which summarizes the mass loss rates on the main sequence and in the red supergiant stage is given by Nieuwenhuijzen & de Jager (1990),

$$\dot{M} = 9.6 \times 10^{-15} \left(\frac{L}{\mathrm{L}_\odot}\right)^{1.42} \left(\frac{M}{\mathrm{M}_\odot}\right)^{0.16} \left(\frac{R}{\mathrm{R}_\odot}\right)^{0.81} \ \mathrm{M}_\odot\ \mathrm{yr}^{-1} \tag{1.82}$$

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 $\sim 10^{-5}$ $\mathrm{M}_\odot$ 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 main sequence phase, with a mass loss rate of $\sim 10^{-6}$ $\mathrm{M}_\odot$ yr^{-1} lasts for $\sim 10^6$ yrs, the RSG phase for $\sim 10^5$ yrs and the WR phase for a comparable period, and that the mass loss rates in these two phases are $10^{-5} - 10^{-4}$ $\mathrm{M}_\odot$ yr^{-1}, means that stellar winds will have a major influence on the evolution.

1.11 Evolution in the HR diagram

Mass loss is crucial both for the observational properties the appearance of the star and for the internal structure. In addition, as we will see later, the resulting supernova properties also depend strongly on the mass of the hydrogen and helium envelope. Because mass loss is increasingly important with mass, the effects increase strongly with mass. In Fig. 9 we show the evolution in the HR diagram of a 60 $\mathrm{M}_\odot$ 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 evolutionary

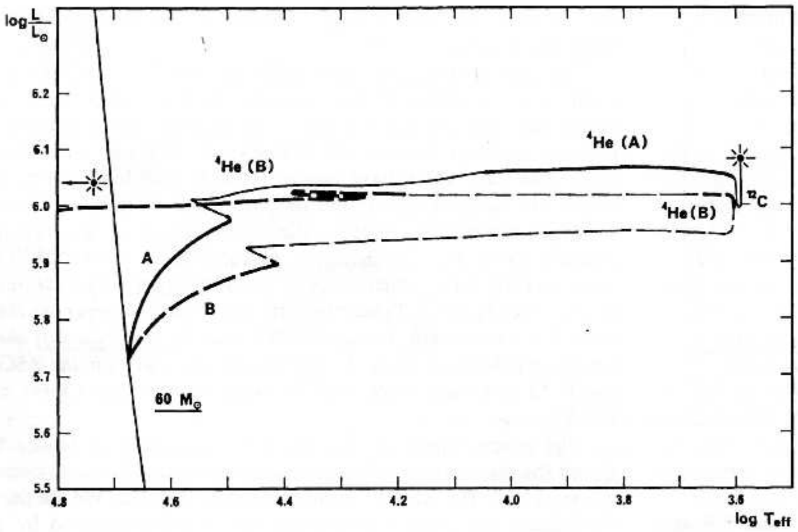

Figure 9: Evolution of a 60 $M_\odot$ star with (dashed line) and without (solid line) mass loss (Maeder 1981)

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

$$\mathrm{O} \rightarrow \mathrm{BSG} \rightarrow \mathrm{RSG} \rightarrow \mathrm{WR} \tag{1.83}$$

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.

$$\mathrm{O} \rightarrow \mathrm{LBV} \rightarrow \mathrm{WR} \tag{1.84}$$

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. 10.

1.12 Rotation

Observationally, it is known that massive stars on the main sequence are rotating rapidly. Typical surface velocities are ~ 200 km s^{-1}. This is not

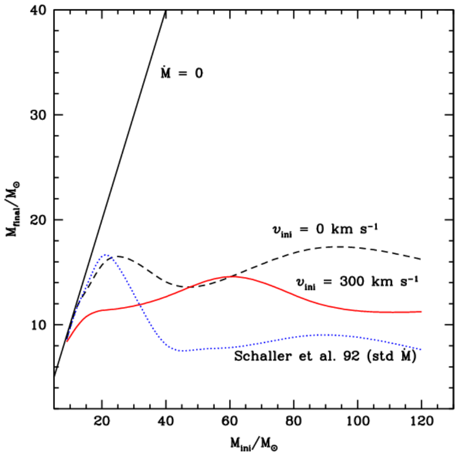

Figure 10: The effects of rotation and mass loss on the final mass of massive stars. The solid line gives the final mass for a ZAMS rotational velocity of 300 km s^{-1}, while the dashed line gives the zero velocity result, only including mass loss (Meynet & Maeder 2003)

very far from break-up velocity, given approximately by

$$V_{\rm c} \approx \left(\frac{GM}{R}\right)^{1/2} \tag{1.85}$$

Because of convection and the Eddington-Sweet circulation, rigid rotation is likely to be established on the main sequence. If this would persist into the advanced burning phases, it could seriously affect the final stages, including the collapse phase.

Because of convection and in radiative layers the Eddington-Sweet circulation, angular momentum can, however, be transported outwards. If the star in addition loses mass in the form of a wind, as is nearly always the case, this wind could transport angular momentum away from the star.

Rotation affects the evolution in several ways. First, rotation induces circulation and mixing of the stellar interior in a way resembling that of increased convection. Secondly, the mass loss rates are affected by the rotation. Because of the centrifugal force the star becomes oblate. The larger radius in the equatorial plane leads to a lower effective temperature and flux compared to the polar direction (by von Zeipel's theorem). Because the mass loss rate increases with the luminosity, this will therefore be higher in the polar direction than in the equatorial.

The conservation of angular momentum implies that the angular velocity decreases dramatically as the star evolves to a RSG. In Fig. 11 this can clearly be seen in the H-envelope at $M \gtrsim 4\ \mathrm{M}_\odot$, where the angular velocity drops by several orders of magnitude after H-depletion, as the star evolves into a RSG. The core, however, increases its angular velocity because of contraction, and may end up rotating close to the critical velocity. This is especially interesting in connection to the rotational velocities of pulsars, where the periods are down to milliseconds. As we have remarked, angular momentum may, however, be lost in especially the red supergiant stage and in the Wolf-Rayet phase due to mass loss. The efficiency of this depends sensitively on the magnetic field.

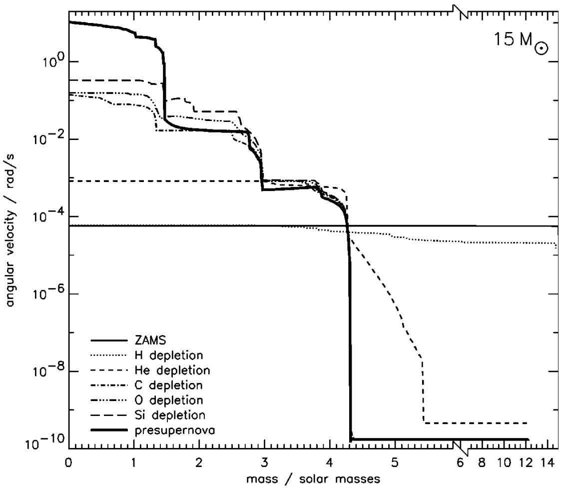

Figure 11: Angular velocity as function of mass for a 15 $\mathrm{M}_\odot$ star in different evolutionary stages. The initial surface rotation was 200 km s^{-1}. Note the drop in angular velocity of the hydrogen envelope after the helium burning stage, as a result of the evolution to the red supergiant stage.

In Fig. 12 we show the evolutionary tracks in the HR diagram with and without rotation. In both cases mass loss is included with the same, most recent rates. It is here apparent that especially for the most massive stars the late evolution is substantially altered by the rotation. For non-rotating stars the lower mass limit to evolve to a WR-star is $\sim 37\ \mathrm{M}_\odot$. With rotation, even a 22 $\mathrm{M}_\odot$ star evolves all the way to the Wolf-Rayet stage, rather than ending its life as a red supergiant. This is also the case for stars more massive than this. The most massive stars, with $M \gtrsim 60\ \mathrm{M}_\odot$, may with rotation skip even the LBV phase, and evolve directly from the MS to the

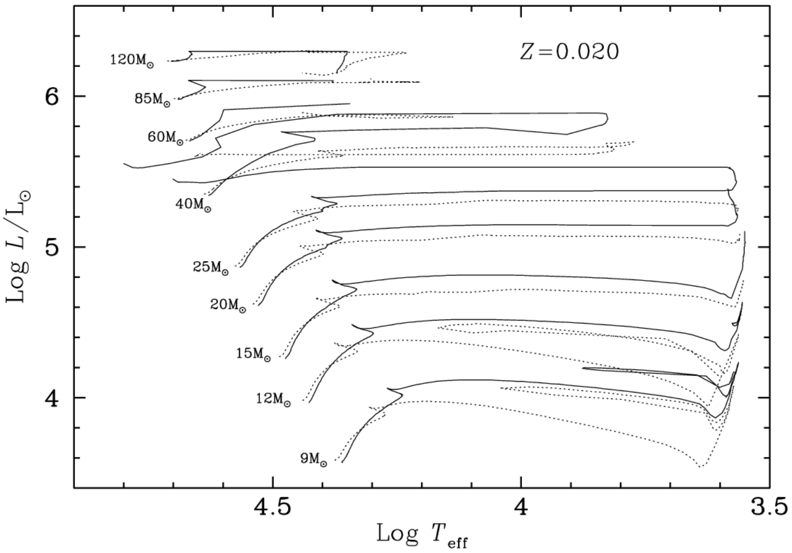

Figure 12: Evolutionary tracks for stars more massive than 9 $M_\odot$ with (solid lines) and without rotation. (Meynet & Maeder 2003)

WR phase.

Because of the increased mass loss rate in the rotational case, the final masses of the stars at the beginning of carbon burning will in general be smaller than in the non-rotational case, as is seen in Fig. 10. The life time in the Wolf-Rayet phase will also increase by a factor of about two. For a non-rotating 60 $M_\odot$ star the total life time is 4.0 Myrs, of which it spends 0.37 Myrs as a Wolf-Rayet star. The corresponding numbers for a star with ZAMS rotational velocity of 300 km s^{-1} are 4.67 Myrs and 0.75 Myrs, respectively.

A further observationally important effect is that because of rotational mixing, nuclear burning products may become visible at the surface already in the hydrogen and helium burning stages. This is most apparent from the presence of CNO burning products already on the main sequence and helium burning stages. Fig. 13 shows the evolution of the surface abundances as function of the total mass of the star as it loses mass, or equivalently as function of time. We note the more efficient mixing of the rotating model which shows CNO processing products (^{14}N) already in the BSG stage. Just before collapse it has evolved to a Wolf-Rayet WC star, with a large ^{12}C abundance from He-burning in the rotating case, while the non-rotating case only shows enhanced ^{14}N, although also a Wolf-Rayet star (in this case a

WN star).

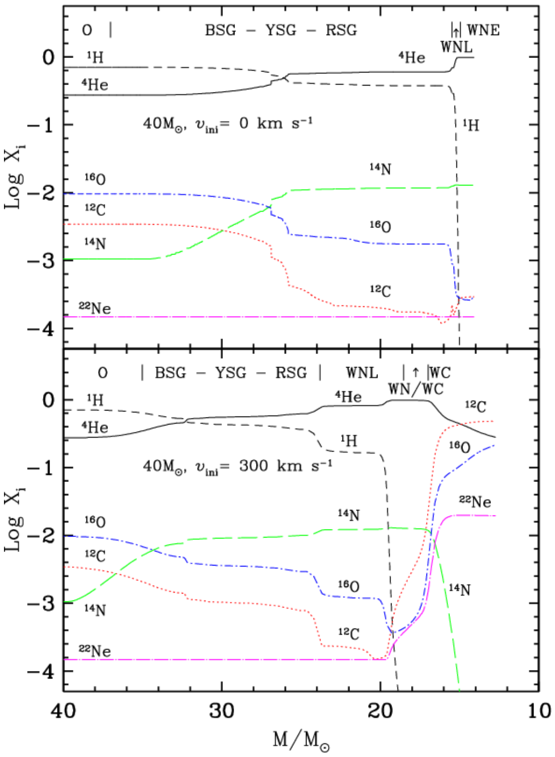

Figure 13: Surface abundances for a 40 M$_\odot$ ZAMS star including mass loss and rotation. On the x-axis is the total mass of the star and on the y-axis the surface abundance at these epochs. The different positions in the HR diagram are indicated by its spectral class. (Meynet & Maeder 2003)

1.13 Observational indications of CNO burning

The most characteristic signatures of the nucleosynthesis during the hydrogen burning phase are the presence of enhanced helium abundances and signatures of CNO processing, as well as the ratio of blue to red supergiants.

The most important chain in the CNO cycle is CN cycle

$$^{12}\mathrm{C}(p,\gamma)^{13}\mathrm{N}(e^+,\nu)^{13}\mathrm{C}(p,\gamma)^{14}\mathrm{N}(p,\gamma)^{15}\mathrm{O}(e^+,\nu)^{15}\mathrm{N}(p,\alpha)^{12}\mathrm{C} \tag{1.86}$$

(see Fig. 14). In this cycle the ^{14}N nucleus is most resistant to destruction, because of a high Coulomb barrier and the fact that an electromagnetic

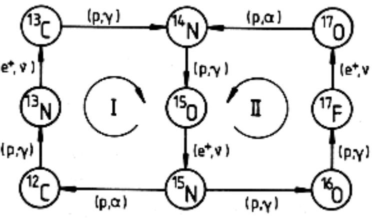

Figure 14: The CNO bi-cycle. (Rolfs & Rodney 1988)

reaction is involved, rather than a strong. Therefore, this nucleus will be the 'bottle neck' in the cycle, and if the cycle reaches equilibrium most of the ^{12}C is converted into ^{14}N. In about 0.1 % of the cases ^{15}N instead of an α decay undergoes a proton capture in which case $^{15}\mathrm{N}(p,\gamma)^{16}\mathrm{O}$. In addition to the CN cycle, there is a second cycle which also involves ^{16}O.

$$^{16}\mathrm{O}(p,\gamma)^{17}\mathrm{F}(e^{+},\nu)^{17}\mathrm{O}(p,\alpha)^{14}\mathrm{N} \tag{1.87}$$

The importance of this cycle is that it creates some less abundant isotopes, like ^{17}F and ^{17}O, and in particular that it converts ^{16}O into ^{14}N. The efficiency of this depends on how long the cycle runs. In Fig. 15 the abundances of the most important isotopes are shown as function of the number of protons consumed per initial CNO nucleus, or equivalently as function of time. The latter scale does, however, depend on the mass of the star. It is here seen that while most ^{12}C is converted into ^{14}N fairly quickly, this takes a much longer time for the ^{16}O.

The end-result of the CNO bi-cycle is that most of the ^{12}C is converted to ^{14}N, and if it runs sufficiently long also the ^{16}O.

Already on the main sequence one observes enhanced N/C ratios, indicative of CNO processing. This is not expected unless both rotation and mass loss are important, as Fig. 13 shows. Also in supergiants and Wolf-Rayet stars an increased N/C is present.

In connection to supernovae, signatures of CNO processing have also seen in the spectra of several objects. As an example, we show in Fig. 16 a spectrum taken with HST of SN 1998S. A number of strong emission lines of C III-IV and N III-V are seen. The fact that these lines are much narrower (FWHM $\lesssim 500$ km s^{-1}) than from the supernova ejecta itself, with a velocity of ~ 7000 km s^{-1}, shows that they come from circumstellar gas lost by the supernova progenitor before the explosion. From an analysis of the line ratios one finds a ratio of N/C$\sim$ 6, which is a factor of $\sim$ 24 higher than the cosmic abundance ratio, and therefore shows that CNO

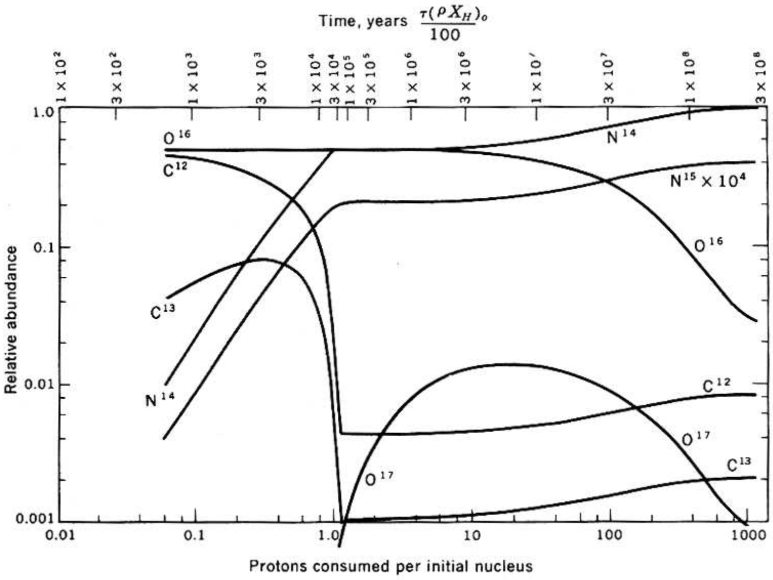

Figure 15: The CNO bi-cycle. (Caughlan 1965)

products were abundant in the star before the explosion. We will discuss the implications of this for the different supernova types later.

1.14 Uncertainties from nuclear rates and convection

While most nuclear reaction rates are well determined experimentally, there are a few rates where the uncertainties are so large that they affect the qualitative outcome of the evolution and nucleosynthesis. The prime example is the $^{12}\mathrm{C}(\alpha,\gamma)^{16}\mathrm{O}$ rate. This reaction, which is mainly important during helium burning, directly determines the relative abundances of carbon and oxygen in both low mass and high mass stars. This reaction proceeds through some resonances at very low energy, and is therefore very difficult to measure experimentally. Extrapolations from higher energies are therefore necessary. Currently, the S factor at 300 keV is only determined to be within the range 100–200 keV barn (see WHW02 III A.2 for a detailed discussion). Because this rate determines the extent of the oxygen zone, and indirectly all zones inwards from the carbon zone, it not only affects the carbon and oxygen abundances, as can be seen in Fig. 17, but the whole structure of the inner core. In particular, the mass of the iron core is larger by $0.05-0.1\ \mathrm{M}_\odot$ for a $^{12}\mathrm{C}(\alpha,\gamma)^{16}\mathrm{O}$ rate larger by a factor of ~ 3. Although not very large,

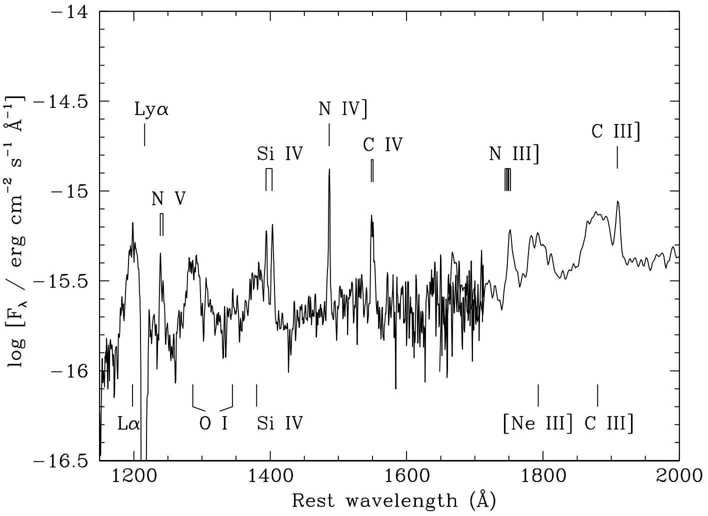

Figure 16: Spectrum of the Type IIn supernova SN 1998S taken with HST. Note the narrow, very strong N III], N IV] and N V lines, and the weaker C III] and C IV lines.

this difference may be crucial for getting a successful explosion or not.

There are additional uncertainties connected to the treatment of convection, especially that of semi-convection, when the distinction between the Schwarzschild criterion and the Ledoux criterion is important. As we saw in Chap. 1.3, this occurs where there are gradients in the chemical composition, i.e., at the interfaces between different burning shells.

There are at least three evolutionary stages when semi-convection matters. Fig. 18 shows the semi-convective regions during the hydrogen burning stage as extensions from the convective central core. During the evolution the convective region recedes, and a gradient in the helium to hydrogen is created, which results in a semi-convective region. This mixing is important for the ignition of the hydrogen shell burning. This in turn affects whether the star becomes a blue or red supergiant. The ratio of these populations may therefore serve as a test of the treatment of semi-convection. Unfortunately, the conclusion from this is not obvious.

Semi-convection may also be important during helium burning and silicon burning . In the former case it affects the size of the carbon-oxygen core, and therefore the whole nucleosynthesis. During the silicon burning there may be a chemical gradient from the electron concentration. The way this is treated affects the size of the final iron core, and therefore the whole

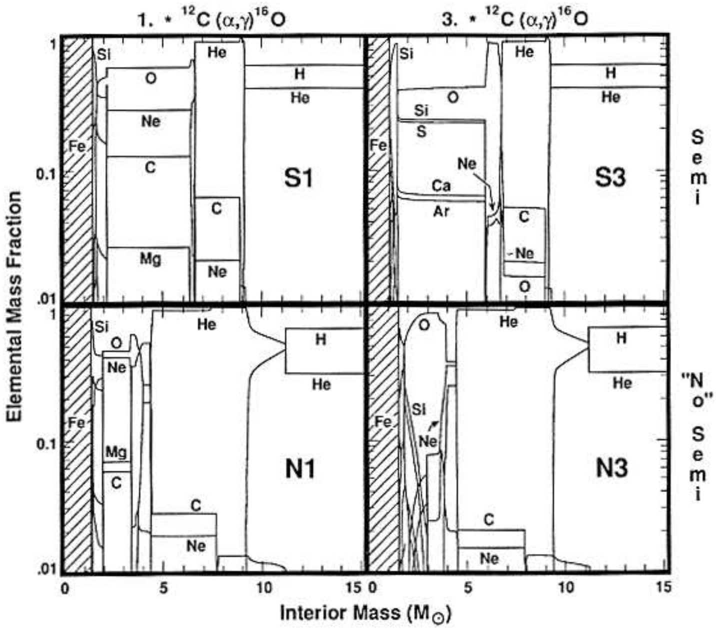

Figure 17: The effects of different rates for $^{12}\mathrm{C}(\alpha,\gamma)^{16}\mathrm{O}$ reaction (changing this rate by a factor of three), as well as the effect of including semi-convection (from Woosley et al. 1995)

explosion result, which is very sensitive to the iron core mass.

In Fig. 17 we see that the mass of the oxygen core is much larger with semi-convection included at the expense of the mass of the helium shell.

Finally, we remark that the use of mixing length theory in especially the final burning stages is doubtful. In these stages the convective and nuclear time scales become comparable, which makes the assumptions behind the mixing length prescription highly questionable (see Bazan & Arnett 1998).

1.15 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. 20 we show the abundance structure of a 15 $\mathrm{M}_\odot$ and a 25 $\mathrm{M}_\odot$ star shortly before core collapse. These are the same models for which the convective evolutions are shown in Figs. 18 and 19.

The density structure before core collapse is important for the properties

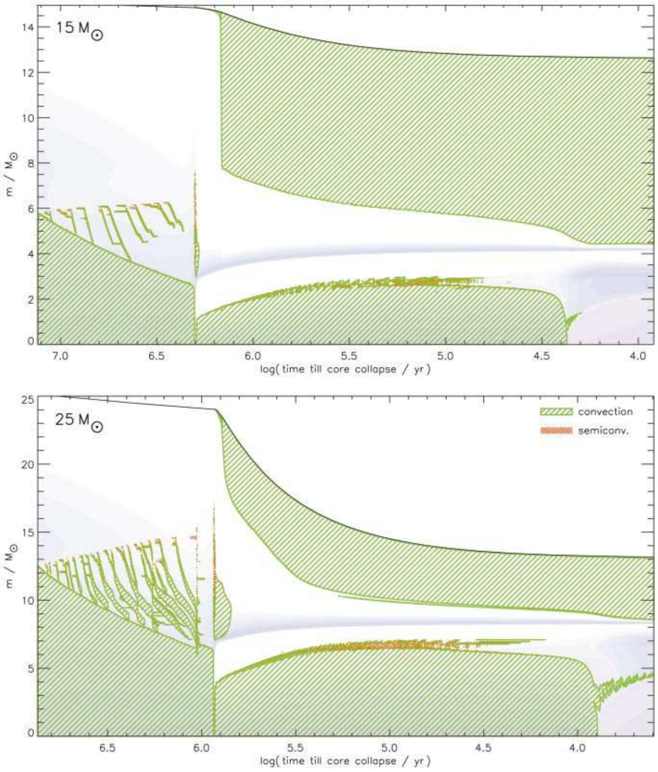

Figure 18: Convective regions for a 15 M$_\odot$ and 25 M$_\odot$ during hydrogen and helium burning stages. (WHW02)

of the supernova explosion. Fig. 21 shows the density as function of mass from the center for stars of different masses, while Fig. 22 shows the radius as function of mass. 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 ~ 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 ~ 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 ap-

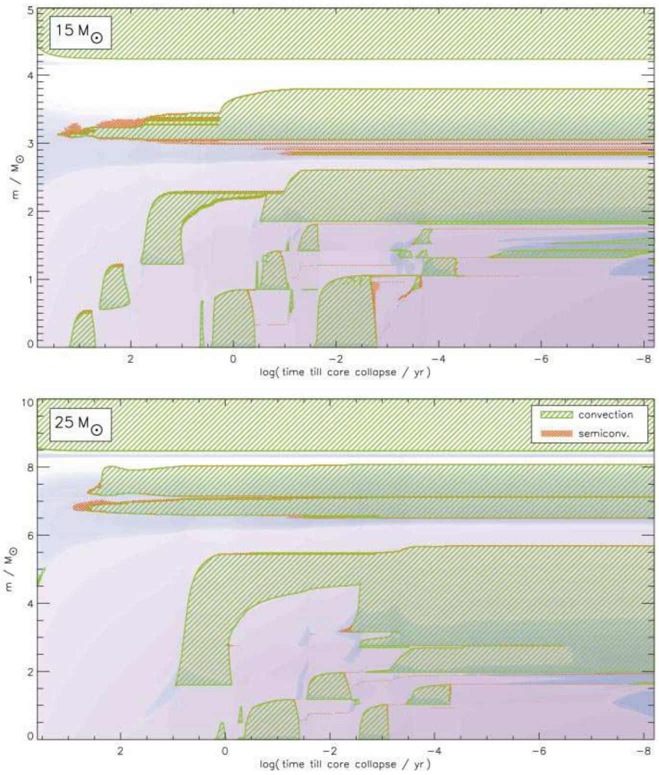

Figure 19: Convective regions for a 15 $M_\odot$ and 25 $M_\odot$ at carbon burning to core collapse. (WHW02)

parent 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.

In Fig. 23 we show the position from the center of the different burning shells for different ZAMS masses. From this we see that there is a well defined relation between the He core mass and the ZAMS mass,

$$M(\mathrm{He}) \approx 0.39[M(\mathrm{ZAMS}) - 13.] + 3.3 \ \ \mathrm{M}_\odot. \tag{1.88}$$

Consequently, the luminosity of the star before the explosion is completely determined by the He-core mass.

The mass of the Fe-core is crucial for whether the explosion will be successful or not. As we see in Fig. 23, this mass is only weakly dependent on the ZAMS mass in this interval, $13-25$ $M_\odot$. The small difference between

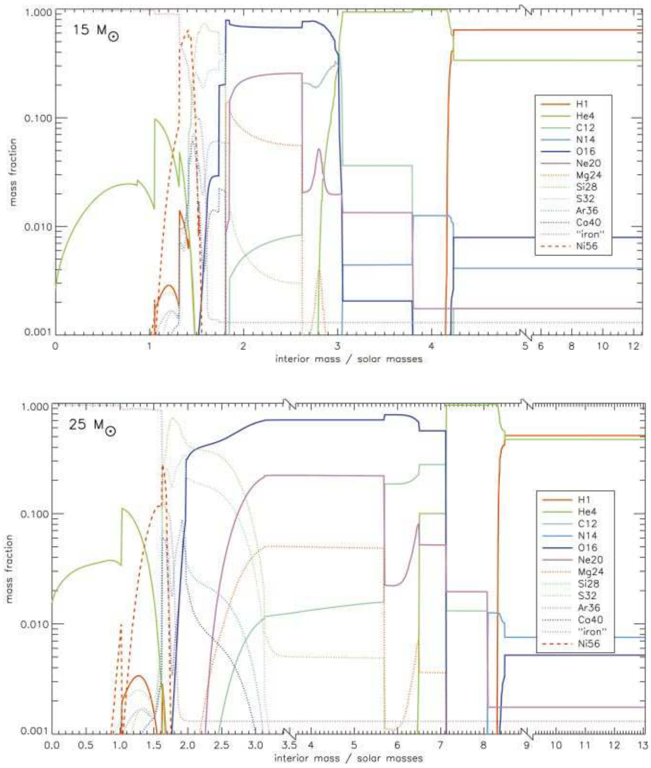

Figure 20: Abundance structure of a 15 M$_\odot$ and 25 M$_\odot$ star shortly before core collapse. (WHW02)

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.

A special case is the low mass range of the massive stars, $\lesssim 12$ M$_\odot$. Up to carbon burning they evolve in the same way as the more massive stars above. When neon burning sets in the O– Ne – Mg core is, however, degenerate in contrast to the more massive stars. Neon burning therefore becomes explosive and proceeds all the way up to nuclear statistical equilibrium. Rather than leading to a thermonuclear explosion as for Type Ia supernovae (discussed later), this results in a collapse of the core to a neutron star. After this the process is similar to that of the more massive stars, although the iron core mass is smaller. This is important for the outcome, and it is in fact this kind of progenitors which are the only ones where a successful explosion has been obtained.

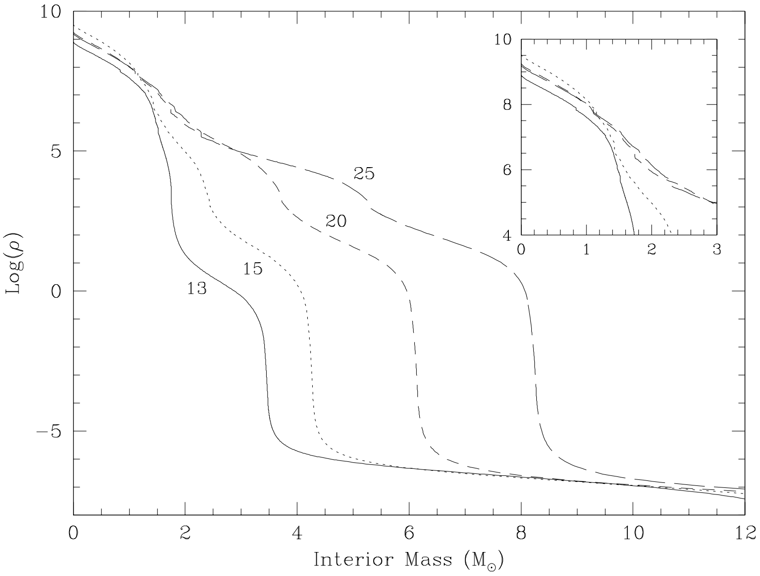

Figure 21: Density structure of four stars with ZAMS 13 − 25 M$_\odot$ star. (Limongi et al. 2002)

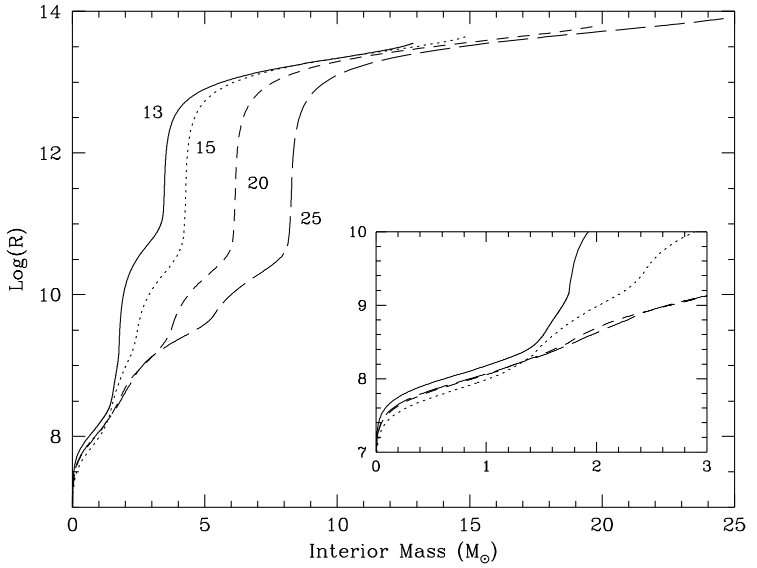

Figure 22: Radius vs mass for the same stars as in Fig. 21. (Limongi et al. 2002)

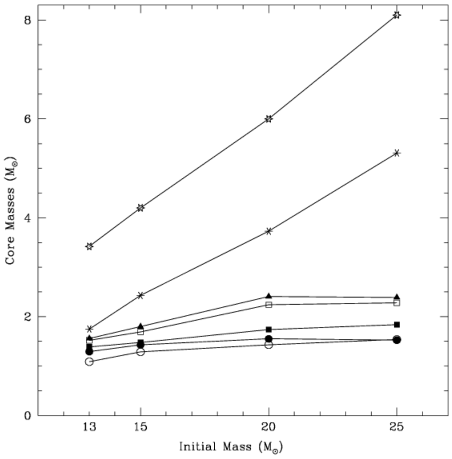

Figure 23: Mass from center of the different nuclear burning shells as function of the ZAMS mass. From top to bottom H, He, C, Ne, O, Si, Fe. (Limongi et al. 2002)

2 Core collapse and explosion

2.1 Dynamic stability and triggering of collapse

A self-gravitating body is stable to perturbations as long as its adiabatic index is larger than 4/3 (see Hans lectures (Padmanabhan 3.7.1), KW Chap. 38 or ST Chap. 6.7). 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 $\sim$ 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 and on bound protons in nuclei (K-capture),

$$e^- + p \rightarrow \nu_{\rm e} + n \tag{2.1}$$

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, $t_{\rm dyn} \sim R/v$, Assuming that the core collapses with the free-fall velocity, V_{ff}, we get

$$t_{\text{dyn}} = \frac{R}{V_{ff}} = \left(\frac{R^3}{2GM}\right)^{1/2} \tag{2.2}$$

In terms of the density this is

$$t_{\text{dyn}} = \left(\frac{3}{8\pi G\rho}\right)^{1/2} \tag{2.3}$$

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

$$t_{\text{dyn}} = 0.13\ \rho_8^{-1/2}\ \text{s} \tag{2.4}$$

and the collapse time scale is therefore of the order of a hundred milliseconds or less. 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. (2.4) is sufficient.

2.2 Neutrino trapping during core collapse

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,

$$\nu + n \rightarrow \nu + n \tag{2.5}$$

and

$$\nu + p \rightarrow \nu + p \tag{2.6}$$

and

$$\nu + (Z, A) \rightarrow \nu + (Z, A). \tag{2.7}$$

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. In the last process the long wavelength of the neutrinos compared to the nucleus makes the scattering of the neutrinos coherent. Because of the coherence, the cross section of the last process is not only proportional to A, but to A^2. An analogy is Rutherford scattering where the cross section is proportional to the charge squared of the particles.

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

$$\nu + e^- \rightarrow \nu + e^- \ . \tag{2.8}$$

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 $m_{\rm n}c^2 \sim 1$ GeV the cross section for the nucleon scattering is

$$\sigma_\nu = \frac{1}{4}\sigma_0 \left(\frac{E_\nu}{m_{\rm e}c^2}\right)^2 \tag{2.9}$$

where

$$\sigma_0 = \frac{4G_{\rm F}^2 m_{\rm e}^2}{\hbar^4} = 1.76 \times 10^{-44}{\rm cm}^2. \tag{2.10}$$

The cross section for scattering against bound nucleons is

$$\sigma_\nu = \frac{1}{16}\sigma_0 \left(\frac{E_\nu}{m_{\rm e}c^2}\right)^2 A^2 \left[1 - \frac{Z}{A} + \frac{Z}{A}(4\sin^2\theta_W - 1)\right]^2 \tag{2.11}$$

for $E_\nu \ll 300A^{-1/3}$ MeV. Here, θ_W is the Weinberg angle, where $\sin\theta_W \approx 0.23$. With the approximation $\sin\theta_W \approx 1/4$, the bracket becomes $(A - Z)/A = N/A$, where $N \approx A/2$ is the neutron number. Therefore, we have

$$\sigma_\nu \approx \frac{1}{64}\sigma_0 \left(\frac{E_\nu}{m_{\rm e}c^2}\right)^2 A^2 \tag{2.12}$$

for this cross section. Note again the A^2 factor from the coherence.

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

$$\lambda_\nu \approx 1.0 \times 10^8 [\frac{N^2}{6A} X_b + X_{\rm n}]^{-1} \rho_{12}^{-1} E_{\rm MeV}^{-2} {\rm cm} \tag{2.13}$$

where ρ_{12} is the density scaled by 10^{12} g cm^{-3} and $E_{\rm MeV}$ the neutrino energy in MeV. N is the number of neutrons in the average atomic nucleus A. The first term accounts for the coherent scattering of the bound nucleons with abundance X_b, while the second accounts for the scattering by the free nuclei with abundance $X_{\rm n}$. At a density of 10^{12} g cm^{-3} the nuclei have $N \sim 50$ (because of K-captures) and $N/A \sim 0.6$, so

$$\lambda_\nu \approx 2 \times 10^5 \left(\frac{E_\nu}{10 \text{ MeV}}\right)^{-2} \rho_{12}^{-1} \text{ cm} \tag{2.14}$$

The typical neutrino energy is $\sim$ 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 ΔR is

$$t_{\rm diff} = \frac{\Delta R^2}{3\lambda_\nu c} \tag{2.15}$$

Another way of seeing this more intuitively is as a random walk, where each step is λ_ν. After N step the distance it has diffused from its origin is $\Delta R = N^{1/2}\lambda_\nu$. Each step takes a time λ_ν/c, so the total time is $t_{\rm diff} = N\lambda_\nu/c$. Therefore, $\Delta R = (ct_{\rm diff}/\lambda_\nu)^{1/2}\lambda_\nu = (ct_{\rm diff}\lambda_\nu)^{1/2}$, or $t_{\rm diff} \sim \Delta R^2/(\lambda_\nu c)$.

For a uniform sphere of radius R, the distribution of the density of the diffusing particles is given by $\sin \pi R/(\pi R)$, and the diffusion time scale is given by

$$t_{\rm diff} = \frac{3R^2}{\pi^2 \lambda_\nu c} \tag{2.16}$$

(see e.g., Arnett, App. D). If we assume a uniform density sphere of mass 1.4 $M_\odot$ we have $R \sim 8.7 \times 10^6 \rho_{12}^{-1/3}$ cm, and with Eq. (2.14) and Eq. (2.16) we get

$$t_{\rm diff} = 3.9 \times 10^{-3}\ \rho_{12}^{1/3} \left(\frac{E_\nu}{10 \text{ MeV}}\right)^2 \text{ s.} \tag{2.17}$$

As an estimate of the neutrino energy we take $E_\nu \approx E_{\rm F} = 50(Y_{\rm e}\rho_{12})^{1/3}$ MeV, where we have used Eq. (1.4) and $E_{\rm F} \approx p_{\rm F}c$). With $Y_{\rm e} \approx 0.4$ we get $E_\nu \approx E_{\rm F} = 36.8\ \rho_{12}^{1/3}$ MeV. We therefore get

$$t_{\rm diff} = 5.2 \times 10^{-2}\ \rho_{12} \text{ s} \tag{2.18}$$

The diffusion time scale should be compared to the dynamical time scale, $t_{\rm dyn}$, from Eq. (2.4),

$$\frac{t_{\rm diff}}{t_{\rm dyn}} = 40\ \rho_{12}^{3/2} \tag{2.19}$$

Therefore, we find that *above a density of* $\sim 10^{11}$ g cm^{-3} *the neutrinos are 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). Rather than becoming a hot, collapsing neutron star, the core has a fairly normal lepton fraction of $Y_l \sim 0.4$, where the lepton fraction is the fraction of neutrinos plus electrons. Because of the neutrino trapping an equilibrium between reactions like $e^- + p \leftrightarrow n + \nu_\mathrm{e}$ is set up, and therefore Y_l rather than Y_e is the relevant quantity.

2.3 Equation of state

Because of neutrino trapping and the low entropy, the collapse proceeds almost adiabatically. The entropy of the nuclei is much lower than that of the free nucleons, and the nucleons remain in the nuclei during the collapse, rather than being dissolved. This in turn has the consequence that the pressure is dominated by the degenerate, relativistic electrons and the trapped neutrinos, so $p \propto Y_\mathrm{e}^{4/3}\rho^{4/3}$. The adiabatic index therefore stays at $\gamma \sim 4/3$, and the collapse proceeds until nuclear density, $\rho_\mathrm{nuc} \sim 2 \times 10^{14}$ g cm^{-3}, is reached.

Once the density is above nuclear, the equation of state suddenly becomes very stiff due to the nuclear repulsive short range forces, and the core bounces. Typically, the adiabatic index changes from $\gamma \sim 4/3$ to $\gamma \sim 2 - 3$ (see Fig. 24). This occurs at a density of $\sim 2 \times 10^{14}$ g cm^{-3}. The exact form of the equation of state is, unfortunately, not well determined above nuclear density. This is for the explosion one of the most severe uncertainties.

2.4 Collapse

As we have just seen, because of neutrino trapping and the low entropy, the collapse proceeds almost adiabatically. Numerical calculations of the collapse show that the velocity close to the center is nearly proportional to velocity, $v \propto r$ (Fig. 25). The $v \propto r$ region extends roughly out to where the collapse becomes supersonic.

Schematically, the situation is that shown in Fig. 26. The $v \propto r$ collapse is what is known as a homologous collapse, since it preserves the shape, and only changes the scaling. The reason for this is, as Goldreich & Weber (1980) have shown, that the collapse is nearly adiabatic and close to $\gamma = 4/3$. They in fact showed that a similarity solution, describing the hydrodynamics, could be found for this case. The homologous region can only extend as far

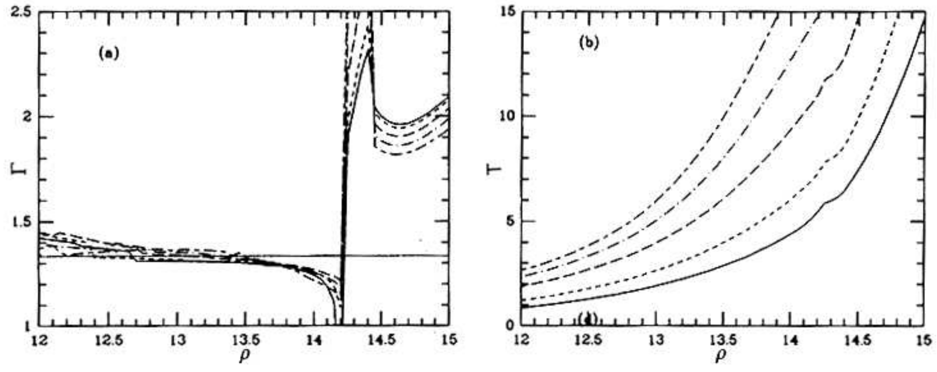

Figure 24: Left: Adiabatic index as function of density for different isentrops. Right: Temperature in MeV for different isentrops (Cooperstein & Baron 1990)

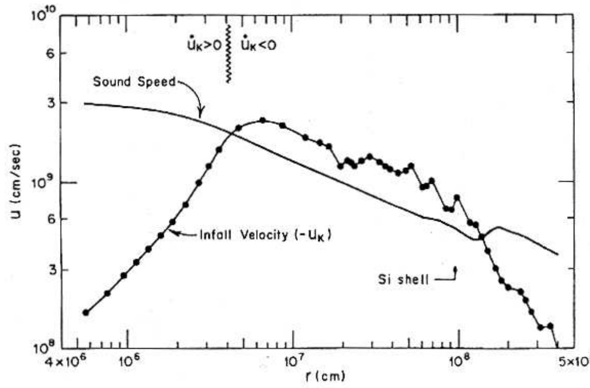

Figure 25: Velocity profile in the core in the homologous phase. The subsonic inner region extends to ~ 0.5 M$_\odot$, where the infall velocity is close to $v \propto r$. The shock will form close to the sonic point. (Arnett 1978)

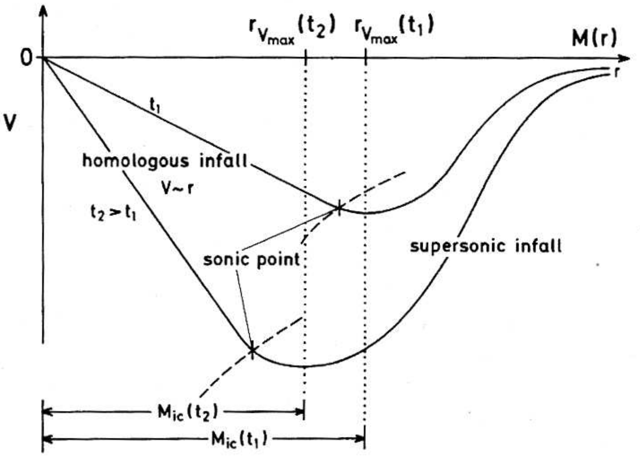

Figure 26: Schematic velocity profile in the core in the homologous phase. (Müller 1991)

as a sound wave can propagate from the center. In particular, the sound waves are trapped at the point where the inflow becomes supersonic, which is just inside of maximum infall velocity. Outside of the sonic point the collapse proceeds supersonically, at close to free fall velocity.

In Fig. 27 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 last epoch when the core is still homologous. The central density is $\sim 4.8 \times 10^{13}$ g cm^{-3}. Note that this figure only corresponds to the inner 50 km in Fig. 25. At the next epoch the central density is $\sim 2.6 \times 10^{14}$ g cm^{-3}, and homology is no longer valid. The compression at the center generates sound waves which propagate outwards. As they reach the sonic point they accumulate there, and steepen to a shock wave. Curve (c) corresponds to maximum density, $\sim 9.7 \times 10^{14}$ 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 $\sim 6.9 \times 10^{14}$ g cm^{-3}, and the core is adjusting to its final density $\sim 4 \times 10^{14}$ g cm^{-3}. As we will 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 homologous core, at a mass of ~ 0.5 M$_\odot$ from the center, or ~ 20 km.

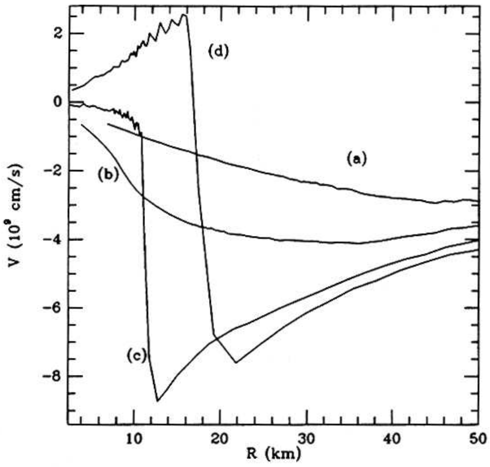

Figure 27: Velocity profile in the core at four epochs during collapse and explosion. Curve a) corresponds to the last epoch of homologous infall, with $v \propto r$ ('last good homology'). b) The density at center is close to nuclear. The matter outside is falling in with increasing velocity. c) The 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)

2.5 Energy losses in the shock

The total energy of the shock, as it is launched outside the homologous core, is close to the binding energy of this, $\sim 5 \times 10^{51}$ ergs. While the initial energy of the shock is therefore 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 shock 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 homologous core $\sim 1.5 \times 10^{51}$ ergs is lost by this process.

In Fig. 28 we show one example of an successful explosion and one unsuccessful. In the upper one it can be seen that a shock wave forms at $\sim$ 25 ms. This propagates through the iron core, and succeeds in reaching escape velocity at the time it has passed the boundary of the iron core, and no more losses take place. The lower figure shows an unsuccessful explosion.

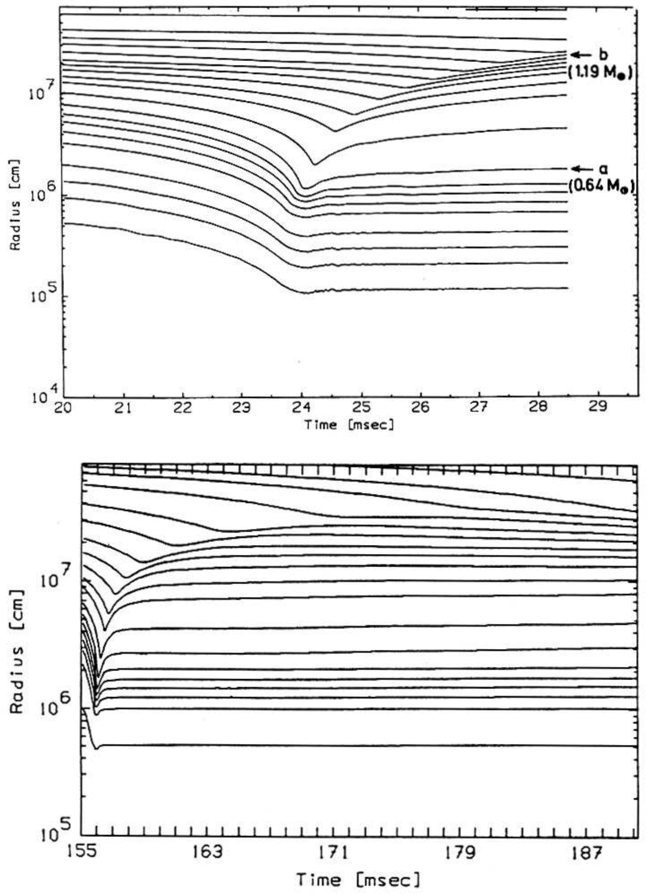

Figure 28: Trajectories of different mass shells as function of time. In the upper panel a successful explosion is shown, where the mass shell labeled b, coinciding with the shock, has reached escape velocity, and the core has settled into a hydrostatic configuration. The lower figure shows an unsuccessful simulation. (Hillebrandt 1984, 1987)

Although a shock does form, this loses so much energy that it stalls at ~ 300 km. However, outside of this accretion continues, and unless there is new energy input it will re-collapse, resulting in a black hole.

While neither of these simulations include all necessary physics, they illustrate the problem of overcoming the losses connected to traversing the high density iron core. Because the iron core mass increases slowly with increasing ZAMS mass, (Sect. 1.15 and Fig. 23) it is easier to explode a star with a mass close to the lower limit of the massive range.

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 ~ 1.2 M$_\odot$. As we have seen earlier, this is sensitive to factors like the ^{12}C$(\alpha, \gamma)^{16}$O reaction rate and the treatment of semi-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$_\odot$.

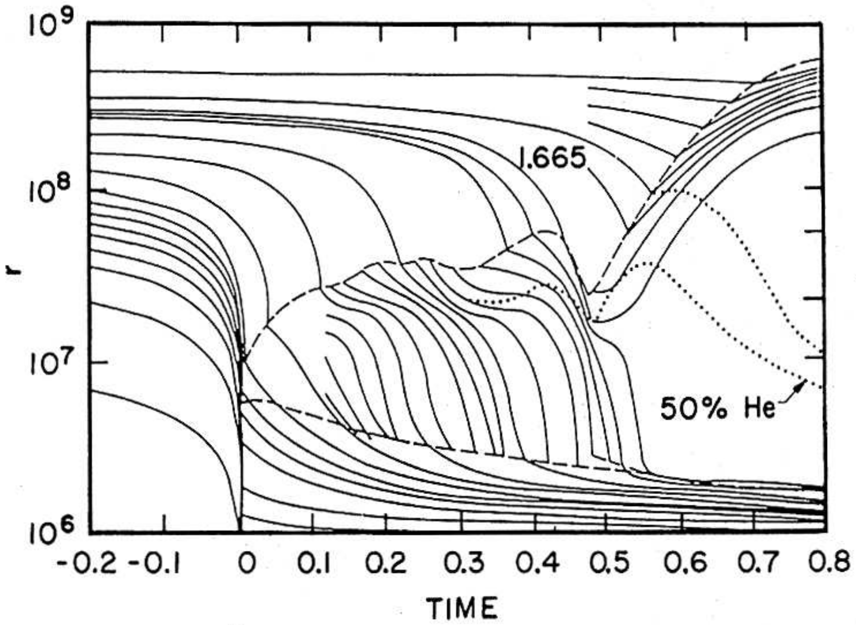

Figure 29: Trajectories of different mass shells for Jim Wilson's successful explosion model. The upper dashed line gives the position of the shock, while the lower is that of the neutrino photosphere. The dotted line marks the region where the abundance of He is 50%. Note the revival of the shock at 0.55 s due to neutrino heating (Bethe & Wilson 1985).

2.6 Neutrino heating

In a now classical computer run, Jim Wilson allowed one of his apparently unsuccessful simulations to 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. 29 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 a the result is a healthy explosion.

Although successful, these simulations had several deficiencies. One was an unrealistic equation of state above nuclear density, considered as too soft. Another is that Wilson's group assumed a vigorous convection caused by 'neutron fingers' in the newly formed neutron star. This instability is similar to the 'salt finger instability' occurring when warm, salty water rests on top of cold, fresh water. Because the heat diffusion is much more rapid than the salt diffusion the temperature will become the same as the surroundings. It will, however, have a larger salt content than the surrounding water and therefore be more dense, and it will sink further, causing a flow of salty water downwards. The analogue of the salinity is in the collapse case the lepton number, which in the proto-neutron star can decrease outwards, while the entropy increases outward.

Finally, the radiative transfer treatment of the neutrinos was in Wilson's calculations rather simplified. 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 $t_{\rm diff}$, given by Eq. (2.18), 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 the photosphere, define a *neutrino-sphere*, with radius R_ν, where the optical depth to infinity is $\tau_\nu = 2/3$.. As the neutrinos propagate out through the shocked gas, they will scatter and be absorbed by the nucleons.

The most important heating processes are

$$\bar{\nu} + p \rightarrow e^+ + n \tag{2.20}$$

and

$$\nu + n \rightarrow e^- + p \tag{2.21}$$

The neutrino heating rate at radius r is given by

$$H_\nu(r) = \frac{4\pi}{(ch)^3} \int c\sigma(E_\nu) E_\nu W(r) f(E_\nu) E_\nu^2 dE_\nu \tag{2.22}$$

Here $f(E_\nu)$ is the distribution function of the neutrinos. We assume this is given by the Fermi distribution

$$f(E_\nu) = \frac{1}{e^{E_\nu/kT\nu} + 1} \tag{2.23}$$

$W(r)$ is the dilution factor, given by $W(r) = \frac{1}{2}[1 - \sqrt{1-(R_\nu/r)^2}]$. Note that the neutrinos have only one helicity state, and there is therefore not the usual factor 2 multiplying the integral, as for electrons and photons.

Now, using Eq. (2.9) for the neutrino cross section and using

$$\int_0^\infty \frac{x^5}{e^x+1}dx = \frac{31\pi^6}{252} \approx 118.3 \tag{2.24}$$

we get

$$H_\nu(r) = \frac{31\pi^7}{252}\frac{m_\mathrm{e}^4c^6}{h^3}\sigma_0 W(r)\left(\frac{kT_\nu}{m_\mathrm{e}c^2}\right)^6 \tag{2.25}$$

The cooling rate for the inverse processes to Eq. (2.20) and Eq. (2.21) is

$$C_\nu(r) = \frac{8\pi}{(ch)^3}\int c\sigma_{e^-+p\rightarrow\nu+n}(E_\mathrm{e})E_\mathrm{e}f(E_\mathrm{e})p_\mathrm{e}^2dp_\mathrm{e} \tag{2.26}$$

where $f(E_\mathrm{e})$ is now the electron distribution, given by the Fermi distribution at the matter temperature T at r

$$f(E_\mathrm{e}) \approx \frac{1}{e^{E_\mathrm{e}/kT}+1} \tag{2.27}$$

We have in Eq. (2.26) included the factor two for the two spin states. We further assume the electrons are relativistic, so that $E_\mathrm{e} \approx p_\mathrm{e}c$, and that they are non-degenerate. This is in general a good approximation in the neutrino heating region.

The cross section $\sigma_{e^-+p\rightarrow\nu+n}$ can be obtained by detailed balance from the reverse of the above reaction. Detailed balance says that the cross sections for a reaction $i \rightarrow f$ and the reverse reaction $f \rightarrow i$ are related as

$$\sigma_{i\rightarrow f}\ g_ip_i^2 = \sigma_{f\rightarrow i}\ g_fp_f^2 \tag{2.28}$$

(e.g., Landau & Lifshitz 3, §144). Note that $g_ip_i^2$ and $g_fp_f^2$ are just the phase space factors. Because $g_\nu = 1$, $g_\mathrm{e} = 2$, $p_\nu = E_\nu/c$, and $p_\mathrm{e} \approx E_\mathrm{e}/c$, since the electrons are relativistic, we get

$$E_\nu^2\ \sigma_{\nu+n\rightarrow e^-+p} = 2E_\mathrm{e}^2\ \sigma_{e^-+p\rightarrow\nu+n}\ . \tag{2.29}$$

Further, $\sigma_{\nu+n\rightarrow e^-+p} = 1/4\sigma_0(E_\nu/m_\mathrm{e}c^2)^2$ (Eq. (2.9)), and $E_\mathrm{e} \approx E_\nu$ we get $\sigma_{e^-+p\rightarrow\nu+n} = 1/8\sigma_0(E_\mathrm{e}/m_\mathrm{e}c^2)^2$, and

$$C_\nu(r) = \frac{\pi\sigma_0c}{2(ch)^3}\int E_\mathrm{e}^5f(E_\mathrm{e})dE_\mathrm{e} \tag{2.30}$$

Using Eq. (2.24) again we obtain

$$C_\nu(r) = \frac{31\pi^7}{252} \frac{m_{\rm e}^4 c^6}{h^3} \sigma_0 \left(\frac{kT}{m_{\rm e} c^2} \right)^6 \ . \tag{2.31}$$

Therefore, the net heating rate becomes

$$H_\nu(r) = \frac{31\pi^7}{252} \frac{m_{\rm e}^4 c^6}{h^3} \sigma_0 [W(r) T_\nu^6 - \frac{1}{2} T^6] \tag{2.32}$$

At the neutrino sphere, $W \approx 1/2$ and $T \approx T_\nu$ so that the net heating is zero. For $r \gg R_\nu$ the dilution factor is $W \approx R_\nu^2/4r^2$.

Exercise: Skip this!

Motivate the different terms in Eq. (2.22), and explain the steps leading up to Eq. (2.32).

Eq. (2.32) shows that for $T > T_\nu (R_\nu/2r)^{1/3}$ there will be a net cooling, and vice verse.

The temperature behind the shock is set by the disintegration of Fe. For each nucleon this costs a binding energy E_{bind}. The total energy per nucleon to be dissipated by the shock is the gravitational binding energy at r. This is divided between the disintegration energy and the thermal energy of the (non-degenerate and non-relativistic) nucleons

$$\frac{GMm_{\rm p}}{r} = E_{bind} + \frac{3}{2} kT \tag{2.33}$$

If iron is completely dissociated into nucleons this costs 8.8 MeV per nucleon. Dissociation into α particles costs ~ 2 MeV per nucleon. Taking this as the typical case, we find

$$T \approx \frac{15}{r_7} - 1.3 \text{ MeV} \tag{2.34}$$

where $r_7 = r/10^7$ cm. The shock tends to stall at $300 - 500$ km, so the temperature will be $1 - 3$ MeV, decreasing outwards. Close to the proto-neutron star it will be fairly constant. The matter temperature is shown schematically in Fig. 30

Because T decreases faster than $T_\nu (R_\nu/2r)^{1/3}$, there will be a radius where heating and cooling balance, given by

$$R_{\rm gain} \approx \left(\frac{T_\nu}{T} \right)^3 R_\nu \tag{2.35}$$

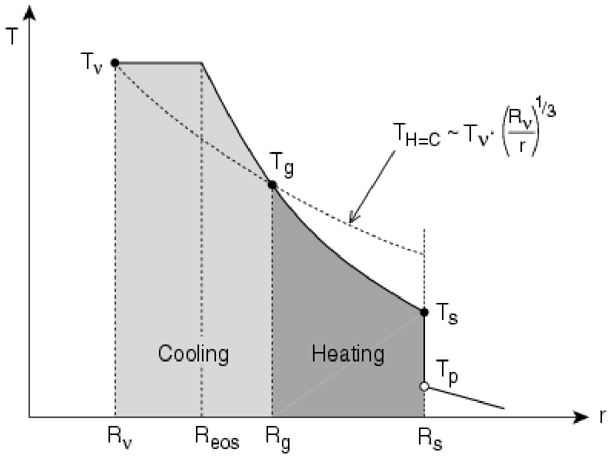

Figure 30: Schematic figure of matter temperature (solid line) and the temperature for which neutrino cooling and heating balance. The crossing of these defines the gain radius. (Janka 2001)

This is 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. 31.

This late neutrino heating certainly helps in getting an explosion, as already Jim Wilsons simulations showed. Even including this it has, however, been extremely difficult to get succesful explosions, except for stars close to the lower limit for supernova explosions. In Fig. 32 we show one of the few cases where this has been possible in a simulation of an 11 $M_\odot$ star. This had an O–Ne–Mg core with a very steep density gradient outside the core, which is one of the reasons why the shock survived the propagation through the core. For more massive stars it has been considerably more difficult to get an explosion and in Fig. 33 we show the result of the collapse of a 15 $M_\odot$ star. Although a shoch is launched after the bounce and initially expands it stalls after 0.15 s and the end result is a black hole.

The neutrino heat input has also important consequences for the dynamics of the gas behind the shock. Just outside the gain radius the matter is heated to a high temperature, decreasing outwards. This therefore induces a strong entropy gradient outwards, towards the shock. Such a gradient in turn leads to convection by the usual Schwarzschild criteria. The large scale convective motions transport entropy to the region close to the shock and can therefore re-energize the shock (see Fig. 29 and Fig. 36). At the same

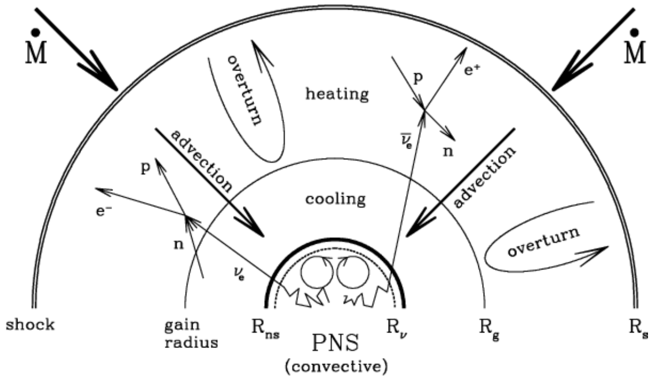

Figure 31: Schematic picture of the different regions close to the neutron star (Janka 2001)

time low entropy gas sinks inward, inside the gain radius. There it will be heated by the neutrinos. This compensates the iron disintegration 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. In terms of the neutrino temperature of the different neutrino species this is given by

$$L_{\nu_i} = 4\pi \frac{7}{8} \sigma_B T_{\nu_i}^4 R_{\nu_i}^2 \tag{2.36}$$

where both T_{ν_i} and R_{ν_i} may differ between the six neutrinos. The factor 7/8 comes from the fact that the neutrinos are fermions, while the usual Stefan-Boltzmann constant, σ_B, is defined for photons.

Exercise:

Motivate the factor 7/8 in Eq. (2.36)!

The neutrino emission is a consequence of the de-leptonization of the proto-neutron star. At the time of collapse the temperature in the core is fairly constant and ~ 20 MeV. After the collapse most of the binding energy is stored not as thermal energy, but as chemical potential energy of the degenerate electrons and neutrinos. With $\rho_c \sim 6 \times 10^{14}$ g cm^{-3} and $Y_e \sim 0.28$ this gives $\mu_e \sim 287$ MeV. The neutrinos, with $Y_\nu \sim 0.08$, have a somewhat lower $\mu_e \sim 238$ MeV. Because the temperature is only ~ 5 MeV, the thermal energy is negligible for the neutrino production.

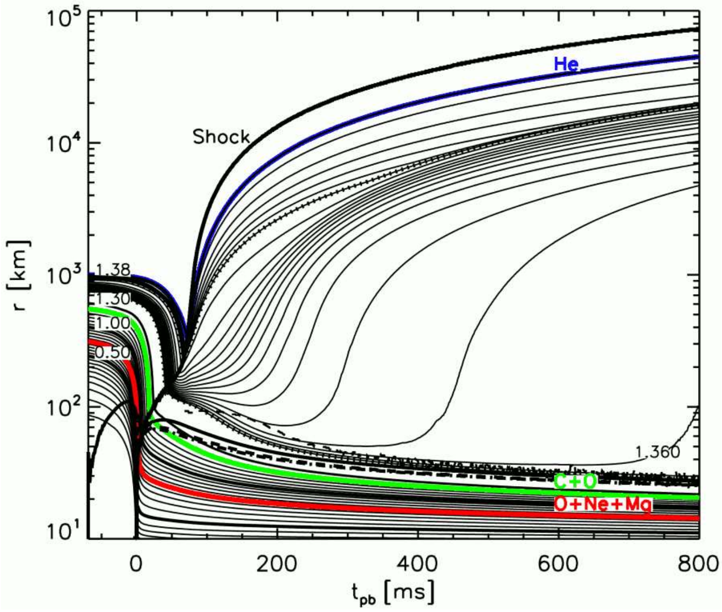

Figure 32: One dimensional simulation of the collapse and explosion of an 11 $M_\odot$ progenitor. This may represent the explosion of the Crab supernova (Kitakura et al. 2006)

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.

In Fig. 34 we show a two-dimensional simulation of the deleptonization of the proto-neutron star. Convection set up by the gradient in electron fraction, Y_e, causing rising bubbles of high Y_e ('light') bubbles to the surface of the neutron star, where the neutrinos are emitted, decreasing the Y_e. This matter now has a lower Y_e and will therefore sink back into the core. In this way there is a deleptonization wave moving in through the core, as can be seen from a comparison of the two epochs.

This is shown more quantitatively in Fig. 35, where one sees the rapidly decreasing lepton fraction from the surface inwards. The total process takes a couple of seconds until the proto-neutron star has got rid of its leptons, and been converted into a proper neutron star of nearly zero lepton fraction. These convective motions will increase the deleptonization rate by a substantial factor, and at the same time also increase the neutrino luminosity.

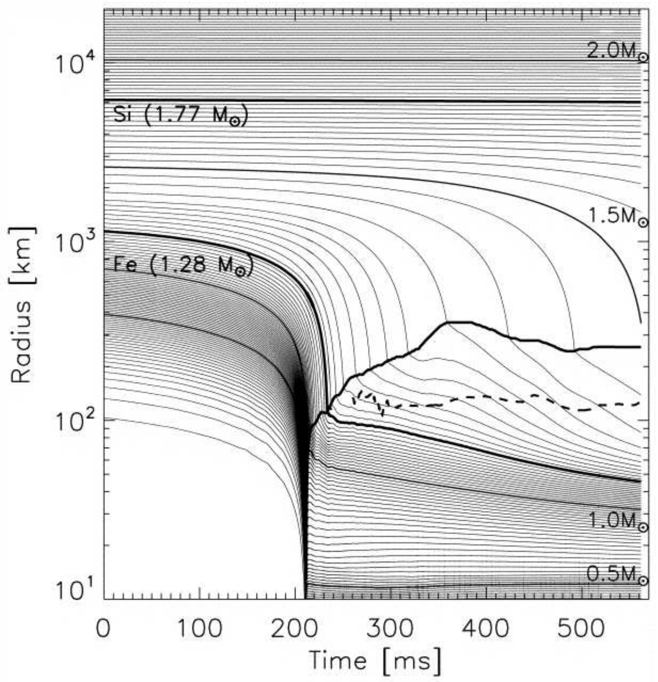

Figure 33: Simulation of the collapse and explosion of a 15 $M_\odot$ explosion. In this case the shock stalls and the end result is a complete collapse with the creation of a black hole. (Rampp & Janka 2000).

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 to early to make any conclusions of the success of this mechanism. The radiative transfer is in particular a demanding task, with scattering, absorption and emission in both energy and space.

In addition, there are other ingredient. 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.

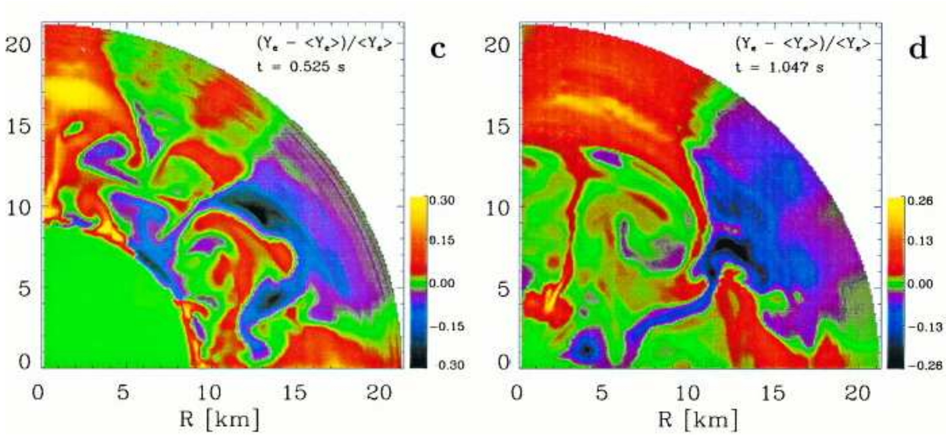

Figure 34: Two-dimensional simulation of the neutrino convection in the proto-neutron star. The plot shows the electron fraction, Y_e, at 0.525s and 1.047 s. Note the down drafts of deleptonized matter which sinks into the core, while high lepton number bubbles rise. (Keil, Janka & Müller 1996)

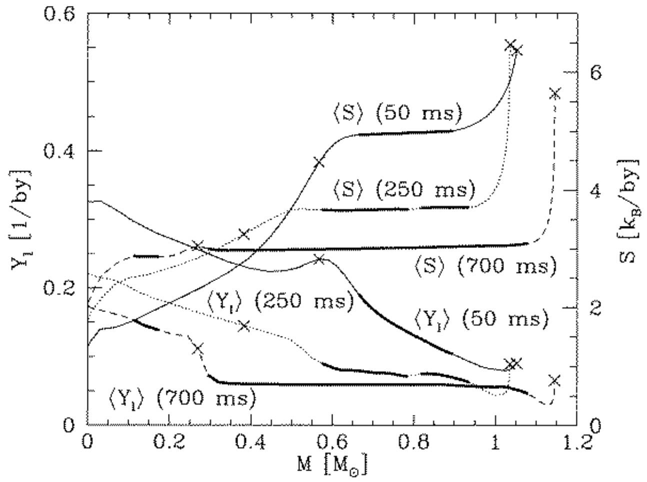

Figure 35: Lepton fraction and entropy at 0.05 s, 0.25 s and 0.70 s. Note the heating of the inner core (increasing entropy), and the deleptonization wave moving in from the surface. (Keil, Janka & Müller 1996)

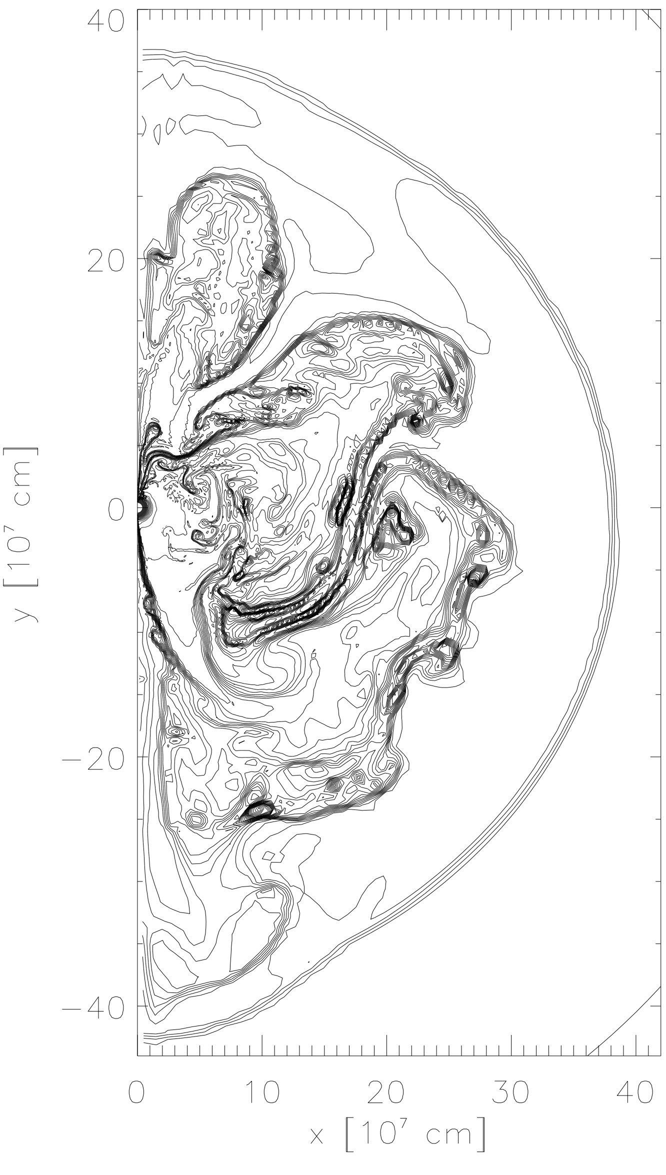

Figure 36: Two-dimensional simulation of the explosion at 0.38 s after core bounce. The contours show the entropy distribution. The shock is at 3800 km. Note the bubbles of neutrino heated gas, and the down drafts of sinking cooler gas. (Janka & Müller 1996)

2.7 SASI

A qualitatively new development has come with the realization that the shock outside the proto-neutron star can be seen as a standing accretion shock. One knows that in spherical geometry such shocks are subject to instabilities. Initial analytical and numerical studies of spherical accretion shocks also confirmed that this wasd the case. Further simulations including realistic neutrino physics and eqation of state in two dimensions showed that a large scale dipolar ($l = 1, 2$) instability develops in a generic way. This SASI instability (Standing Accretion Shock Instability) leads in general to an accretion plume in one region and an outflow in others. In Fig. 37 we show the result of such a simulation. Although these simulations show that these low mode instabilities favor a succesful explosion, it is only in models at the lower mass end, in particular for stars with O–Ne–Mg cores, one finds explosions. A problem is, however, that these simulations have only been performed in 2D. This constrains instabilities etc. in an artificial way compared to 3D, and only for such simulations can one trust the outcome.

An interesting side result is that these large scale modes may be able to explain the observed high neutron star velocities of the pulsars, with velocities up to ~ 1000 km s^{-1}. The asymmetry of the explosion naturally leads to a momentum in the same direction, and the simulations show that velocities in the range observed can indeedbe obtained in the explosion models. These are, however, based on parameterized neutrino luminosities from the neutron star, and it is likely that the quantitative results will change as more ab initio models become available. Nevertheless, this is a very promising avenue for the solution of this long lasting problem.

A further recent interesting result is that the same SASI instability also tends to develop spiral modes ($m = 1$). This will lead to an angular momentum of the neutron star and simulations show that this effect may explain the observed rapid spin of the pulsars, even for non-rotating progenitors.

A modified version of the SASI mechanism starts with the SASI oscillations with a period of $\sim 15 - 30$ ms during the first 200 ms. The SASI mechanism by itself does not lead to an explosion. However, the accretion plumes onto the protoneutron star induces gravity mode (radial) oscillations with a peiod of ~ 3 ms in this. These grow and after $\sim 0.5 - 1$ s these become so violent that an explosive outflow results. The explosion now resembles an anisotropic wind more than a spherical outflow. As with the SASI instability this will naturally explain the large pulsar velocities. Whether this is the final solution to the problem remains to be seen. History tells us that it is probably not!

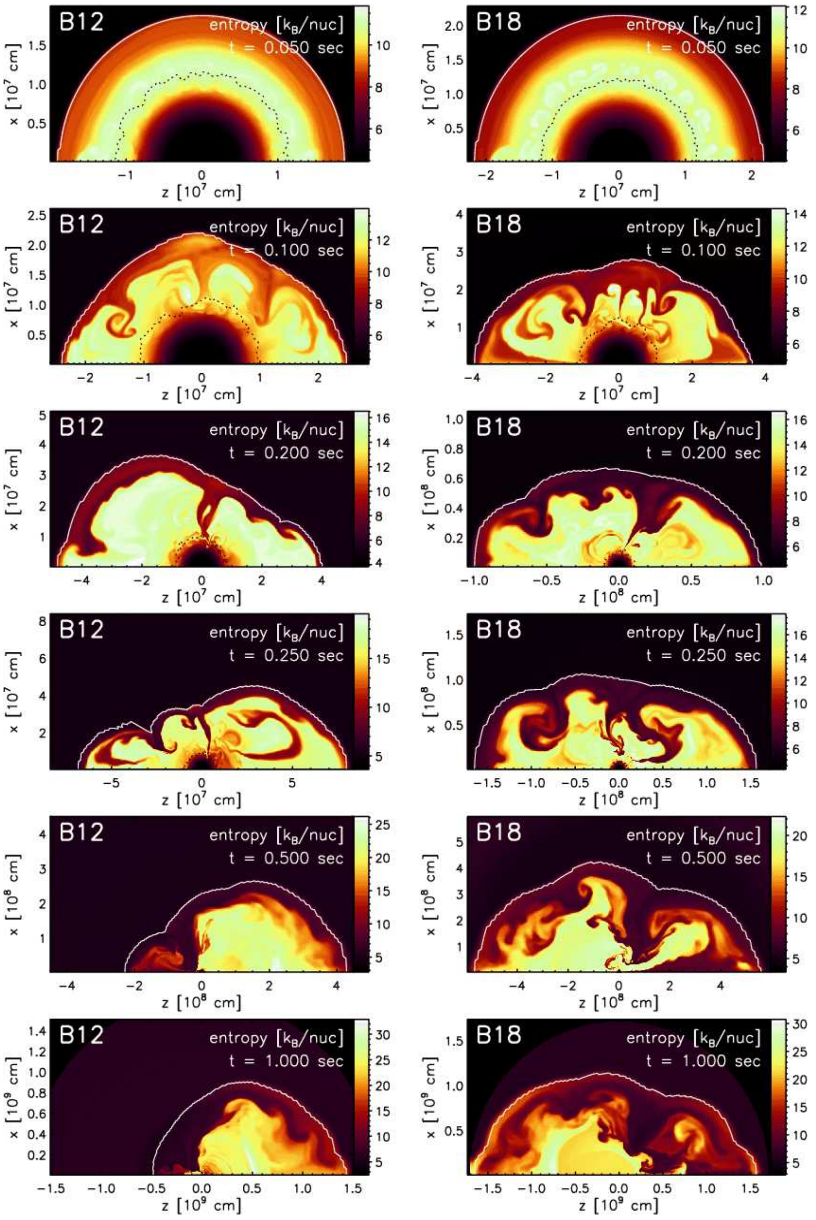

Figure 37: Two-dimensional simulation of the explosion at several epochs after core bounce, and for two different progenitor models. The contours show the entropy distribution. Note the bubbles of neutrino heated gas, and the down drafts of sinking cooler gas. (Scheck et al. 2006)

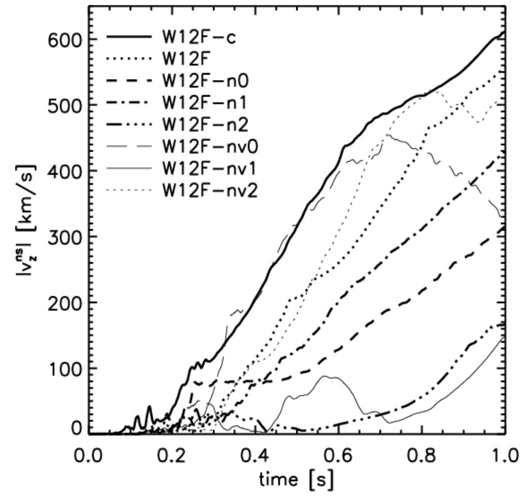

Figure 38: Neutron star kick velocity for different simulations ans . (Scheck et al. 2006)

2.8 Shock propagation

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.

The journey to the surface is, however, not just a smooth ride. The density in the progenitor decreases in steps as the various burning zones are passed. For an adiabatic shock wave $p = (\gamma - 1)E/V \propto E/r^3$, where E is the constant, total energy and r the shock radius. The pressure behind the shock is given by the ram pressure swept up by the shock, $\rho v^2 \approx p$. Therefore, the shock speed will be given by $v \propto (p/\rho)^{1/2} \propto E/(\rho r^3)$. We therefore conclude that if ρr^3 increases, the shock will slow down, and vice versa.

In Fig. 41 we show the shock speed as function of time, and also the quantity ρr^3. As we see, the shock speed behaves just as we expect from the analysis above. We note that the shock decelerates strongly in both the He core and in the H envelope. This deceleration has the consequence that there will be positive pressure gradient behind the shock, which slows down the gas behind it. Gas far enough behind the shock will not have time to react

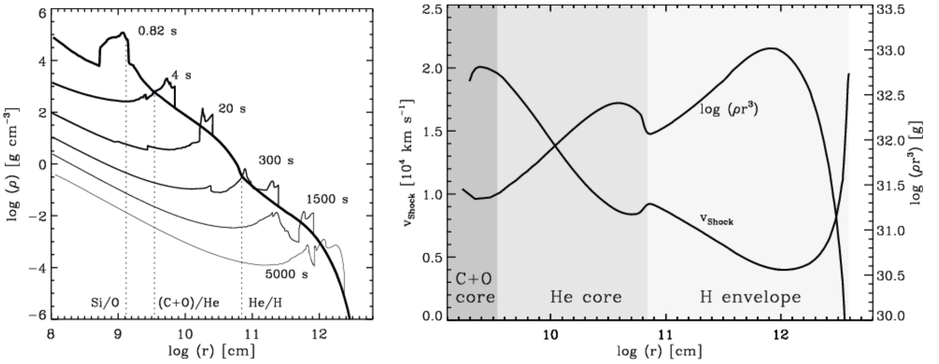

Figure 39: Propagation of the shock through the progenitor from the center to the shock break-out. The left panel gives the density structure from 0.8 s to 5000 s, while the right panel gives the shock velocity together with ρr^3. Note the deceleration of the shock in regions with increasing ρr^3 (Kifonidis et al. 2003)

to this increased pressure, and will thus shock the gas, i.e, create a reverse shock wave. These shocks, which have formed at the Si/O, (C+O)/He and He/H shells, can be seen in the left panel of Fig 39 at 0.82 s, 20 s, and 1500 s, respectively.

The reverse shock results in pressure and density gradients with opposite signs at the composition interfaces between the Si/O, (C+O)/He and He/H shells. As can be shown from a stability analysis, the growth rate of perturbations in the density is

$$\omega^2 = \frac{p}{\rho}\frac{\partial \ln p}{\partial r}\frac{\partial \ln \rho}{\partial r} \tag{2.37}$$

Therefore, if the gradients have different signs, the layer is unstable to perturbations. This is just the classical Rayleigh-Taylor instability, but with the gravity replaced by a pressure gradient.

Exercise: Skip this!

Derive the condition for the Rayleigh-Taylor instability for two fluids of different densities in a gravitational field on top of each other. How does this condition relate to Eq. (2.37)? (See e.g. Spitzer or Padmanabhan, Vol. I)

In Fig. 40 we show the pressure, density, velocity and composition for the one-dimensional model in Fig. 39. at 20 s and 300 s. At 20 s the reverse

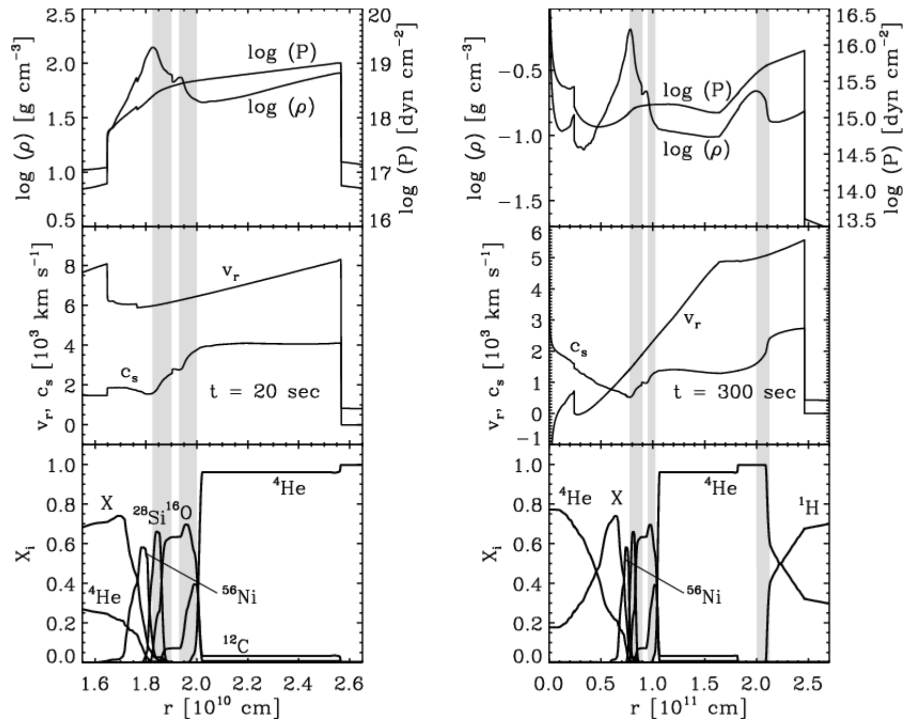

Figure 40: Pressure, density, velocity and composition for the one-dimensional model in Fig. 39 at 20 s and 300 s after explosion. The shaded regions show the locations of R-T unstable regions. (Kifonidis et al. 2003)

shock is at 1.6×10^{10} cm and at 3×10^{10} cm at 300 s. From the pressure and density curves at 300 s one can clearly identify the unstable regions at at 0.8×10^{11} cm, 1.0×10^{11} cm and at 2.0×10^{11} cm. These coincide with the density jumps at the Si/O, (C+O)/He and He/H interfaces, respectively. Based on this one dimensional analysis, one can therefore expect instabilities at these positions.

This is confirmed by the two-dimensional simulations shown in Fig. 41. and Fig. 42, which shows that the whole core is broken up by instabilities, extending almost out to the He/H interface.

Exercise:

Discuss the evolution of the two-dimensional simulations in Figs. 41-43. (See Kifonidis, K., et. al. 2003, A&A, 408, 621)

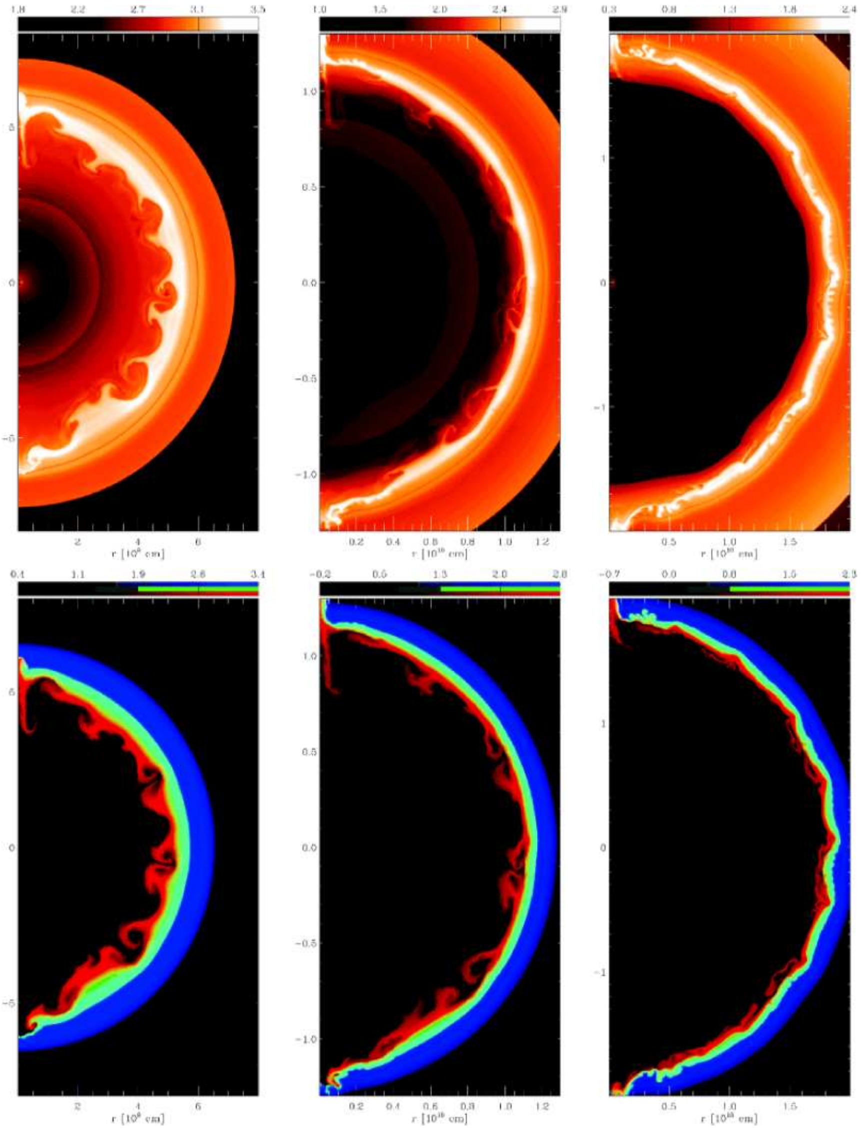

Figure 41: Evolution of the explosion at 4 s, 10s, and 20 s in two dimensions. The upper panels show the density and the lower the composition with O (blue), Si (green), Ni (red). Note the rapid development of the Rayleigh Taylor instabilities. the neutron star (Kifonidis et al. 2003)

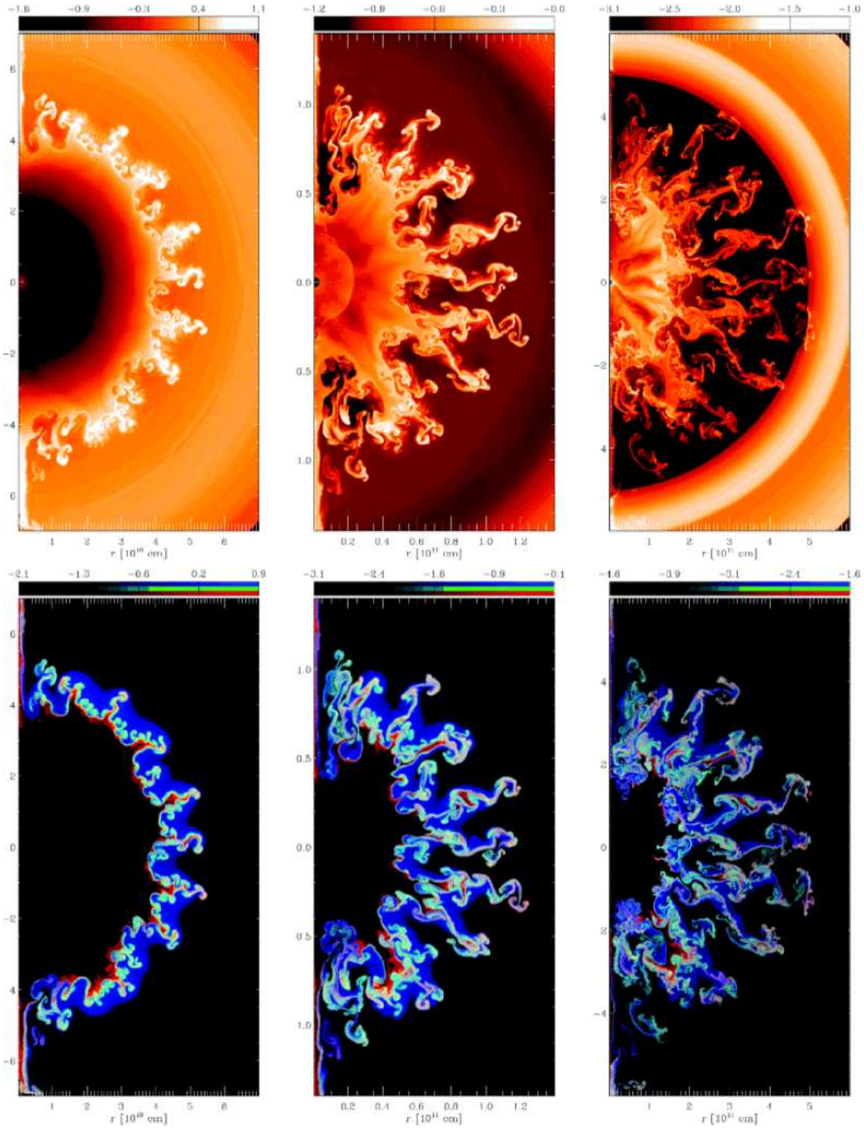

Figure 42: Same as Fig. 41 but at 100s, 300 s and 1500 s. Note the difference in scale (Kifonidis et al. 2003)

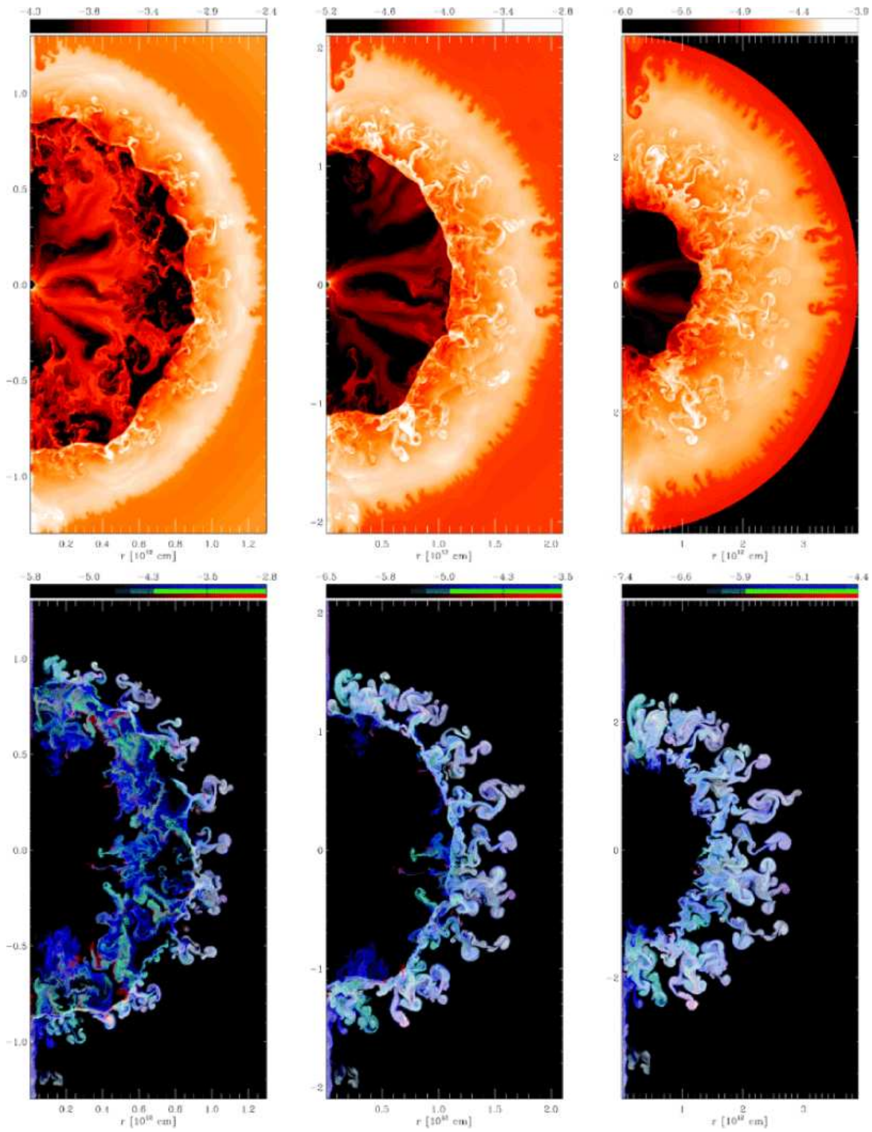

Figure 43: Same as Fig. 41 but now at 5000 s, 10000 s and 20000 s. Note the difference in scale (Kifonidis et al. 2003)

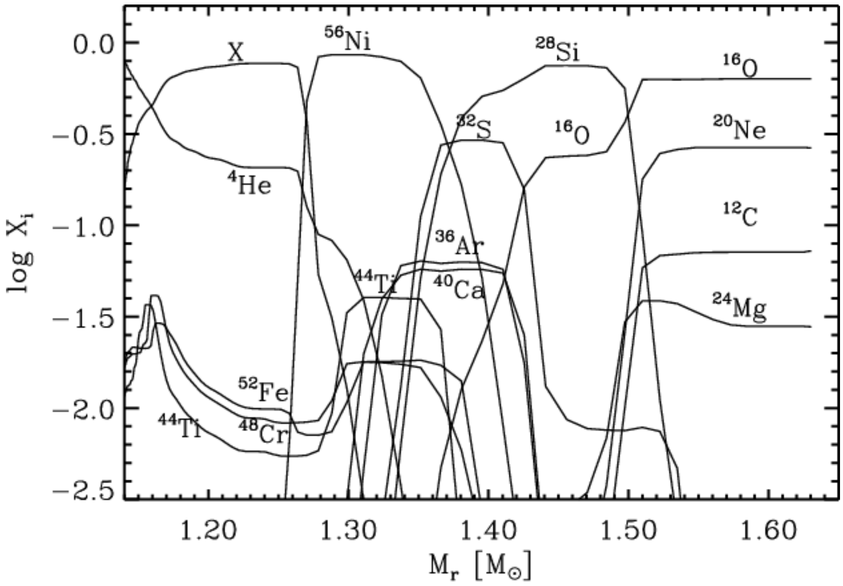

Figure 44: Explosive nucleosynthesis in the central region (Kifonidis et al. 2003).

2.9 Explosive nucleosynthesis

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. 44).

During the first seconds after the core bounce some from the observational point of view 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 §1.7, 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 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 ~ 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. While ^{57}Ni (not included in Fig. 44) can be produced in NSE, ^{44}Ti needs a more fine tuned version of this, known as an 'α-rich freeze-out'.

The α-rich freeze-out occurs when the explosive burning reaches NSE, but the medium is then cooled so fast that the free α-particles, created mainly by the photo-disintegration of the iron peak elements, does not have time to fuse back by the 3α reaction. For long enough time the α's are consumed in the chain

$$\alpha(2\alpha,\gamma)^{12}\mathrm{C}(\alpha,\gamma)^{16}\mathrm{O}(\alpha,\gamma)......^{52}\mathrm{Fe}(\alpha,\gamma)^{56}\mathrm{Ni} \tag{2.38}$$

However, the 3α reaction is comparatively slow, and if the temperature falls rapidly this acts like a bottle-neck, resulting in a large fraction of free α particles. Thus, its name, α-rich freeze-out. Calculations (e.g., Woosley & Hoffman 1991) show that this results in an excess of especially ^{44}Ti.

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.

2.10 The r-process

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 relays 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. The first step is therefore,

$$(Z,A)+n\rightarrow(Z,A+1)+\gamma \tag{2.39}$$

This may be followed either by another neutron capture

$$(Z,A+1)+n\rightarrow(Z,A+2)+\gamma \tag{2.40}$$

or by a β-decay

$$(Z,A+1)\rightarrow(Z+1,A+1)+e^{-}+\nu \tag{2.41}$$

The time scale for β-decay varies between seconds and days. This is by ordinary evolutionary standards short, and leads for normal stars in connection to e.g., the He-burning phase to the so called *s-process*. However, in connection to supernova explosions the neutron flux may be extremely large during the explosion phase which is only of the order of seconds at most. In this case the neutron capture (Eq. 2.40) may occur before the β-decay, and this may continue up to very neutron rich nuclei, $(Z, A + k)$. This is the beginning of the *r-process*

This can, however, not continue infinitely. The neutron capture cross section decreases with an increasing number of neutrons, and at the same time the photodissociation cross section $(Z, A+k) + \gamma \rightarrow (Z, A+k-1) + n$ increases. There will therefore be an isotope $(Z, A + k)$ for which the n-capture rate and photodissociation rate balances. At this point the β-decay may become important $(Z, A + k) \rightarrow (Z + 1, A + k) + e^{-} + \nu$. The process may now continue for this new element until a new equilibrium isotope is formed.

The neutron capture cross section decreases just above the neutron numbers corresponding to closed neutron shells, $N = 8, 20, 28, 50, 82, 126, 184$ (see Fig. 45). This leads to a pile up of nuclei with this number of neutrons and the typical r-process path is shown in Fig. 46

As the nuclei produceesed by the r-process β-decay back to the stable nuclei this pile up for the magical neutron numbers is directly reflected in the abundances of the r-process elements, shown in Fig. 47. The peaks are now shifted to N=46, 76 and 116. In the s-process the β-decays brings back the nuclei to the stability valley and the peaks in abundances occur at the same neutron number as the peaks in the neutron cross section.

Although the r-process was proposed already 1957 by Burbidge, Burbidge, Fowler and Hoyle, the site of this has always been a mystery. Because of the need for a very strong neutron exposure under a short time interval supernova explosions have for a long time been a main candidate. The precise location and conditions for this has, however, been a puzzle. It is clear that the general conditions just outside of the newborn neutron star are ideal for this. The neutron density is very high, because of inverse β-decays. Also the presence of a large density of seed nuclei in the form of iron peak elements is very favorable. At the moment the region close to the shock wave seems like the most promising, but it is not clear that the observed distribution of abundances can be reproduced in detail.

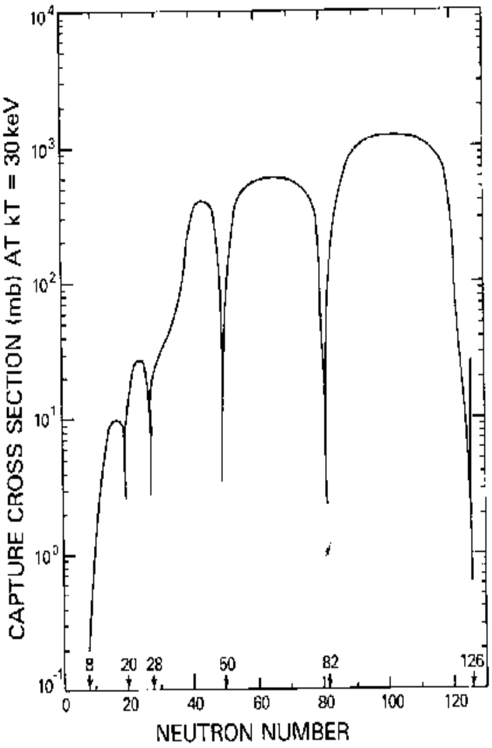

Figure 45: Cross section for neutron capture as function of the neutron number. (Rolfs & Rodney 1988)

3 Observables of core collapse supernovae

3.1 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,

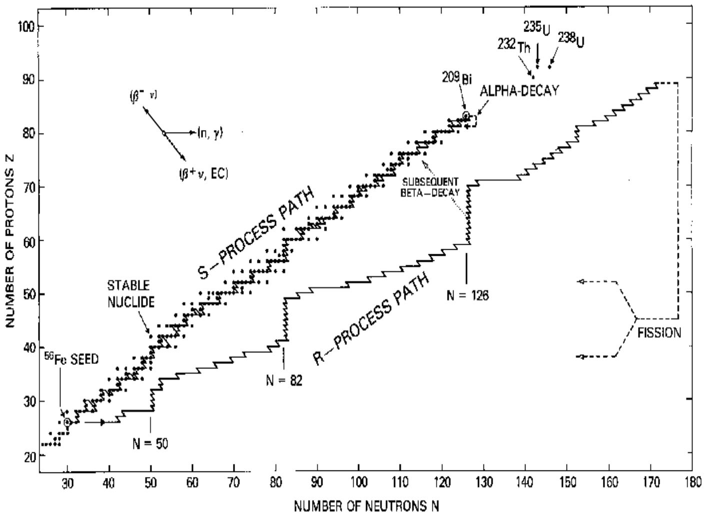

Figure 46: Typical path for the r-process. Note the position for the magical neutron numbers where a pile up occurs. (Rolfs & Rodney 1988)

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. 48 shows the energies of the individual electrons produced by the neutrinos, which approximately corresponds to a neutrino energy given by $E_\nu \approx E_{\rm e} + 2$ 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

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

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.

Exercise:

Derive Eq. (3.1)!

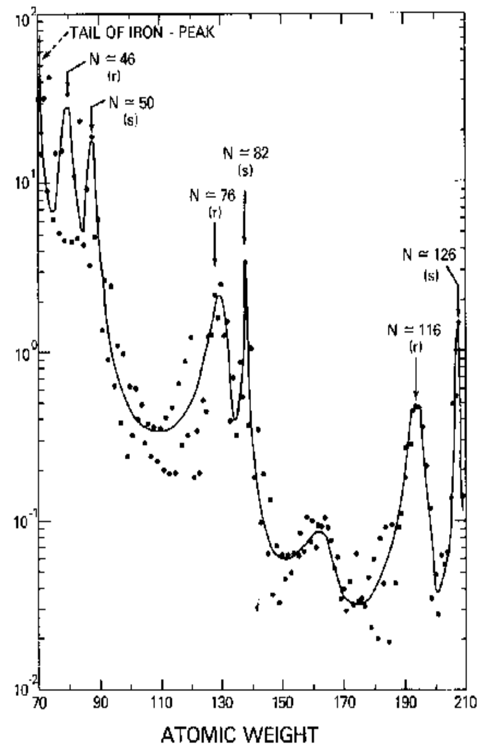

Figure 47: Abundances of the heavy elements above iron. Note the prominent r-process peaks, as well as the s-process peaks. (Rolfs & Rodney 1988)

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. In the proto-neutron star the neutrino energies are so high that the coherence involved in the neutrino scattering against the nuclei, discussed in connection to Eq. (2.13), is lost. Therefore the bracket in Eq. (2.13) should be replaced by ~ 1. The mean free path 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 $\sim 1.4 \ \mathrm{M}_\odot$, 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. (2.16), we get

$$t_{\text{diff}} \approx 0.2 \left(\frac{R}{30 \text{ km}} \right)^{-1} \left(\frac{E_\nu}{100 \text{ MeV}} \right)^{2} \text{ s.} \tag{3.2}$$

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

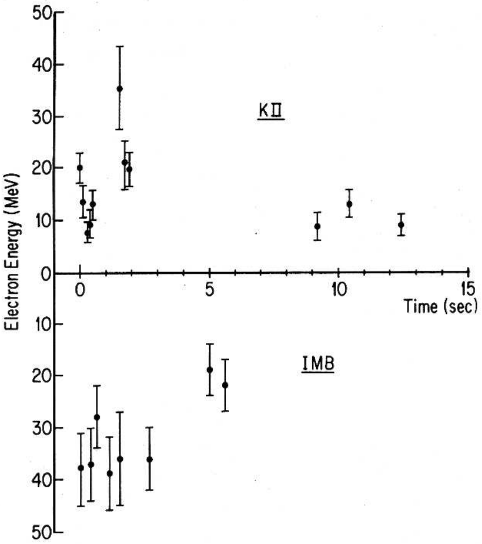

Figure 48: Energy versus arrival time for the detected neutrinos in the Kamiokande II and IMB detectors. (Burrows 1988).

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.

While the main feature of the neutrino emission can be understood from these simple arguments, a more detailed study requires numerical simulations. It then turns out that one can divide the evolution into two main phases (see Fig. 49). The first of these is extremely brief, only of the order a millisecond, and is connected to the breakout of the shock through the neutrinosphere. The shock forms below this, but within $\lesssim 1$ ms it passes the neutrinosphere into the optically thin region. The nucleons in the extremely hot gas behind the shock undergo electron captures, $e^- + p \rightarrow \nu_{\rm e} + n$, which is an important energy loss process for the shock. The result of this is a very large density of $\nu_{\rm e}$'s behind the shock. As the gas becomes optically thin to the neutrinos, this energy density is released. This produces a very short, a few ms, burst of $\nu_{\rm e}$'s (Fig. 50). The luminosity is $\sim 5 \times 10^{53}$ erg s^{-1} and the total energy in the burst $(1-3) \times 10^{51}$ ergs. Compared to the total energy budget, this is a small fraction, and it is unlikely that any of the SN

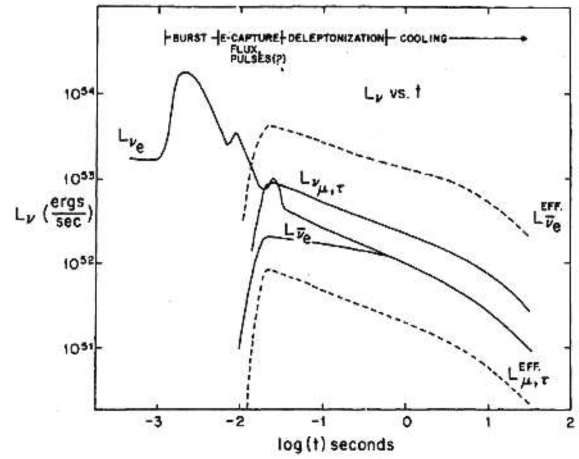

Figure 49: Approximate neutrino luminosities of the different neutron species. Note the logarithmic time scale. (Burrows 1984).

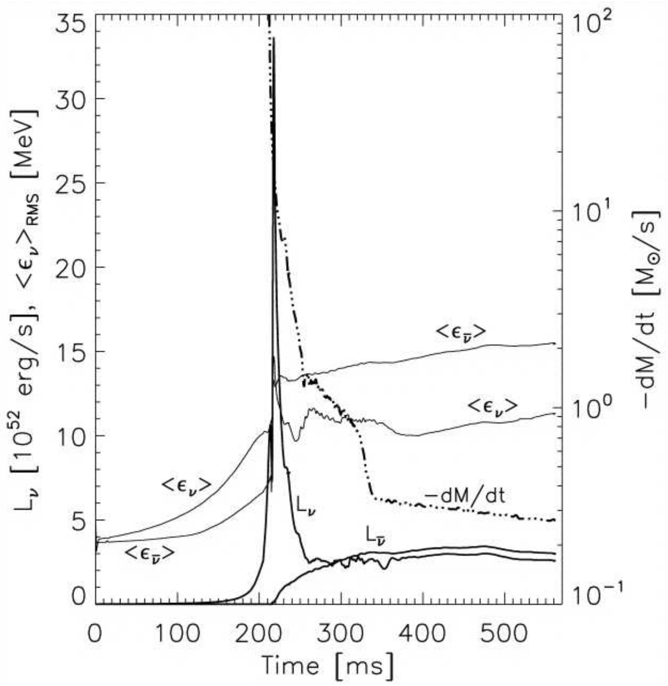

Figure 50: Luminosity and mean neutrino energy as function of time during the shock breakout phase. (Rampp & Janka 2000).

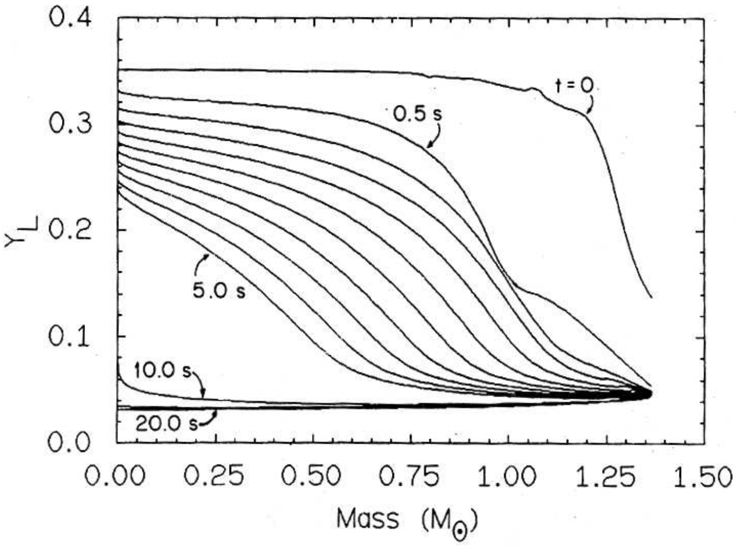

Figure 51: The lepton fraction ($Y_\mathrm{e} + Y_\nu$) through the proto-neutron star for different times after the core bounce. Note the deleptonization wave from the surface inwards. (Burrows & Lattimer 1986).

1987A neutrinos came from this burst. For a Galactic supernova this would, however, be extremely interesting as a diagnostic of the shock.

The next phase is connected to the deleptonization of the proto-neutron star, and has a much longer time scale, 10 – 20 s. As we discussed in §2.6, the deleptonization proceeds as a wave from the neutrinosphere to the center (see Fig. 2.6). In Fig 51 we show the same thing on a longer time scale, from 0–20 s. At the last epoch $Y_L \sim 0.$, showing that the neutron star is formed.

The neutrino and electron losses at the surface has the consequence that the pressure support of the neutrinos and electrons in the region just outside the core is removed. This region therefore contracts towards the center onto the core of the neutron star. As a consequence, the neutrinosphere moves in from ∼ 100 km to 10–20 km. This contraction heats the matter in the outer parts to $\sim 30 - 50$ MeV. In addition, accretion from the outside further enhances the heating and the neutrino luminosity.

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 $< E_\nu > \sim 3\ T_\nu \sim 10 - 15$ MeV (Fig. 50). The reason that the $\bar{\nu}_\mathrm{e}$'s have a

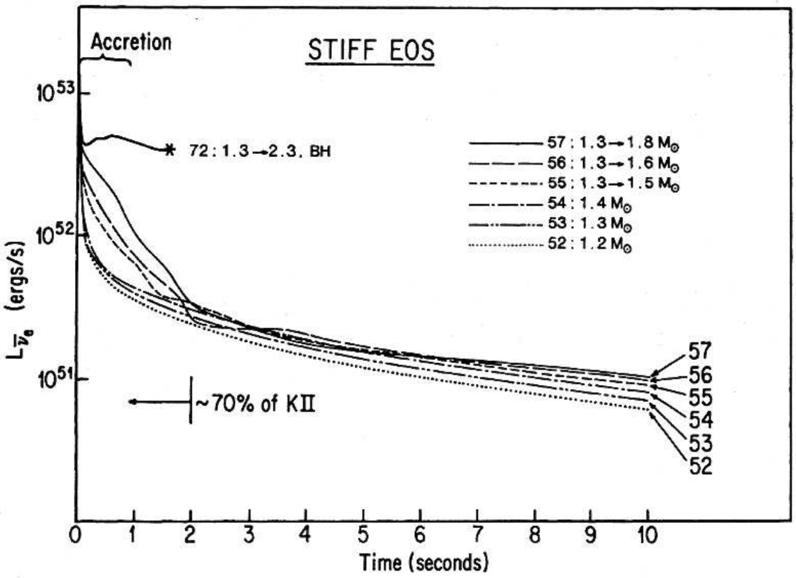

Figure 52: Luminosity as function of time for $\bar{\nu_{\rm e}}$ and different equations of state. Note the upper curve (model 72), which shows the short burst accompanying the formation of a black hole. (Burrows 1988).

higher temperature than the $\nu_{\rm e}$'s is that the opacity for the $\bar{\nu_{\rm e}}$'s is lower than for the $\nu_{\rm e}$'s. This is in turn a consequence of the lower number of protons ($\bar{\nu_{\rm e}} + p \rightarrow e^+ + n$) compared to neutrons ($\nu_{\rm e} + n \rightarrow e^- + p$) as the surface is deleptonized. One therefore 'sees' to deeper layers with $\bar{\nu_{\rm e}}$'s, where the temperature is higher.

Because of the high temperature and trapping an approximately equal number of all six neutrino species, $\nu_{\rm e}, \bar{\nu_{\rm e}}, \nu_\mu, \bar{\nu_\mu}, \nu_\tau$, 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

$$\bar{\nu}_{\rm e} + p \rightarrow e^+ + n \tag{3.3}$$

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

$$\nu_i + e^- \rightarrow \nu_i + e_i^- \tag{3.4}$$

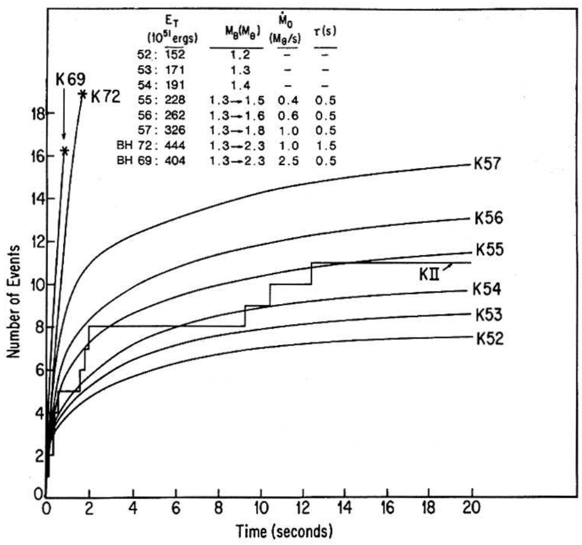

Figure 53: The cumulative number of anti-neutrino detections for the Kamiokande II detector compared to models with different equations of state and accretion rates. (Burrows 1988).

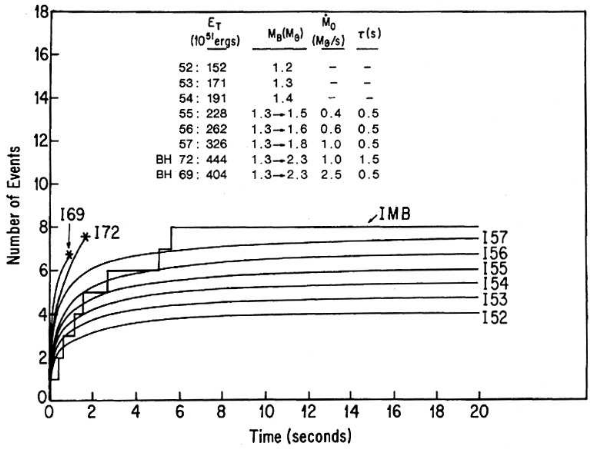

Figure 54: Same as Fig. 53 for the IMB detector. (Burrows 1988).

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. The results of this are shown in Figs. 53 and 54. Because of the small numbers (19 neutrinos in total) one compares the predicted curves with the cumulative number of neutrinos detected. The different curves correspond to different equations of state. As we see there is a general good agreement, but because of the statistics it is difficult to draw any far-reaching conclusions about e.g., the equation of state or the explosion mechanism, e.g., the late neutrino eating mechanism.

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, $< E_\nu >\sim 3\ T_\nu$, is shown in Fig. 50. The shorter duration of the IMB signal, which has a higher threshold, actually gives some indication that the source is cooling. The signal is best fitted with two exponentials with $\tau \lesssim 1$ s and $\tau \sim 4$ s, respectively.

Although there has been a flood of papers claiming evidence for rotational modulation, neutrino oscillations etc., there is in these few events no evidence for this.

In addition to the clear detections by Kamiokande II and IMB, there is also a marginal detection from a detector in Baksan. The claims for a detection with a much less sensitive detector under Mont Blanc about five hours before the others should be taken with extreme suspicion.

3.2 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. 55 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 Hα. A closer examination,

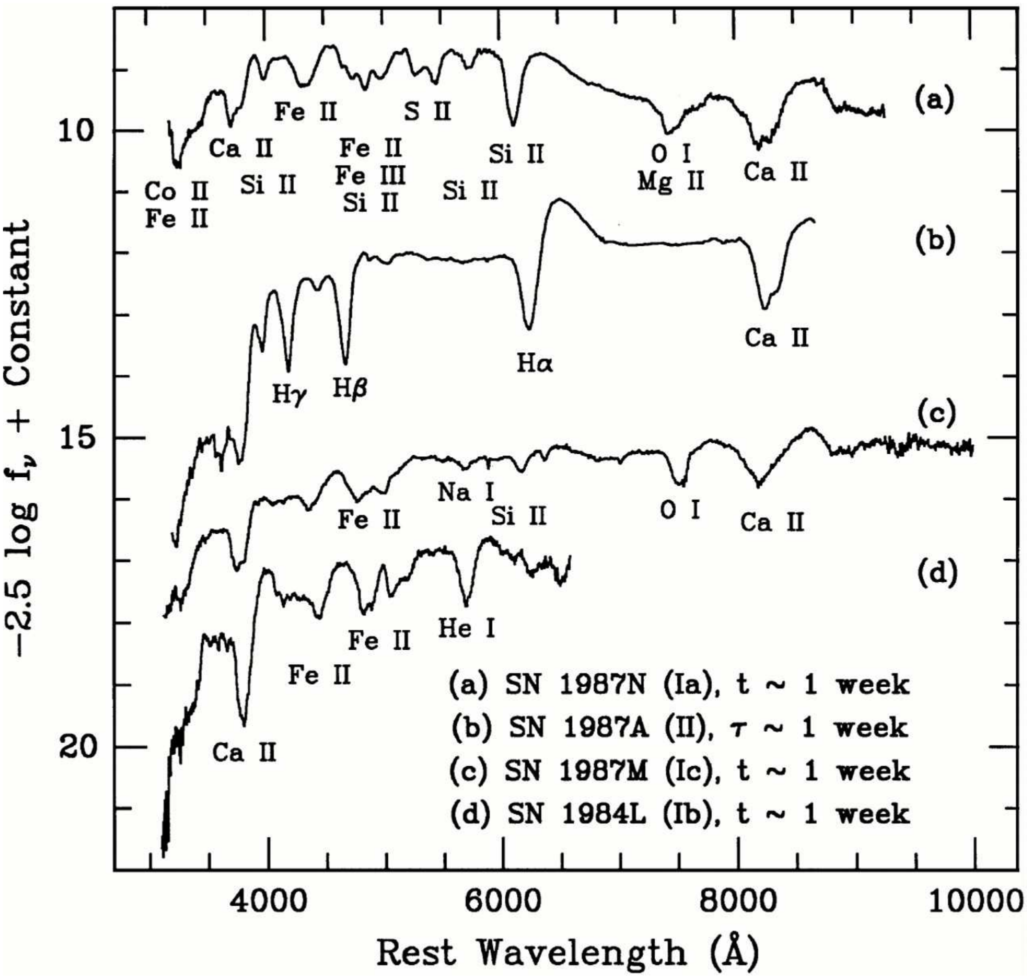

Figure 55: Spectra of different supernova types one week after explosion (Filippenko 1997).

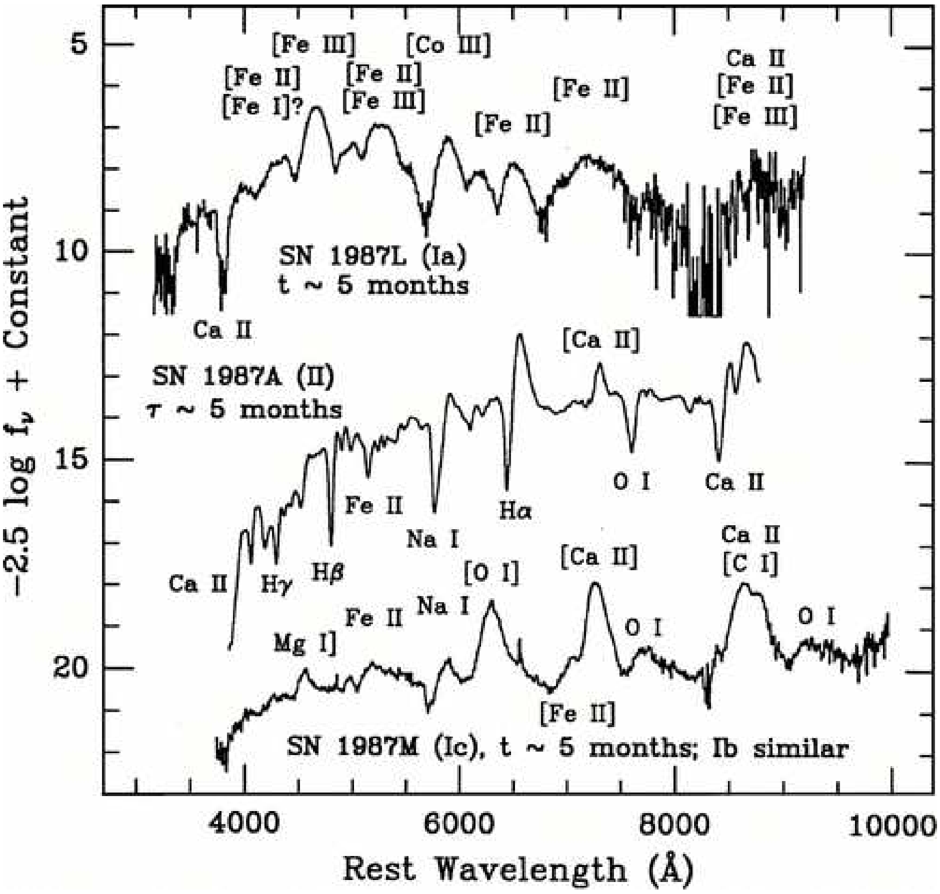

Figure 56: Spectra of different supernova types 5 months after explosion (Filippenko 1997).

however, shows that there are important differences within both of these classes. The most important are between the extremely heterogenous Type I's. These are shown as the top and lower two spectra in Fig. 55. While neither of them have any trace of Hα, the observational distinction is that the Type Ia's have a strong Si II λ6150 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 I b/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. 56 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 treason 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 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. 57 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

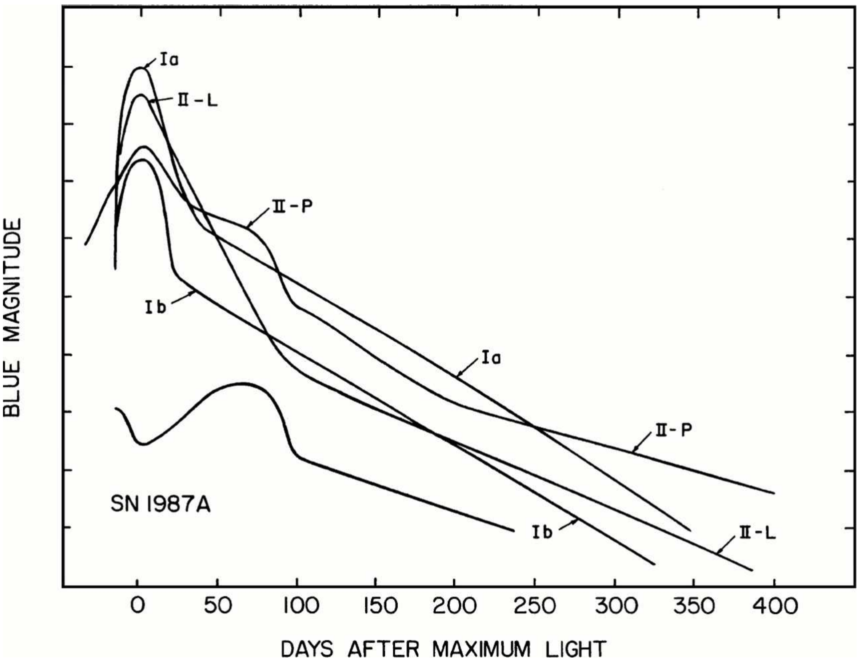

Figure 57: Light curves of different supernova types (Filippenko 1997).

evolution. However, while the Type Ia's 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 $\sim$ 50 days differ considerably.

While the Type II's have a fairly large range of light curves, one can distinguish two main types. The IIP's 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 $\sim$ 100 days. This is the reason for the P = plateau. After this there is a drop by a magnitude or more, and then they decline with a fairly uniform rate of $\sim$ 1 magnitude per 100 days.

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

In addition to theses 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. 58 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.

3.3 Radioactivity

As we saw in §2.9, 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

$$^{56}\mathrm{Ni} \rightarrow {}^{56}\mathrm{Co} + \gamma \ . \tag{3.5}$$

In this process it emits gamma-rays with energies 0.158 – 0.812 MeV (see Fig. 59). The ^{56}Co isotope resulting from this decay is, however, not stable either, but decays by electron capture or by positron decay according to

$$\begin{aligned} ^{56}\mathrm{Co} &\rightarrow {}^{56}\mathrm{Fe} + \gamma && (3.6)\\ &\rightarrow {}^{56}\mathrm{Fe} + e^{+} && (3.7) \end{aligned}$$

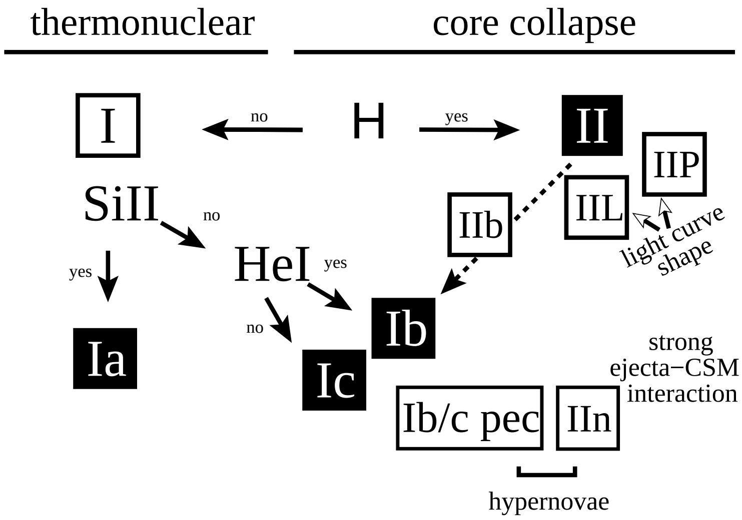

Figure 58: Classification scheme based on spectra and light curves of different kinds of supernovae (Turatto, 2003).

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

$$^{57}\mathrm{Ni} \rightarrow {}^{57}\mathrm{Co} + \gamma \tag{3.8}$$

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

$$^{57}\mathrm{Co} \rightarrow {}^{57}\mathrm{Fe} + \gamma \tag{3.9}$$

with $\tau = 390$ days.

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

$$^{44}\mathrm{Ti} \rightarrow {}^{44}\mathrm{Sc} + \gamma \tag{3.10}$$

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

$$\begin{aligned} ^{44}\mathrm{Sc} &\rightarrow {}^{44}\mathrm{Ca} + \gamma && (3.11)\\ &\rightarrow {}^{44}\mathrm{Ca} + e^{+} && (3.12) \end{aligned}$$

(see Fig. 60).

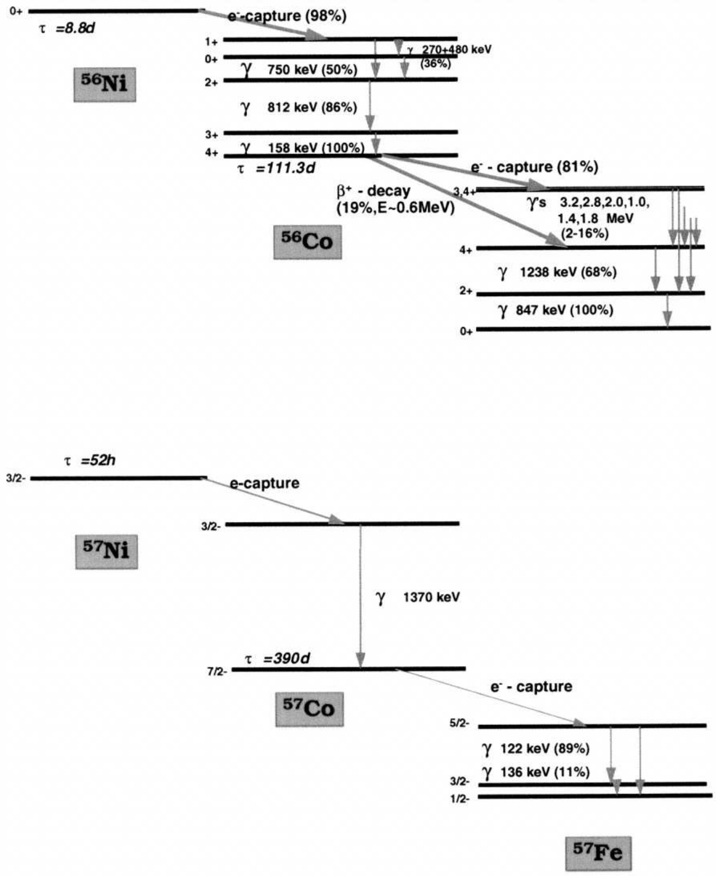

Figure 59: Transitions in the ^{56}Ni and ^{57}Ni decays (Diehl & Timmes 1998).

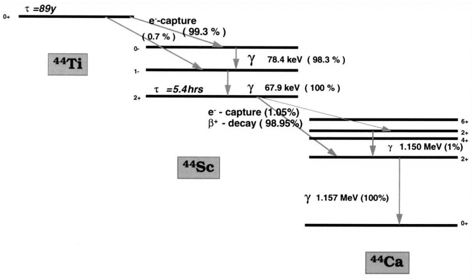

Figure 60: Transitions involved in the ^{44}Ti and ^{44}Sc decays (Diehl & Timmes 1998).

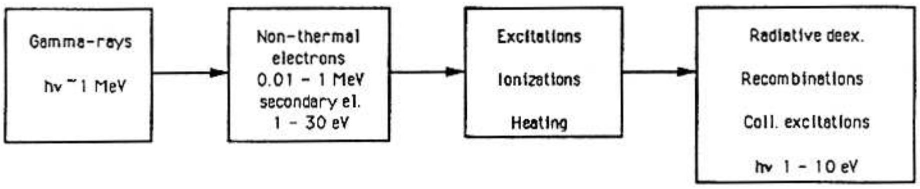

Figure 61: The different steps in the thermalization of the gamma-rays. Positrons undergo a similar thermalization history, except for the first step.

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. 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 energy lost in each Compton scattering is transferred to the electron, which therefore have energies ~ 100 keV - 1 MeV. These primary electrons lose their energy by ionizations and excitations of the atoms in the ejecta, and by Coulomb scattering against the free electrons. The secondary electrons resulting from the ionizations by the primary electron will in turn lose their energy by the same three processes. The result is therefore a cascade of secondary electrons. While the primary electron loses its energy mainly in ionizations and excitations (unless the medium is strongly ionized), the secondary electrons have an energy of $\lesssim 30$ eV, and can therefore at most give rise to one or two more ionizations. Finally when the energy of the secondary electrons is below that of the excitation energy of the lowest level in the most abundant ions, it loses all their energy by inelastic Coulomb scattering (Fig. 62).

Summarizing this, one fraction of the energy goes into heating of the thermal electrons, one into ionizations of neutral and singly ionized atoms, and one into excitations of discrete levels in these. The exact fraction depends on the ionization. At high ionization $\gtrsim 0.1$? most is going into heating, while at very low ionization approximately one third is going into each

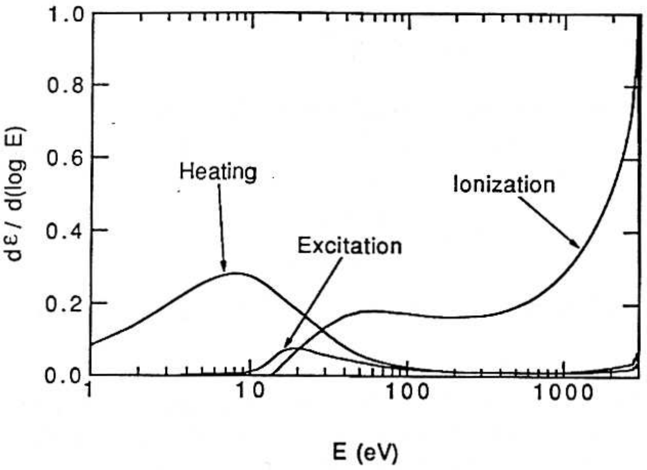

Figure 62: The energy going into the different channels as function of the electron energy for pure oxygen.

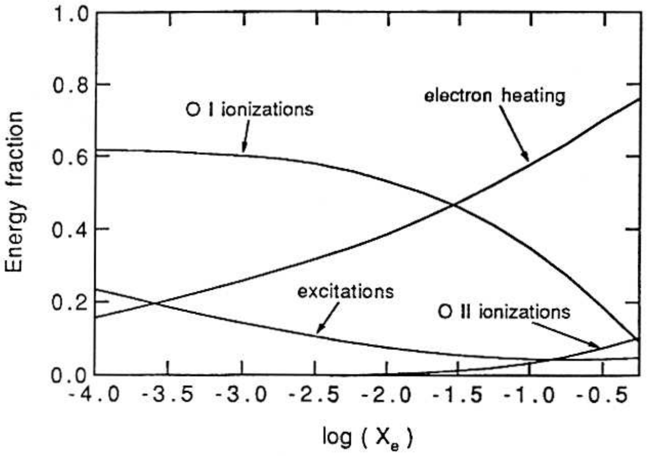

Figure 63: The fractions going into heating, ionization and excitation as function of the electron fraction for a pure gas of oxygen.

channel. A proper determination requires either a Monte Carlo calculation of the cascade, or solving the Boltzmann equation for the electron distribution (see e.g., Kozma & CF 1992). As an example we show in Fig. 63 the different fractions for the gamma-ray thermalization as function of the electron fraction for a pure oxygen plasma.

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

$$L_\gamma = \quad 1.27 \times 10^{42} \left(\frac{M(^{56}\mathrm{Ni})}{0.1\ \mathrm{M}_\odot} \right) e^{-t/111.3^\mathrm{d}} + \tag{3.13}$$

$$6.9 \times 10^{38} \left(\frac{M(^{57}\mathrm{Co})}{5. \times 10^{-3}\ \mathrm{M}_\odot} \right) e^{-t/390.^\mathrm{d}} + \tag{3.14}$$

$$4.1 \times 10^{36} \left(\frac{M(^{44}\mathrm{Ti})}{10^{-4}\ \mathrm{M}_\odot} \right) e^{-t/89.^\mathrm{yrs}} \mathrm{erg\,s}^{-1}, \tag{3.15}$$

and the positron input by

$$L_+ = \quad 4.44 \times 10^{40} \left(\frac{M(^{56}\mathrm{Ni})}{0.1\ \mathrm{M}_\odot} \right) e^{-t/111.3^\mathrm{d}} + \tag{3.16}$$

$$1.3 \times 10^{36} \left(\frac{M(^{44}\mathrm{Ti})}{10^{-4}\ \mathrm{M}_\odot} \right) e^{-t/89.^\mathrm{yrs}} \mathrm{erg\,s}^{-1} \tag{3.17}$$

3.4 Light curves

3.4.1 Shock breakout

As the shock reaches the surface of the star the hot radiation will be able to escape. This is the first time we usually observe the explosion as a supernova.

At the time of shock breakout approximately half of the total energy is thermal. Because the shock is still radiation dominated $E_\mathrm{tot}/2 \approx 4\pi/3R^3aT^4$, and the temperature behind the shock is therefore

$$T \approx \left(\frac{3E_\mathrm{tot}}{8\pi a R^3} \right)^{1/4} \approx 2 \times 10^5 \left(\frac{E_\mathrm{tot}}{10^{51}\mathrm{ergs}} \right)^{1/4} \left(\frac{R}{10^{13}\mathrm{cm}} \right)^{-3/4} \mathrm{K} \tag{3.18}$$

A compact star, like a Wolf-Rayet star, will therefore have an effective temperature $\gtrsim 5 \times 10^6$ K, while a red supergiant with $R \sim 10^{14}$ cm will have an effective temperature $\lesssim 5 \times 10^4$ K.

For SN 1987A, with a progenitor radius of $\sim 2 \times 10^{12}$ cm and an energy $\sim 2 \times 10^{51}$ ergs, the effective temperature was therefore $\sim 4 \times 10^5$ K. In reality, the scattering of the radiation, as well as other effects, increases the radiation temperature by a factor of about two over the effective temperature.

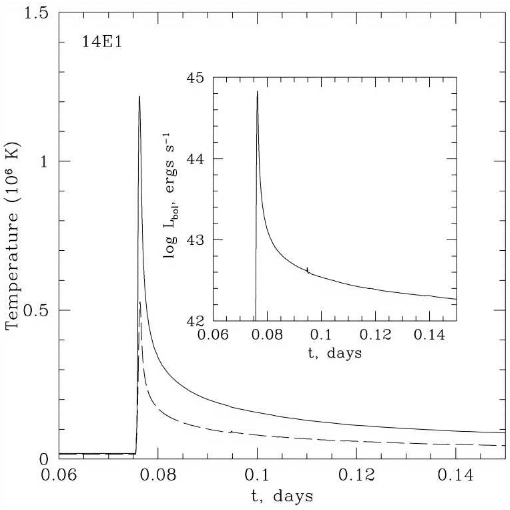

Figure 64: Temperatures and luminosity for the shock breakout for SN 1987A. The solid line shows the radiation temperature from a blackbody fit and the dashed the effective temperature. (Blinnikov et. al. 2000).

Especially for compact stars the duration will be very short. For SN 1987A the time scale was only 3-4 min. The peak luminosity was $\sim 10^{45}$ erg s^{-1}, and the total energy emitted in the burst $\sim (1-2)\times 10^{47}$ ergs (Fig. 64).

Because of the short time scale, direct observations of the breakout burst is difficult. For SN 1987A one, however, had the fortunate situation of having some circumstellar material in the form of the well known ring, lost by the progenitor, outside of the supernova. This was ionized by the radiation from the shock breakout, and then recombined on a time scale of years. From observations of the strengths of the emission lines one can determine both the state of ionization in the ring and its temperature. One then finds that to explain the ionization of the ring the burst should have a radiation temperature close to that estimated above. In this way one can indirectly confirm the reality of this extremely short event.

3.4.2 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. (2.16), which we write as

$$t_{\rm diff} = \frac{3R^2\rho\kappa}{\pi^2 c} \tag{3.19}$$

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

$$\frac{t_{\mathrm{diff}}}{t_{exp}} = 1.9 \left(\frac{M}{\mathrm{M}_\odot}\right) \left(\frac{V}{10^4\ \mathrm{km\ s^{-1}}}\right) \left(\frac{R}{10^{15}\mathrm{cm}}\right)^{-2} \qquad (3.20)$$

For a typical mass of 10 $\mathrm{M}_\odot$ we therefore find that not until the supernova has expanded to $R_{peak} \sim 4 \times 10^{15}$ 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 §2.2, 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

$$E_{int} = (\gamma - 1)^{-1} pV = 3K\rho^{4/3}V \propto \rho^{1/3} \propto R^{-1} \qquad (3.21)$$

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 §3.3, the radioactive isotopes created in the explosion give rise to gamma-rays and positrons as they decay. These are loosing 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.

Under LTE conditions hydrogen becomes neutral at ~ 5000 K. Because the cooling occurs from the surface (photosphere), there will be a recombination wave propagating inwards in the ejecta. The recombination front is at approximately constant radius (but decreasing velocity since $V = r/t$). Because the recombination temperature is constant this means that the luminosity is also constant.

The duration of the plateau is approximately given by

$$t_{plat} = 99 \left(\frac{\kappa R_{0\_500} M_{10}^3}{E_{51} T_{ion,5050}^4} \right)^{1/6} \text{ days} \tag{3.22}$$

(Popov 1993). From this expression it is evident that to have an appreciable plateau phase the hydrogen envelope mass must be large. In addition, the initial radius R_0 has to be of the order of a red supergiant.

It is therefore likely that the Type IIL's, which lack a plateau, have a low envelope mass.

3.4.3 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 angle and energy averaged opacity one can for ^{56}Ni and ^{56}Co use $\kappa_\gamma = 0.06 Z/A$ cm^{2}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 except for epochs later than $\sim$ 500 days they can be considered to be stopped and annihilate on the same spot as the radioactive decay. If there is a non-radial magnetic field this becomes an even better approximation. Further, since we neglect diffusion we are only considering epochs later than $\sim$ 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 then 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

$$L_{bol} = 1.27 \times 10^{42} \left(\frac{M(^{56}\text{Ni})}{0.1 \text{ M}_\odot} \right) e^{-t/111.3^d} [(1 - e^{-\tau_\gamma}) + 0.035] \text{ erg s}^{-1}. \tag{3.23}$$

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

$$\begin{aligned} \tau_\gamma &= \kappa_\gamma \rho R = \kappa_\gamma \frac{3}{4\pi}\frac{M}{V^2 t^2} && (3.24)\\ &= 0.38\frac{Z}{A}\left(\frac{M}{\mathrm{M}_\odot}\right)\left(\frac{V}{10^4\ \mathrm{km\ s^{-1}}}\right)^{-2}\left(\frac{t}{100\ \mathrm{days}}\right)^{-2} && (3.25)\end{aligned}$$

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 $M \sim 4\ \mathrm{M}_\odot$ and for the expansion velocity of the core ~ 2000 km s^{-1}. We than get $\tau_\gamma \sim 40(t/100\ \mathrm{days})^{-2}$. The gamma-rays are therefore in this case trapped up to $\sim$ 600 days. As another extreme case we can take a Type Ic supernova, with $M \sim 6\ \mathrm{M}_\odot$, $Z/A \sim 0.5$ and an expansion velocity $\sim 10,000$ km s^{-1}. In this case we get $\tau_\gamma \sim 1.2(t/100\ \mathrm{days})^{-2}$, and the ejecta is therefore transparent already at $\sim$ 100 days, or earlier for higher ejecta velocities.

Eq. (3.23) 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 (§3.2).

From Eq. (3.23) 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.

3.4.4 The bolometric light curve of SN 1987A

As an example of the usefulness of the bolometric light curve we take SN 1987A. In Fig. 65 we show this during the first 1000 days. After the diffusion phase, which ends by day $\sim$ 130, and up to day $\sim$ 400, the light curve closely follows the predicted linear relation expected for full trapping, $M_{bol} = -t2.5\log e/111.3^d + \mathrm{const} = -\mathrm{t}/102.5^{\mathrm{d}} + \mathrm{const}$. After day 400 there is an increasing deviation from the full trapping case, well fit by Eq. (3.23) with $\tau_\gamma = 30(t/100^d)^{-2}$, showing that some of the gamma-rays now escape the

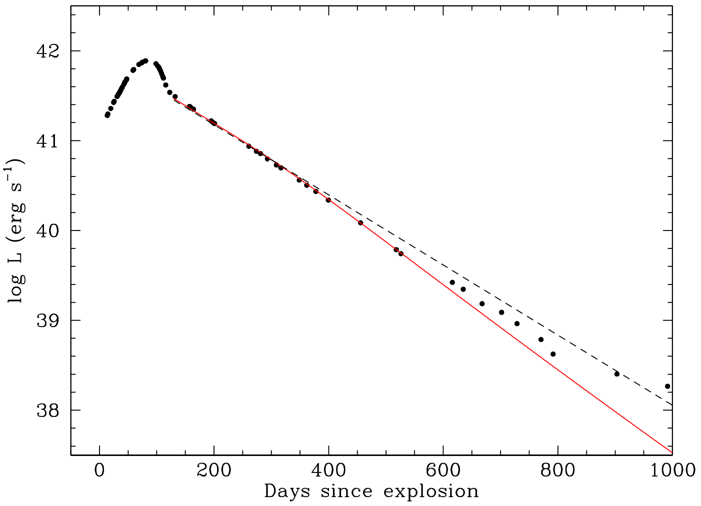

Figure 65: Bolometric light curve from ESO data for SN 1987A during the first 1000 days. The solid line gives the bolometric light curve from Eq. (3.23) with $M(^{56}\mathrm{Ni}) = 0.07\ \mathrm{M}_\odot$ and $\tau_\gamma = 30(t/100^d)^{-2}$. The dashed line shows the total radioactive energy, including that escaping the ejecta. (Data from Bouchet et al. 1990).

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 $M(^{56}\mathrm{Ni}) = 0.07\ \mathrm{M}_\odot$. 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. At this point the simple model above becomes questionable. The reason is that some of the radiative processes balancing the radioactive input, in particular the recombination, becomes slower than the radioactive time scale. This makes a time dependent model, including all the atomic physics of the recombination and cooling processes. This can be done and one then finds a good agreement with $M(^{57}\mathrm{Ni}) = 3.3 \times 10^{-3}\ \mathrm{M}_\odot$ (Fig. 66).

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 $\sim 25\%$ 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

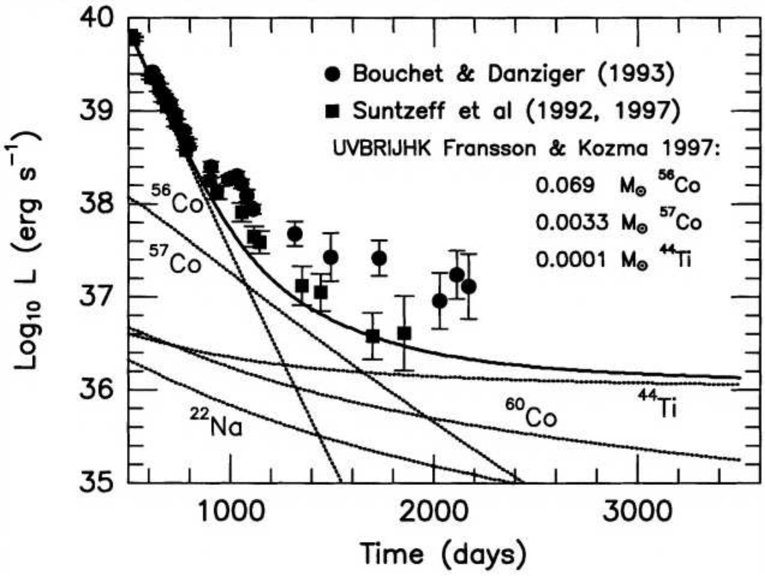

Figure 66: Bolometric light curve for SN 1987A compared to theoretical light curves from different isotopes. (Diehl & Timmes 1998).

time scale means that the light curve will be essentially flat after this epoch.

Exercise:

Calculate the epochs when the energy input due to ^{56}Ni, ^{57}Ni and ^{44}Ti dominate, respectively, for $M(^{56}\mathrm{Ni}) = 0.07$ M$_\odot$, $M(^{57}\mathrm{Ni}) = 3.3 \times 10^{-3}$ M$_\odot$, and $M(^{44}\mathrm{Ti}) = 1.0 \times 10^{-4}$ M$_\odot$. First assume an ejecta mass of 4 M$_\odot$ and a velocity of 2000 km s^{-1} (SN 1987A), and then 1.4 M$_\odot$, a velocity of 10,000 km s^{-1}, and $M(^{56}\mathrm{Ni}) = 0.7$ M$_\odot$, $M(^{57}\mathrm{Ni}) = 3.3 \times 10^{-3}$ M$_\odot$, and $M(^{44}\mathrm{Ti}) = 1.0 \times 10^{-4}$ M$_\odot$ (Type Ia). When do the positrons dominate?

While the bolometric light curve in principle is very appealing to use for the determination of these isotopes, there are some major observational and theoretical problems connected to this. We have already mentioned the problems connected with the delayed recombination and cooling. This can, however, be handled by more detailed modeling. Worse is the fact that after $\sim$ 500 days most of the luminosity comes out in the mid- and far-IR. This is due to both the cooling of the ejecta, and in the case of SN 1987A, dust formation in the ejecta (see later). Especially the spectral region above $\sim 20\mu$ is very difficult or impossible to observe from the ground. For SN

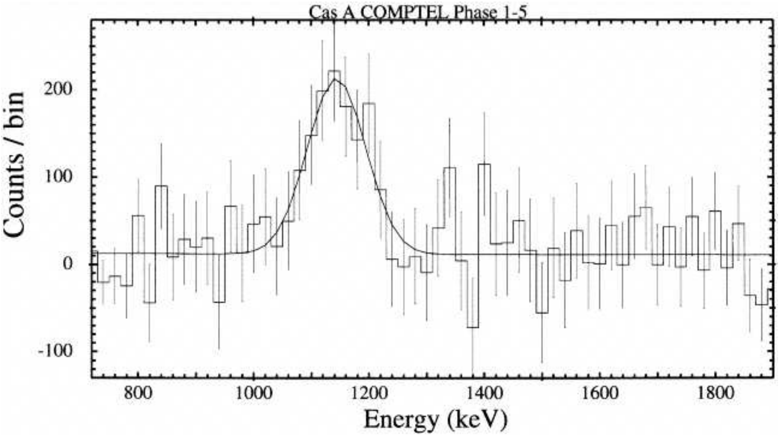

Figure 67: Gamma-ray spectrum of Cas A obtained with COMPTEL on the Compton Gamma Ray Observatory. (Iyudin et al. 1994).

1987A one was, thanks to its low distance, fortunate to be able to observe it with the Kuiper Airborne Observatory (KAO). One could therefore up to ~ 1400 days correct the ground based observations for the far-IR contribution. Unfortunately this is difficult for more distant supernovae, although the newly launched Spitzer observatory will help in this respect.

One way of circumventing these problems is to use broad band optical and near-IR observations, and from modeling of these calculate the bolometric correction and thus the isotopic masses. This has been used for SN 1987A, where it is found that a mass of $(1-2)\times10^{-4}$ $M_{\odot}$ of ^{44}Ti was formed in the explosion (Kozma & CF 2002). As we discussed in §3.3, 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. 67). The strength of the line corresponds to $M(^{44}\mathrm{Ti}) \sim 1.7 \times 10^{-4}$ $M_{\odot}$, close to that inferred for SN 1987A.

3.5 Spectra

The spectra of supernova involve virtually all complications a radiative transfer problem can have. This includes spherical geometry, a high velocity field with the consequent blending of lines, extreme NLTE effects, and non-thermal gamma-ray input. However, paradoxically, the high velocities

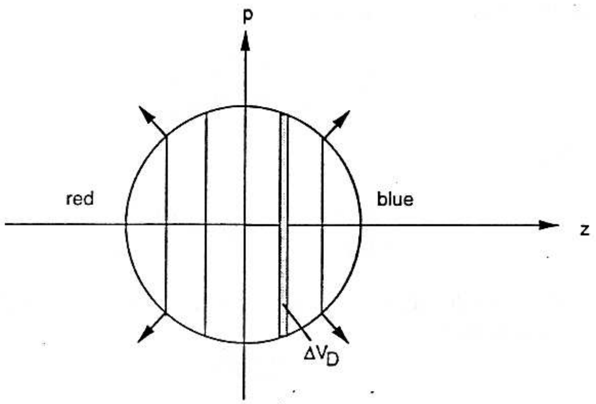

Figure 68: Surfaces of constant frequency along the line of sight, for a velocity $v(r) \propto r$. The shaded area indicates schematically the region within which a photon emitted at a given frequency interacts.

also offer some simplifications, and unique possibilities.

One of the most important of these is that the homologous expansion of the ejecta allows us to probe the conditions in the envelope from observations of spectral line profiles.

An important concept for high velocity flows is the surfaces of constant velocity along the line of sight (SCV). This is defined by

$$\frac{\nu - \nu_0}{\nu_0} = \frac{V_z}{c} = \frac{z}{r}\frac{V(r)}{c} = \text{constant} \tag{3.26}$$

where z is the coordinate along the line of sight, and V_z the corresponding velocity (Fig. 68). For a homologously expanding ejecta $V(r) = V_0(r/R_0) = r/t$, where V_0 is the velocity at a reference radius R_0, and one obtains

$$\frac{z}{R_0} = \frac{ct}{R_0}\frac{(\nu - \nu_0)}{\nu_0} \equiv x \tag{3.27}$$

Therefore, the SCV's are surfaces of constant z, perpendicular to the line of sight (Fig. 68). As one shifts in frequency from red to blue in the line, one probes the ejecta from the back in towards the center and out to the front.

An important simplification is possible because when the expansion velocity is much larger than the thermal width of the line, only a narrow interval in depth contributes to a given frequency (Fig. 68). This is the basis of the Sobolev approximation. A simple derivation of the optical depth in this approximation starts with

$$\tau_\nu = \sigma_\nu n_1 \, \delta l \tag{3.28}$$

where σ_ν is the cross section at frequency ν, n_1 the population in the lower state and δl the thickness of the layer with which the photon interacts. For a box line profile we have

$$\sigma_\nu = \frac{\pi e^2}{mc} \frac{f_{12}}{\Delta\nu_D} = \frac{A_{21} g_2 \lambda^2}{8\pi g_1 \Delta\nu_D}, \tag{3.29}$$

where f_{12} is the absorption oscillator strength of the line, and A_{21} the transition probability. A photon emitted at a radius r interacts with the surrounding gas within a region corresponding to a velocity difference by one Doppler width of the line, Δv_D. Regions further away are redshifted too much to come within the line width. As in an expanding universe with velocity proportional to radius, this corresponds to a spherical region with velocity radius Δv_D. The escape probability is therefore the same in all directions, and can be evaluated in the radial direction. The radius of the interaction region is then

$$\delta l = \frac{\partial r}{\partial V} \Delta v_D = \left[\frac{\partial}{\partial r} \left(\frac{V_0}{R_0} r \right) \right]^{-1} \Delta v_D = \frac{R_0}{V_0} \Delta v_D = \frac{\Delta\nu_D}{\nu} ct, \tag{3.30}$$

Including the usual stimulated emission factor we get

$$\tau = \frac{A_{21} \lambda^3 g_2}{8\pi g_1} \left(n_1 - \frac{g_1}{g_2} n_2 \right) t. \tag{3.31}$$

Since τ is isotropic, the escape probability averaged over the region is given by

$$\beta = \frac{1}{2\tau} \int_0^\tau \int_{-1}^1 e^{-\tau'} d\mu \, d\tau' = \frac{1 - e^{-\tau}}{\tau}. \tag{3.32}$$

Knowing the optical depth and escape probability, one can calculate the line profile for a given source function, which in terms of the level populations is given by

$$S = \frac{2h\nu^3}{c^2 (n_2 g_1 / n_1 g_2 - 1)} \tag{3.33}$$

For simplicity, we assume that there is no background continuum. This is a good approximation for supernovae at late times, but is not valid in the early stages.

Using cylindrical coordinates, with p perpendicular to the line of sight (Fig. 68), the flux is given by

$$F_x = 2\pi \int\limits_{p_{\rm min}}^{p_{\rm max}} [S(p, z = xR)(1 - e^{-\tau}) + e^{-\tau} I_{\rm cont}] p\, dp \tag{3.34}$$

where the last term is the attenuated continuum contribution from the background, ususally the photosphere.

Using $p\, dp = r\, dr$ for constant z (i.e., constant x), this can be written as

$$F_x = 2\pi \int\limits_{r_{\rm min}}^{R} [S(r)(1 - e^{-\tau}) + e^{-\tau} I_{\rm cont}] r\, dr \tag{3.35}$$

where $r_{\rm min} = xR$, and x is given by Eq. (3.27). If the level populations are known, the line profile can now be calculated.

The source function has two contributions. One part is from the photons locally created at the point by radiative de-excitation, and the other from scattering of photons from the background continuum. In the simple, but interesting case of two discrete levels, the source function is given by

$$S = \frac{(1 - \epsilon)\beta W I_{\rm cont} + \epsilon B_\nu}{(1 - \epsilon)\beta + \epsilon} \tag{3.36}$$

where ϵ is the probability for collisional de-excitation, and $W(r) = \frac{1}{2}[1 - \sqrt{1 - (R_{\rm phot}/r)^2}]$ is the dilution factor (see J. I. Castor, 1970, MN 149, 111 for a derivation).

As an illustrative example of a line profile calculation we consider the case of an optically thick line, $\tau \gg 1$, and pure scattering, $\epsilon = 0$ and . The source function is then given by $S \approx W I_{\rm cont}$, i.e., simply the background continuum emission at the wavelength of the line times the fraction of the total hemisphere covered by the photosphere.

$$F_x = 2\pi \int\limits_{r_{\rm min}}^{r_{\rm max}} [W(r) I_{\rm cont} + e^{-\tau} I_{\rm cont}] r\, dr \tag{3.37}$$

For convenience we scale the frequency to the maximum frequency shift, i.e., to the radius of the supernova, $R_{\rm max} = V_{\rm max} t$. With this scaling the dimensionless frequency shift $x \equiv c/V_{\rm max}(\nu - \nu_0)/\nu_0 = z/R_{\rm max}$ is then in the range $-1 \leq x \leq 1$. The integration is over the SCV, i.e., $z = R_{\rm max} x =$

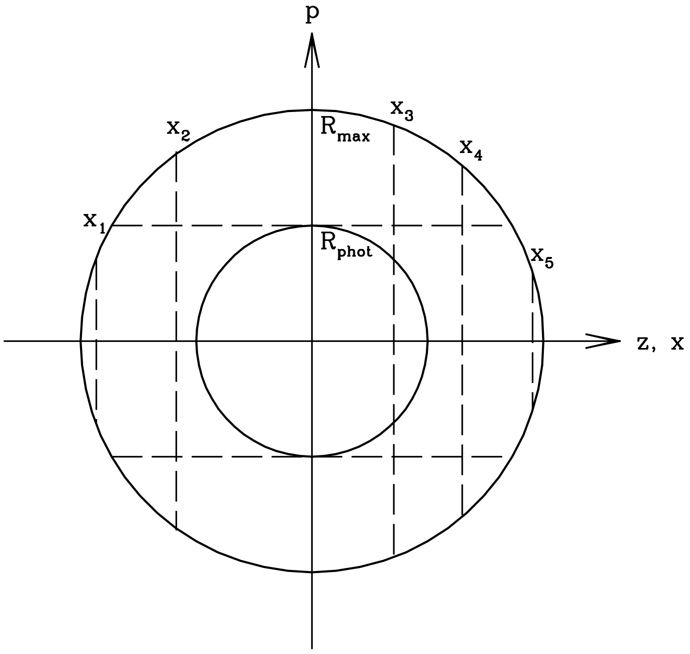

Figure 69: Surfaces of constant frequency along the line of sight, for a velocity $v(r) \propto r$.

constant. Depending on the frequency we now have to consider five different possibilities, depending on to what extent the scattering surface scatters the photospheric photons out of the line of sight, i.e., the $e^{-\tau} I_{\rm cont}$ term in Eq. (3.37) (Fig. 69).

To determine these cases we use the relation between the radius, impact parameter, p, and frequency x, given by $r^2 = p^2 + z^2 = p^2 + x^2 R_{\rm max}^2$. For $-1 \leq x \leq -[1 - (R_{\rm phot}/R_{\rm max})^2]^{1/2}$ (Case 1 in Fig. 69) the scattering surface is occulted by the photosphere and we only receive the photons from the photosphere. For $-[1 - (R_{\rm phot}/R_{\rm max})^2]^{1/2} \leq x \leq 0$ (Case 2) we still receive the full contribution from the photosphere, but now also the part of the scattering contribution which is not occulted.

For positive x, the scattering surface may occult part of the photosphere. In Case 3 which occurs for $0 \leq x \leq R_{\rm phot}/R_{\rm max}$ there is both a photospheric contribution and a scattering contribution. For Case 4 when $R_{\rm phot}/R_{\rm max} \leq x \leq [1 - (R_{\rm phot}/R_{\rm max})^2]^{1/2}$ the scattering surface occults the photosphere completely, and the only contribution is that from scattering. Finally for Case 5 in the range $[1 - (R_{\rm phot}/R_{\rm max})^2]^{1/2} \leq x \leq 1$ the central part of the photosphere is occulted and the photospheric contribution is decreased by this factor, while the scattering part is unaffected. To further complicate the situation, if $R_{\rm max}/R_{\rm phot} \leq \sqrt{2}$ then $[1 - (R_{\rm phot}/R_{\rm max})^2]^{1/2} \leq R_{\rm phot}/R_{\rm max}$ and the limits between the three cases above are changed according to this (Cases 3a-5a).

To calculate the photospheric contribution one calculates the area perpendicular to the line of sight, which is not occulted by the scattering surface. In Case 5 , e.g., this is $\pi I_{\rm cont}[R_{\rm phot}^2 - (R_{\rm max}^2 - z^2)] = \pi R_{\rm phot}^2 I_{\rm cont}[1 - (R_{\rm max}/R_{\rm phot})^2(1 - x^2)]$. The scattering integral is in this case calculated from $r_{\rm min} = z = R_{\rm max} x$ to $R_{\rm max}$. In Table 4 we give the expresions for the two contributions for the different cases.

The integral over the dilution factor can be integrated analytically to give

$$F_x = F_x^{\rm cont} + 2\pi \int_{r_{\rm min}}^{R_{\rm max}} W(r) r dr = \pi R_{\rm phot}^2 I_{\rm cont} \left[\frac{F_x^{\rm cont}}{\pi R_{\rm phot}^2 I_{\rm cont}} + f(\alpha) - f(\frac{r_{\rm min}}{R_{\rm phot}}) \right] \tag{3.38}$$

where $\alpha \equiv R_{\rm max}/R_{\rm phot}$ and

$$f(y) = \frac{1}{2}\{y^2 - y(y^2 - 1)^{1/2} + \ln[y + (y^2 - 1)^{1/2}]\}. \tag{3.39}$$

This completes the analytical solution.

Table 4: **Contributions to the line profile. Note $\alpha \equiv R_{\rm max}/R_{\rm phot}$.**

Case	Frequency range	$F_x^{\rm cont}/\pi R_{\rm phot}^2 I_{\rm cont}$	$r_{\rm min}/R_{\rm phot}$
1	$-1 \le x \le -(1-\alpha^{-2})^{1/2}$	1	0
2	$-(1-\alpha^{-2})^{1/2} \le x \le 0$	1	$\left[1+\alpha^2 x^2\right]^{1/2}$
3	$0 \le x \le \alpha^{-1}$	$1-\alpha^2 x^2$	1
4	$\alpha^{-1} \le x \le (1-\alpha^{-2})^{1/2}$	0	αx
5	$(1-\alpha^{-2})^{1/2} \le x \le 1$	$1-\alpha^2(1-x^2)$	αx
3a	$0 \le x \le (1-\alpha^{-2})^{1/2}$	$1-\alpha^2 x^2$	1
4a	$(1-\alpha^{-2})^{1/2} \le x \le \alpha^{-1}$	$2-\alpha^2$	1
5a	$\alpha^{-1} \le x \le 1$	$1-\alpha^2(1-x^2)$	αx

In Fig. 70 we show the result for $\alpha = R_{\rm max}/R_{\rm phot} = 1.2, 1.5, 2.0, 4.0$ and 10.0. There are several interesting things to note from this example. First we note that the total equivalenth width of the line is close to zere, as is expected for a pure scattering line, where the photons scattered out of the line of sight from the approaching side are compensated by photons scatteried into the line of sight from the sides and the receding part of the ejecta. Further we note that as the ratio of $R_{\rm max}/R_{\rm phot}$ decreases more and more of the recedng part of the ejecta are occulted by the photosphere. This occurs for $x = -(1-R_{\rm phot}/R_{\rm max})^2)^{1/2}$. Finally, we note that the minimum of the intensity occurs for $x = (1-R_{\rm phot}/R_{\rm max})^2)^{1/2}$, i.e., the same frequency-shift as where the occultation is complete. Another interesting result is that for $\alpha < \sqrt{2}$ the blue part of the line has a section with a flat bottom of the line, as is sometimes observed.

Another interesting case for supernovae is that of a geometrically thin shell at a radius r_s. In this case we can take $S(r) = S_0\delta(r-r_s)$ and obtain

$$F_x = 2\pi S_0(1-e^{-\tau_s})r_s \qquad \text{if } |x| \le \frac{r_s}{R} \tag{3.40}$$

and zero otherwise, which means that the resulting line profile is flat out to a velocity corresponding to the shell velocity.

Exercise:

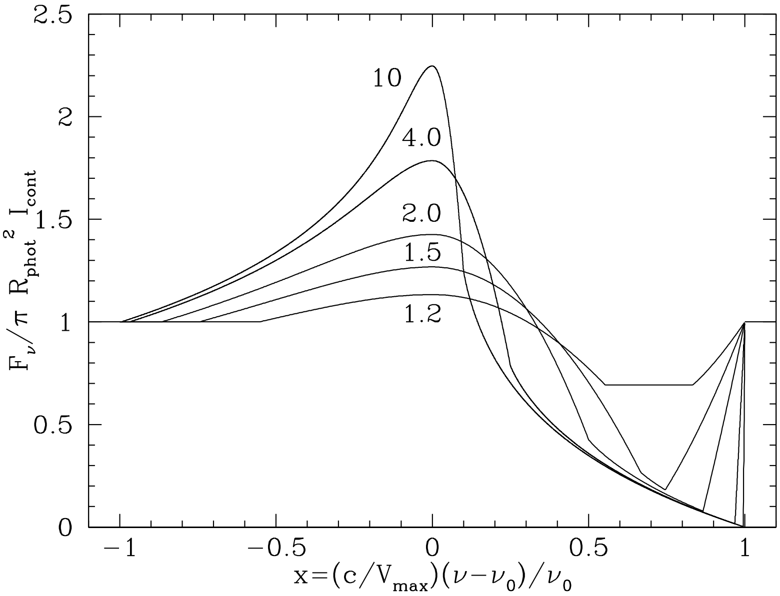

Figure 70: Line profiles for different values of $\alpha = R_{\max}/R_{\rm phot}$. Pure scattering in a velocity $v(r) \propto r$ ejecta is assumed.

a) Calculate the SCV's for a shell with constant velocity V_0 with inner radius R_{in} and outer R_{out}, and show that these are given by $\theta =$ constant, where θ is the angle to the line of sight of the observer.

b) Show that the optical depth in this case is given by $\tau = \tau_0/(1-\mu^2)$, where $\mu = \cos\theta$.

c) Show that the line profile is given by

$$F_x = 2\pi(1-x^2)\int_{R_{\rm in}}^{R_{\rm out}} S(r)(1-e^{-\tau_0(r)/(1-\mu^2)})r\,dr, \tag{3.41}$$

where $x = c(\nu-\nu_0)/V_0\nu_0$.

d) Calculate the line profile from the shell, assuming it is optically thin.

e) Same as d), but for an optically thick line. Compare with the $v \propto r$ case.

Another important case is that of an optically thin line. In this case

$$F_x = 2\pi \int_{r=xR}^{R} j(r) r \, dr. \tag{3.42}$$

From an observational point of view it is of more interest to invert this relation.

$$j(r = xR) = \frac{-1}{2\pi x R} \frac{dF_x}{dx}. \tag{3.43}$$

The line profile therefore directly gives the emissivity, and therefore the energy input, as a function of radius in the remnant.

A special feature of supernova spectra is the extreme degree of line blending, which is a result of the large expansion velocity. At early times the velocity may be up to 3×10^4 km s^{-1}, and the probability for fluorescence between various transitions is very large. Examples of this has been discussed by Lucy et al. (1991). The most extreme case is the ultraviolet part of the spectrum. Here a large number of ions, like C I-II, O I, Si I-II, Ti I-III, Fe I-III, and Co I-III, have strong transitions from low excitation levels, and the spacing between the lines is small. Using Eq. (3.31), with the density in the hydrogen envelope estimated by $n \approx 1.2 \times 10^9 (t/300 \text{ days})$ cm^{-3}, we can estimate the optical depth of a typical resonance line from an ion with relative abundance X_i in the envelope

$$\tau_{21} = 2.7 \times 10^3 \left(\frac{\lambda}{2500 \text{ Å}}\right)^3 \left(\frac{A_{21}}{10^7 \text{ s}^{-1}}\right) \left(\frac{X_i}{10^{-5}}\right) \left(\frac{t}{300 \, \text{days}}\right)^{-2} \tag{3.44}$$

Even low abundance elements in the envelope and core can therefore have large optical depths in the UV.

Consider now two lines with wavelengths λ_1 and λ_2, with $\lambda_1 < \lambda_2$. If a photon is emitted from a point with wavelength λ_1, it will relative to another point be redshifted to $\lambda_1' = (1 + v_r/c)\lambda_1$, where v_r is the *relative* velocity between two points. As already discussed, v_r is constant on a sphere with radius $d = R v_r / V$, so that $\lambda_1' = \lambda_2$ for

$$\frac{d}{R} = \frac{c}{V} \frac{(\lambda_2 - \lambda_1)}{\lambda_1}. \tag{3.45}$$

The requirement for scattering is that $d < R$, and that the optical depth is larger than one.

After the absorption, the photon will be emitted isotropically from the resonance point. It will consequently perform a random walk from line to line with step length d (varying from line pair to line pair), and at the same time be redshifted. This will continue until either there is a) a wavelength gap between the optically thick lines larger than $\lambda V/c$, b) the photon escapes in a transition to an intermediate excited level in the optical range, c) it is destroyed by continuum absorption, or d) escapes from the optically thick region of the supernova. If we for simplicity assume that there is a constant velocity spacing, $\Delta v = c\Delta\lambda/\lambda$ between the optically thick lines, the total number of scatterings for a photon until it reaches the boundary is $N \sim (R/d)^2 \sim (V/\Delta v)^2$. The Doppler shift after N scatterings is $N\Delta v \sim V^2/\Delta v$, which can be much larger than V, if $\Delta\lambda$ is small.

As an illustrative toy model we consider a uniform density ejecta with an expansion velocity of 3000 km s^{-1} and abundances typical for the core. In Fig. 71 we show the distribution of lines with optical depth larger than one at an age of 500 days. The lines have been selected from the line list by Kurucz and Peytremann (1975) and Kurucz (1981), and we assume an equal mixture of neutral and singly ionized ions. This is a very rough approximation, but because of the large depths, this is not expected to introduce too large errors in the fluxes. More serious are any incompletenesses in the line list. The most interesting feature is that most of the UV up to $\sim$ 3000 Å is covered by a forest of optically thick lines. Above $\sim$ 3000 Å there are increasingly wide gaps between the lines. Also at shorter wavelengths, there are some important gaps in the line distribution, the most obvious at $\sim$ 1200 Å, $\sim$ 1500 Å and at $\sim$ 2900 Å.

To schematically see the scattering effects of these lines we assume that the UV flux from the core is emitted in four emission lines, at 912 Å, 1356 Å, 1640 Å, and 2335 Å, with equal fluxes. The line emissivities are in an actual case determined by the γ-ray input to the core, and the resulting recombination cascade. We then calculate the scattering process with a Monte-Carlo code. Although not very realistic for the resulting spectrum, it illustrates the random walk aspect well, and the result of this experiment is shown in the lower panel of Fig. 71. As expected, there is an inverse correlation between the gaps in the line distribution and the emission from the supernova. The photons emitted in e.g. the O I $\lambda\lambda 1302-56$ lines scatter $\sim$ 10 times until they escape in the gap at $\sim$ 1500 Å. The photons emitted at 2335 Å (corresponding to either C II or Si II) do not find any gaps before $\sim$ 3300 Å, so most photons emitted between $\sim$ 2000 Å and $\sim$ 3300 Å emerge in this gap, if they have not been destroyed by continuum absorption or photon splitting, before escaping.

Resonance scattering is probably the most important factor for the strong UV deficiency in both Type IIP's, e.g., SN 1987A, and Type Ia and Ib/c su-

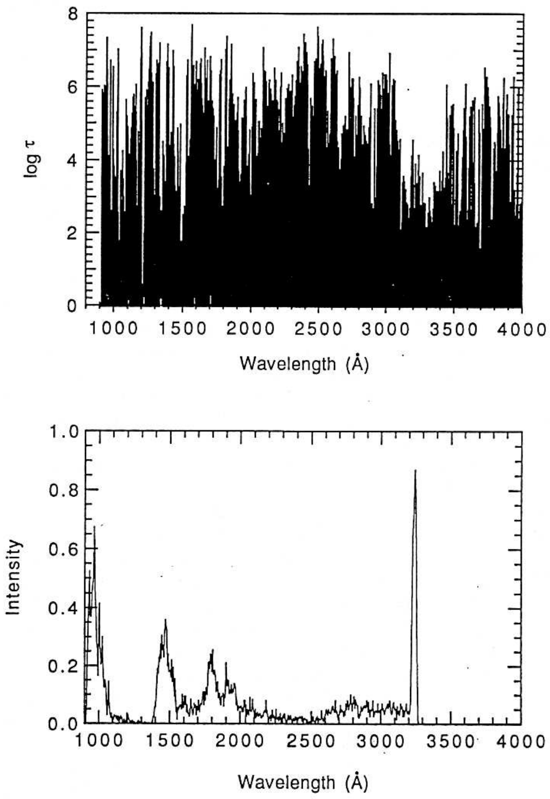

Figure 71: Upper panel: Distribution of lines with optical depth larger than one, for an expansion velocity of 3000 km s^{-1} and for a typical core composition, at 500 days. Note the gaps at $\sim$ 1200, 1500, and above 3000 Å. Lower panel: Emerging spectrum for the same parameters as in the upper panel. The emitting lines are assumed to be at 912, 1302, 1640, and 2335 Å. We note that the radiation escapes preferentially in the opacity gaps in the line distribution, and the wavelengths of these features have little to do with their emission wavelengths.

pernovae. Because of the high optical depths, solar abundances are sufficient to give this effect (Eq. 3.44), especially at early times when the velocities are higher, making blanketing even more efficient. It is interesting to note that the IUE spectra of SN 1987A even after a year showed a residual continuum flux below ~ 2500 Å, with the strongest features at 1300–1400 Å and 1800–2000 Å. These were very broad without any obvious correspondence to emission lines. It is likely that they corresponded to line emission from the core, e.g. [O I] $\lambda\lambda 1302, 1356$, which managed to escape in the wavelength regions where the number of optically thick lines is a minimum.

The total path length for the photon is $Nd \sim R^2/d \sim RV/\Delta v$, where d was defined in Eq. (3.45). This can be much larger than R, so that continuum absorption may be effective, even if the radial optical depth in the continuum is considerably less than one. The photons in Fig. 71 have on the average been scattered a total distance corresponding to 10–30 times the radial distance. This is important for both the Balmer continuum, and for dust absorption in the ejecta.

Also continuum scattering can be important. If a photon is scattered coherently in the rest frame of the scattering particle, as for Thompson scattering, it will like resonance scattering, be redshifted for an observer at rest. If the scattering optical depth is τ_s, the probability to be scattered N times is τ_s^N, if $\tau_s < 1$, and there will be a red extension to the line profile to $\Delta\lambda_{\rm tot} \sim N\lambda V/c$. Electron scattering has been proposed as an explanation for the observed line asymmetries at early epochs in SN 1987A.

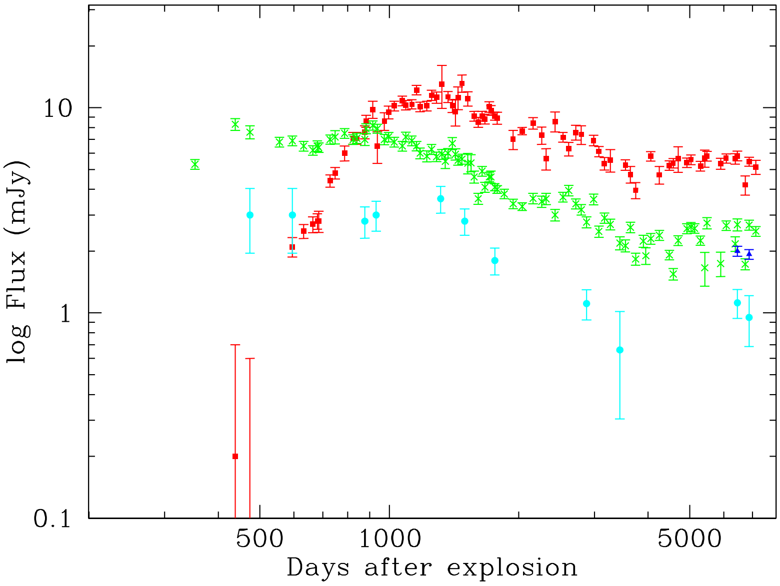

Figure 72: Radio light curves for SN 1979C from VLA. Squares 1.465 GHz,, crosses 4.88 GHz, filled circles 14.94 GHz (Data from Weiler et al.)

4 Interaction with a circumstellar medium

4.1 Observational evidence for CSI

The first strong evidence for circumstellar interaction (CSI) came from the radio. Although not seen eight days after explosion with VLA, the Type IIL SN 1979C showed nearly a year after explosion a strong flux at 4.88 GHz and 14.94 GHz (see Fig. 72). Later it also showed a radio turn-on at 1.465 GHz. This frequency dependent turn-on is characteristic for most radio supernovae observed to date. After the turn-on it has only decreased by a factor of ~ 2 over ~ 20 years.

After SN 1979C a large number of core collapse SNe of all types have been observed in the radio. When well sampled in time, they are all first seen at high frequencies and later at longer, strongly indicating some time dependent absorption mechanism. The details, like the rate at which they turn-on, as well as their decline rate, however, vary considerably.

Although it was from radio CSI was first noted, the optical spectra often show strong indications of CSI. Especially at late epochs when the radioactivity has died off CSI provides an additional energy source.

The best and most well studied example of strong CSI is SN 1993J in M81 at only 3.6 Mpc. In some respects it plays the same role for this class of SNe as SN1987A for the understanding of the core collapse and ejecta. In Fig. 73 we show the spectral evolution of this SN from day 3 after explosion up to day 433. This is remarkable in several ways. At early epochs it displayed a clear Hα line, and was therefore initially classified as a Type II SN. After about a month this line, however, disappeared, and instead a strong [O I] $\lambda\lambda 6300, 6364$ line, together with He I lines appeared. There was therefore a transition from a Type II to a Type Ib SN! Therefore it is usually classified as a Type IIb SN.

It is obvious this SN did not have much of hydrogen in its envelope. In fact models of the light curve and spectra showed that the mass of the hydrogen envelope was $\lesssim 0.5$ M$_\odot$. The most accepted model is that the SN was in a binary system, and that because of Roche lobe overflow it had lost most of its hydrogen envelope.

This was, however, not the end of the story. After about a year Hα returned, but now with a completely different line profile compared to the early epochs. Instead of a centrally peaked profile, the line was box-like, as seen in e.g., the day 433 spectrum in Fig. 73. As we saw in §3.5, this is characteristic of a line coming from a thin shell. It is therefore obvious that most of the action occurs not in the core, but at the outer edge of the SN. The velocity of this shell was $\sim 13,000$ km s^{-1}, decreasing slowly with time.

SN 1993J showed already from the first days a strong radio flux (left panel of Fig. 74). As with other radio supernovae, this first turned on at high frequencies and later at longer. Compared to SN 1979C, the turn-on was, however, considerably slower. In Fig. 84 we show the excellent radio light curves from VLA. We will discuss the modeling of these in §4.8.

That the emission was coming from a shell at high velocity was obvious from the spectacular VLBI images of SN 1993J which have been obtained by two different groups. In Fig. 75 we show a time sequence from day 50 up to day 1893. Perhaps the most surprising aspect of these images is how close to spherical the SN really is. In fact most of the emission is coming from a shell with thickness $\sim 30\%$ of the radius. The fluctuations seen in the images is at a level of $\sim 20?\%$, so a 'spherical cow' approximation is an extremely good one. This is especially remarkable because it is believed that the SN occurred in a binary system, with a lot of angular momentum.

SN 1993J has also been observed in X-rays with ROSAT, ASCA, COMPTON/GRO, Chandra and XMM (right panel of Fig. 74). During the first two months it displayed a hard thermal spectrum with $kT \sim 100$ keV. It

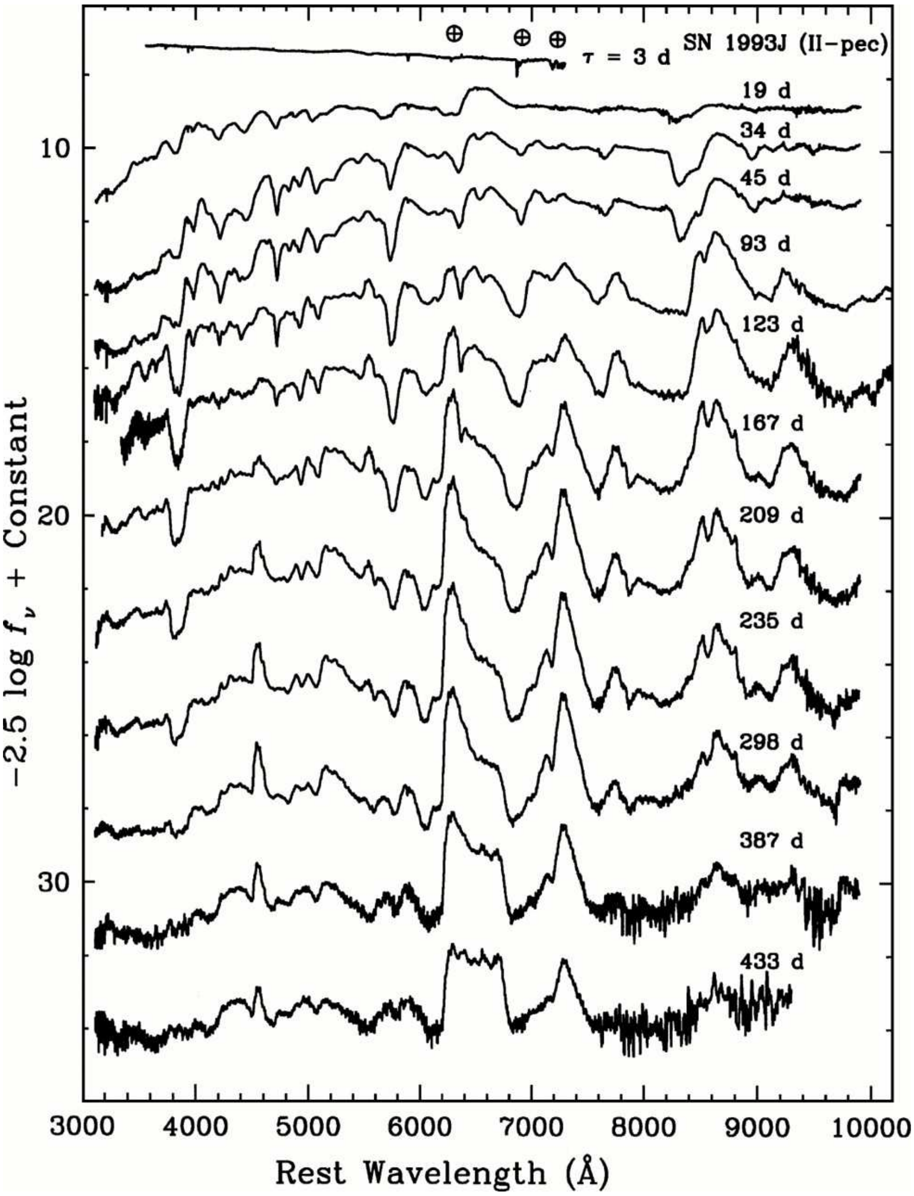

Figure 73: Spectral evolution of SN 1993J. Note the gradual change of the Hα line, defining it as a Type IIn supernova. (Filippenko 1997)

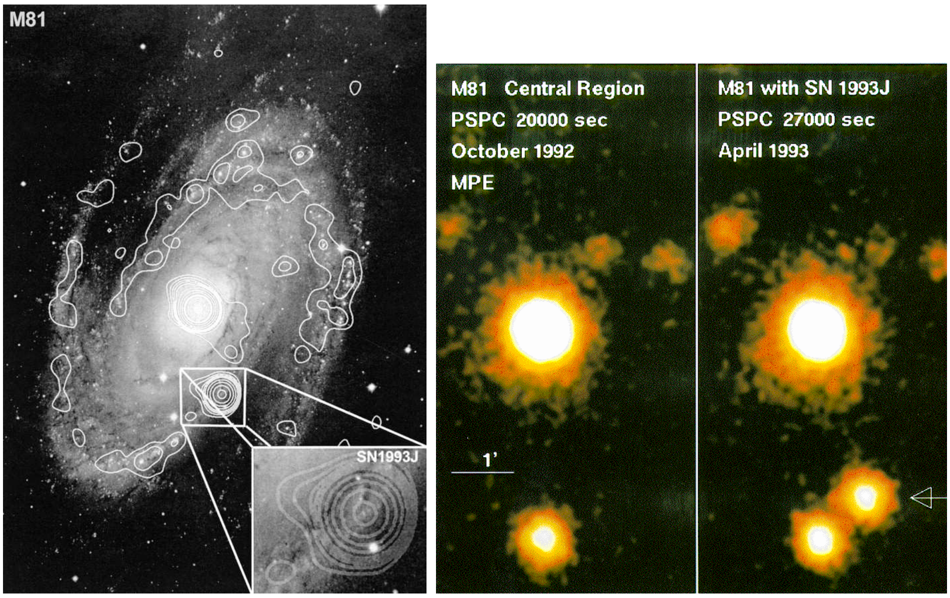

Figure 74: Left: VLA (left and ROSAT (right) images of SN 1993J. The scales are different in the two images. (VLA: Bartel et al. ROSAT: Zimmermann et al)

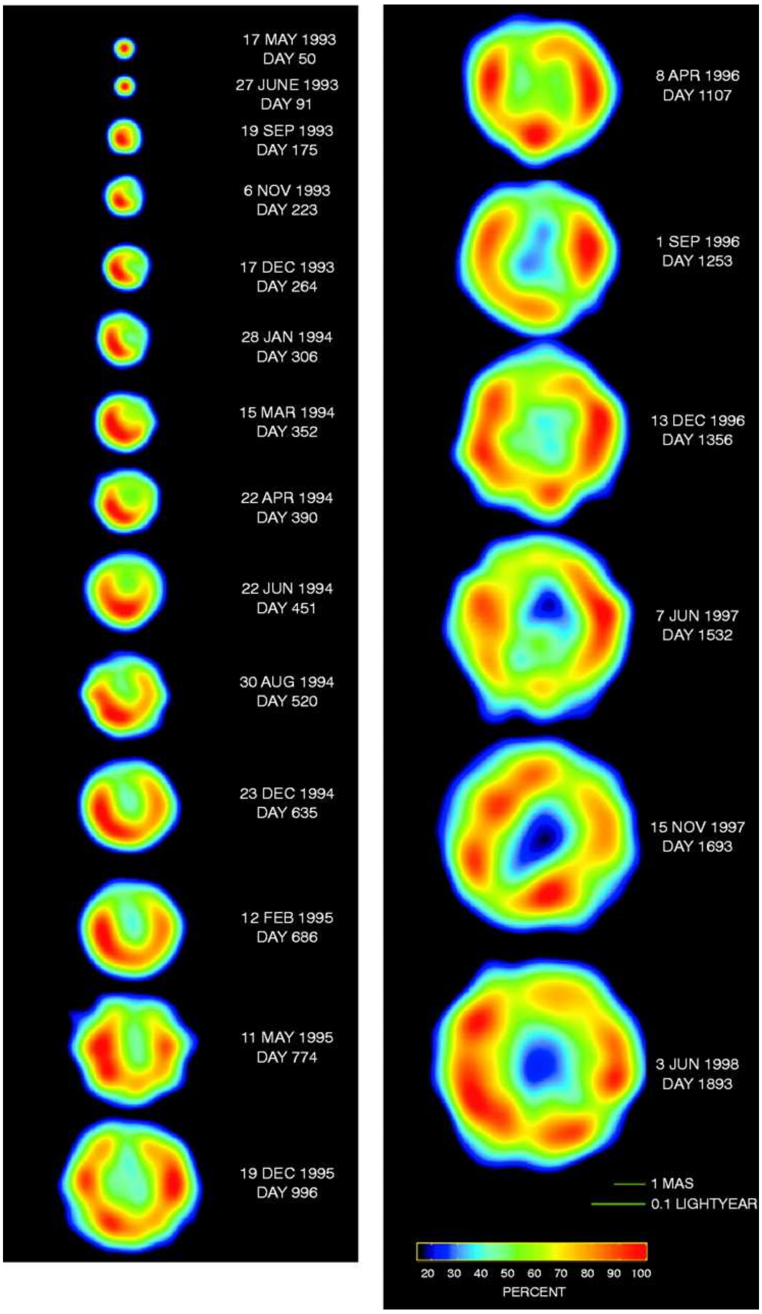

Figure 75: VLBI sequence of SN 1993J. (Bartel et al.)

then came close to the sun and could not be re-observed until ~ 200 days after explosion. The spectrum had then changed completely and was now a very soft with $kT \sim 1$ keV.

Type IIn's show the most extreme cases of CSI. Although CSI became very dominant in the optical spectrum after about a year, the Type IIn's are dominated by this immediately after explosion. Fig. 76 shows several examples of Type IIn's. The most prominent of the lines is Hα. Compared to e.g., SN 1993J, these lines are, however, very different. Instead of a box-like profiles the lines are usually peaked close to zero velocity. Chugai has modeled this by electron scattering in a dense, optically thick envelope. This again confirms the high density CSM of these SNe.

All types of core collapse SNe show CSI to a varying degree. Type IIP's only show weak indications of it, primarily in the radio and X-rays, where there is no background from the SN ejecta. Type IIL's, like SN 1979C are usually strong radio and X-ray emitters. The Type IIn's are probably extreme examples of this.

Type Ib/c SNe nearly all show radio and usually X-ray emission. Compared to the Type IIL's their radio emission is fairly weak, and decays fast. There is little indication of CSI in their optical spectra.

4.2 The standard model

The general scenario for the CSI is shown schematically in Fig. 77. The SN ejecta expand out into the CSM produced by the mass loss of the progenitor. Because the velocity of the wind is much less than that of the ejecta a strong shock will form at the interface. The pressure behind this will be large enough for driving a reverse shock back into the ejecta. Because the density of the ejecta is usually much larger than that of the wind, the reverse shock will be much slower than the outgoing. The temperature behind this will therefore be lower. Because of the high density and low temperature the reverse shock is most often radiative (cooling). This has the consequence that a cool, dense shell will form between the reverse shock and the contact discontinuity. As we will see, this cool, dense shell can have several observationally important consequences.

We will now discuss the details of this scenario.

4.3 Ejecta structure

The density structure of a supernova is set up during the first days after the explosion. Over this time scale, the pressure forces resulting from the

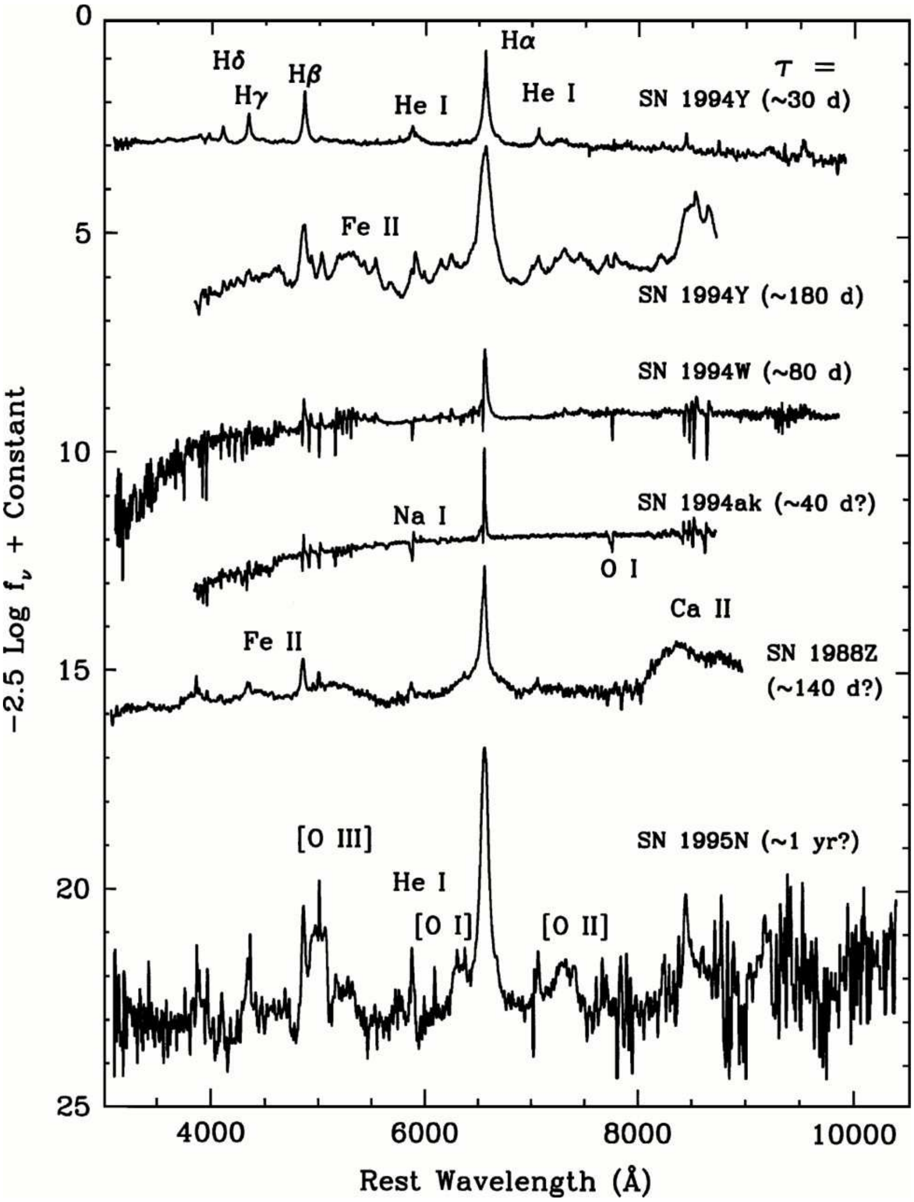

Figure 76: Collection of Type IIn spectra (Filippenko 1997)

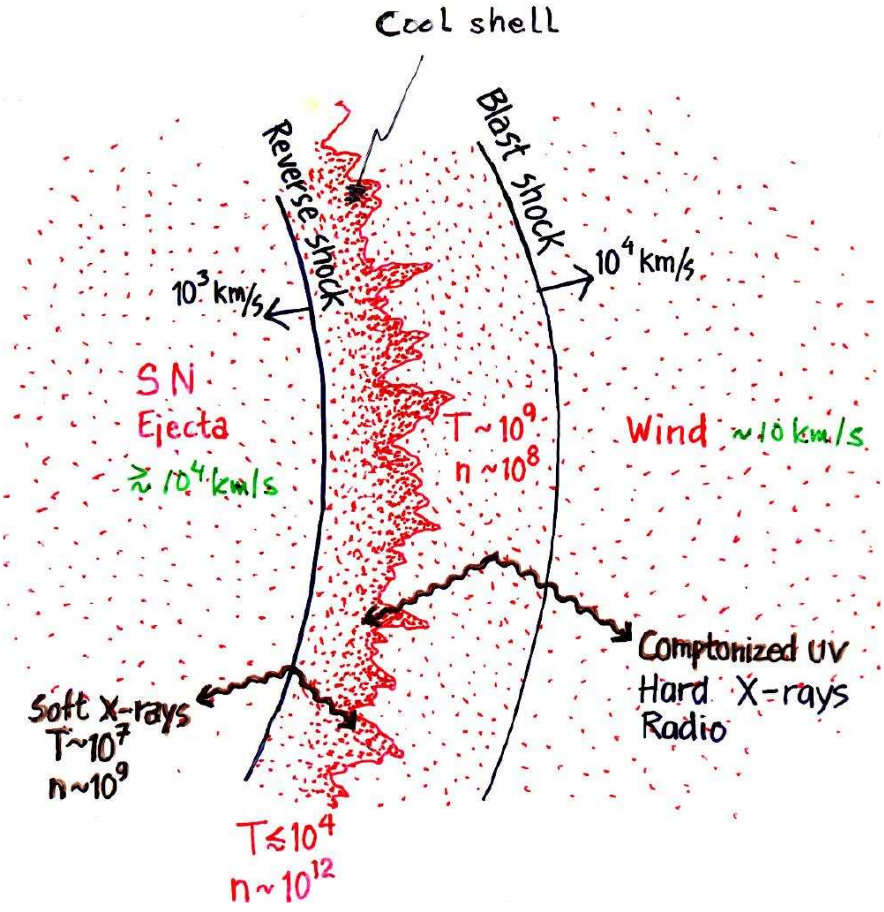

Figure 77: Schematic picture of the shock structure.

initial explosion, and some later power input from radioactivity, become too small to change the supernova density structure. The velocity profile tends toward that expected for free expansion, $v = r/t$, and the density of a gas element drops as t^{-3}. The pressure drops to a sufficiently low value that it is not a factor in considering shock waves propagating in the ejecta as a result of later interaction.

The density structure of the ejecta is a more complex problem. For core collapse supernovae, where all the explosion energy is generated at the center of the star, the explosion physics lead to an outer, steep power law density profile along with an inner region with a relatively flat profile. The outer profile is produced by the acceleration of the supernova shock front through the outer stellar layers with a rapidly decreasing density. This part of the shock propagation does not depend on the behavior in the lower layers and the limiting structure is described by a self-similar solution. The structure depends on the initial structure of the star and thus depends on whether the progenitor star has a radiative or a convective envelope. In the radiative case, which applies to Wolf-Rayet stars and to progenitors like that of SN 1987A, the limiting profile is $\rho \propto r^{-10.2}$. In the convective case, which applies to red supergiant progenitors, the limiting profile is $\rho \propto r^{-11.7}$. These are the limiting profiles and the density profile over a considerable part of the supernova might be described by a somewhat flatter profile. Numerical calculations of the explosion of SN 1987A have indicated $\rho \propto r^{-(8-9)}$ in the outer parts of the supernova, as Fig. 78 shows.

The overall result of these considerations is that the outer part of a core collapse supernova can be approximated by a steep power law density profile, or $\rho_{ej} \propto r^{-n}$ where n is a constant. After the first few days the outer parts of the ejecta expand with constant velocity, $V(m) \propto r$ for each mass element, m, so that $r(m) = V(m)t$ and $\rho(m) = \rho_o(m)(t_o/t)^3$. Therefore

$$\rho_{\rm ej} = \rho_o \left(\frac{t}{t_o}\right)^{-3} \left(\frac{V_o t}{r}\right)^{n} . \tag{4.1}$$

This expression takes into account the free expansion of the gas.

4.4 Stellar mass loss

In §1.10 we have already discussed the fact that most massive stars have winds of different character in the different evolutionary stages. If the mass loss parameters stay approximately constant leading up to the explosion,

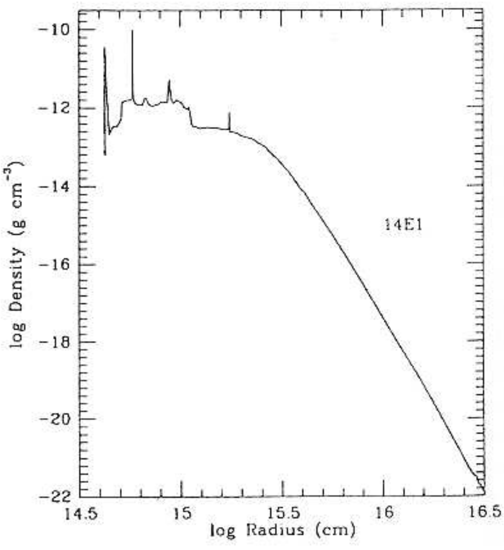

Figure 78: Density as function of radius for an ejecta model of SN 1987A at an age of 100 days. Note the power law density dependence outside of $\sim 3 \times 10^{15}$ cm, corresponding to ~ 3500 km s^{-1} (Shigeyama & Nomoto 1990).

the circumstellar density is given by

$$\rho_{\rm w} = \frac{\dot{M}}{4\pi u_{\rm w} r^2}. \tag{4.2}$$

In the late stellar evolutionary phases, the evolution of the stellar core occurs on a rapid time scale, but the stellar envelope has a relatively long dynamical time, which can stabilize the mass loss properties.

The mechanisms by which mass is lost in the red supergiant phase are poorly understood, but some insight into the wind properties can be gained by considering observations of the winds. VY CMa, thought to have a zero-age main sequence mass of $30-40$ $\mathrm{M}_\odot$, a mass loss rate $\sim 3\times10^{-4}$ $\mathrm{M}_\odot$ yr^{-1}, and $u_{\rm w} \approx 39$ km s^{-1}, is an especially well-observed case. *HST* imaging shows that the density profile is approximately r^{-2} over the radius range $3\times10^{16} - 1.4\times10^{17}$ cm but that there is considerable structure superposed on this profile, including knots and filaments. Observations of masers imply the presence of clumps with densities up to $\sim 5\times10^{9}$ cm^{-3} and suggest the presence of an expanding disk seen at an oblique angle. VY CMa is an extreme mass loss object, but there is evidence for irregular mass loss in α Orionis and other red supergiants.

The fact that especially the most massive stars evolve from blue supergiant to red supergiant and finally to a Wolf-Rayet star mean that the CSM of these stars can be extremely complex. In particular, the fast wind from the Wolf-Rayet star can create a bubble in the surrounding medium, which is typically the slow wind from a previous evolutionary phase, and the resulting shells have been observed around a number of Wolf-Rayet stars. Their typical radii are a few pc.

4.5 Hydrodynamics

When the radiation dominated shock front in a supernova nears the stellar surface, a radiative precursor to the shock forms when the radiative diffusion time is comparable to the propagation time. There is radiative acceleration of the gas and the shock disappears when optical depth ∼unity is reached. The fact that the velocity decreases with radius implies that the shock will re-form as a viscous shock in the circumstellar wind. This occurs when the supernova has approximately doubled in radius.

The interaction of the ejecta, expanding with velocity $\gtrsim 10^4$ km s^{-1}, and the nearly stationary circumstellar medium results in a reverse shock wave propagating inwards (in mass), and an outgoing circumstellar shock. The density in the circumstellar gas is given by Eq. (4.2). As discussed above, hydrodynamical calculations show that to a good approximation the ejecta density can be described by Eq. (4.1). A useful similarity solution for the interaction can then be found. Here we sketch a simple derivation.

Assume that the shocked gas can be treated as a thin shell with mass $M_{\rm s}$, velocity $V_{\rm s}$, and radius $R_{\rm s}$. Balancing the ram pressure from the circumstellar gas and the impacting ejecta, the momentum equation for the shocked shell of circumstellar gas and ejecta is

$$\frac{d}{dt}(M_{\rm s}V_{\rm s}) = 4\pi R_{\rm s}^2[\rho_{{\rm e}j}(V-V_{\rm s})^2 - \rho_{\rm w}V_{\rm s}^2]. \tag{4.3}$$

Here, $M_{\rm s}$ is the sum of the mass of the shocked ejecta and circumstellar gas. The swept up mass behind the circumstellar shock is $M_{cs} = \dot{M}R_{\rm s}/u_{\rm w}$, and that behind the reverse shock $M_{rev} = 4\pi\ t_o^3V_o^n(t/R_{\rm s})^{n-3}/(n-3)$, assuming that $R_{\rm s} >> R_{\rm p}$, the radius of the progenitor. With the ejecta velocity at the shock given by $V = R_{\rm s}/t$ we obtain

$$\frac{d}{dt}\left\{\left[\frac{\dot{M}}{u_{\rm w}}R_{\rm s} + \frac{4\pi\ \rho_o\ t_o^3\ V_o^n\ t^{n-3}}{(n-3)\ R_{\rm s}^{n-3}}\right]\frac{dR_{\rm s}}{dt}\right\} =$$
$$4\pi R_{\rm s}^2\left[\frac{\rho_o\ t_o^3\ V_o^n\ t^{n-3}}{R_{\rm s}^n}\left(\frac{R_{\rm s}}{t} - \frac{dR_{\rm s}}{dt}\right)^2 - \frac{\dot{M}}{4\pi\ u_{\rm w}R_{\rm s}^2}\left(\frac{dR_{\rm s}}{dt}\right)^2\right]. \tag{4.4}$$

This equation has the power law solution

$$R_{\rm s}(t) = \left[\frac{4\pi(n-5)\rho_o\ t_o^3\ V_o^n\ u_{\rm w}}{(n-3)(2n-7)\ \dot{M}}\right]^{1/(n-2)} t^{(n-3)/(n-2)}. \tag{4.5}$$

The form of this similarity solution can be written down directly by dimensional analysis from the only two independent quantities available, $\rho_o\ t_o^3\ V_o^n$ and $\dot{M}/u_{\rm w}$. The solution applies after a few expansion times, when the initial radius has been 'forgotten.' The requirement of a finite energy in the flow implies $n > 5$. More accurate similarity solutions, taking the structure within the shell into account, are given by Chevalier (1982). In general, these solutions differ by less than $\sim 30\%$ from the thin shell approximation.

The maximum ejecta velocity close to the reverse shock depends on time as $V = R_{\rm s}/t \propto t^{-1/(n-2)}$. The velocity of the circumstellar shock, $dR_{\rm s}/dt$, in terms of V is $V_{\rm s} = V(n-3)/(n-2)$ and the reverse shock velocity, $V_{rev} = V - V_{\rm s} = V/(n-2)$.

From the shock relation

$$T_{\rm cs} = \frac{3\mu}{16k}V_s^2 \tag{4.6}$$

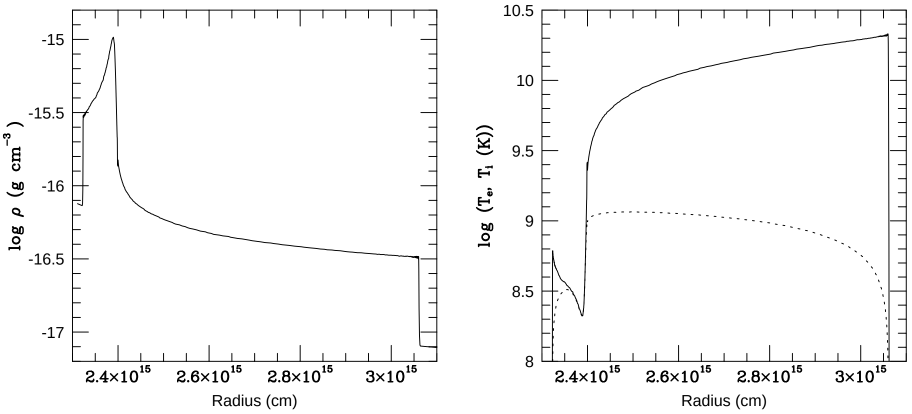

Figure 79: Density and temperature structure of the reverse and circumstellar shocks for $n = 7$ and a velocity of 2.5×10^4 km s^{-1} at 10 days. Both shocks are assumed to be adiabatic. Because of the slow Coulomb equipartition the electron temperature (dotted line) is much lower than the ion temperature (solid line) behind the circumstellar shock.

where μ is the mean mass per particle (see e.g., Spitzer), we can now obtain the temperature in the shocked circumstellar gas. Assuming cosmic abundances and equipartition between ions and electrons, the temperature is

$$T_{\rm cs} = 1.36 \times 10^9 \ \left(\frac{n-3}{n-2}\right)^2 \left(\frac{V}{10^4 \ {\rm km \ s^{-1}}}\right)^2 \ {\rm K} \tag{4.7}$$

and at the reverse shock

$$T_{\rm rev} = \frac{T_{\rm cs}}{(n-3)^2}. \tag{4.8}$$

The time scale for equipartition between electrons and ions is

$$t_{\rm eq} \approx 2.5 \times 10^7 \ \left(\frac{T_{\rm e}}{10^9 \ {\rm K}}\right)^{1.5} \ \left(\frac{n_{\rm e}}{10^7 \ {\rm cm^{-3}}}\right)^{-1} \ s. \tag{4.9}$$

One finds that the reverse shock is marginally in equipartition, unless the temperature is $\gtrsim 5 \times 10^8$ K. The ion temperature behind the circumstellar shock is $\gtrsim 6 \times 10^9$ K for $V_4 \gtrsim 1.5$, and the density a factor $\gtrsim 4$ lower than behind the reverse shock. Ion-electron collisions are therefore ineffective, and $T_{\rm e} << T_{\rm ion}$, unless efficient plasma instabilities heat the electrons collisionlessly (Fig. 79).

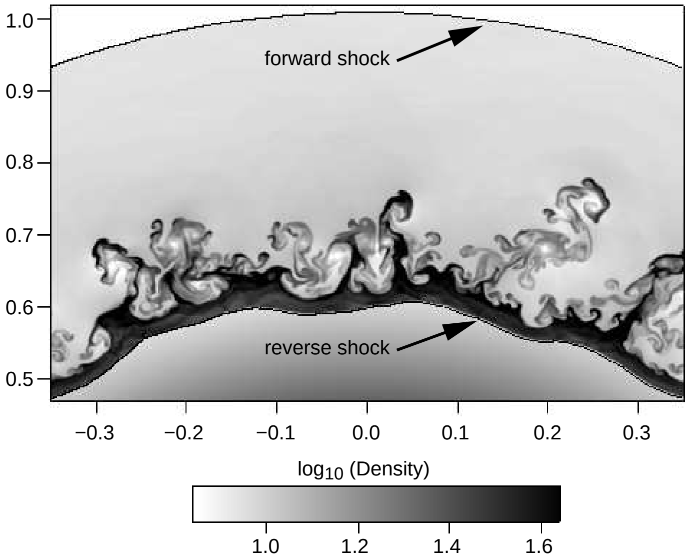

Figure 80: Two-dimensional calculation of the shock structure for a supernova with $n = 6$ in a stellar wind (courtesy John Blondin).

For typical parameters, the electron temperatures of the two shocks are very different, $\sim (1-3)\times10^9$ K for the circumstellar shock and $10^7-5\times10^8$ K for the reverse shock, depending on n. The radiation from the reverse shock is mainly below $\sim$ 20 keV, while that from the circumstellar shock is above $\sim$ 50 keV.

The density behind the circumstellar shock is $\rho_{\rm cs} = 4\rho_{\rm w}$, while that behind the reverse shock is

$$\rho_{\rm rev} = \frac{(n-4)(n-3)}{2}\rho_{\rm cs} \tag{4.10}$$

and is much higher than behind the circumstellar shock for $n \gtrsim 7$. There is a drop in density across the contact discontinuity, moving from the shocked ejecta to the circumstellar medium (see Fig. 79). The fact that low density gas is decelerating higher density gas leads to a Rayleigh-Taylor instability. Chevalier & Blondin have calculated the structure using a two-dimensional PPM hydrodynamic code. They indeed find that instabilities develop, with dense, shocked ejecta gas penetrating into the hotter, low density shocked circumstellar gas (Fig. 80). The instability mainly distorts the contact surface, and does not seriously affect the general dynamics.

In addition to asymmetric winds, there is evidence for supernova shock waves interacting with clumps of gas in the wind, as have been observed in some red supergiant winds. In some cases the clumps can be observed by their very narrow lines in supernova spectra, as in Type IIn supernovae. The velocity of a shock wave driven into a clump, $v_{\rm c}$, can be estimated by approximate pressure balance $v_{\rm c} \approx v_{\rm s}(\rho_{rms}/\rho_{\rm c})^{1/2}$, where $v_{\rm s}$ is the shock velocity in the smooth wind with density ρ_{rms} and $\rho_{\rm c}$ is the clump density. The lower shock velocity and higher density can lead to radiative cooling of the clump shock although the main shock is non-radiative. Optical line emission of intermediate velocity observed in Type IIn (narrow line) supernovae like SN 1978K, SN 1988Z, and SN 1995N can be explained in this way.

The presence of many clumps can affect the hydrodynamics of the interaction. Jun, Jones, & Norman (1996) found that propagation in a region with clumps gives rise to widespread turbulence in the shocked region between the forward shock and the reverse shock, whereas the turbulence is confined to a region near the reverse shock for the non-clump case (Fig. 2). Their simulations are for interaction with a constant density medium, but the same probably holds true for interaction with a circumstellar wind.

Exercise:

In analogy with the discussion leading up to Eq. (4.4), using the thin shell approximation, derive the equation determining the evolution of an adiabatic point explosion with total energy E_o and adiabatic index γ in a constant density medium with density ρ_o. Neglect the ejecta mass compared to the swept up mass by the forward shock. Solve this equation for the radius as function of time to get

$$R_{\rm s}(t) \approx \left[\frac{75(\gamma - 1)E_o}{16\pi\rho_o}\right]^{1/5} t^{2/5}. \tag{4.11}$$

This is the famous Sedov-Taylor solution. An exact calculation gives a somewhat different numerical coefficient, but the same dependence on E_o/ρ_o and time (see e.g., Shu or Landau-Lifshitz).

Use the jump condition

$$\frac{\rho_2}{\rho_1} = \frac{(\gamma + 1)}{(\gamma - 1)} \tag{4.12}$$

(see e.g., Spitzer) to show that the thickness of the shell is $\Delta R/R \approx (\gamma - 1)/[3(\gamma + 1)]$. For $\gamma = 5/3$ this gives $\Delta R/R \approx 1/12$, while for $\gamma = 4/3$ $\Delta R/R \approx 1/21$.

Discuss the weak points of this derivation, and in which way this situation differs from the circumstellar shock discussed in this section.

4.6 Emission from the hot gas

During the first month, radiation from the supernova photosphere is strong enough for Compton scattering to be the main cooling process for the circumstellar shock. After this free-free and line emission dominate the cooling.

One can estimate the free-free luminosity from the circumstellar and reverse shocks from

$$L_{\rm i} = 4\pi \int \Lambda_{\rm ff}(T_{\rm e}) n_{\rm e}^2 r^2 dr \approx \Lambda_{\rm ff}(T_{\rm i}) \frac{M_{\rm i}\rho_{\rm i}}{(\mu_{\rm e} m_H)^2}. \qquad (4.13)$$

where the index i refers to quantities connected either with the reverse shock or circumstellar shock. The density behind the circumstellar shock is $\rho_{\rm cs} = 4\ \rho_0 = \dot{M}/(\pi u_{\rm w} R_{\rm s}^2)$. The swept up mass behind the circumstellar shock is $M_{\rm cs} = \dot{M} R_{\rm s}/u_{\rm w}$ and that behind the reverse shock $M_{\rm rev} = (n-4)M_{\rm cs}/2$. With $\Lambda_{\rm ff} = 2.4 \times 10^{-27} \bar{g}_{\rm ff}\ T_{\rm e}^{0.5}$, we get

$$L_{\rm i} \approx 3.0 \times 10^{39}\ \bar{g}_{\rm ff}\ C_{\rm n} \left(\frac{\dot{M}_{-5}}{u_{\rm w1}}\right)^2 \left(\frac{t}{10 \text{ days}}\right)^{-1} \ \text{erg s}^{-1}, \qquad (4.14)$$

where $\bar{g}_{\rm ff}$ is the free-free Gaunt factor, including relativistic effects. For the reverse shock $C_{\rm n} = (n-3)(n-4)^2/4(n-2)$, and for the circumstellar shock $C_{\rm n} = 1$. This assumes electron-ion equipartition, which is highly questionable for the circumstellar shock (see Fig. 79). Because of occultation by the ejecta, only half of the above luminosity escapes outward.

At $T_{\rm e} \lesssim 2 \times 10^7$ K, line emission increases the cooling rate and $\Lambda \approx 3.4 \times 10^{-23}\ T_{e7}^{-0.67}$ erg s^{-1}cm^3. If the temperature of the reverse shock falls below $\sim 2 \times 10^7$ K, a thermal instability may occur and the gas cools to $\lesssim 10^4$ K, where photoelectric heating from the shocks balances the cooling. Using $t_{cool} = 3kT_{\rm e}/\Lambda$, one obtains for the cooling time

$$t_{cool} = \frac{605}{(n-3)(n-4)(n-2)^{3.34}} \left(\frac{V_{\rm ej}}{10^4 \text{ km s}^{-1}}\right)^{5.34} \left(\frac{\dot{M}_{-5}}{u_{w1}}\right)^{-1} \left(\frac{t}{\text{days}}\right)^2 \ \text{days}, \qquad (4.15)$$

assuming solar abundances. From this expression it is clear that the cooling time is very sensitive to the density gradient, as well as the shock velocity and mass loss rate. SNe with high mass loss rates, like SN 1993J, generally have radiative reverse shocks for $\gtrsim 100$ days, while SNe with lower mass loss rates, like the Type IIP SN 1999em, have adiabatic shocks from early times.

The most important effect of the cooling is that the cool gas may absorb most of the emission from the reverse shock. Therefore, in spite of the higher intrinsic luminosity of the reverse shock, little of this will be directly observable. The column density of the cool gas is given by $N_{cool} = M_{\rm rev}/(4\pi R_{\rm s}^2 m_{\rm p})$, or

$$N_{cool} \approx 1.0 \times 10^{21} (n-4) \left(\frac{\dot{M}_{-5}}{u_{\rm w1}}\right) \left(\frac{V}{10^4 \ {\rm km \ s^{-1}}}\right)^{-1} \left(\frac{t}{100 \ {\rm days}}\right)^{-1} \ {\rm cm}^{-2}. \tag{4.16}$$

Because the threshold energy due to photoelectric absorption is related to N_{cool} by $E(\tau = 1) = 1.2(N_{cool}/10^{22} \ cm^{-2})^{3/8}$ keV, it is clear that the emission from the reverse shock is strongly affected by the cool shell, and a transition from optically thick to optically thin is expected during the first months, or year. As an illustration, we show in Fig. 81 the calculated X-ray spectrum at 10 days and at 200 days for SN 1993J. At early epochs the spectrum is dominated by the very hard spectrum from the circumstellar shock, which reaches out to $\gtrsim 100$ keV. At later epochs the soft spectrum from the reverse shock penetrates the cool shell, and the line dominated emission from the cooling gas dominates.

If cooling, the total energy emitted from the reverse shock is

$$\begin{aligned} L_{\rm rev} &= 4\pi R_{\rm s}^2 \ \frac{1}{2} \ \rho_{ej} V_{\rm rev}^3 = \frac{(n-3)(n-4)}{4(n-2)^3} \ \frac{\dot{M} V^3}{u_{\rm w}} \\ &= 1.6 \times 10^{41} \ \frac{(n-3)(n-4)}{(n-2)^3} \dot{M}_{-5} u_{\rm w1}^{-1} V_4^3 \ \ {\rm erg \ s^{-1}}. \end{aligned} \tag{4.17}$$

For high $\dot{M}/u_{\rm w}$ the luminosity from the reverse shock may contribute appreciably, or even dominate, the bolometric luminosity.

Because $V \propto t^{-1/(n-2)}$, $L_{\rm rev} \propto t^{-3/(n-2)}$ in the cooling case. Although the total luminosity is likely to decrease in the cooling case, the increasing transparency of the cool shell, $\tau_{cool} \propto t^{-1}$, can cause the observed flux in energy bands close to the low energy cutoff, $E(\tau = 1)$, to increase with time, as was seen, e.g., in SN 1993J (Fig. 3).

Because of the low temperature the spectrum of the reverse shock is dominated by line emission from metals (Fig. 3). An important point is that the observed spectrum is formed in gas with widely different temperatures, varying from the reverse shock temperature to $\sim 10^4$ K. A spectral analysis based on a one temperature model can be misleading.

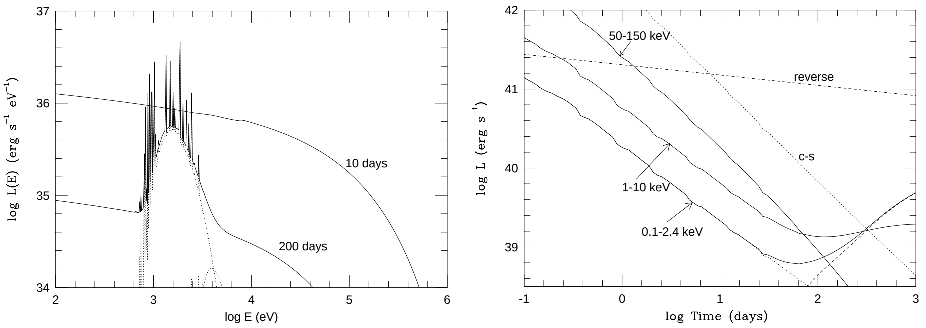

Figure 81: X-ray spectrum of SN 1993J at 10 days and at 200 days. At 10 days the free-free emission from the outer shock dominates, while at 200 days the cool shell is transparent enough for the line dominated spectrum from the reverse shock to dominate instead. Right: The solid lines give the luminosity in the 0.1-2.4, 1-10, and 50-100 keV bands, corrected for absorption, as a function of time, while the dotted lines give the total emitted luminosity from the reverse and circumstellar shocks.

Chugai has proposed that the X-ray emission from the Type IIn SN 1986J is the result of the forward shock front moving into clumps, as opposed to the reverse shock emission. One way to distinguish these cases is by the width of line emission; emission from the reverse shock wave is expected to be broad.

4.7 Radiative heating and re-emission

4.7.1 Soft X-ray burst and circumstellar gas

The earliest form of circumstellar interaction occurs at shock break-out. As the shock approaches the surface, radiation leaks out on a time scale of less than an hour. The color temperature of the radiation is $\sim (1-5) \times 10^5$ K and the energy $\sim (1-10) \times 10^{48}$ *ergs* (see §3.4.1).

The radiative effects of the soft X-ray burst were most clearly seen from the ring of SN 1987A, where a number of narrow emission lines from highly ionized species, like N III-N V, were first seen in the UV. Later, a forest of lines came to dominate also the optical spectrum. Imaging with HST showed that the lines originated in the now famous circumstellar ring of

SN 1987A at a distance of ~ 200 light days from the SN. The presence of highly ionized gas implied that the gas must have been ionized and heated by the radiation at shock break-out. Because of the finite light travel time across the ring, the observed total emission from the ring is a convolution of the emission at different epochs from the various part of the ring. Detailed modeling shows that while the ionization of the ring occurs on the time scale of the soft X-ray burst, the gas recombines and cools on a time scale of years, explaining the persistence of the emission decades after the explosion. The observed line emission provides sensitive diagnostics of both the properties of the soft X-ray burst, and the density, temperature and abundances of the gas in the ring. In particular, the radiation temperature must have reached $\sim 10^6$ K, in good agreement with the most detailed recent modeling of the shock break-out. Narrow emission lines are not unique to SN 1987A, but have also been observed for several other SNe, in particular several Type IIn SNe, such as SN 1995N and SN 1998S.

The X-ray emission from the shocks ionizes and heats both the circumstellar medium and the SN ejecta. Observationally, these components are distinguished easily by the different velocities. The circumstellar component is expected to have velocities typical of the progenitor winds, i.e., $\lesssim 1000$ km s^{-1}, while the ejecta have considerably higher velocities. The density is likely to be of the order of the wind density $10^5 - 10^7$ cm^{-3}, or higher if clumping is important. The ionizing X-ray flux depends strongly on how much of the flux from the reverse shock can penetrate the cool shell. The state of ionization in the circumstellar gas is characterized by the value of the ionization parameter,

$$\zeta = \frac{L_{cs}}{r^2 n} = 10^2 \frac{L_{\rm cs}}{10^{40}\ {\rm erg\ s^{-1}}} \left(\frac{r}{10^{16}\ cm}\right)^{-2} \left(\frac{n}{10^6\ {\rm cm}^{-3}}\right)^{-1} \tag{4.18}$$

The comparatively high value of $\zeta \approx 10 - 10^3$ explains the presence of narrow coronal lines of [Fe V-XI] seen in objects like SN 1995N.

4.7.2 SN ejecta

The ingoing X-ray flux from the reverse shock ionizes the outer parts of the ejecta. The state of highest ionization therefore is close to the shock, with a gradually lower degree of ionization inwards. Unless clumping in the ejecta is important, the ejecta density is $\sim 10^6 - 10^8$ cm^{-3}. In the left panel of Fig. 82 we show temperature and ionization structure of the ejecta, as well as the emissivity of the most important lines. The temperature close to the shock is $\sim 3 \times 10^4$ K. Calculations show that most of the emission here is emitted as UV lines of highly ionized ions, like Lyα, C III-IV, N III-V, and O III-VI. Inside the ionized shell there is an extended partially ionized zone, similar to that present in the broad emission line regions of AGN's. Most of the emission here comes from Balmer lines.

As we have already discussed, the outgoing flux from the reverse shock is to a large extent absorbed by the cool shell between reverse shock and the contact discontinuity if radiative cooling has been important. The whole region behind the reverse shock is in approximate pressure balance, and the density of this gas is therefore be a factor $\sim 4T_{rev}/T_{cool} \approx 10^3 - 10^4$ higher than that of the ejecta. Because of the high density, the gas is only be partially ionized and the temperature only $(5 - 8) \times 10^3$ K. Most of the emission comes out as Balmer lines, Mg II and Fe II lines (Fig. 82, right panel). The thickness of the emitting region is also very small, $\sim 3 \times 10^{12}$ cm. In one dimensional models, the velocity is marginally smaller than the highest ejecta velocities. Instabilities in the shock are, however, likely to erase this difference.

An important diagnostic of the emission from the cool shell and the ejecta is the Hα line. This line arises as a result of recombination and collisional excitation. It can be shown that $\sim 1\%$ of the reverse shock luminosity is emitted as Hα, fairly independent of density and other parameters. Observations of this line permit us to follow the total luminosity from the reverse shock, complementary to the X-ray observations. In SN 1993J, the Hα line had the box-like shape that is expected for shocked, cooled ejecta (Fig. 73). The top of the line showed structure that varied with time; this could be related to hydrodynamic instabilities of the reverse shocked gas.

4.7.3 Line profiles

The line profiles of circumstellar interactors can have several characteristic appearences. Lines coming from the circumstellar medium will be narrow, with a width given by the expansion velocity of the wind. For a red supergiant progenitor this is $10 - 100$ km s^{-1}.

The lines can be seen in either emission or in absorption if the line is optically thick. This, however, requires the wind not to be compleatly ionized by the radiation from the shock breakout, or if this occurs, that the wind is sufficiently dense to recombine. A slow RSG wind will have a considerably higher density than a fast WR wind. In addition, the radiation temperature of the shock breakout burst will be lower (see Sect. 3.4.1). Therefore, line emission from the wind will more likely be seen for a RSG progenitor than a WR progenitor, for which the wind is likely to be fully

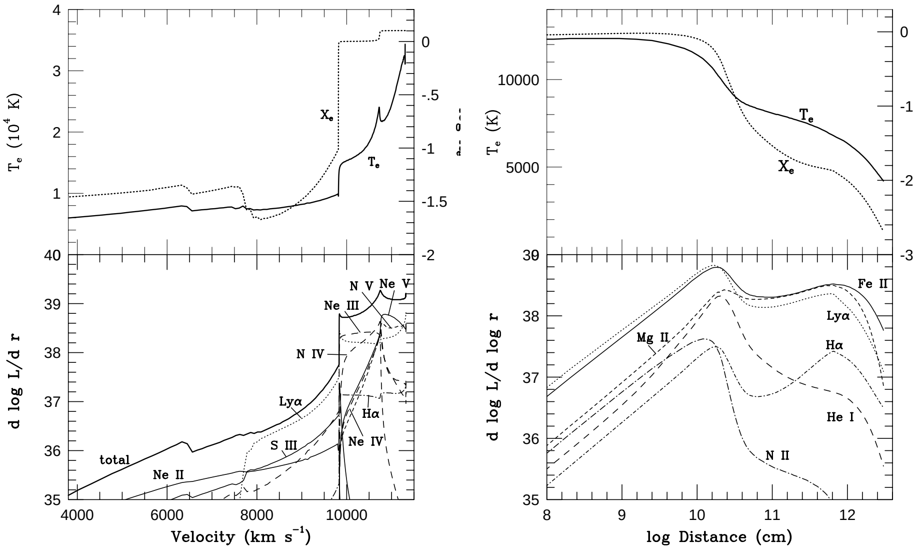

Figure 82: Structure of the ejecta and cool shell ionized by the reverse shock at 500 days for parameters appropriate to SN 1993J ($\dot{M} = 5 \times 10^{-5}$ M$_\odot$ yr^{-1} for $u = 10$ km s^{-1}). Upper panels show the temperature and ionization of the ejecta (left panels) and the cool shell (right panels), while the lower panels show the corresponding luminosities per unit distance. Note the different length scales in the two panels. The ejecta region has low density, high temperature and ionization, while the cool shell has a high density, is extremely thin, has a low temperature, and is only partially ionized.

ionized.

Some of the optical lines, like Hα and the Ca II lines, as well as most of the UV resonance lines, may be optically thick. In this case the line profile may show a typical P-Cygni appearence, with a velocity width given by the velocity of the wind.

The velocity of the wind may be changed by the strong burst of radiation from the supernova. In the simple case of electron scatterng the momentum transfered to a spherical shell of radius r, thickness dr and electron density $n_{\rm e}$ by the radiation is given by

$$\frac{dp}{dt} = \frac{1}{c} L(t) \sigma_T n_{\rm e} dr \ . \tag{4.19}$$

Now if the burst is short compared to the dynamical time scale r/v then we can integrate Eq. (4.19) over time, keeping the radius constant to get the total change in momentum by the shell. Therefore,

$$\Delta p \approx \frac{1}{c}\sigma_T n_{\rm e} dr \int L(t)dt \ . \tag{4.20}$$

Further, $\Delta p = dM\Delta V = 4\pi\mu m_{\rm u} n_{\rm e} r^2 dr \Delta V$. Therefore,

$$\Delta V \approx \frac{\sigma_T \int L(t)dt}{4\pi\mu m_{\rm u} c r^2} \ . \tag{4.21}$$

or

$$\Delta V \approx 1.1\times 10^3 \left(\frac{E}{10^{48}\ {\rm erg\ s^{-1}}}\right)\left(\frac{r}{10^{14}\ cm}\right)^{-2}\ {\rm km\ s^{-1}}, \tag{4.22}$$

where E is the total energy in the burst. Because $R = Vt = 8.6\times 10^{13}(V/10^4\ {\rm km\ s^{-1}})$(t/days) cm, radiative pre-acceleration of the gas ahead of the shock can give velocities of $10 - 1000$ km s^{-1} in the circumstellar gas during the first days after explosion. Absorption by resonance lines, similar to that ocurring in early type stars, may increase the velocity further.

Lines coming from the ejecta and shock region can have several different characteristics. In the simplest case the line will aise in a geometrically thin shell with inner velocity V_{in} and outer velocity V_{out}. If $V_{in} \sim V_{out}$ we have already in Sect. 3.5 found that if the Sobolev approximation is valid and the line is optically thin the line profile will be box-like. This is also the case in the optically thick case if the velocity is proprtional to the radius. qqqqqIn the case If the line is optically

Box profiles, M-shaped lines,

4.8 Relativistic particles

Unambiguous evidence for the presence of relativistic electrons comes from radio observations. A characteristic is the wavelength-dependent turn-on of the radio emission, first seen at short wavelengths, and later at longer wavelengths (§4.1). This behavior is interpreted as a result of decreasing absorption due to the expanding emitting region.

Depending on the magnetic field and the density of the circumstellar medium, the absorption may be produced either by free-free absorption in the surrounding thermal gas, or by synchrotron self-absorption by the same electrons that are responsible for the emission. The relativistic electrons are believed to be produced close to the interaction region, which provides an ideal environment for the acceleration of relativistic particles. The details of the acceleration and injection efficiency are still not well understood. Here we just parameterize the injection spectrum with the power law index $p_{\rm i}$

and an efficiency, η, in terms of the post-shock energy density. Without radiation or collisional losses the spectral index of the synchrotron emission will then be $\alpha = (p-1)/2$, where flux $\propto \nu^{-\alpha}$. Diffusive acceleration predicts that $p_{\rm i} = 2$ in the test particle limit. If the particle acceleration is very efficient and nonlinear effects are important, the electron spectrum can be steeper.

For free-free absorption, the optical depth $\tau_{\rm ff} = \int_{R_{\rm s}}^{\infty} \kappa_{\rm ff} n_{\rm e} n_{\rm i} dr$ from the radio emitting region close to the shock through the circumstellar medium decreases as the shock wave expands, explaining the radio turn-on. Assuming a fully ionized wind with constant mass loss rate and velocity, so that Eq. (4.2) applies, the free-free optical depth at wavelength λ is

$$\tau_{\rm ff}(\lambda) \approx 7.1 \times 10^2 \lambda^2 \left(\frac{\dot{M}_{-5}}{u_{w1}}\right)^2 T_5^{-3/2} V_4^{-3} t_{\rm days}^{-3} \tag{4.23}$$

where $\dot{M}_{-5}$ is the mass loss rate in units of 10^{-5} $\mathrm{M}_\odot$ yr^{-1}, u_{w1} the wind velocity in units of 10 km s^{-1}, and T_5 the temperature of the circumstellar gas in 10^5 K. From the radio light curve, or spectrum, the epoch of $\tau_{\rm ff} = 1$ can be estimated for a given wavelength, and from the line widths in the optical spectrum the maximum expansion velocity, V, can be obtained. Because the effects of the radiation from the supernova have to be estimated from models of the circumstellar medium, the temperature in the gas is uncertain. Calculations show that initially the radiation heats the gas to $T_{\rm e} \approx 10^5$ K. $T_{\rm e}$ then decreases with time, and after a year $T_{\rm e} \approx (1.5 - 3) \times 10^4$ K. In addition, the medium may recombine, which further decreases the free-free absorption. From $t[\tau(\lambda)_{\rm ff} = 1]$ the ratio $\dot{M}/u_{\rm w}$ can be calculated. Because $\dot{M}/u_{\rm w} \propto T_{\rm e}^{3/4} x_{\rm e}^{-1}$, errors in $T_{\rm e}$ and $x_{\rm e}$ may lead to large errors in $\dot{M}$. If the medium is clumpy, Eq. (4.23) may lead to an overestimate of $\dot{M}/u_{\rm w}$.

Under special circumstances (see below), synchrotron self-absorption (SSA) by the same relativistic electrons emitting the synchrotron radiation may be important. The emissivity of the synchrotron plasma is given by $j(\lambda) = \text{const. } \lambda^{\alpha} B^{1+\alpha} N_{\rm rel}$ (see e.g., Rybicki & Lightman), while the optical depth to self-absorption is given by

$$\tau_{\rm s} = const.\ \lambda^{(5/2)+\alpha} B^{(3/2)+\alpha} N_{\rm rel}\ . \tag{4.24}$$

Here $N_{\rm rel}$ is the column density of relativistic electrons and B the magnetic field. The flux from a disk with radius $R_{\rm s}$ of relativistic electrons is, including the effect of SSA, given by

$$F_\nu(\lambda) \propto R_{\rm s}^2 S(\lambda)[1 - e^{-\tau_{\rm s}(\lambda)}], \tag{4.25}$$

where $S(\lambda) = j(\lambda)/\kappa(\lambda) = const.\ \lambda^{-5/2}B^{-1/2}$ is the source function. In the optically thick limit we therefore have $F_\nu(\lambda) \approx const.\ R_{\rm s}^2\ \lambda^{-5/2}B^{-1/2}$, independent of $N_{\rm rel}$. A fit of this part of the spectrum therefore gives the quantity $R_{\rm s}^2\ B^{-1/2}$. The break of the spectrum determines the wavelength of optical depth unity, $\lambda(\tau_{\rm s} = 1)$. Eq. (4.24) therefore gives a second condition on $B^{3/2+\alpha}\ N_{rel}$. If $R_{\rm s}(t)$ is known in some independent way, one can determine both the magnetic field and the column density of relativistic electrons, independent of assumptions about equipartition, etc. In some cases, most notably for SN 1993J, the shock radius, $R_{\rm s}$, can be determined directly from VLBI observations. If this is not possible, an alternative is from observations of the maximum ejecta velocity seen in, e.g., the Hα line, which should reflect the velocity of the gas close to the shock. Because the SN expands homologously, $R_{\rm s} = V_{max}t$. A fit of the spectrum at a given epoch can therefore yield both B and $N_{\rm rel}$ independently. From observations at several epochs the evolution of these quantities can then be determined.

Although the injected electron spectrum from the shock is likely to be a power law with $p_{\rm i} \approx 2$ ($\alpha \approx 0.5$), the integrated electron spectrum is affected by various loss processes. Most important, the synchrotron cooling time scale of an electron with Lorentz factor γ is $t_{syn} \approx 9 \times 10^3 \gamma^{-1} B^{-2}$ days. This especially affects the high energy electrons, steepening the index of the integrated column density electron spectrum by one unit, $p = p_{\rm i} + 1 \approx 3$ ($\alpha \approx 1$). Note that it really is the distance from the shock it radiates which decreases as γ^{-1}. Inverse Compton losses have a similar effect as synchrotron losses. At low energy, Coulomb losses may be important, causing the electron spectrum to flatten.

The best radio observations of any SN were obtained for SN 1993J. This SN was observed from the very beginning until late epochs with the VLA at wavelengths between 1.3 – 90 cm, producing a set of beautiful light curves. In addition, the SN was observed with VLBI (see Fig. 75), resulting in an impressive sequence of images in which the radio emitting plasma could be directly observed. These images showed a remarkable degree of symmetry and clearly resolved the shell of emitting electrons. The evolution of the radius of the radio emitting shell could be well fitted by $R_{\rm s} \propto t^{0.86}$, implying a deceleration of the shock front.

From a fit of the observed spectra for the different epochs the magnetic field and number of relativistic electrons could be determined for each epoch, as described above. In Fig. 83 we show the evolution of B and $N_{\rm rel}$, plotted as a function of the shock radius. The most remarkable thing is the smooth evolution of these quantities, showing that $B \approx 6.4(R_{\rm s}/10^{16}\ cm)^{-1}$ G, and $n_{\rm rel} \propto \rho V^2 \propto t^{-2}$, the thermal energy density behind the shock. The

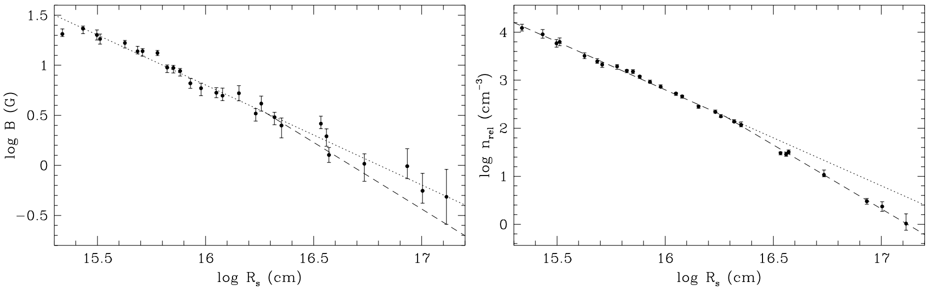

Figure 83: Magnetic field (left) and density of relativistic electrons (right) as a function of the shock radius for SN 1993J. The dashed lines show the expected evolution if the magnetic energy density and relativistic particle density scale with the thermal energy density, $B^2/8\pi \propto \rho V_{\rm s}^2 \propto n_{\rm rel} \propto t^{-2}$, while the dotted lines show the case when $B \propto r^{-1}$ and $n_{\rm rel} \propto r^{-2}$.

magnetic field is close to equipartition, $B^2/8\pi \approx 0.14\ \rho V_{\rm s}^2$, much higher than expected if the circumstellar magnetic field, of the order of a few mG, was just compressed, and strongly argues for field amplification, similar to what has been seen in simulations. Contrary to earlier, simplified models for SN 1993J based on free-free absorption only, the circumstellar density was consistent with $\rho \propto r^{-2}$.

In Fig. 84, we show the excellent fit of the resulting light curves. The high values of B implied that synchrotron cooling was important throughout most of the evolution for the electrons responsible for the cm emission, and also for the 21 cm emission before $\sim$ 100 days. At early epochs, Coulomb losses were important for the low energy electrons. The injected electron spectrum was best fitted with $p_{\rm i} = 2.1.$ acceleration.

The form of the light curves can be understood if, for simplicity, we *assume* equipartition, so that $B^2/8\pi = \eta\rho V_{\rm s}^2$. With $\rho \propto (\dot{M}/u)\ R_{\rm s}^{-2}$ and $V_{\rm s} \propto R_{\rm s}/t$, we find that $B \propto (\dot{M}/u)^{1/2} t^{-1}$. The optically thick part is therefore given by

$$F_\nu(\lambda) \propto\ R_{\rm s}^2\ \lambda^{-5/2} B^{-1/2} \propto (\dot{M}/u)^{-1/4} \lambda^{-5/2} t^{(5n-14)/2(n-2)}, \qquad (4.26)$$

since $R_{\rm s} \propto t^{(n-3)/(n-2)}$. For large n, we get $F_\nu(\lambda) \propto t^{5/2}$. An additional curvature of the spectrum is produced by free-free absorption in the wind, although this only affects the spectrum at early epochs.

In the optically thin limit, $F_\nu(\lambda) \propto R_{\rm s}^2 j(\lambda) \propto R_{\rm s}^2 \lambda^\alpha B^{1+\alpha} N_{\rm rel}$. If losses are unimportant, $N_{{\rm rel},tot} = 4\pi R_{\rm s}^2 N_{\rm rel}$, the total number of relativistic electron, may either be assumed to be proportional to the total mass, if a fixed

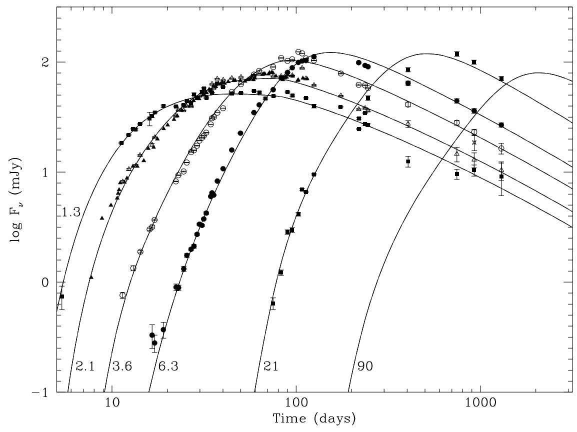

Figure 84: Observed and model radio light curves of SN 1993J.

fraction of the shocked electrons are accelerated, or be proportional to the swept up thermal energy. In the first case, $N_{\mathrm{rel},tot} \propto \dot{M} R_{\mathrm{s}}/u_w$, while in the second $N_{\mathrm{rel},tot} \propto \dot{M} R_{\mathrm{s}} V_{\mathrm{s}}^2/u_w$, so that in general $N_{\mathrm{rel},tot} \propto \dot{M} R_{\mathrm{s}} V_{\mathrm{s}}^{2\epsilon}/u_w$, where $\epsilon = 0$ or 1 in these two cases. Therefore, $F_\nu(\lambda) \propto \dot{M}/u \; R_{\mathrm{s}} V_{\mathrm{s}}^{2\epsilon} \lambda^\alpha B^{1+\alpha}$. If the B-field is in equipartition, as above, and using $V = (n-3)/(n-2) R_{\mathrm{s}}/t \propto t^{-1/(n-2)}$ we find

$$F_\nu(\lambda) \propto (\dot{M}/u_w)^{(3+\alpha)/2} \; \lambda^\alpha t^{-\alpha-(1+2\epsilon)/(n-2)}. \tag{4.27}$$

If synchrotron cooling is important, a similar type of expression can be derived. The main thing to note is, however, that the optically thin emission is expected to be proportional to the mass loss rate, $\dot{M}/u_w$, and that the decline rate depends on whether the number of relativistic particles scale with the number density or the thermal energy of the shocked gas, as well as spectral index. Observations of the decline rate can therefore test these possibilities.

Although a self-consistent model can be developed for SN 1993J and other radio supernovae, the modeling of SN 1986J and related objects has been unclear. Chugai & Belous propose a model in which the absorption is by clumps. The narrow line optical emission implies the presence of clumps, but they are different from those required for the radio absorption. The possible presence of clumps and irregularities introduces uncertainties into models for the radio emission, although rough estimates of the circumstellar density can still be obtained.

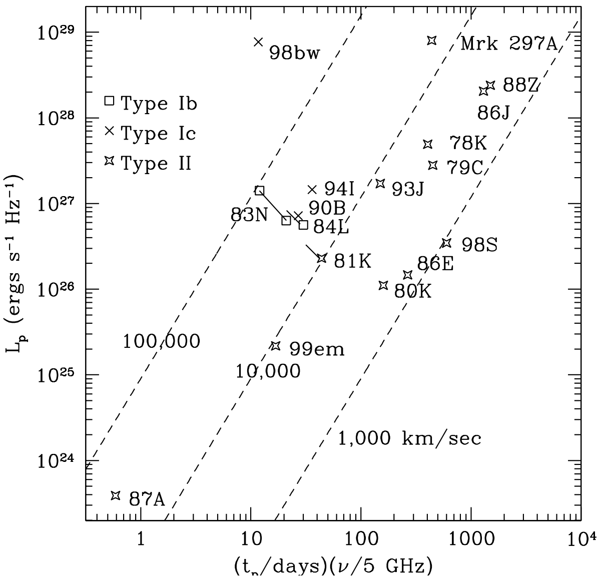

Figure 85: Peak luminosity and corresponding epoch for the well-observed radio SNe. The dashed lines give curves of constant expansion velocity, *assuming* SSA.

The relative importance of SSA and free-free absorption depends on a number of parameters. The most important of these are the mass loss rate, $\dot{M}/u_w$, the shock velocity, V_s, and the circumstellar temperature, T_e. In general, a high shock velocity and a high circumstellar temperature favor SSA, while a high mass loss rate favors free-free absorption.

Assuming SSA, an interesting expression for the velocity of the shock can be derived which can be tested against the observations. If we assume equipartition, $N_{\rm rel} \propto \rho V_s^2 R_s$ and $B^2/8\pi \propto \rho V_s^2$ we have $N_{\rm rel} \propto B^2 R_s$. Using this in Eq. (4.24) we get $\tau_s = const\ \lambda^{5/2+\alpha} B^{7/2+\alpha}\ R_s$. The peak in the light curve is given by $\tau_s \approx 1$. If we approximate the flux at this point by the optically thick expression Eq. (4.26) and solve for B we get $B \propto F_\nu(\lambda)^{-2}\ R_s^4\ \lambda^{-5}$. Inserting this expression in the condition $\tau_s \approx 1$, we find $R_s^{15+4\alpha} F_\nu(\lambda)^{-7-2\alpha}\ \lambda^{-15-4\alpha} \approx$ const. With $V_s = (n-3)/(n-2) R_s/t$ we finally have

$$V_s \approx \text{const}\ F_\nu(\lambda)^{(7+2\alpha)/(15+4\alpha)}\ \lambda\ t^{-1} \tag{4.28}$$

where all parameters refer to their values at the peak of the light curve.

Using this expression, we can plot lines of constant shock velocity into a diagram with peak radio luminosity versus time of peak flux, *assuming* that SSA dominates (Fig. 85). The positions of the lines depend only weakly on the equipartition assumption. Each SN can now be placed in this diagram to give a predicted shock velocity. If this is lower than the observed value (as measured by VLBI or from line profiles) SSA gives a too low flux and should therefore be relatively unimportant and free-free absorption instead dominate. The most interesting point is that most Type Ib/Ic SNe, SN 1983N, SN 1994I and SN 1998bw, fall into the high velocity category, while Type IIL SNe, like SN 1979C and SN 1980K, as well as the Type IIn's SN 1978K, SN 1988Z, SN 1998S fall in the free-free group. SN 1987A is clearly special with its low mass loss rate, but is most likely dominated by SSA.

4.9 The ring of SN 1987A

Hydrodynamics, SHock breakout, Echo, Binary evolution, Ring collision, ...

4.10 Conclusions

Circumstellar interaction of supernovae gives an important window on the nature of stars that explode and their evolution leading up to the explosion. Mass loss rates for the red supergiant progenitors of Type II supernovae range from $\sim 2 \times 10^{-6}$ $\mathrm{M}_\odot$ yr^{-1} for SN 1999em to $\gtrsim 2 \times 10^{-4}$ $\mathrm{M}_\odot$ yr^{-1} for SN 1979C and SN 1986J. Evidence for CNO processing has been found in a number of supernovae, including SN 1979C, SN 1987A, SN 1995N, and SN 1998S (see §1.13). In some cases, the reverse shock appears to be moving into gas that is H poor and O rich, e.g., SN 1995N; this relates to the total amount of mass loss before the supernova. The complex circumstellar environment of SN 1987A has become clear because of it proximity

The evidence on circumstellar interaction is especially useful when it can be combined with information from other aspects of the supernovae, such as their light curves and stellar environments. For example, from the pre-supernova stellar environment of SN 1999em, Smartt et al. deduced an initial mass of 12 ± 1 $M_\odot$. The supernova was of the plateau type, implying that hydrogen envelope was largely intact at the time of the supernova. This is consistent with the relatively low rate of mass loss deduced for the supernova progenitor.

In addition to information on the evolution of massive stars and their explosions, circumstellar interaction provides an excellent laboratory for the study of shock wave physics. Compared to older supernova remnants, the shock velocities are higher and the time evolution gives an additional dimension for study, although there is little spatial information in most cases. VLBI observations can, however, in this respect be extremely valuable, as demonstrated by SN 1993J. An object where both the spatial and time dimensions are accessible is SN 1987A, which has turned out to be an excellent source for the study of shock waves.

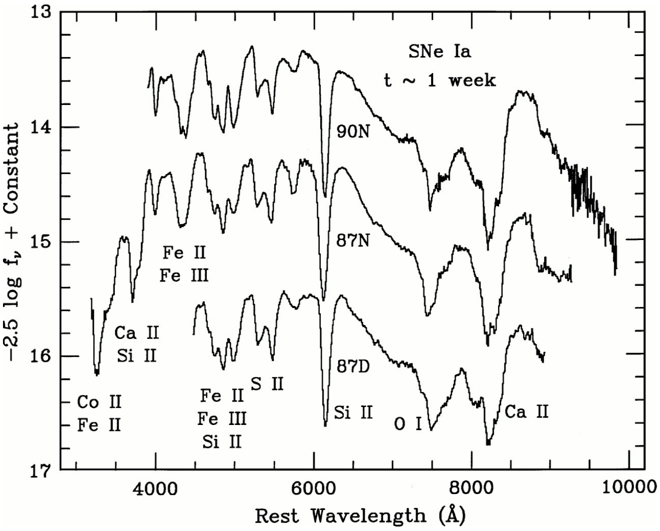

Figure 86: Compilation of three different Type Ia spectra at an age of about one week (Filippenko 1997).

5 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 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. 86 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 particularily 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. Sometimes, instead of Δm_{15} on uses a 'stretch factor', which strethes the light curve in the time direction. Using this the, dispersion in absolute magni-

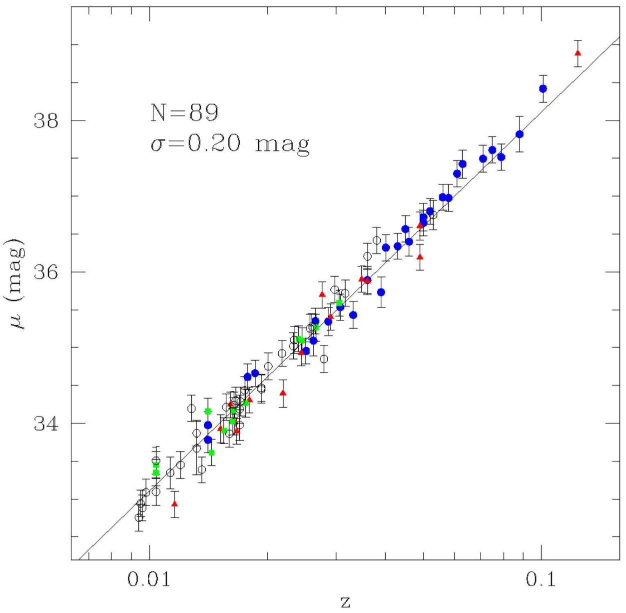

Figure 87: Hubble diagram based on Type Ia supernovae. The distance module μ is the difference in apparent and absolute magnitude, $\mu = m - M$, and z is the redshift. Note the small scatter illustrating the usefulness of the lightcurve as a distance measure. The dispersion is $\sigma = 0.2$ magn. (From Prieto et al 2006).

tude 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. Type Ias are,however, also found in spirals with ongoing star formation and there are indications that there are two different populations of Type Ia 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 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, $10^{-7} - 10^{-6}$ $M_\odot$ 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. **??**), and the density will therefore increase. This will heat up the degenerate core and when the density reaches 1.5×10^9 g cm^{-3} and the temperature $\sim 8 \times 10^8$ 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. 88)The nuclear burning will at the center go all the way to nuclear statistical equilibrium, mainly ^{56}Ni. As the explosion proceeds and the WD expands the nuclear burning occurs at lower densities and only partial burning will 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. 89). Because of the small radius of the white dwarf, the time scale until the burning front reaches the surface is only of the order of seconds.

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. 89). 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 stars 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

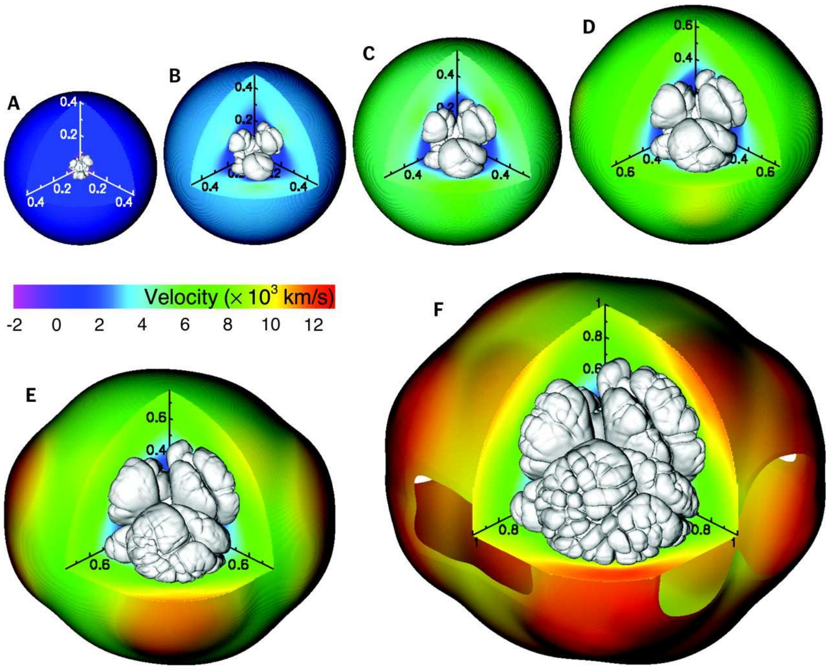

Figure 88: The development of the deflagration flame seen at (A) 1.26, (B) 1.49, (C) 1.57, (D) 1.65, (E) 1.76, and (F) 1.90 s after ignition. The color code shows the velocity. (Gamezo et al. 2003)

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 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 Tychos 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.

6 Gamma-ray Bursts

6.1 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. 90). CGRO also found that the bursts could roughly be divided 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 GRB's 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. 85). Also its optical luminosity was an order of magnitude higher than the typical Type Ic luminosity, indicating $M(^{56}\mathrm{Ni}) \sim 0.7\ \mathrm{M}_\odot$, 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 GRB's.

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.

6.2 Summary of observations

6.2.1 Prompt phase

In Fig. 91 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 profiles.

The duration of the bursts vary from 10^{-2} s to 10^{3} s. As is seen from Fig. 91, 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. 92). 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 index. At energies less than the peak energy, $E_{\rm p}$, $\alpha \approx 1 \pm 1$, while above $E_{\rm p}$ 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 GRB's 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 $E_{\rm p}$ is very large ranging from MeV's down to tens of keV's. 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 in the 100-300 keV channel divided by that in the 50 – 100 keV channel of the BATSE instrument (Fig. 93). Clearly, the short bursts have considerably harder spectra. The transition between the two groups is, however, smooth.

6.3 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 $f_{\rm p}$ of the photons have energies above the pair creation threshold, $\sim 2m_{\rm e}c^2$, the density of energetic photons is $n_\gamma = f_{\rm p}L/4\pi R^2 m_{\rm e}c^3$. The optical depth to pair production is therefore

$$\tau_{\gamma\gamma} = \sigma_T n_\gamma R = \frac{\sigma_T f_{\rm p} L}{4\pi R m_{\rm e} c^3} = \frac{\sigma_T f_{\rm p} F D^2}{R\Delta t m_{\rm e} c^3} = \frac{\sigma_T f_{\rm p} F D^2}{(\Delta t c)^2 m_{\rm e} c^2} \tag{6.1}$$

As a typical value for the fluency we take $F \sim 10^{-6}$erg cm^{-2} and $D \sim 2$ Mpc, corresponding to a total energy of 5×10^{50} ergs. If we take $\Delta t \sim 0.01$ s we get

$$\tau_{\gamma\gamma} = 3 \times 10^{14} f_{\rm p} \frac{F}{10^{-6}\text{erg cm}^{-2}} \left(\frac{\Delta t}{0.01\text{s}}\right)^{-2} \tag{6.2}$$

Therefore for any reasonable values of $f_{\rm p}$ 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 therefore the time when it will be observed is $t_{2\ obs} = t_{2\ em} + (D - r_2 \cos\theta)/c = [D - (r_1 + v dt_{em})\cos\theta]/c$. The time interval it will be received in is therefore

$$dt_{obs} = dt_{em} - \beta dt_{em} \cos\theta = dt_{em}(1 - \beta\cos\theta) \tag{6.3}$$

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. (6.3) as

$$dt_{obs} = \frac{dt_{em}}{2\Gamma^2} \tag{6.4}$$

The true size is therefore not cdt_{obs} but $2\Gamma^2 cdt_{obs}$.

Note that Eq. (6.4) 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 factor Γ. 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. 94 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

$$E \sim \Gamma M c^2 \tag{6.5}$$

which means that

$$M \approx 5 \times 10^{-6} \left(\frac{\Gamma}{10^3}\right)^{-1} \left(\frac{E}{10^{52}\mathrm{ergs}}\right) \mathrm{M}_\odot \tag{6.6}$$

This means that the explosion has to have a very small fraction of baryons to photons.

7 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_0c^2. To explain the large energies involved $\sim 10^{51} - 10^{52}$ 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.

7.1 The supernova - GRB connection

Supernovae have from the theoretical point of view for a long time been proposed as an origin for GRB's. 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 $60,000$ 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 ~ 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 lead 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 GRB's 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. 95). 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. 96). 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 GRB's where there is strong evidence for an underplaying supernova. With GRB980425, GRB030329 and these other cases, the SN/GRB connection is now firmly established,

Note, however, that the optically identified GRB's all belong to the long GRB's. 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.

7.2 Collapsars

The collapsar model for GRB's 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 GRB's. From the beaming angle, corresponding to a solid angle of $\Omega \sim 0.03$, we observe only every $\lesssim 200$ of all GRB's. 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 ~ 33 Gpc^{-3} 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. 97). The density contrast between the equatorial disk, where the density is $\sim 10^9$ g cm^{-3}, and the polar direction will therefore be very large.

The accretion rate in the disk is considerably lower, ~ 0.07 M$_\odot$ s^{-1}, than that of free-fall, $MR/V_{ff} \sim M\rho^{1/2}/450 \sim 1$ M$_\odot$ s^{-1}. With a total mass in the disk of $0.1 - 1$ M$_\odot$, the accretion will be sustained for 10–20 s.

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, k, $\nu + \bar{\nu} \rightarrow e^- + e^+$, 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 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. 98 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 $\sim 30°$.

In Fig. 99 we show the density and Lorentz factor at the final epoch of the model above. Although only ~ 20 at the time of the jet breakout, 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.

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 (XRF's), 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.

7.3 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 lead to the suggestion that merging neutron stars may have something to do with GRB's, and was for a long time the most popular GRB model. As we have seen, there is now compelling evidence that the long GRB's are connected to supernovae. This evidence does, however, not apply to the short GRB's. In particular, this models has some properties which can easier explain the short time scales connected with this class of GRB's.

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. 100). While most of the mass results in a black hole of mass ~ 2.5 $M_\odot$, a substantial fraction, $\sim 0.1-0.2$ $M_\odot$, will stay in the form of en 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 ~ 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.

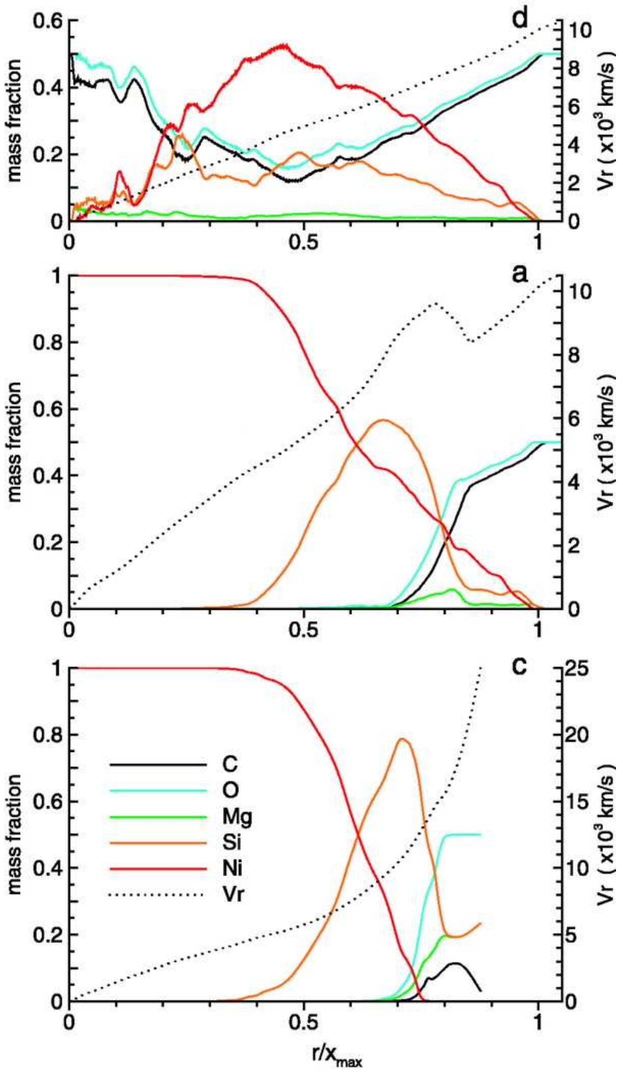

Figure 89: Abundances in one pure deflagration model (top) and two delayed detonation models started at 1.62 s (b) and at 1.51 s (c). Note the absence of oxygen and carbon in the delayed detonation models (Gamezo et al. 2005)

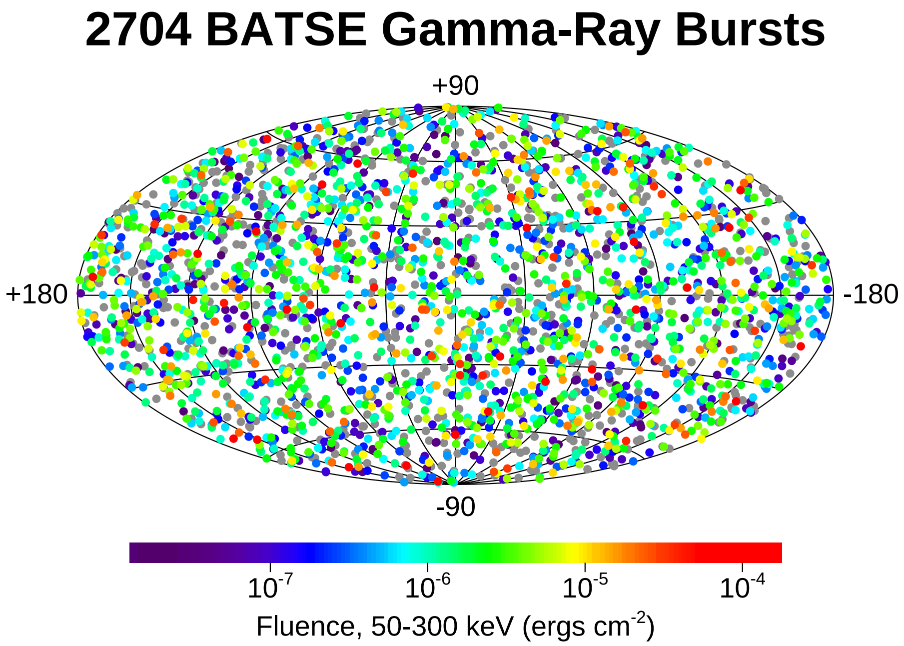

Figure 90: GRB distribution on the sky for bursts observed with BATSE

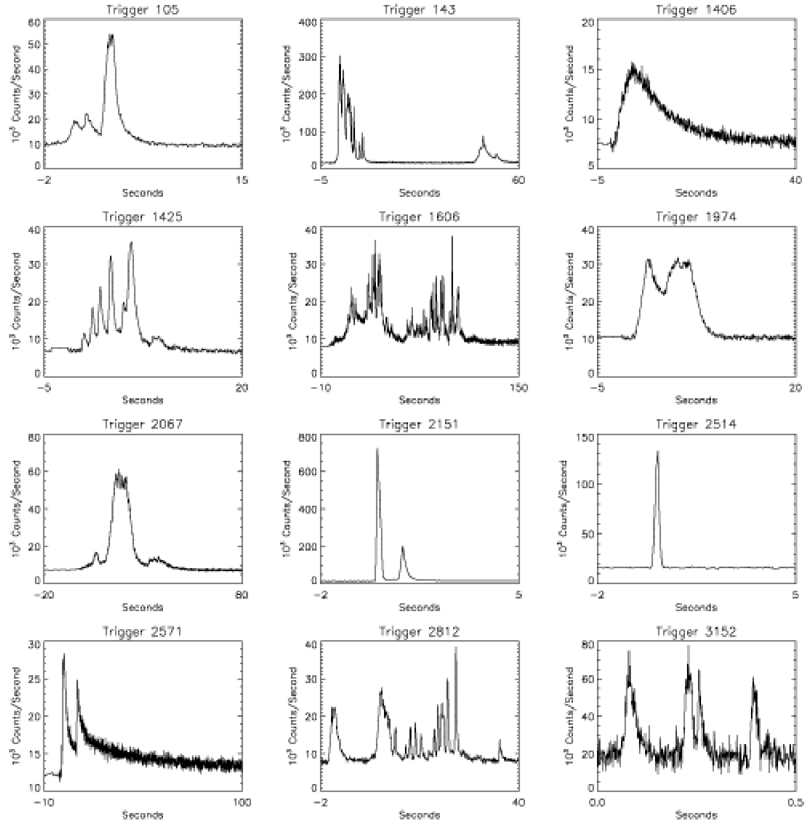

Figure 91: Examples of light curves observed with BATSE.

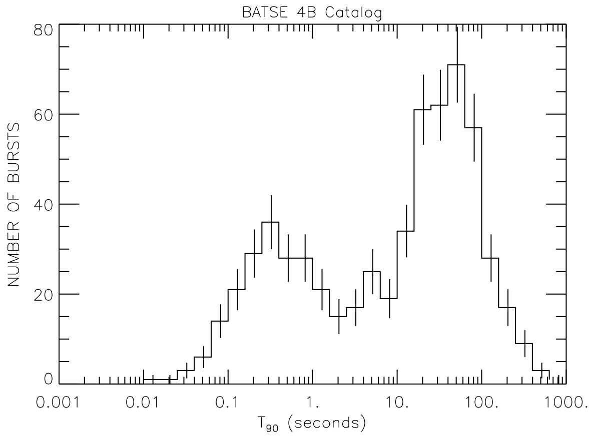

Figure 92: Distribution of durations of gamma-ray bursts detected by BATSE. The duration is defines as the time, T90, between when 5% and 95% of the total number of counts are measured.

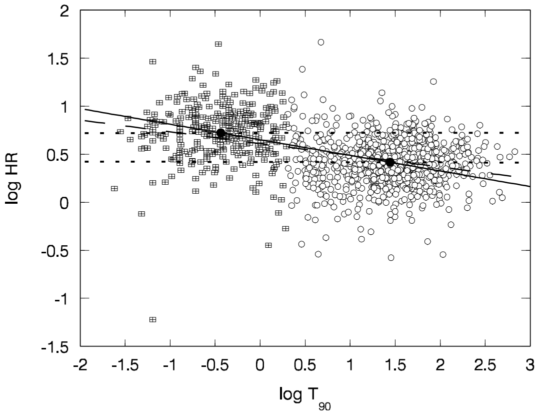

Fig. 1

Figure 93: 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)

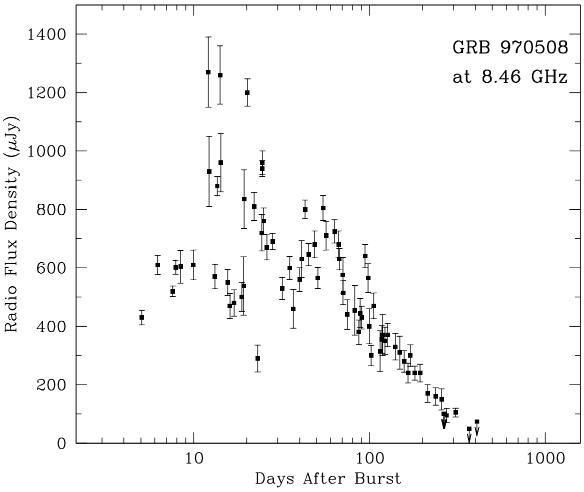

Figure 94: Radio light curve for GRB970508 at 8.46 GHz. Note the rapid fluctuations due to interstellar scintillations in the light curve during the first ~ 50 days (Frail 2003).

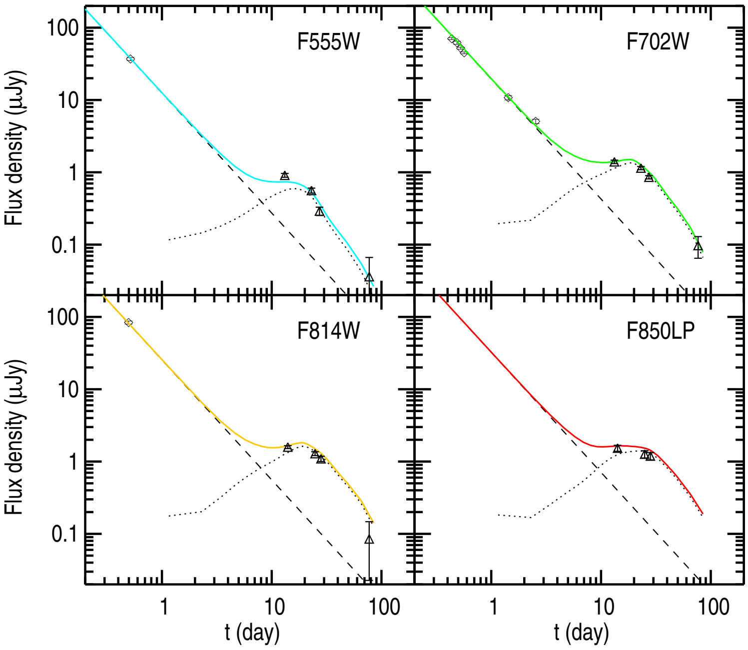

Figure 95: Light curves of the afterglow of GRB011121 obtained with HST (triangles) and ground based telescopes (diamonds). Note the bump in the light curve at 10-30 days, consistent with that from a of SN 1998bw, dimmed by $\sim 55\%$ (Bloom et al 2002).

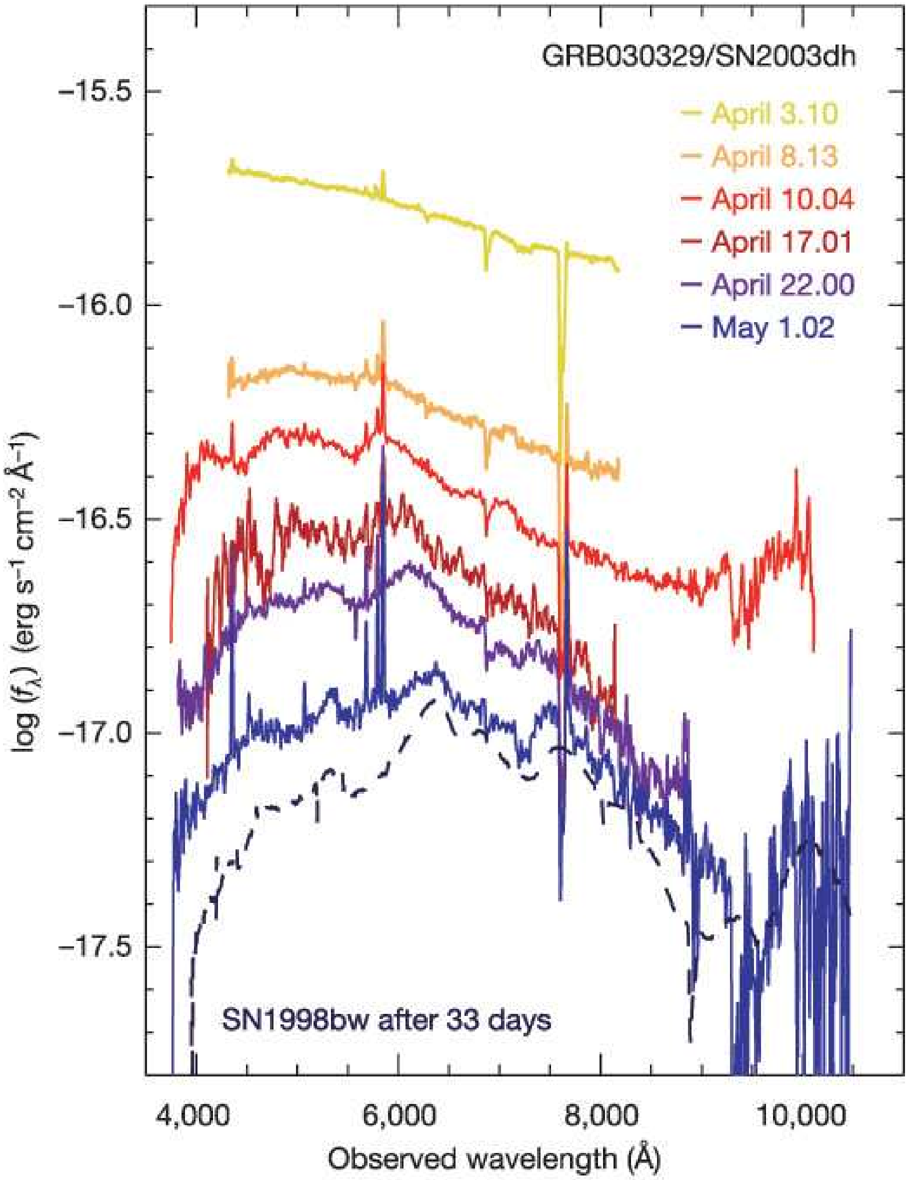

Figure 96: Spectral sequence of GRB030329/SN2003dh with VLT (Hjorth et al 2003). Note the power law spectrum on April 3 and the gradually stronger supernova contribution. The dashed line shows the spectrum of SN 1998bw at an age of 33 days.

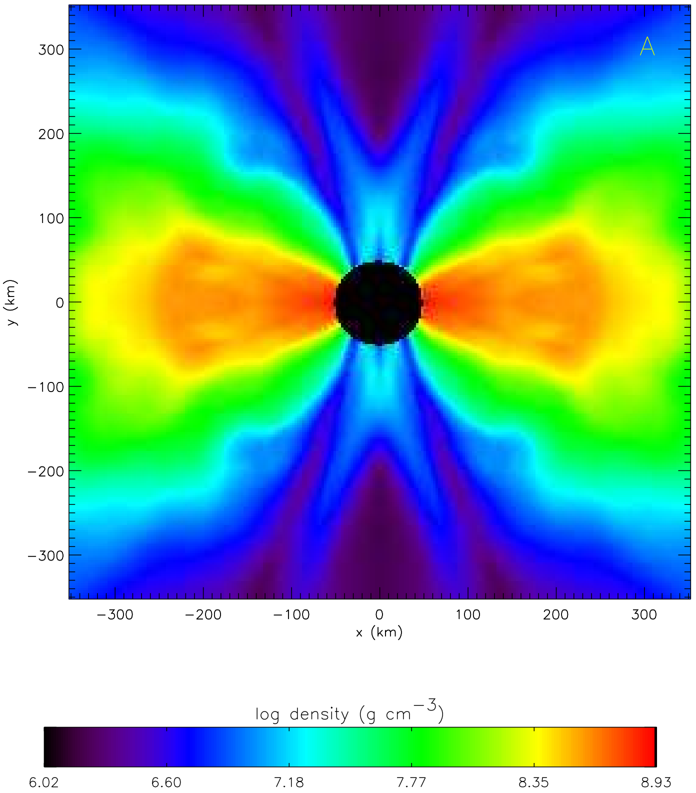

Figure 97: Density in the center ∼ 7 s after collapse. Inside of ∼ 200 km the centrifugally supported torus can be seen. In the polar direction the density is very low because the lack of centrifugal support has emptied this region. (MacFadyen & Woosley 1999)

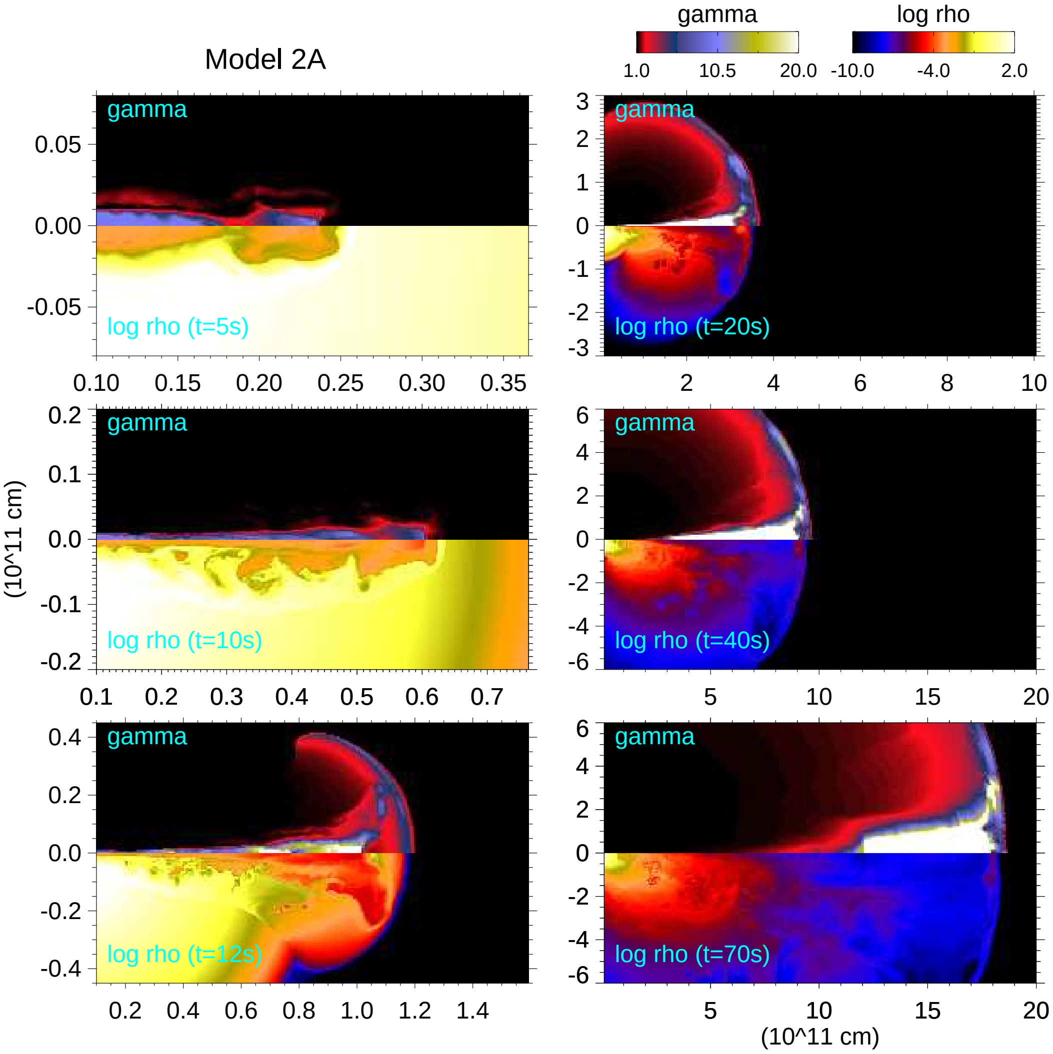

Figure 98: Simulation of jet propagation and breakout. The different panels show the Lorentz factor and density at six epochs, 5, 10, 12, 20, 40, and 70 s. (from Zhang, Woosley, Heger 2003)

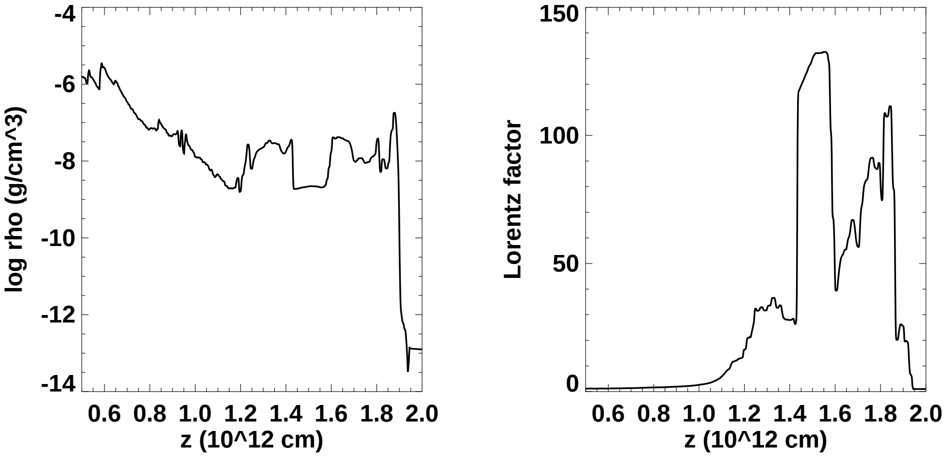

Figure 99: Density and Lorentz factor in the jet in Fig. 98 at 70 s. Note the highly variable density and especially Lorentz factor (Zhang, and Woosley 2002).

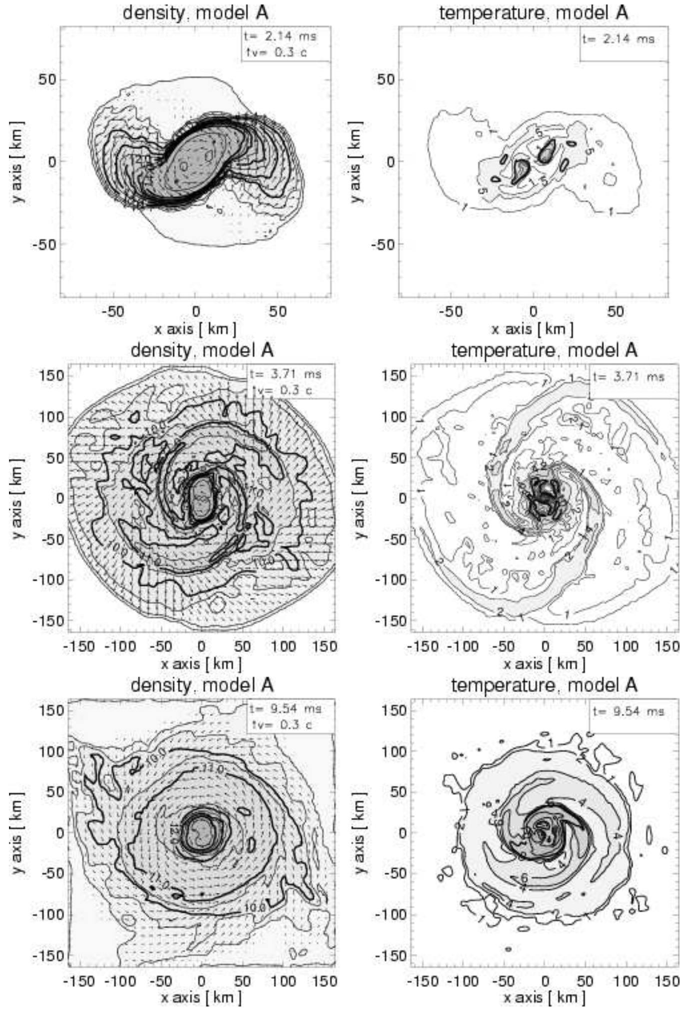

Figure 100: Evolution of the neutron star binary at different epochs after the start of the simulation. The contours show the density and temperature, while the arrows show the velocity field. (Ruffert & Janka 2001).