ISAS - INTERNATIONAL SCHOOL
FOR ADVANCED STUDIES

Relativistic Hydrodynamics for Gamma-Ray Burst Models

Thesis submitted for the degree of
Doctor Philosophyæ

Astrophysics Sector

Candidate:

FUKUN LIU

Supervisors:

DR. ANTONIO LANZA

PROF. JOHN C. MILLER

April 1999

SISSA - SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI

TRIESTE
Via Beirut 2-4

TRIESTE

Acknowledgement

In the past more than four years, I have been living and working in the very good institute SISSA in the nice city Trieste. I met many people there and enjoyed my daily life and working with them. Without their support, it would have been impossible for me to finish this thesis. I have been waiting for and would like to take this opportunity to give my deep thanks to all of them.

Dr. Antonio Lanza and Prof. John Miller, my supervisors, are the first two people whom I am greatly indebted to for their constant encouragement, stimulating discussions, and very kind help whenever I had difficulties not only in the work but also in my daily life. They are really two good supervisors. Their painstaking reading, especially by Prof. John Miller, have made this thesis more readable. Very special thanks are still due to Dr. Antonio Lanza who introduced me this exciting topic "Gamma Ray Burst" and organized the financial support for the last several months. I also wish to thank Prof. Dennis W. Sciama, until recently the head of the Astrophysical Sector of SISSA. When I just arrived, he suggested me to talk to all the professors and doctors in Astrophysics Sector of SISSA and I am benefitted a lot from this suggestion. However, my thanks are not only due to this but also for his long-standing supports and fascinating discussions.

My deep thanks are also given to the other members of the Astrophysics Sector of SISSA: Dr. Hara Papathanassiou, Dr. Ewa Szuszkiewicz (now at Torun, Poland), Prof. Aldo Treves (now at Como University, Milan), Prof. Luigi Danese, Dr. Riccardo Valdarnini, Dr. Paolo Salucci, Dr. Annalisa Celotti, and Dr. Koji Uryu. They are always very kind, whenever I discussed with them. Their help means a lot for me.

I have been getting a lot of help from many people outside SISSA during my PhD study. Among them, I am particularly very grateful to Dr. Friedrich Meyer and Dr. Emmi Meyer-Hofmeister at MPA (Garching, Germany), Prof. Marek Abramowicz in Chalmers University of Technology (Göteborg, Sweden), Prof. Xie Guangzhong at Yunnan Observatory (Kunming, China) and Prof. Cao Shenglin in Beijing Normal University (Beijing, China). Dr. Friedrich Meyer and Dr. Emmi Meyer-Hofmeister were my advisors when I was in MPA, and are always ready to help me whenever I need. Prof. Xie Guangzhong was the person who addressed my attention from physics

to astrophysics. I learned from him not only on astrophysics.

It is my great pleasure to give my many thanks to the SISSA computer staff members and the SISSA secretaries, in particular Mrs. Alexandra Meehan Poretti, Mr. Riccardo Iancer and Mr. Andrea Parma, for their kind help. I still remember the embarrassing moment when I had trouble with my visa in Sweden. I very much appreciate their help and their efforts to provide in SISSA a very efficient environment. How essential their work is for us!

Last, but not the least, I would like to thank all my friends, from whom I got a lot of help and shared my good times during the past years. I could not list all of them, otherwise this long thesis would become even longer. However, they are so important that it cannot be imagined without them.

This thesis is dedicated to my parents, brothers, sisters, and their families.

Table of Contents

1 Introduction

In 1960s, the US Defense Department worried that the Soviet Union or China could explode nuclear weapons in space and launched a series of Vela spy satellites aimed at detecting gamma-ray emission associated with clandestine nuclear test. These spacecraft were launched and operated in pairs with two identical satellites of a circular orbit 250,000 kilometers in diameter so that no part of the Earth was shielded from direct observations. The satellites carried x-ray, gamma-ray, and neutron detectors as a basic instrumentation.

Of course, clandestine nuclear explosions were not found, but occasional flashes of gamma rays, lasting a few seconds, were indeed seen nearly simultaneously by detectors on at least two different, widely separated spacecraft. The bursts were from random directions in space and therefore came neither from the Earth nor from the sun. The first cosmic gamma-ray burst event (GRB) was recorded on July 2, 1967, but the discovery was announced only four years later, 1973, by Klebesadel, Strong and Olson in a paper entitled *Observations of Gamma-ray Bursts of Cosmic Origin* (Klebesadel, Strong, Olson 1973), which discussed 16 short bursts of photons in energy range between 0.2 and 1.5 MeV observed by Vela 5 and Vela 6 satellites between July 1969 and July 1972.

In the scientific literatures, the sources of the bursts are sometimes called "γ-ray burster," while the observed phenomena themselves are called γ-ray bursts. In many situations, the term "γ-rays" is applied only to quanta with energies larger than several hundred - perhaps 511 - keV, the electron rest mass energy. In the context of GRBs, it is customary to designate all quanta with energies of about 20 keV and above as "γ-rays", as one finds a continuous spectrum all the way from about 20 keV to several tens of MeV, with the energy output reaching a maximum in the MeV region. A pure radiation energy condensation region (fireball) is opaque due to electron-positron pair production, if its temperature is high enough. When the local temperature drops around 20 keV, the number of pairs becomes sufficiently small and the plasma become optically thin to the photons. Usually a GRB burst is denoted by GB or GRBijklmn, where ij, kl and mn refer to the year, the month and the date on which the burst occurred. If more than one burst was seen on a given date, an additional symbol a, b,

c, ... is added in the order (in time) in which they occurred. If the source direction of a GRB is determined to accuracy of better than 1^{deg}, the source is also referred to as GBS ijklmn where ij, kl refer to right ascension (hours, minutes) and mn to declination (degrees).

Following the announcement of the discovery of GRBs, groups all over the world built dedicated instrument to study this exciting phenomena, and these have since been flown in geocentric and heliocentric orbiting spacecraft and on planetary probes in order to have improved direction information. The most important instrument among them is the Burst and Transient Source Experiment (BATSE) (Fishman *et al.* 1989). Some of the other important satellites or detectors are KONUS experiments of the Leningrad group (Mazets and Golenetskii 1981), SIGNE experiments of the Franco–Soviet collaboration (Barat *et al.* 1981), PVO the Pioneer Venus Orbiter Experiment (Evans *et al.* 1981), SIGMA a French–Soviet high resolution detector (Paul *et al.* 1990), PHEBUS a Russian-French experiment (Terekhov *et al.* 1994), WATCH (Brandt, Lund, Castro-Tirado 1994), Ulysses (Hurley, *et al.* 1992). The following important satellites or detectors are also used in the observations of GRBs: ISEE-3 the International Sun-Earth Explorer experiment (Klebesadel, Evans & Laros 1981), SMM the Solar Maximum Mission Experiment (Forrest *et al.* 1980), HEAO-1 (Knight, Matteson and Peterson 1981), GINGA a Japanese satellite experiment (Murakami *et al.* 1989), BeppoSAX a Italian-Dutch satellite (Boella *et al.* 1997), RXTE the Rossi X-ray Timing Explorer. The other current and future experiments which will be dedicated to or can be used in the observations of GRBs are HETE-II High Energy Transient Explorer-II (Riker 1999), Current IPN (3rd IPN) Interplanetary Network (Ulysses, Near Earth Asteroid Rendezvous (NEAR), RXTE, Wind, and BeppoSAX), CATSAT Cooperative Astrophysics and Technology Satellite (Forrest, *et al.* 1995), INTEGRAL International Gamma-ray Laboratory (Kretschkmar *et al.* 1999), Future IPN (MSO Mars Surveyor Orbiter, Near Earth Asteroid Prospector, INTEGRAL, BATSE, and Ulysses), MIDEX (Hurley 1999). The four experiment aboard the Compton Gamma-ray Observatory (CGRO) are providing a major fraction of the current γ-ray burst data, although several other experiments are also working. BATSE detected about 6 bursts per week. The latest published fourth BATSE catalogue of GRBs contains 1637 events (Paciesas *et al.* 1998).

Astrophysicists are understandably excited about gamma-ray burst. Immediately after the discovery of GRBs, there was a plethora of models to explain their origin. After two years of the discovery, 21 models had been proposed (Ruderman 1975). The candidate objects were supernovas, neutron stars, flare stars, antimatter effects, relativistic dust, white holes, and some even more bizarre options. Most theoretical astrophysicists tend to twist rather strenuously to convince themselves and others that a new phenomena fits into their chosen specialties. The number of proposed models

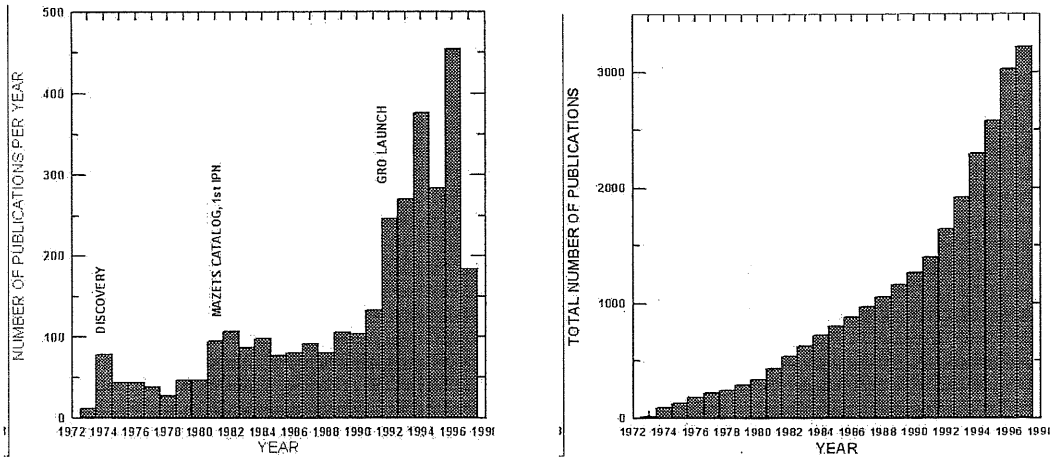

Figure 1.1: Number of gamma-ray burst publications as a function of year. (Left) for number of publication per year and (Right) for number of accumulative publications. The cutoff date is mid-1997. Some milestones are indicated which reflect the sudden increase in the publication rate. From Hurley 1998

continued increasing with time and reached more than one hundred before the end of 1992 (Nemiroff 1994). At a time there were more theories than bursts. The number of articles published each year since 1973 is shown in figure 1.1 (From Hurley 1998). Starting with a modest article per month in 1973, it had reached more than one per day in 1996. There are more than 3000 papers published until the middle 1997.

During 1970s and 1980s, the models were generally divided into three classes: those in which the bursts were respectively in the Galactic Disc at distances of a few hundred parsecs, in the halo at distances of tens of kilo-parsecs, and at cosmological distances. The candidate objects ranged from Fermi-sized primordial black holes to 10 km-sized neutron stars. There was a consensus until 1992 that GRBs originate somehow on or near neutron stars with high magnetic fields, probably due to magneto-spheric phenomena or glitches.

However, a surprise came with the launch, in April 1991, of the Compton Gamma-ray Observatory (CGRO) satellite, whose Burst and Transient Source Experiment (BATSE) offered systematic all-sky coverage with good sensitivity over the photon energy range 30 keV – 1.9 MeV. The most remarkable BATSE result is the unambiguous evidence that the burst are highly isotropic over the sky and lack of faint bursts which deviate from the classic Euclidean distribution. The "non-Euclidean" counts imply that the surveys are probing to distances where the sources are, for some reason, thinning out. This results revolutionize GRB observations and consequently some of our basic ideas on the origin of GRBs. The lack of any enhancement either towards the plane of the Galaxy, or towards the Galactic Center, is a very severe constraint on the hypothesis that bursts come from the Galaxy. The magnetized neutron star models

within the Galactic plane were therefore ruled out.

Since then, gamma-ray burst research has switched from a field with very little data and detailed theoretical models to one with a lot of data and models with little or no detail. The consensus opinion of the locale of the sources of gamma-ray bursts changed from a fraction of galactic scale height to either an extended galactic halo with typical distance about 300 kpc (Lamb 1995) or cosmological distances with a typical redshift of order unity (Paczyński 1995). The gamma-ray burst enigma appeared to be as great as it was twenty years ago (Ruderman 1975). A wealth of data on time profiles, spectral characteristics, and burst distributions had failed to provided conclusion evidence on the distance scale, central objects, or emission mechanism for classical gamma-ray bursts.

The immediate implication was that the total energy involved for a typical burst with observed fluence 10^{-5}ergs/cm^2 is of order 10^{52}ergs or more if cosmological, or of order 10^{48}ergs if they were in the extended halo. The rapid variability of GRBs suggests a compact source, which will have a huge optical depth due to pair productions by the high energy γ-rays. However, the observed spectrum extends far above 500 keV and the spectrum is non-thermal. Both indicate that the sources are optically thin. To solve the compactness problem, relativistically moving regions with a Lorenz factor $\gamma \geq 100$ are needed. This led to the fireball model, according to which GRBs are produced when a relativistic flow is slowed down and its kinetic energy is extracted and converted to radiation. The important properties of the fireball model is that it deals only with the emitting regions and leave open the questions what is the source that produces the relativistic flow and how the flow is produced. The emitting regions are distant and separate from the source.

An immediate prediction of the fireball models is the appearance of afterglow, produced on much longer time scale and at longer wavelengths. The second GRB revolution occurred on 28 February 1997 when the Italian-Dutch satellite BeppoSAX discovered X-ray emission that accompanied GRB970228 (Costa *et al.* 1997). The X-ray emission continued for a few hours. The accurate X-ray position enabled follow-up observations that detected a decaying optical source. GRB afterglow was discovered and an optical and X-ray counterparts were found. Since then, about a dozen afterglows that accompanied GRBs were observed in X-ray, optical, IR, and radio wavelengths. Cosmological redshifts were measured in several cases, confirming the cosmological origin of GRBs. Evidence for relativistic motion and relativistic shocks has been found and the fireball model has been practically confirmed.

The emerging picture is now that GRBs are extremely luminous objects with extremely relativistic motion ever discovered: (1) GRBs involve relativistic motion at a Lorenz factor of 100 or larger; (2) most current GRB models involve the formation of a black hole in one way or another; (3) if binary neutron star mergers are the sources

of GRBs, then GRBs are also associated with gravitational radiation signals; and (4) GRBs observed to high redshift might be used to measure cosmological parameters and to teach us about the epoch of galaxy formation.

The discovery of long-wavelength afterglow emission lasting for days to months constitutes a turning point in GRB research. However, despite these exciting new data and over 2300 publications on the observations and theoretical models of GRBs, most of the basic questions on GRBs remains unanswered. For fireball models, though the consistency between the observation of afterglows and predication of the model is good, the consistency between the prediction and the observations of GRBs is not satisfying.

Our project is to use general relativistic hydrodynamics and then general relativistic radiation hydrodynamics to construct a fireball model. We try to investigate the whole evolution process of the fireball from optically thick through the transitional to optically thin phases and from the initial acceleration through the saturation to the end of the slowing down stage via shocks. The time-dependent and spectrum-independent radiation will be calculated.

As the observations of GRBs, afterglows and the related phenomena are the base of any theoretical work, we will completely, hopefully, review the observations in Chapter (2) after this introduction. When we observe the phenomenon, the first and probably most important question we should want to ask the origin: where they are from and what kind of astronomical objects produce it. This question is most difficult on the study of GRBs. Many candidate models, which consist of the most of the models on GRBs, are trying to answer it, but no one has given very promising predications or explained the observations well. The two revolutions on the research of GRBs happened on this aspect. The current models, except the one on the merging of binary neutron stars, are very schematic. In Chapter (3), we will discuss the models, according to their distance scales to us. After the question about the candidate objects to produce GRBs, we naturally ask how they produce such kind of bursts with very extreme features. In general, it is very difficult to address this question before we find out the burster. The fireball models isolate the hydrodynamic processes from the candidate objects for the central engines and enable us to study the process in very detail and to get fruitful results, which are what our work wants to compare. So, before we start discussing our work in Chapters (5)-(7), we want to see the present status of the fireball models and survey their results in Chapter (4).

We start presenting our work on relativistic hydrodynamics on fireball models from the deriving all the general relativistic hydrodynamic equations in for continuous fluid and for shock and contact discontinuities in Chapter (5). As the equation system is hyperbolic, it will be written in characteristic form. The characteristic equations will be used when we treat boundary and discontinuity surfaces. These discontinuities may be left from the initial discontinuities between the fireball and the external medium,

when a fireball initially at rest is set in an external medium. A fireball expanding into external medium is a special Riemann problem, which we will discuss in more detail.

In Chapter (6), we will use the basic equations obtained in Chapter (5) to study two fireball systems: free expansion fireball and fireball expanding into ISM. As the conditions in the fireball models are very extreme, we have to use some techniques to treat the fluid, boundaries and the discontinuity surface. The techniques will be discussed in detail in this chapter. Our results on gamma-ray burst fireball models will be given after the discussions on the techniques.

As the last of the main body of the thesis, Chapter (7) talk about our work on jets and accretion disk based on the collections of observational data, which are beyond the fireball model but not the completely beyond gamma-ray bursts. In the popular central engine models on gamma-ray bursts, torus or thick disk–black hole system is essential. Jets and accretion disk–black hole system are symbiotic. A fireball may be created in such a system with mildly relativistic speed. However, our work in Chapter (7) is on jets and accretion disk in active galactic nuclei. We compile a list of extragalactic radio jets and give some statistic results. However, as the jet catalogue is too large, we do not include it in the thesis. It is believed that jets are produced in the inner region of the accretion disk. A slim accretion disk with a high viscosity parameter α is unstable and show limit cycle variation. The variation occurs in the inner region of the accretion disk and may probably influence the jets and make it observable due to the amplification of the jets. Therefore, such variation is must likely observed in BL Lac objects. We collect observational data for two BL Lac objects and study the historical light curves. We indeed find that periodic variation in the flux exist in these two AGNs. The periodicity in this two objects and in others will be explained based on the slim accretion disk models.

Polytropic models are important in astrophysics. When people simulate the merging of binary neutron stars, which is probably the favorite model for the central engine of GRBs, they make polytropy assumption. We have worked out an analytic approximate solutions to the polytropic gas sphere. However, as it is a bit too exotic to our topic, we put the results in the appendix.

When we do calculation, we will use $H_0 = 65\text{km} \cdot \text{sec}^{-1} \cdot \text{Mpc}^{-1}$ throughout the thesis. In the thesis, we also use the follwoing abbreviations

GCN: the Gamma-ray Coordinates Network (GCN)

ROTSE: the Robotic Optical Transient Search Experiment

FRED: Fast Rise, Experiment Decay

2 The Observations of GRBs

2.1 Burst rate

Since the 2B catalog, the significant decrease in the rate of solar flare triggers has improved the exposure. The burst detection rate rise from 0.82 bursts per day for 1B through 0.88 bursts per day for 2B to 0.96 bursts per day for 3B catalog period (Meegan *et al.* 1995). For 3B catalog, the probability of detecting a burst above threshold is about 47.7% for 3B catalog (Paciesas *et al.* 1998). This corresponds to $\simeq$ 1000 bursts per year. For cosmological sources with no source evolution, this corresponds to $\sim 2 * 10^{-6}$ (long) event per galaxy per year (independent of H_0) (Cohen & Piran 1995), corresponding to a mean comoving galaxy density of $n_0 = 2.7 \times 10^{-3}$ Mpc^{-3} for $\Omega = 1$ and $H_0 = 65 \mathrm{km\,sec}^{-1}\,\mathrm{Mpc}^{-1}$ (Kirshner *et al.* 1983). The rate of short bursts is comparable. The total GRB rate is $\sim 4 \times 10^{-6} \mathrm{yr}^{-1}$. It goes without saying that if the bursts are beamed with an opening angle θ then the event rate should increase by a factor of $4\pi/\theta^2$ relative to this rate.

As GRBs are probably associated with star formation region and if the GRB rate follows the massive star formation rate (Totani 1997), which was recently found to have about 10 times higher rate at $z \approx 1$ than it is at $z = 0$ (Madau *et al.* 1996, 1998), the comoving GRB rate will increase dramatically with the cosmological distance (Sahu *et al.* 1997a).

In any case, GRBs are very rare compared to ordinary supernova.

2.2 Burst morphology

The burst have a complicated and irregular time profiles which vary drastically from one burst to another.

The morphology of γ-ray bursts is very diverse, which together with the large range of the durations hampers us to classify γ-ray bursts into well-defined types based on their time profiles. Figures 2.1, 2.2, 2.3, 2.4 give some examples. Another difficulty is that there appear to be no other observational parameters that have a distinct correlation with the temporal morphologies. However, several attempts have made in

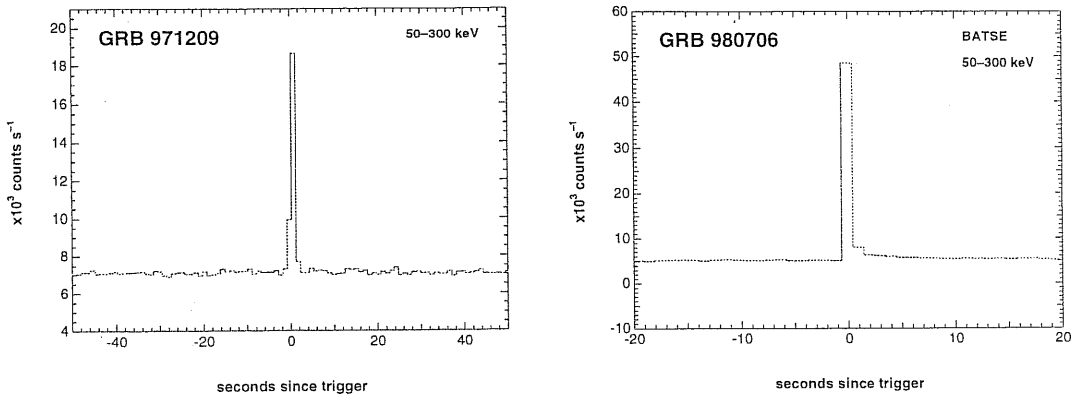

Figure 2.1: Examples of single or spike events.

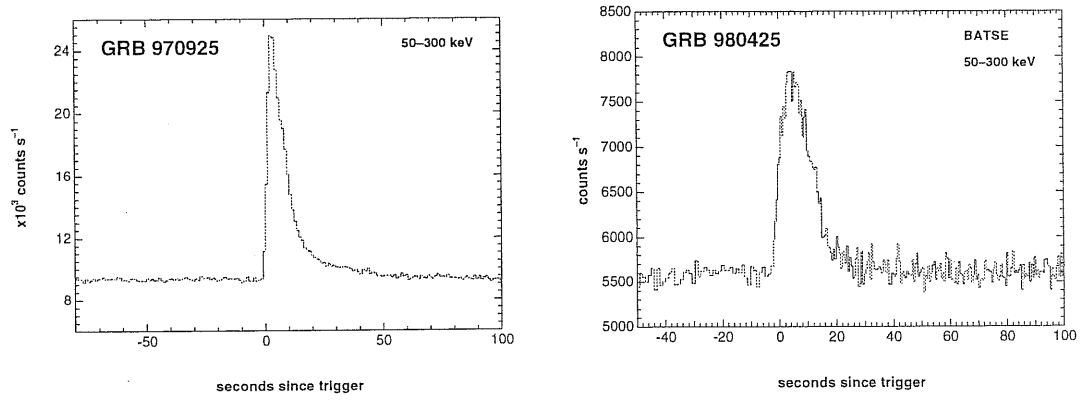

Figure 2.2: Examples of strong bursts which show no structures on fine scales.

the past to categorize γ-ray burst morphologies. Fishman and Meegan (1995) show four possible classes of γ-ray burst profiles from the BATSE Gamma-Ray Burst Catalogs:

(1). **single pulse or spike events** (Figure 2.1);

(2). **smooth, either single or multiple,** *well-defined* **peaks** (Figure 2.2);

(3). **distinct, well-separated episodes of emission** (Figure 2.3); and

(4). **very erratic, chaotic, and spiky bursts** (Figure 2.4).

The single pulse bursts of **Class 1** have little or no additional structure. A particular subclass of these events are the so-called FREDs (Fast Rise, Exponential Decay), in which the rise time is much shorter than the fall time. The decaying portion is not strictly an exponential function but has a continually decaying slope. In many FREDs, there are one or more minor increase during the decaying portion. The spectra of these γ-ray bursts almost always evolve from hard to soft, and the secondary peaks, if present, are usually much softer than the primary peak.

In many of the smooth, multiple-peak events of the **class 2**, the rise-times and the fall-times tend to be similar.

Class 3 events exhibits relatively long periods of time between peaks in which there

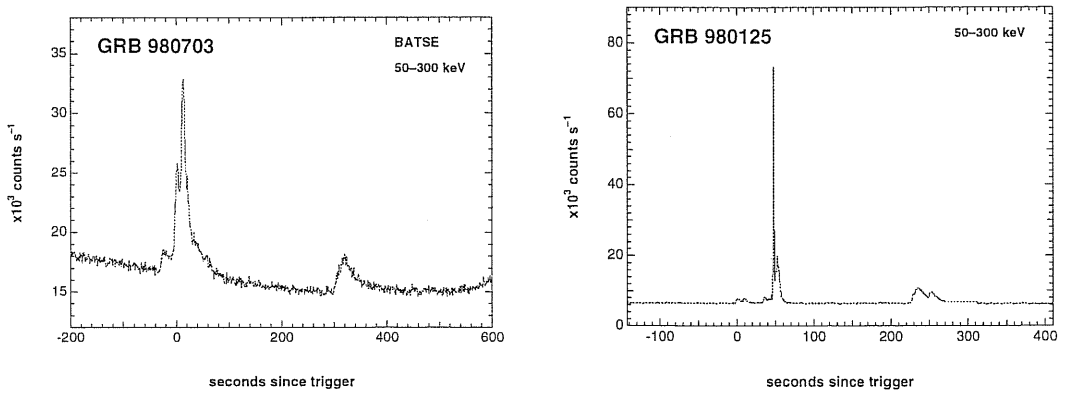

Figure 2.3: Some bursts showing distinct, well-separated episodes of emission.

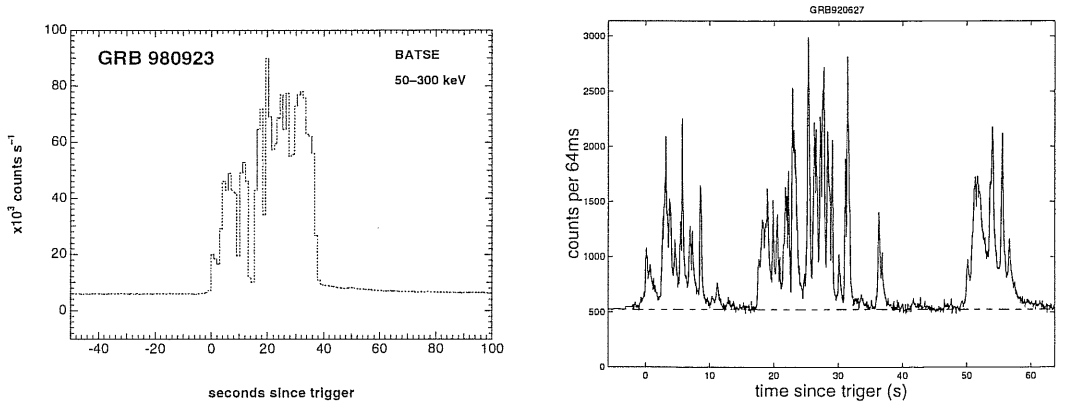

Figure 2.4: Examples of GRBs with extremely complex temporal structures.

is no detectable emission. This type of bursts is ill-defined in weak bursts, because there is less certainty in distinguishing periods of emission.

The majority of **class 4** γ-ray bursts are highly structured, with many complex, overlapping peaks and spikes (spikes are narrow, intense peaks with durations $\lesssim$0.1 s). In many of these bursts, there appears to be an underlying "envelope" of emission from which the peaks and spikes arise.

Another general property of the γ-ray burst time profiles is that at higher energies the overall burst duration are shorter and sub-pulses with a burst tend to have shorter rise times and fall times (sharper spikes) (references). Most bursts also show an asymmetry, with shorter leading edges than trailing edges (Fishman *et al.* 1994; Nemiroff *et al.* 1994; Link & Epstein 1996 and references therein). However, some bursts show no asymmetry, even when high counting rates would permit such measurements to be made.

No short-duration (microsecond) flares or bunching of photons in γ-ray bursts have been found. These high-energy, short bursts might be expected to come from relatively nearby Galactic evaporating black holes remaining from the early Universe (Hawking

1974; Page & Hawking 1976).

The most bursts, the typical variation takes places on a time-scale δT significantly smaller than the total duration of the burst, T. In a minority of the bursts, there is only one peak with no substructure and in this case $\delta T \sim T$. It turns out that the observed variability provides an interesting clue to the nature of GRBs. We define the ratio $N \equiv T/\delta T$ which is a measure of the variability. Fig. 2.4 depicts the total observed counts (at $E > 25$ keV) from the GRB920627. The bursts lasted $T \sim 100$ sec and it had peaks of width $\delta T \sim 1$ sec, leading to $N = 100$.

2.3 Burst Duration

The duration, like the burst morphology, is difficult to quantify because it is dependent on the intensity and background and somewhat dependent on the time resolution of the experiment. The BATSE group defined two measure: the T_{90} (T_{50}) measure, a time over which 90% (50%) of the burst energy fluence is observed, eliminate the initial and final 5% (25%) of the total burst emission and yield durations that are, to first order, independent of intensity. The duration of γ-ray bursts range from about 5 ms (Bhat *et al.* 1992) to over 1000 s (Fishman & Meegan 1995, Meegan *et al.* 1996). The duration is energy-dependent and the duration is smaller in higher energy band (Fenimore *et al.* 1995; Norris *et al.* 1996; Mitrofanov *et al.* 1997). The bursts GRB961027a, GRB961027b, GRB961029a and GRB961029b occurred from the same region in the sky within two days (Meegan *et al.* 1996a). If this "gang of four" is considered as a single very long burst, then the longest duration so far is two days! These observations may indicate that some sources display a continued activity at a variable level over a period of days (Katz 1997).

The duration distribution (T_{90}) exhibits bi-modality (Kouveliotou *et al.* 1993; Lamb *et al.* 1993; Mao *et al.* 1994; Fishman *et al.* 1994; Meegan *et al.* 1996), with peaks at $\sim$ 0.5 s and $\sim$ 30 s and a minimum at around 2 s (Meegan *et al.* 1996). The observed bimodal temporal behavior results from properties of the pulsed emission of GRBs (Mitrofanov *et al.* 1998). However, the distribution of total duration T_{90} is fitted better with tri-modality (Fig.2.5) (Mukherjee *et al.* 1998; Horváth 1998), with peaks at short $\sim$ 0.45 s (< 5 s), intermediate $\sim$ 4.4 s (2.3 - 8) and long $\sim$ 33 s (> 4 s). The bursts belonging to the three log-normal duration distributions in Fig.2.5 have different spectra and fluence (Mukherjee *et al.* 1998) and can be classified as Class I with long-bright-soft burst, Class II with short-faint-hard bursts, and Class III with intermediate-intermediate-soft bursts.

The distribution drop-off at short durations is partially due to an instrumental bias – the minimum time scale over which the BATSE can trigger on bursts is 64 ms. Some bursts are extremely long and in one case (GRB940217) high energy (up to 20 GeV)

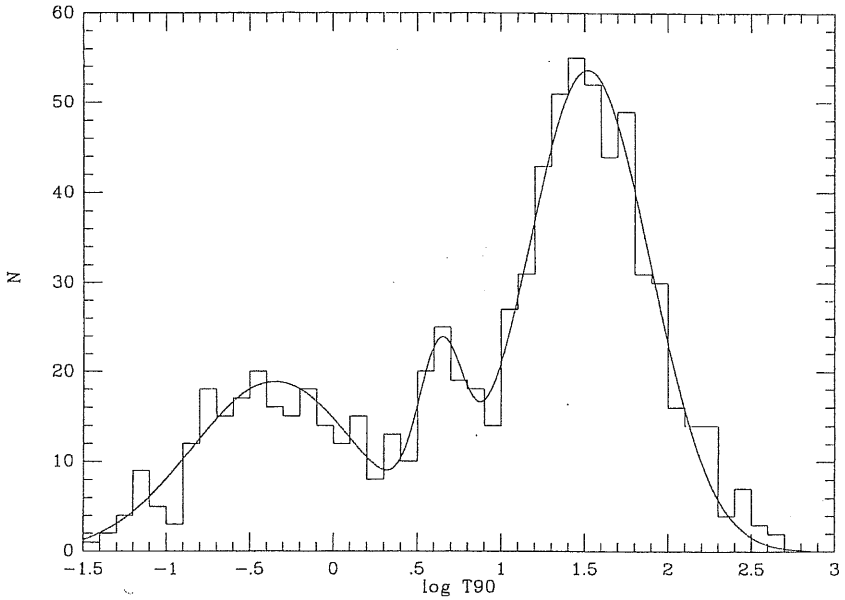

Figure 2.5: Distributions of $\log T_{90}$ for 797 bursts from the 3B catalog. The T_{90} measure represents 90% of the total burst fluence. The solid line represent a fits of three log – normals. (From Horváth 1998)

photons have been observed several hours after the main pulse (Hurley 1994). About 3 percent of the bursts are preceded by a precursor with a lower peak intensity than the main burst (Koshut *et al.* 1995). Several bursts observed by GINGA detector showed significant and apparently thermal X-ray emission (Yoshida *et al.* 1989, Murakami *et al.* 1991). These are probably pre-discovery detections of the X-ray afterglow observed now by BeppoSAX and other X-ray detectors.

In the average peak-aligned time profile of GRBs (Mitrofanov *et al.* 1995), post-peak time history obey an unique and simple analytic law (Stern 1996)

$$I = e^{-(t/t_0)^{1/3}}$$

where t is time measured from the peak of the event and t_0 is a constant ranging from 0.3 s for strong bursts and to $\sim$ 1 s for weak bursts. Gamma-ray burst time histories often consist of many peaks and these peaks tend to be narrower at higher energy. The average pulse width follow a power law

$$w \propto E^{-\alpha}, \quad \alpha \sim 0.4.$$

2.4 Continuum Spectrum

Almost all of the power of γ-ray bursts is emitted above 50 keV with character emission in the few hundred keV ranges with a non-thermal spectrum. A spectrum from the

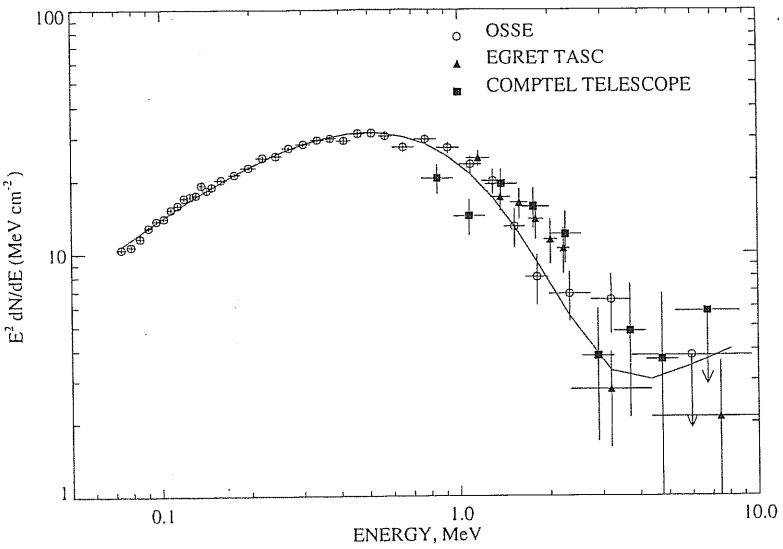

Figure 2.6: The composite time-integrated spectrum for the burst GRB910601 as measured by three experiments on the CGRO, multiplied by the square of the photon energy E to depict the emitted power per logarithmic energy bandwidth. The spectrum shows a low-energy portion which nearly a $\nu^{4/3}$ power law and a peak νF_ν, around 500 keV.

bright burst GRB910601, using data from the four experiments on CGRO, is shown in Figure 2.6 (Schaefer *et al.* 1998). In several cases, the spectrum was observed simultaneously by several instruments. The COMPTEL experiment on the Compton Observatory has observed the peak power from many bursts to be in several hundred keV to MeV energy range. Burst GRB960622, for example, was observed simultaneously by BATSE, COMPTEL and Ulysses. The time integrated spectrum on those detectors, which ranges from 25 keV to 10 MeV agrees well with a Band spectrum with $E_p = 457 \pm 30$ keV, $\alpha = -0.86 \pm 0.15$ and $\beta = -2.5 \pm 0.07$ (Greiner *et al.* 1994). Schaefer *et al.* (1998) present a complete spectrum from 2 keV to 500 MeV for three bright bursts.

Several bursts have been observed downward to 1 keV, but the prompt X-ray emission is very weak – only a few percent of the energy is emitted below 10 keV (Yoshida *et al.* 1989, Murakami *et al.* 1991). Soft X-ray precursor emission was observed in one γ-ray burst out of 17 studied. Extended and precursor emission at energies of about 10 keV in at least several γ-ray bursts (Castro-Tirado *et al.* 1994). An almost prompt emission in X-ray (Feroci *et al.* 1999), optical (Akerlof *et al.* 1999), IR (Bloom *et al.* 1999a), and radio (Frail & Kulkarni 1999) is detected in GRB990123 (Kippen *et al.* 1999; Feroci *et al.* 1999). The total fluence in the soft X-ray band 1.5–26 keV is a few percent of the γ-ray fluence and the initial optical flash contains most the optical fluence with about $\sim 10^{-3}$ of the γ-ray fluence. The optical emission may be from the reverse shock in a fireball (Sari and Piran 1999, 1999a). For other bursts, prompt

emission at lower energies has not been observed.

Most bursts are accompanied by a high energy tail which contains a significant amount of energy - $E^2N(E)$ is almost a constant. EGRET has seen significant flux and power into the GeV energy range (Schneid *et al.* 1992, Hurley 1994). The EGRET spectra are consistent with a single power law spectrum at high energy, when integrated over the burst. Typical spectral indices range between -1.7 and -2.5 for these power-law. Many of these high-energy photons (> 100 MeV) are delayed with respect to the bulk of the lower energy emission (Hurley *et al.* 1994). At the time of the delayed high-energy photons, there was no emission at lower energies. GRB940217, for example, had a high energy tail up to 18 GeV 1.5 hours after the burst (Hurley *et al.* 1994). 57 of the 3rd BATSE GRBs are in the field of Tibet air shower array and several show a excess of 10 TeV gamma-ray events with time-scale of $\sim$ 10 s with a statistical significance of about 6σ (Amenomori *et al.* 1996). the HEGRA AIROBICC Cherenkov array (Padilla *et al.* 1998) detected gamma-ray-like excesses with 5.4σ above 20 TeV and the events are associated with WATCH GRB920925c with a 99.7% confidence level. The TeV emission needs to be confirmed.

All the spectra are no thermal. The spectra have been fitted with broken power laws, log-normal distributions, and power laws joined by exponential spectra. The simplest power law fit to the spectrum gives

$$N(E)dE = AE^{-\alpha}dE \tag{2.1}$$

with a typical power-law index of $\alpha \simeq 1.8 - -2$ in the energy range of 50 – 300 keV. In many cases, a flexible function introduced by Band *et al.* (1993)

$$N(E) = N_0 \begin{cases} (h\nu)^{\alpha} \exp(-\frac{h\nu}{E_0}) & \text{for } h\nu \leq H \quad , \\ [(\alpha-\beta)E_0]^{\alpha-\beta}(h\nu)^{\beta} \times \exp(\beta-\alpha) & h\nu > H \quad , \end{cases} \tag{2.2}$$

where $H \equiv (\alpha - \beta)E_0$, has been used to a comprehensive examination of BATSE spectral data. There is no particular theoretical model that predicts this spectral shape. This function provides an excellent fit to most of the observed spectra over the range $\sim$ 20 keV to several MeV and is sometimes called Band spectrum. When $E_0 = 0$, it becomes a single power law. It is characterized by two power laws joined smoothly at a break energy H. For most observed values of α and β, $\nu F_\nu \propto \nu^2 N(\nu)$ peaks at $E_p = (\alpha + 2)E_0 = [(\alpha + 2)/(\alpha - \beta)]H$. The "typical" energy of the observed radiation is E_p. That is E_p is where the source emits the bulk of its luminosity. The break energy of the spectra, E_0, ranged from less than 100 keV to over 1 MeV. α is about -1 and β is about -2.

Fig. 2.7 shows the distribution of observed values of H in several samples (Band *et al.* 1993; Mallozi *et al.* 1995; Cohen *et al.* 1998). Most of the bursts are in the range 100keV $< H <$ 400keV, with a clear maximum in the distribution around $H \sim$ 200keV.

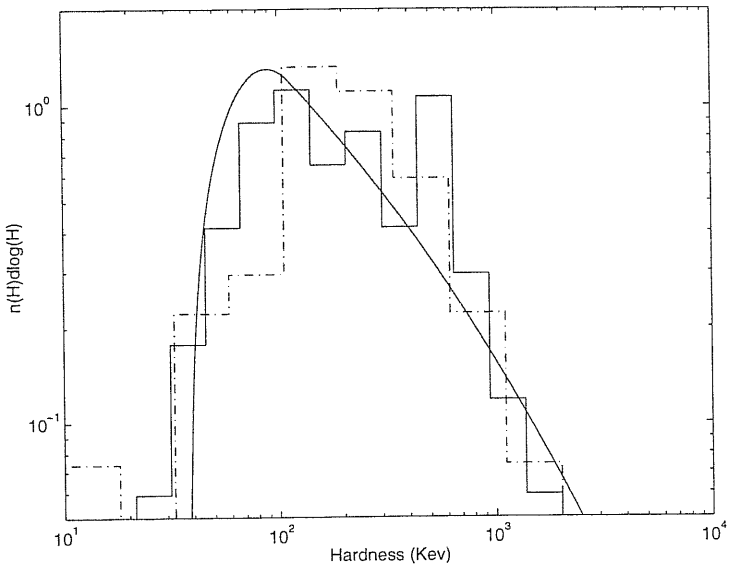

Figure 2.7: The number of bursts $N(H)$ with hardness H in the Band *et al.* (1993) sample (dashed-dotted line) and in the Cohen *et al.* (1998) sample (solid line) together with a theoretical fit of a distribution above $H = 120$ keV with $\gamma \sim -0.5$.

There are not many soft GRBs – that is, GRBs with peak energy in the tens of keV range. This low peak energy cutoff is real as soft burst would have been easily detected by current detectors. However, it is not know whether there is a real paucity in hard GRBs and there is upper cutoff to the GRB hardness or it just happens that the detection is easiest in this few hundred keV band. BATSE triggers, for example, are based on the count rate between 50 keV and 300 keV. BATSE is, therefore, less sensitive to harder burst that emit most of their energy in MeV range. Using BATSE observation alone, one cannot rule out the possibility that there is a population of harder GRBs that emit equal power in total energy which are not observed because of this selection effect. More generally, a harder burst with the same energy as a soft one emits fewer photons. Furthermore, the spectrum is generally flat in the high energy range and it decays quickly at low energies. Therefore, it is intrinsically more difficult to detect a harder burst. A study of the SMM (the Solar Maximum Mission) data (Harris 1997) suggests that there is a deficiency of GRBs with hardness above 300 MeV by at least a factor of 5, relative to GRBs peaking at $\sim 0.5 MeV$, but this data is consistent with a population of hardness that extends up to 2 MeV. The reason is that the bursts in the subclass with significant fluence above 300 keV are about one order of magnitude more luminous than the other subset with a marked lack of fluence above 300 keV (Pendleton *et al.* 1997). The intensity distributions are different for the burst in the different subclasses (Belli 1995; Pendleton *et al.* 1994, 1997). The burst may be

classified into three subclasses according to spectrum hardness, bright, and duration (Horváth 1998; Mukherjee *et al.* 1998).

Overall, the spectrum is non-thermal. This indicates that the source must be optically thin. The spectrum deviates from a black body in both the low and the high energy ends: The X-ray paucity constraint rule out optically thick model in which the γ-rays could be effectively degraded to X-rays (Imamura & Epstein 1987). The high energy tails with energy > 500keV lead to another strong constraint on physical models. These high energy photons escape freely from the source without producing electron positron pairs! As we show later, this provides the first and most important clue on the nature of GRBs.

The low energy part of the spectrum behaves in many cases like a power law: $F_\nu \propto \nu^\alpha$ with $-\frac{1}{2} < \alpha < \frac{1}{3}$ (Katz 1994; Cohen *et al.* 1997). This is consistent with the low energy tail of synchrotron emission from relativistic electrons – a distribution of electrons in which *all* the population, not just the upper tail, is relativistic. This is a direct indication for the existence of relativistic shocks in GRBs. More than 90% of the bright burst studies by Schaefer *et al.* (1998) satisfy this limit. However, there may exist such bursts whose low energy tail is steeper. Such a spectrum cannot be produced by a simple synchrotron emission model and it is not clear how it is produced.

2.5 Spectral Evolution

Observations by various detectors have shown that the spectrum varies during the bursts. Although within most busts there is a hard-to-soft spectral evolution, different trends have been found. Norris *et al.* (1986) investigated ten bursts seen by instruments on the Solar Maximum Mission (SMM) satellites. They found that individual intensity pulses evolve from hard to soft with the hardness peaking earlier than the intensity. This was supported by more recent BATSE data. Liang and Kargatis (1996) note that within most bursts in their sample there existed a hard-to-soft spectral evolution, resulting in that the rate of decrease of the characteristic photon energy is proportional to the instantaneous γ-ray luminosity of the source. The peak energy E_{pk} at which the power output νF_ν is maximum for each pulse decreases linearly with photon fluence (Crider *et al.* 1998). Golenetskii *et al.* (1983) examined two channel data from five bursts observed by the KONUS experiment and found a correlation between the effective temperature and the luminosity, implying that the spectral hardness is related to the luminosity. Similar results were obtained by Mitrofanov *et al.* (1984). Ford *et al.* (1995) who analyzed 37 bright BATSE bursts found that the spectral evolution is a mixture of those found by Golenetskii *et al.* (1983) and by Norris *et al.* (1986): The peak energy either rises with or slightly proceeds the major increases of intensity and softens for the remainder of the pulse. For bursts with multiple peak emission, later

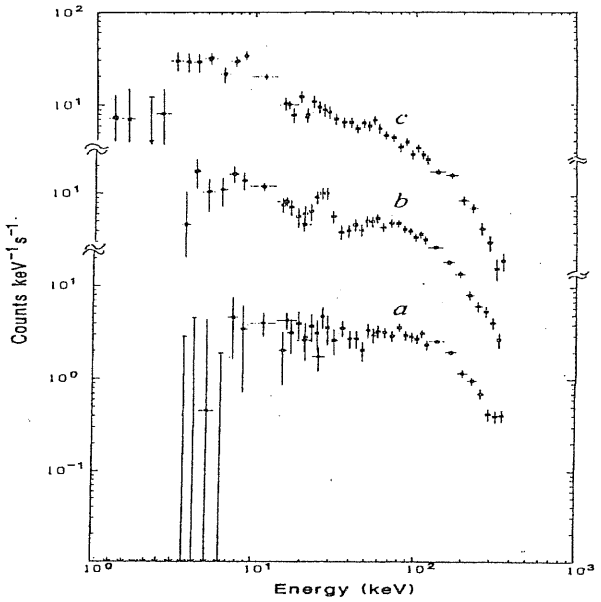

Figure 2.8: The photon energy spectrum of GRB880205 observed by the GINGA experiment (Murakami *et al.* 1988). Two absorption lines, with a ratio of 2.0 in their energies, are clearly seen. The letters a, b, and c indicate different observed time.

spikes tend to be softer than earlier ones.

Related but not similar trend is shown by the observations that the bursts are narrower at higher energies with $T(\nu) \propto \nu^{-0.4}$ (Fenimore *et al.* 1995). This behavior is consistent with synchrotron emission.

2.6 Spectral Lines

The detection of both absorption and emission features have been reported by the GRB detectors prior to the launch BATSE. Broad absorption lines in the 20 – 40 keV range in the spectra of bright GRBs have been observed by several experiments (e.g. Mazets *et al.* 1980; Mazets & Golenetskii 1981; Murakami *et al.* 1988, 1990; Fenimore *et al.* 1988). GINGA has discovered several cases of lines with harmonic structure (Murakami *et al.* 1988; Fenimore *et al.* 1988). Figure 2.8 shows the two absorption lines in GRB880205 detected by GINGA as an example (Murakami *et al.* 1988; Murakami 1990). These lines were interpreted as Landau harmonic cyclotron lines in a magnetic field of $\simeq 10^{12}$ Gauss, a typical value for magnetic field on the surface at the polar region of a neutron star, and provided one of the strongest arguments in favor of the galactic neutron star model. Emission features have been claimed in other bursts and usually occur in the energy range 330 – 850 keV (Mazets *et al.* 1981). Most of them are at 400 – 460 keV, which have been interpreted as red-shifted 511 keV annihilation lines

with a corresponding red-shift of $\simeq$ 20% due to the gravitational field on the surface of the neutron star. These provided additional evidence for the galactic neutron star model.

So far, BATSE has not found any of the spectral features (absorption or emission lines) reported by earlier satellites (Palmer *et al.* 1994; Band *et al.* 1996). Some line candidates were reported recently by the automatic line-search program. However, there are apparent inconsistencies between the data collected from different BATSE spectroscopy detectors (Briggs *et al.* 1999). So, it is necessary to have more careful study to the candidates. This is not a serious discrepancy between GINGA and BATSE results (Palmer *et al.* 1994; Band *et al.* 1996) and, given the rate of observed lines in previous experiments, it is possible that the two sets of data are consistent with each other at a level 20% (Band *et al.* 1997). The discrepancy can also be interpreted as a problem with previous observations or with the difficult analysis of the observed spectra.

Recently, Piro *et al.* (1999) report a possible detection of an iron emission line in the X-ray afterglow spectrum of GRB970508. The line is consistent with an iron K_α line redshifted to the rest frame of the candidate host galaxy of redshift $z = 0.835$ (Metzger *et al.* 1997). The line has flux (equivalent widths) $F_{Fe} = 2.8 \times 10^{-13}\,\mathrm{ergs\,cm^{-2}s^{-1}}$ (EW $\sim$ 1 keV. A line may also be detected in the X-ray afterglow of GRB980828 (Yoshida *et al.* 1999). This GRB has no measured redshift and identification of the feature with the same iron line would imply a redshift $z \sim 0.33$. The line of this burst has flux (equivalent widths) $F_{Fe} = 1.5 \times 10^{-13}\,\mathrm{ergs\,cm^{-2}s^{-1}}$ (EW $\sim$ 3 keV. The dedection of strong iron lines can severely put constraints on the possible central engines (Lazzati *et al.* 1999). The possible detection of emission line in X-ray afterglow of GRBs has been discussed recently (Böttcher *et al.* 1999; Ghisellini *et al.* 1999; Mészáros & Rees 1998a) and in gamma-ray (Mészáros & Rees 1998) based on different central engine models. It seems that only supra-novae model (Vietri & Stella 1998) can easily explain the strong iron line (Lazzati *et al.* 1999)

2.7 Angular Distribution

The most direct evidence of the spatial distribution of the sources of γ-ray bursts comes from their observed angular and intensity distributions. Since the launch of the Compton Observatory, burst locations have been available for a large sample of bursts. One of the most significant findings of BATSE was the observation that the angular distribution of GRBs on the sky is perfectly isotropic. The distribution in galactic coordinates of 2109 bursts observed by BATSE is shown in Figure 2.9. The isotropic distribution (Briggs *et al.* 1996) has severely challenged the hypothesis that the bursts originate within the Galaxy. While observations of GRBs by previous detectors have

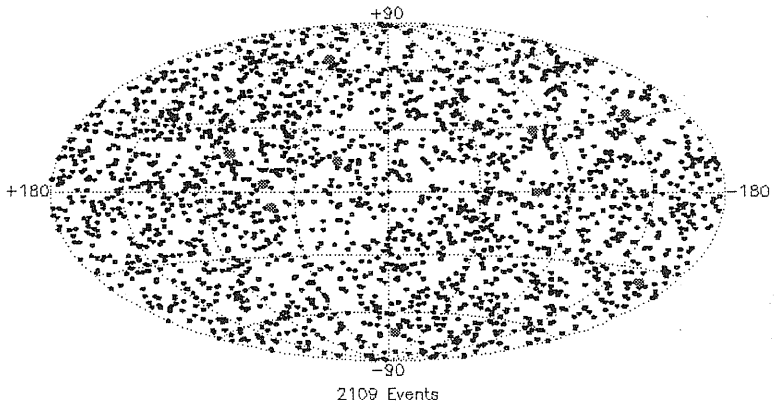

Figure 2.9: The distribution of 2109 GRBs on the sky as seen by BATSE until the end of April 1999. No clustering or anisotropies are seen.

also shown an isotropy, the BATSE results (Meegan *et al.* 1992, 1996; Fishman *et al.* 1994; Briggs *et al.* 1996) show that this isotropy extends to the weakest bursts – those that lie below the break from the $-3/2$ slope in the intensity distribution. In all Galactic disk models, a deviation in slope below $-3/2$ should be expected by a strong concentration of sources toward the Galactic plane (Mao & Paczyński 1992).

There are a number of statistics available to characterize the angular distribution and to search for anisotropy. Two of them have become standard for the comparison of burst distributions to Galactic models (Briggs 1993): in the galactic coordinate, the average value $\langle \cos\theta \rangle$ measuring the concentration toward the Galactic center, and the average value $\langle \sin^2 b - 1/3 \rangle$ measuring the concentration to the Galactic plane or the Galactic pole, and the Bingham statistic B and the Rayleigh-Watson statistic W which, respectively, provide coordinate-system-independent measures of the quadrupole and dipole moments. The statistic $\langle \cos\theta \rangle$ and $\langle \sin^2 b - 1/3 \rangle$, where θ is the angle between vectors from us to the location of a burst and to the Galactic center and b is the galactic latitude, are coordinate-based and most powerful, if a dipole or quadrupole anisotropy originates in the chosen coordinate system. Based on the unit vector of a burst $\vec{r}_i = (x_i, y_i, z_i)$ for $i = 1, N$ with N the total number of GRBs, the Rayleigh-Watson statistic W can be defined as

$$W = \frac{3}{N} \left(\sum_{i=1}^{N} \vec{r}_i \right)^2 \quad , \tag{2.3}$$

and the Bingham statistic B can be defined as

$$B = \frac{15N}{2} \sum_{k=1}^{3} \left(\lambda_k - \frac{1}{3} \right)^2 \quad , \tag{2.4}$$

where λ_k is the kth eigenvalue of the "orientation" matrix

$$M_N = \frac{1}{N}\sum_{i=1}^{N}\begin{bmatrix} x_i x_i & x_i y_i & x_i z_i \\ y_i x_i & y_i y_i & y_i z_i \\ z_i x_i & z_i y_i & z_i z_i \end{bmatrix} \quad . \tag{2.5}$$

For 1122 BATSE bursts the observed dipole and quadrupole (corrected to BATSE sky exposure) relative to the galaxy are: $\langle \cos\theta \rangle = -0.002 \pm 0.017$ and $\langle \sin^2 b - 1/3 \rangle = -0.003 \pm 0.009$. This mean galactic dipole and quadrupole moments deviate from the expected values -0.013 and -0.005 for perfect isotropy by 0.6σ and 0.3σ, respectively (Meegan *et al.* 1996). The coordinate independent Watson dipole $W = 0.5 \pm 2.4$ and Bingham quadrupole $B = 5.8 \pm 3.2$ is different from complete isotropy both by 1.2σ. Recent released 4th BATSE burst catalog (Paciesas *et al.* 1998; Kippen *et al.* 1998a) sources gives similar galactic dipole and quadrupole moments.

Quashnock & Lamb (1993) noted that the medium intensity bursts in the *1st BATSE Catalog* exhibited unexpectedly large dipole and quadrupole moments. The effect is not present in subsequent data (Briggs *et al.* 1996). Other proposed subclasses according to the duration (Kouvelioutou *et al.* 1993; Belli 1995; Dezalay *et al.* 1996) or to the fluence (Pendleton *et al.* 1997; Tavani *et al.* 1998) also distribute isotropically.

Balázs *et al.* (1998) reported a quadrupole anisotropy in the distribution of short GRBs. However, this anisotropy is not associated with the Galaxy. The short burst are still cosmological and the anisotropy needs careful explanation.

2.8 Burst Repetition and Soft Gamma-ray Repeaters

The question of whether GRB sources repeat on times scales of a year or less has been under much debate since the public release of the first BATSE catalog (Fishman *et al.* 1994). Information concerning such repetition would enormously constrain most GRB models. In particular, models based on "once in lifetime" catastrophic event would be ruled out by any observation of repetition, as they cannot plausibly repeat on such time-scales. Models of burst sources located in a galactic halo, on the other hand, are expected to repeat on such time scales.

Quashnock and Lamb (1993) suggested that there is evidence from the data in BATSE 1B catalog for repetition of bursts from the same source and for a possible excess of pairs of GRBs clustered in both time and space (Wang & Lingenfelter 1995). Some studies of data in 2B catalog provided marginal evidence for both temporal and angular clustering (Petrosian & Efron 1995). An analysis of a larger independent sample satisfying the same selection criteria as Quashnock and Lamb failed to confirm the result (Briggs *et al.* 1996). No evidence for repetition has been found from 2B catalog (Brainerd *et al.* 1995), from the revised 1B catalog, or from 3B catalog of 1122

bursts (Meegan *et al.* 1996; Bennett & Rhie 1996, Tegmark *et al.* 1996). Strohmayer *et al.* (1994) show that an upper limit 20% can be imposed to the fraction of the repetition sources in 1B GRB catalog. A similar upper limit can be derived for the WATCH GRB sample (Gorosabel *et al.* 1998). Recent studies confirm the lack of repetition in 4B catalogue and lead to an upper limit to the repetition rate of 0.04 burst source^{-1} yr^{-1} (Hakkila *et al.* 1997).

A unique group of four burst - "the gang of four" - emerge from the same position on the sky within two days (Meegan *et al.* 1996a). The third burst GRB961029a of them are extremely strong and is one of the strongest observed by BATSE so far. Consequently it was observed by the IPN network as well, and its position is known accurately. The other three (GRB961027a, GRB961027b, and GRB962029d) were detected only by BATSE. The precise position of one of the bursts is within 1σ circles of the three other bursts. However, two of the bursts are almost 3σ away from each other. Is this a clear cut case of repetition? It is difficult to assign a unique statistical significance to this question as the significance depends critically on a priori hypothesis that one tests. Furthermore, the time difference between the first and the last burst is less than two days. This is only one order of magnitude longer than the longest burst observed beforehand. It might still be possible that all those bursts came from the same source and that they should be considered as one long burst.

Among more than a thousand GRBs, there is a unique group of four, possible five, bursts including the famous 1989 March 5th event (SGR0526-66) and the most intense gamma-ray burst SGR1900+14 (Hurley *et al.* 1999) ever detected, that are different from all classical bursts. These bursts, referred to as soft Gamma-ray repeaters (SGRs), are characterized by their repetition, short duration (generally < 1 s), simple time profiles, and very soft, time-independent energy spectra. The four SGRs have been identified to associate with neutron stars in young supernova remnants. It is generally believed that SGRs are different from classical GRBs (Atteia *et al.* 1987; Kouveliotou *et al.* 1987).

We will not be discussing SGR in this thesis. Reviews of the observations (Hurley 1996) and of theoretical models (Thompson 1996) for SGRs have been given recently.

2.9 Time Dilation

Piran (1992) and Paczyński (1992) suggested that the light curves of the GRBs should be stretched due to the cosmological time dilation. This effect has been studied through the investigation of the correlation of the peak alignment duration of bursts (T_{90}) and the burst brightness (peak flux) (Norris *et al.* 1994, 1995; Nemiroff *et al.* 1994; Mallozzi *et al.* 1995) or time intervals between peaks and the burst brightness (peak flux) (Deng & Schaefer 1998). Due to the extreme complexity of the γ-ray burst time structures and

the wide range of their durations, any dilation effect can only be tested in a statistical sense. The subtleties of data analysis play a role, leading to different groups getting different results using identical method and similar data sets.

Norris *et al.* (1994, 1995) found that the dimmest bursts are longer by a factor of 2.3 compared to the bright ones. This anti-correlation between the pulse width and their intensity is compatible with the finding from the count distribution that $z_{max} \simeq 2$ and $z_{min} \simeq 0.2$ since $(1+z_{max})/(1+z_{min}) \simeq 2.5 \pm 0.8$. The results have been challenged in the scientific literatures. Motrofanov *et al.* (1994, 1996) do not detected the dilation effect in an alternative analysis. Other effects could also be responsible for the observed dilation, such as an intrinsic luminosity-duration correlation (Brainerd 1994, 1997) between pulse width and burst brightness for bursts drawn from a volume limited sample, or an evolutionary effect due to the correlation between duration and frequency: duration is narrower at higher frequency. It has been argued that the observed correlation could arise either from the beaming effect of relativistic jets (Brainerd 1994), or the correlation of total energy in mono-energetic sources (Wijers & Paczyński 1994). Fenimore and Bloom (1995) showed that if the observed time-stretching is caused by time dilation, then the redshift of the dimmest bursts may be as large as $z \sim 6$. According to Stern, Poutanen, and Svensson (1997), some or all of the time-stretching may be intrinsic to the bursts.

2.10 C_{min}/C_{max} and Peak Flux Distribution

In addition to angular isotropic distribution of GRBs on the sky (cf. Section 2.7), the lack of faint GRBs is another significant finding of BATSE. Several measures of burst intensity can be used to derived burst distributions. It has been customary to calculate the value of the $\langle V/V_{max} \rangle$ statistics for the detected bursts. For photon counting experiments like BATSE, the value is not strictly $\langle V/V_{max} \rangle$ but rather $\langle (C_{min}/C_{max})^{3/2} \rangle$, where C_{min} is the threshold count rate and C_{max} is the maximum count rate measured during the burst (for the reasons, see the discussion in Section 3.3.3). Some authors directly interpreted $\langle (C_{min}/C_{max})^{3/2} \rangle$ as $\langle V/V_{max} \rangle$. Among all the measures, peak flux (erg cm^{-2} s^{-1}) is preferable, because it can be more directly related to the trigger criteria typically employed by γ-ray burst detectors and allows comparisons among experiments and comparisons with theoretical models.

The limiting fluence observed by BATSE is $\sim 10^{-7}$ ergs/cm^2. The actual fluence of the strongest bursts is larger by two or three orders of magnitude. If the sources are distributed homogeneously in Euclidean space, $\langle (C_{min}/C_{max})^{3/2} \rangle$ will have an average value 1/2. A sample of 657 bursts (Meegan *et al.* 1996) has $\langle (C_{min}/C_{max})^{3/2} \rangle = 0.33 \pm 0.01$, which is 17σ away from the homogeneous flat space value of 0.5. The $\langle V/V_{max} \rangle$ of 4th BATSE GRBs indicated that GRB source density evolution is weaker

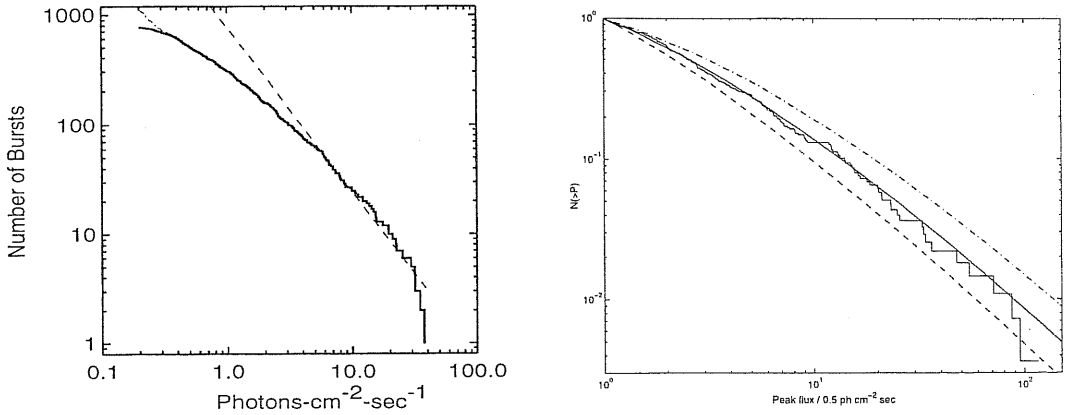

Figure 2.10: (*left panel*) BATSE integral brightness distributions $\log N(>P) - \log P$ in the energy range of 50 – 300 keV and for the 1024 ms triggering time-scale, contain 772 bursts (Meegan *et al.* 1996). The solid curve for uncorrected data, and the dot-dash curve for the data corrected for efficiency, which is over-corrected at low fluxes. Dashed line for the expected $-3/2$ slope for homogeneous sources. (*right panel*) The observed long burst peak flux distribution and three theoretical cosmological distributions with $\Omega = 1$, $\Lambda = 0$, $\alpha = -3/2$, standard candles and no source evolution. Solid line: best fit, dashed line: lower 1% bound, and dashed-dotted line: upper 1% bound.

than expected for star formation (Che *et al.* 1998). This requires a signification modification of star formation rate estimates or a proper GRB luminosity evolution, signification beaming, signification error in standard candle assumptions.

This deviation is very clear in the total $\log N(f_p > P) - \log P$ distribution in Figure 2.10, which is taken from Meegan *et al.* (1996). The integral peak flux distribution of 796 BATSE γ-ray bursts shown in Figure 2.10 is incompatible with a homogeneous population of sources in Euclidean space. The deviation from the form expected $-3/2$ power law is significant and shows that there is a paucity of faint bursts. This distribution is confirmed by larger samples (Kommers *et al.* 1998; Tavani 1998). It is compatible with a cosmological distribution (see Fig. 2.10, but it is also consistent with an Extended Galactic Halo distribution.

It was found that bursts can be divided into two subsets according to their spectral properties (Belli 1995; Kouvelioutou *et al.* 1996; Pendleton *et al.* 1997). Pendleton *et al.* (1997) showed that the subset of GRBs exhibiting a marked lack of fluence above 300 keV (the no-high-energy bursts (NHE)) are about one order of magnitude less luminous than bursts with significant fluence above 300 keV (heigh-energy bursts (HE)). The two subsets have different intensity distribution: the bursts lacking emission above 300 keV appear effectively homogeneous in the intensity distribution and the other bursts have non-homogeneous intensity distribution (Pendleton *et al.* 1997). The simplest explanation for the homogeneity of the NHE burst subclass is that these burst are so intrinsically faint compared with the HE burst that we can see only the nearby

ones that are distributed homogeneously. Tavani (1998) classified the GRBs in 4th BATSE catalog into three subsets: short-hard bursts, long-hard bursts, and long-soft bursts and found that long-hard bursts show the strongest deviation from the three-dimensional Euclidean brightness distribution and the other two sets show little, if any, deviations. The long-hard bursts have very hard spectra of typical photon energy above 1 MeV. Tavani (1998) suggested that BATSE selection effects and/or intrinsic burst properties may cause the strong deviation for long/hard bursts and coexistence of multiple GRB populations remains possible. Belli (1997) also found homogeneous intensity distribution in two subclasses of GRBs.

2.11 Quiescent Counterpart and Afterglow

One of the main obstacles in resolving the GRB mystery was the lack of identified counterparts in other wavelengths. This has motivated numerous attempts to discover counterparts to γ-ray bursts in any other wavelength region before, during, or after the γ-ray event, ever since the initial discovery of γ-ray bursts. This is a difficult task – it was not know what to expect and where and when to look for it.

The searches for counterparts have taken many forms, including searches for statistical associations of known objects with bursts (e.g Quashnock 1996) and can basically be divided into efforts to find a flaring, a fading or a quiescent counterpart. Searches have been made of archival plates and other data bases for transient or unusual objects within the error boxes of well-determined burst locations. Neither a flaring nor a quiescent counterpart to a gamma-ray burst has yet been convincingly identified at any wavelength region (McNamara *et al.* 1996).

However, fading counterparts – afterglow – have been recently discovered by BeppoSAX and as expected this discovery has revolutionized GRB studies. This allowed also the discovery of host galaxies in several cases, which is discussed in Section 2.12. Simultaneously soft X-ray flaring was discovered in several bursts but it is an ambiguous question whether this should be considered as a part of the GRB itself or as it is a separate component.

2.11.1 X-ray Afterglow

GRB observations were revolutionized on February 28, 1997 by the Italian-Dutch satellite BeppoSAX that discovered a X-ray counterpart to GRB970228 (Costa *et al.* 1997). GRB970228 was a double peak GRB. The first peak which lasted $\sim$ 15 sec was hard. It was followed, 40 seconds later, by a much softer second peak, which lasted some $\sim$ 40 sec. The burst was detected by the Gamma-ray Burst Monitor (GRBM) as well as the Wide Field Camera (WFC). The WFC of a $40° \times 40°$ field view detected soft X-rays

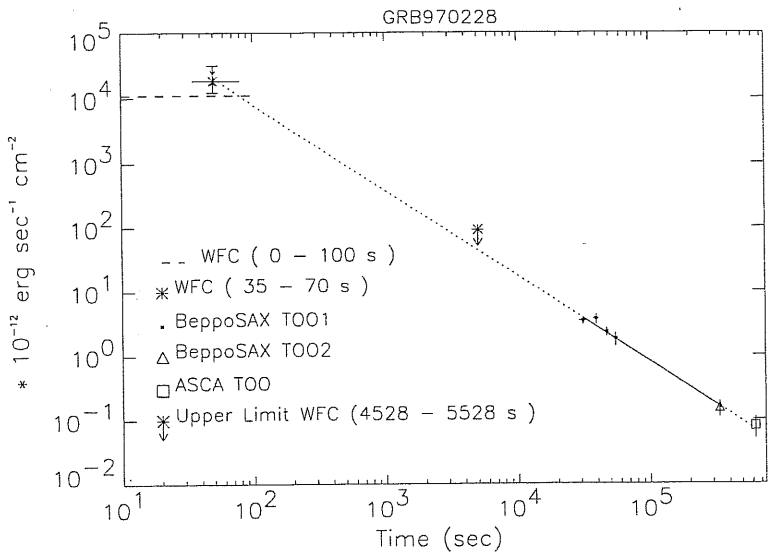

Figure 2.11: Decay of the X-ray afterglow of GRB970228 in the energy range 2 - 10 keV. The data is fitted with a power law $t^{-1.33}$, for which the zero time is taken at the GRB monitor trigger time. From Costa *et al.* (1997).

simultaneously with both peaks. Eight hours later, the Narrow Field Instrument (NFI) was pointed towards the burst directions and detected a continuous X-ray emission.

Following X-ray detections by BeppoSAX (Costa *et al.* 1997; Frontera *et al.* 1997), ROSAT and ASCA revealed a decaying X-ray flux $\propto t^{-1.33\pm0.11}$ (see Fig 2.11). The decaying flux can be extrapolated as a power law directly to X-ray flux of the second peak (even though this extrapolation requires some care in determining when is $t = 0$).

GRB970508 was detected by both BATSE in γ-rays and BeppoSAX in X-rays on 8 May 1997. The γ-ray burst lasted for $\sim$ 15 sec, with a γ-ray fluence of $\sim 3 \times 10^{-6}$ergs/cm^{-2}. Variable emission in X-rays, optical and radio (Frail *et al.* 1997; Taylor *et al.* 1997) followed the γ-rays.

GRB970828 was a strong GRB that was detected by BATSE on August 28, 1997. Shortly afterwards, RXTE focused on the approximate BATSE position and discovered X-ray emission. This X-ray emission determined the position of the burst to within an elliptical error box with $5' \times 2'$. However, in spite of enormous efforts no variable optical counterpart brighter than $R = 23.8$ mag was detected (Groot *et al.* 1998). There was also no indication of any radio emission. Similarly, X-ray afterglow was detected from several other GRBs (GRB970615, GRB970402, GRB970815, GRB980519) with no optical or radio emission.

GRB990123 was a very strong burst detected both by BATSE and BeppoSAX on 1999 January 23.508. It consisted of multi-peaked structure lasting more than 100 seconds. Its total fluence ($> 20keV$) of 3×10^{-4}erg/cm^2 (Kippen *et al.* 1999, Feroci, *et al.* 1999) places it at the top 0.3% of BATSE bursts. It has a multi-wavelength

afterglow ranging from X-ray via optical and IR to radio. The prompt total fluence in the soft X-ray band 1.5 – 26 keV is 7×10^{-6}ergs/cm^2. The NFI observations (Piro *et al.* 1999), beginning approximately six hours after the burst and lasting for 26 hours, correspond to a power law decay with a slope of -1.35 from the prompt observation (about 60 seconds after the burst). The decay from the NFI to ASCA observations obtained after about 2 days and 7 hours after the burst is slightly slower with a slope of ~ -1.1.

Seventeen GRBs have been detected with arc-minute positions by July 22, 1998: fourteen by the WFC of BeppoSAX and three by the All-Sky Monitor (ASM) on board the Rossi X-ray Timing Explorer (RXTE). Of these seventeen bursts, thirteen were followed up within a day in X-rays and all those resulted in good candidates for X-ray afterglow. A short summary of some of their properties are given in Table 2.1.

2.11.2 Optical and Infrared Afterglow

The accurate position of GRB970228 determined by BeppoSAX enabled the identification of an optical afterglow (van Paradijs *et al.* 1997) - a 20 magnitude point source adjacent to a red nebulae. HST (Sahu *et al.* 1997a) observations revealed that the nebulae adjacent to the source is roughly circular with a diameter of $0.8''$. The diameter of the nebula is comparable to the one of galaxies of similar magnitude found in the Hubble Deep Field, especially if one takes into account a possible visual extinction in the direction of GRB970228 of at least one magnitude (Lamb 1997).

The optical emission also depict a decaying flux (Galama *et al.* 1997) (see Fig 2.12). The sources could not be observed from late March 97 until early September 1997. When it was observed again on Sept. 4th by HST (Fruchter *et al.* 1998), it was found that the optical nebula does not decay and the point source shows no proper motion, refuting earlier suggestions. The visual magnitude of the nebula on Sept. 4th was 25.7 ± 0.25 compared with $V = 25.6 \pm 0.25$ on March 26th and April 7th. The visual magnitude of the point source on Sept. 4th was $V = 28.0 \pm 0.25$, which is consistent with a decay of flux as $t^{1.14 \pm 0.05}$ (Fruchter *et al.* 1998). In spite of the extensive efforts, no radio emission was detected and one can set an upper limit of$\sim 10\mu\ Jy$ to the radio emission at 8.6 GHz.

The spectrum of the optical transient of GRB970508 taken by Keck revealed a set of absorption lines with a redshift $z = 0.835$ (Metzger *et al.* 1997). A second absorption line system with an underlying, dim galaxy host. HST images (Pian *et al.* 1998a; Natarajan *et al.* 1997) and Keck observations (Bloom *et al.* 1998b) show that this host is a very faint ($R = 25.72 \pm 0.2$ mag) and compact (≤ 1 sec) dwarf galaxy at $z = 0.835$ and nearly coincident on the sky with the transient.

Optical light curve peaks at around 2 days after the burst. Assuming isotropic emis-

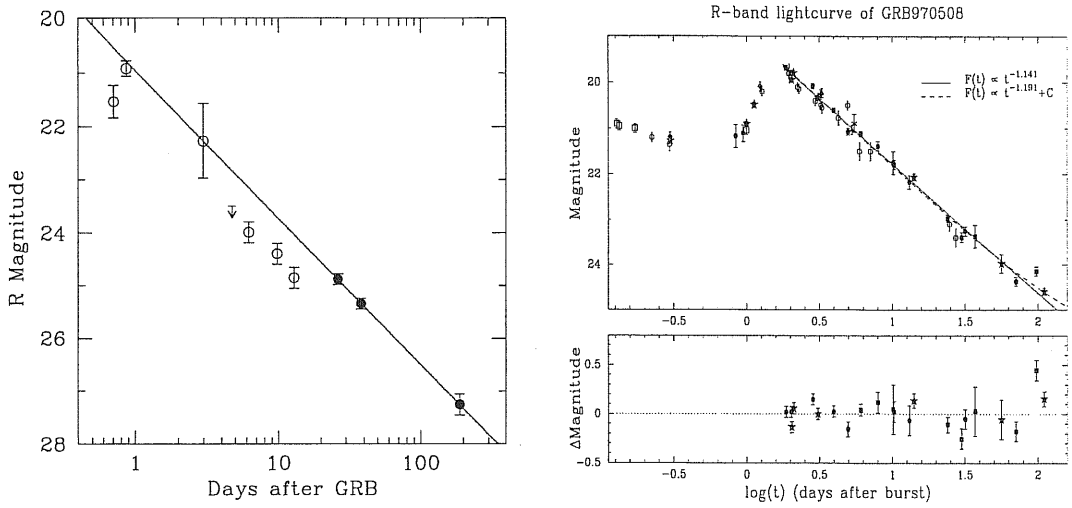

Figure 2.12: Decay of the optical (R-band) afterglow of GRB970228 (left panel), from Fruchter *et al.* (1998) and GRB970508 (right upper panel), from Galama *et al.* (1998). The decay for both objects is power-law. The best fitted power-law function is $f(t) \propto t^{-1.14\pm0.05}$ for GRB970228. The constant C corresponds to a magnitude $R_c = 26.09 \pm 0.36$. Lower right panel is the residual of the data after subtraction of the power-law $f(t) \propto t^{-1.141}$.

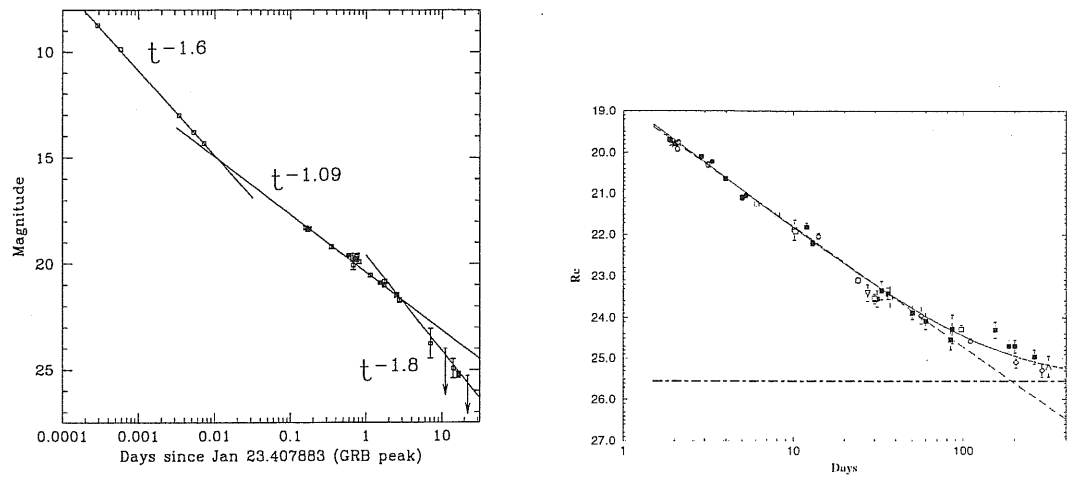

Figure 2.13: Optical light curves of the afterglow of GRB990123 (left panel) after the subtraction of host galaxy background, from Fruchter *et al.* (1999), and GRB970508 (right panel), from Zharikov *et al.* (1998). For GRB980508, The power-law fitted is shown as long-dashed line, while the solid line is a fit by a power law plus a constant host $R_c = 25.55 \pm 0.19$ mag (dot-dashed).

sion and using $z = 0.835$, this peak flux corresponds to a luminosity of $\simeq 10^{46}$ergs/sec. The flux decline shows a continuous power law decay $\propto t^{-1.27\pm0.02}$ (Bloom *et al.* 1998b; Galama *et al.* 1998; Castro-Tirado *et al.* 1998; Pedersen *et al.* 1998). After about 100 days, the light curve began to flatten as the transient faded and become weaker than the host (Pederson *et al.* 1998; Sokolov *et al.* 1998). Integration of the light curve results in an overall emission of $\simeq 10^{51}$ergs in the optical band.

GRB990123 gave us another surprising. An 11.82 magnitude optical flash was detected by ROTSE on the first 5 seconds exposure, 22.18 seconds after the onset of the burst (Kulkarni *et al.* 1999). This was the first observation ever of a prompt optical counterpart of a GRB. Another 5 seconds exposure, 25 second later, revealed a 8.95 magnitude signal ($\sim$ 1 Jy). The optical signal decayed down to 14.53 magnitude in subsequent several 75 second exposures that took place up to 10 minutes after the burst. The five last exposures depict a power law decay with a slope of -1.6 for $t < 0.1$ days in Fig. 2.13, which is taken from Fruchter *et al.* (1999). It is important to note that the measured power-law index at early times depends strongly on the assumed time origin. Figure 2.13 (Fruchter *et al.* 1999) took the time of peak hard X-ray and gamma-ray emission as the time origin, which was independent of energy and therefore of detectors. Sari & Piran (1999) chooses the BATSE trigger time as the zero time and obtained a slope ~ -2.0. The initial optical flash, which probably arise from a reverse shock in a fireball (Sari & Piran 1999; Mészáros & Rees 1999) or from internal shocks as GRB itself (Mészáros & Rees 1999; Liang *et al.* 1999), contains most of the optical fluence: $\sim 2.5 \times 10^{-7}$ergs/cm^2, about 7.7×10^{-4} of the γ-ray fluence.

Optical afterglow follow up by larger telescopes began some 3 hours and 46 minutes later with the observations on Palomar (Odewahn *et al.* 1999), which revealed an 18.2 R magnitude source. The observations in Fig. 2.13 (Kulkarni *et al.* 1999; Fruchter *et al.* 1999, and references therein; Odewahn *et al.* 1999, Gal *et al.* 1999, Bloom *et al.* 1999, Zhu & Zhang 1999, Sokolov *et al.* 1999, Ofek & Leibowitz 1999, Masetti *et al.* 1999, Garnavich *et al.* 1999, Sagar *et al.* 1999, Yadigaroglu *et al.* 1999) showed a power law decay with a slope of ~ -1.1 for $0.1 < t < 2$ days, which is significantly different from the initial slope and can be explained with a standard external shock model (Sari & Piran 1999; Mészáros & Rees 1999). There probably exists a third power-law segment in the light curve (for $t > 2$ days), which has a power-law index ~ -1.8 in Fig. 2.13. However, Galama *et al.* (1999) fit very well the third segment together with the second segment with a single power law. A possibility for the late break in the optical transient light curve to a steeper power law is the evolution from a strongly beamed flow to an isotropic flow (Rhoads 1997; Kulkarni *et al.* 1999). The optical spectra revealed several absorption line system, showing that the redshift of GRB990123 is ≥ 1.6004 (Kelson *et al.* 1999; Hjorth *et al.* 1999, 1999a; Kulkarni *et al.* 1999). HST observations (Bloom *et al.* 1999b; Fruchter *et al.* 1999) revealed that the

host with magnitude $V \simeq 24.20 \pm 0.15$ is offset by 0.6 arc-sec from the optical transient.

In near IR, a counterpart with a $K = 18.3 \pm 0.03$ magnitude was detected on January 24.6356 (Bloom *et al.* GCN240). The decay in IR is also a power-law with a slope of -1.14 ± 0.08. A radio source at 8.46 GHz was detected on Jan. 24.65 by the VLA (Frail & Kulkarni GCN211), but not on the earlier or later observations.

Not all the bursts given in Table 2.1 have detected optical afterglow. Even though, we will discuss not all of them but only mention GRB971214, GRB980425 and GRB980703.

GRB971214 was a rather strong burst. It was detected on 14.9 December 1997 (Heise *et al.* 1997). Its optical counterpart was observed with a magnitude 21.2 ± 0.3 mag on I band by Halpern *et al.* (1997) twelve hours after the burst. It was observed one day later on Dec. 16.47 with I magnitude 22.6 mag. Kulkarni *et al.* (1998) obtained a spectrum of the host galaxy for GRB971214 and found a redshift of $z = 3.418$. With a total fluence of $1.09 \times 10^{-5}\mathrm{ergscm}^{-2}$ (Meegan *et al.* 1998), this large redshift implies, for isotropic emission, an energy release of $\sim 10^{53}$ ergs in γ-ray alone [1] (Kulkarni *et al.* 1998).

GRB980425 detected by BeppoSAX as well as BATSE was a moderately weak burst with a peak flux of $(3 \pm 0.3) \times 10^{-7}\mathrm{ergscm}^{-2}\mathrm{sec}^{-1}$. It was a single peak burst with a rise time of 5 seconds and a decay time of about 25 seconds. The WFC of BeppoSAX obtained a position with error box 8′. Inspection of an image of this error box taken by the New Technology Telescope (NTT) revealed a type Ic supernova SN1998bw that took place more or less at the same time as the GRB (Galama *et al.* 1998). Since the probability for a chance of the SN and the GRB is only 1.1×10^{-4}, it is likely that this association is real. The host galaxy of this supernova (ESO 184-G82) has a redshift of $z = 0.0085 \pm 0.0002$ putting it at a distance of 39 ± 1 Mpc. The corresponding γ-ray energy is 5×10^{47} ergs. With such a low luminosity, it is inevitable that if the association of this burst with the supernova is real it must correspond to a new and rare subgroup of GRBs.

GRB980703 was a very strong burst with an observed γ-ray fluence of $(4.59 \pm 0.42) \times 10^{-5}\mathrm{ergscm}^{-2}$ (Kippen *et al.* 1998b). Keck observations revealed that the host galaxy has a redshift of $z = 0.966$. The corresponding energy released (for isotropic emission) is $\sim 10^{53}$ergs (Djorgovski *et al.* 1998).

2.11.3 Radio Afterglow

Radio emission was observed, first, one week after the burst GRB970508 (Frail *et al.* 1997) (see Fig. 2.14). This emission showed intensive oscillations which were interpreted as scintillation (Goodman 1997). The subsequent disappearance of these oscillations after about three weeks enabled Frail *et al.* (Frail *et al.* 1997) to estimate

[1] This value depends also on the spectral shape of the burst

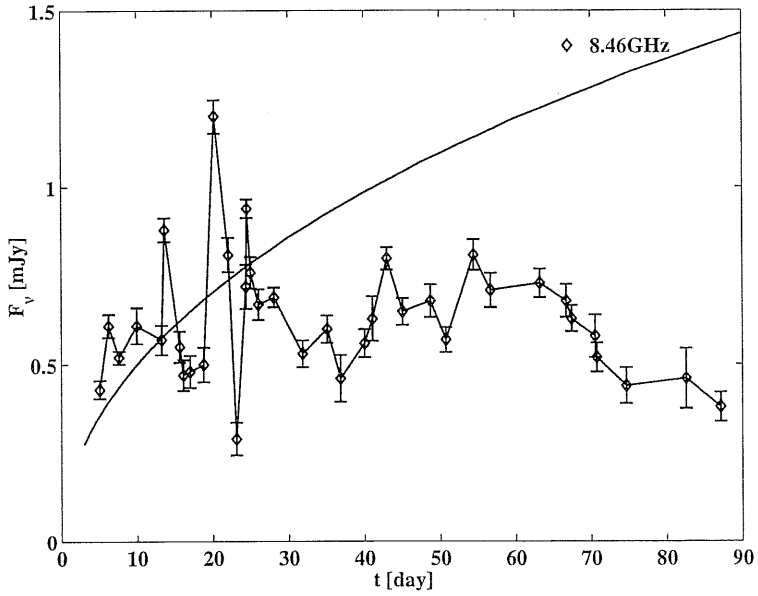

Figure 2.14: Light curve of the radio afterglow of GRB970508, from Frail *et al.* (1997)

the size of the fireball at this stage to be $\sim 10^{17}$ cm. This was supported by the indication that radio emission was initially optically thick (Frail *et al.* 1997), which yields a similar estimate to the size (Katz & Piran 1997). The observational results of radio afterglow of other bursts are summarized in Table 2.1

2.12 Host Galaxies and the "No Host" Problem

Most models for γ-ray bursts at cosmological distances suggest that GRBs are in a host galaxy. If so, deep searches within the small error boxes of some GRBs localized by the IPN system should reveal the host galaxy. Until the discovery of GRB afterglow, have these searches yielded only upper limits on the magnitudes of possible hosts. This has led to what was so-called the "No Host" Problem. This was often used to argue against the cosmological interpretation of cosmic gamma-ray bursts. The fundamental problem with the host galaxy tests is that our understanding of GRBs is too limited for constructing a sensitive measure of the presence of suitable hosts. Schaefer *et al.* conducted searches in near and far infrared using IRAS (Schaefer *et al.* 1987), in radio using VLA (Schaefer *et al.* 1989) and in archival optical photographs (Schaefer 1990) and have found only upper limits and no clear counterpart candidates. Similar results from multi wavelength observations have been obtained by Hurley *et al.* (1994). Vrba, Hartmann & Jennings (1995) have monitored the error boxes of seven bursts for five year. They did not find any unusual objects. As for the "No Host" problem, they as well as Luginbull *et al.* (1995) and Larson, McLean & Becklin (1996) concluded,

using the standard galaxy luminosity function, that there are enough dim galaxies in the corresponding GRB error boxes which could be the hosts of cosmological burst and therefore, there is no "No Host Galaxy" problem.

More recently, Larson & McLean (1997) monitored in the infrared nine of the smallest error boxes of burst localized by the IPN with a typical error boxes of eight arc-min^2. They found in all error boxes at least one bright galaxy with K $\leq$ 15.5. However, the error boxes are too large to discern between the host galaxy and unrelated background galaxies. Schaefer *et al.* (1997) searched the error boxes of five GRBs using the HST. Four of these are smaller boxes with a size of $\sim$ 1 arc-min^2. They searched but did not find any unusual objects with UV excess, variability, parallax or proper motion. They have found only faint galaxies. For the four small error boxes, the luminosity upper limits of host galaxy are 10 – 100 times smaller than the luminosity of L_* galaxy. Band & Hartmann (1998) concluded that the error boxes of Larsen & McLean (1997) are too large to discriminate between the presence or the absence of host galaxies. However, they find that the absence of host galaxies in the Schaefer *et al.* (1997) data is significant, at 2×10^{-6} level. Suggesting that there are no bright hosts.

This situation has drastically changed and the "no host" problem disappeared with the observations of afterglow. These observations have allowed for an accurate position determination and to identification of host galaxies for almost all GRBs associated with detected optical transients (OTs) to within a fraction of an arcsecond after the OT has faded. Most of these host galaxies are dim with magnitude $23.6 < R < 25.8$. This support the conclusion of the earlier studies that GRBs are not associated with bright galaxies and definitely not with cores of such galaxies and rules some models, e.g. nuclear black hole (AGN) type models. These observations are consistent with GRB rate being either a constant or being proportional to the star formation rate (Hogg & Fruchter 1998). According to this analysis, it is not surprising that most host are detected at R $\sim$ 25. However, though these two models are consistent with current data, both predicted the existence of host galaxies brighter than 24 mag, which were not observed so far. This may not be a problem but rather an indication on the nature of the sources.

The four GRBs with measured cosmological redshifts lie in host galaxies with a strong evidence for star formation. The GRBs whose redshifts are obtained from host galaxies are GRB970508 of $z = 0.835$ (Metzger *et al.* 1997; Bloom *et al.* 1998b), GRB971214 of $z = 3.428$ (Kulkarni *et al.* 1998), GRB980703 of $z = 0.966$ (Djorgovski *et al.* 1998), and GRB980613 of $z = 1.0964$ (Djorgovski *et al.* 1999). These galaxies display prominent emission lines associated with star-formation. In all four cases, the strength of those lines is high for galaxies of comparable magnitude and redshift (Hogg & Fruchter 1998; Djorgovski *et al.* 1998; Bloom *et al.* 1998b; Fruchter *et al.* 1998; Bloom *et al.* 1999b). The host of GRB980703, for example, show a star forming rate

of $\sim 10 M_{\odot}\mathrm{yr}^{-1}$ or higher with a lower limit of $\sim 7 M_{\odot}\mathrm{yr}^{-1}$ (Djorgovski *et al.* 1998). For most GRBs with afterglow, the host galaxy was detected but no emission line or absorption lines were found and no redshift was measured. This result is consistent with the hypothesis that all GRBs are associated with star-forming galaxies. For those hosts that are at redshift $1.3 < z < 2.5$, the corresponding emission lines are not observed as for this redshift range no strong lines are found in the optical spectroscopic window (Fruchter *et al.* 1998). At least five GRBs now have offsets between the centroid of the visible host and the optical transient. For example, GRB990123 is offset by 0.6 arcsec from the presumed host galaxy and appears to be coincident with a star-formation region (Bloom *et al.* 1999b; Fruchter *et al.* 1999).

The simplest conclusion of the above observations is that all GRBs are associated with star formation regions. However, one has to keep in mind that those GRBs on which this conclusion was based had a strong optical afterglow, which not all GRBs show. It is possible that the conditions associated star formation regions (such as high interstellar matter density - or the existence of molecular clouds) are essential for the appearance of strong optical afterglow and not for the appearance of the GRB itself.

2.13 Beaming

The observations of afterglow in X-ray, optical and radio wave-bands have confirmed the fireball and its relativistic expansion with a Lorentz factor $10^2 \leq \Gamma \leq 10^3$. A key question is now whether the required relativistic expansion is quasi-isotropic or beamed, mostly into an opening angle $\sim 1/\Gamma$. The prompt GRB could appear similar in either case (Mészáros *et al.* 1998b), although beam-edge effects could also be observable (Panaitescu *et al.* 1998), but the source energetics and required numbers would be modified by factors of Γ^{-2} and Γ^2, respectively.

Whereas the prompt GRB emission is probably due to the initial expansion and the deceleration of the fireball, the subsequent afterglow emission likely results from the later deceleration of the expanding shock in the ambient ISM and therefore likely to be more isotropic than the prompt GRB emission. The detection of GRBs by just their afterglow emission can test the differential beaming: if a significantly larger rate of afterglow events is detected than expected from $\log N - \log S$ for prompt GRBs of same total fluence threshold, then afterglow must have indeed less beaming than prompt emission. Optical wave-band is not very suitable, as optical afterglow is very faint and in many cases undetectable probably due to large dense gas environments. After analyzing the "fast transient" X-ray sources catalogued by Pye and McHardy (1983) from the *Ariel V* survey as well as several other surveys, Grindley (1998) showed that the rate of the candidate afterglow events, 0.15 day^{-1}, of GRBs in X-ray is consistent with the BATSE $\log N(> P) - \log P$ distribution and that the difference of beaming

effect between GRB (in γ-ray) or prompt emission in other wavelengths (e.g. optical) and X-ray afterglow is $\Delta\Gamma\lesssim 3$. The beaming in X-ray is possibly still comparable to the beaming of prompt emission (Rhoads 1997; Dar 1998), while the radio afterglow emission occurs when $\Gamma \sim 1$ and the beaming differences between γ-ray and radio is largest and easiest to be detected. From the total number of a VLA survey with a sensitivity of $\gtrsim$0.1 mJy at 8.44 GHz and the fraction of fading sources at 1.44 GHz, Perna and Loeb (1998) get a crude limit of $\Gamma\lesssim 2.5$ on the differential beaming between radio emission and γ-ray. This results gives a constraint on the beaming Doppler factor of the fireball to $\Gamma\lesssim 3$. In the search for faint X-ray bursts with duration 10 - 300 second with the *ROSAT* PSPC, Vikhlinin (1998) found that 28 events among the 141 flaring sources cannot be identified. If several of the 28 events are GRB afterglow, the beaming factor should be > 100. However, this is unlikely, as 27 of them have optical counterparts obtained years before the X-ray event and which are quite different from the properties of the optical transients associated with BeppoSAX GRB (Vikhlinin 1998). This discussion of beaming is preliminary. To get more restrict constraint on beaming effect in fireball model, we need much larger survey sample.

The smooth power-law decline of optical afterglow, observed over a time scale of days to months for GRB970228 and GRB970508 (Sahu *et al.* 1997; Sokolov *et al.* 1998), suggests that these fireball behaved as if they were spherically symmetric on angular scales ranging from $1/\Gamma(1\text{day}) \sim 1/8$ to $1/\Gamma(1\text{month}) \sim 1/3$ (Rhoads 1997; Waxman 1997). Therefore, in the three well studied GRBs: GRB970228, GRB970508, and GRB980703, there is little evidence for significant beaming (Rhoads 1998) and the beaming angle for GRB980508 can be firmly limited to 30 deg. (Rhoads 1999). However, the observations of GRB990123 probably provides us the first observational evidence of beaming in a GRB (Kulkani *et al.* 1999). Its optical light curve shown in Figure 2.13 probably compose of three broken power laws. The first power law with index ~ -1.6 is probably from the reverse shock and the second can be explained with the normal external shock model of a afterglow. The third power law might not exist and the data in the third segment can be fitted together with the data in the second segment by one power law (Galama *et al.* 1999). The third one of index ~ -1.8, if real, has several possible explanations. The first possible reason is that electrons emitting the optical and IR photons enter the cooling regime. The cooling break has a specific spectral signature (spectral index β defined with $f_\nu \propto \nu^\beta$ steepens by 0.25), which is not consistent with the observations. Second, for a spherical ejecta, a break is expected to mark the transition to a non-relativistic phase (Waxman *et al.* 1998), which for longer wavelength emission occurs with a lag with respective to short waveband emission. The observations seem suggesting an almost simultaneously flux drops in the K and r band fluxes. Kulkarni *et al.* (1999) suggested that the break in the light curve between the second and the third power laws are due to the third

reason (Rhoads 1997): a break will occur, if the ejecta are collimated and the jet enters the "spreading" phase, the sideways expansion, when the Doppler factor Γ of beaming drops below its θ_o^{-1}, where θ_o is the jet opening angle. In this scenario, one expects a steepening from $t^{-1.1}$ to t^p, where p is electron index, typically $p = 2.5$, so the change expected would be more than one power of t. Sari *et al.* (1999) show that the transition took place at 2.04 days implies a beaming factor about 100. Mézsáros & Rees (1999) argue that the break between the second segment and the third segment in the optical light curve of the afterglow implies that the Doppler factor Γ drops below the θ_o^{-1} but the third power law is likely due to jet-edge effects rather than sideways expansion. GRB990123 might be the first burst showing possible evidence for beaming effects and a jet geometry. The fluence of 3×10^{-4}ergs/cm^2 (Kippen *et al.* 1999) places it at the top 0.3% of BATSE bursts and therefore GRB990123 is special. The very fast time decay of optical afterglow of GRB980518 and the observed optical spectrum may indicate a beaming factor $\Gamma \gtrsim 300$ (Sari *et al.* 1999). Another burst with possibly detected jet signal is GRB980326 (Sari *et al.* 1999).

If the association between GRBs and Type Ib-Ic supernova found by Wang and Wheeler (1998) can be proved true, it would strongly imply that the beaming effect in all GRBs is significant, $\Gamma \sim 16$. Unfortunately, their study suffers several deficiencies. The analyses of the corrected version of Wang and Wheeler's sample and other sample show no association between supernova and GRBs (see the discussion in section 2.15.2).

If the TeV emission in some GRBs is confirmed and the emission is from the synchrotron emission of proton in shocks of a fireball, the Doppler emission in such fireballs would be $\gtrsim 500$ (Totani *et al.* 1998).

2.14 Lensed Gamma-Ray Bursts

There is now overwhelming evidence that the majority of GRB sources lie at cosmological distances (Metzger *et al.* 1997). It is possible for γ-ray bursts at cosmological distances to be lensed by intervening galaxies or other large mass concentrations (Paczyński 1987). Their short duration and the transparency of the Universe in gamma-rays make GRBs ideal probes of dark matter in the form of compact objects over the range 10^{-16} to 10^{12} $M_\odot$. Estimates of the frequency of lensed events, or γ-ray burst echoes, under various cosmological scenarios have been made by Mao (1992), Blaes & Webster (1992), Grossman & Nowak (1994), and Nemiroff *et al.* (1993), for an average redshift about $z < 1$, and by Holz & Wald (1998), for an average redshift $z > 1$.

The multiple images of the same burst have separations of order arc-seconds, which are tiny compared to BATSE location errors, and cannot be angularly resolved by present gamma-ray detectors, but they are easily temporally resolved (Paczyński 1987). Thus, the lensed GRBs must be found by searching such events with identical tempo-

ral and spectral signatures rather than by spatially identifying multiple images, only if micro-lensing by stars in the lensing galaxy or cluster is unimportant and the observations are free of noises. Williams and Wijers (1997) showed that micro-lensing by individual stars in the lensing galaxy can smear out the light curves on millisecond time scales. The micro-lensing by an intervening star can also modify significantly the light curve and polarization signal from a GRB afterglow (Loeb & Perna 1997).

The searches for lensed events from the BATSE data base of bursts have thus far been unsuccessful in identifying any credible lensed events (Nemiroff *et al.* 1994; Grossman & Nowak 1994; Marani *et al.* 1998). The best candidate for lensing up to now is GRB990123, whose optical spectra show absorption lines with redshift $z = 1.61$ (Kelson *et al.* 1999) and $z = 0.286$ (Hjorth *et al.* 1999). However, subsequent observations and analysis did not confirm the existence of this foreground galaxy (Yadigaroglu *et al.* 1999). Present data set a limit on the magnification of $\mu < 60$ and likely $\mu < 10$ (Blandford & Helfand 1999).

The null results set an upper limit to the average redshift $\langle z \rangle$ of GRBs (Holz *et al.* 1998). For Robertson-Walker cosmology with a parameter $\Omega = 0.5$ and $\Lambda = 0.0$, $\langle z \rangle < 4.3$ at 95% confidence level. The null results can also give a limit on the cosmological abundance of dark compact objects (Marani *et al.* 1998a) and the much discussed $\Omega = 0.15$ universe dominated by black holes of masses $\sim 10^{6.5} M_{\odot}$ is ruled out at the $\sim 90\%$ confidence level.

2.15 Association with Other Astronomical Objects

2.15.1 Abell Cluster, Quasars and AGNs

Various efforts were made to search any correlation between GRBs and other astronomical objects and failed. For example, Hartmann & Blumenthal (1989) found no angular correlation between GRBs and nearby galaxies. They concluded that if GRBs are cosmological then they must be located at distance larger than 100 Mpc. Otherwise, they would have shown a positive correlation with the galaxy distribution.

Kollat and Piran (1996) claimed a correlation (at 95 % confidence level) between GRBs in 3B catalog and Abell Clusters. This correlation has been recently confirmed by Kompaneetz & Stern (1997) but not by Hurley (1997) and Gorosabel & Castro-Tirado (1997) with more precise GRB localization. Rood & Struble (1996) and Marani *et al.* (1996) showed that the brighter of the GRBs are correlated with Abell clusters. If the correlation is true, comparison of the rich clusters auto-correlation with the cross-correlation found suggests that $\sim 26 \pm 15$ % of the accurate position (bright) GRBs sub-sample members are located within $925 Mpc$. The redshift implied by the correlation would be very small, which seems not consistent with the observational

distribution of redshift.

Recently Schartel *et al.* (1997) found that a group of 134 GRBs with position error radius smaller than 1.8° are correlated with radio quiet quasars. The probability of such correlation by chance coincidence is less than 0.3%. The strongest correlation was detected at the $> 99.7\%$ confidence level for intrinsically bright QSO with $z < 1$. Using 327 GRBs with positions known to better than $\sin 1°$, Burenin *et al.* (1998) showed that a correlation between the GRBs and $M_B < -21$ AGN in the redshift range $0.1 < z < 0.32$ is significant at 99.99% level. The correlation with the all cataloged AGN is found at a lower significance level. They reproduced the results obtained by Schartel *et al.* (1997) and showed that the correlation between GRB and QSO is probably not physically significant.

It should be stressed that the correlation between GRBs and Abell Cluster does not imply that there is a direct association between them, such as would have been if GRBs would have emerged from Abell clusters. All that it means that GRBs are distributed in space like the large scale structure of the universe. Since Abell clusters are the best tracers of this structure, they are correlated with GRBs. Therefore the lack of excess Abell clusters in much smaller IPN error boxes than BATSE (Hurley *et al.* 1997) does not rule out this correlation (Piran 1999)

2.15.2 Supernova

The optical follow-up observations of GRB980425 (Galama *et al.* 1998a) detected a peculiar Type Ib/c supernova SN1998bw with redshift $z = 0.008$ within its spatial error box of 8 minutes provided by Wide Field Camera (WFC) on BeppoSAX (Soffita *et al.* 1998). This Supernova coincides with the revised position of one (S1) of the two X-ray sources (S1 and S2) (Galama *et al.* 1998a,1999a) detected by the narrow field instrument on BeppoSAX. The S2 is $\sim 4'$ away from SN1998bw, and therefore inconsistent with it (Galama *et al.* 1999a). The variable of the two sources is possibly the X-ray afterglow of GRB980425. S1 showed a decay of a factor of about 2 within $\sim$ 6 months, while S2 was no longer detected during the second portion of the first pointing as well as during the successive pointing (Pian *et al.* 1999). Assuming a power-law decay, one can derive, from the data of WFC and NFI, a power-law index of ~ -1.4 for S2, which is consistent with the upper limits determined later on by the NFI and similar to the index found for X-ray afters of previous GRBs (Pian *et al.* 1999). GRB980425 is indistinguishable from other BeppoSAX and BATSE GRBs with respect to duration, time history, spectral shape, peak flux, and fluence (Galama *et al.* 1998a), while SN19980425 is a Type Ib/c supernova with redshift $z = 0.008$.

A physical association between GRB980425 and SN1998bw would be very interesting. Any estimate of the probability that the supernova and the GRB coincided by

chance both spatially and temporally suffers from the problem of *a posteriori* statistics and is very difficult to be evaluated. The probability of catching a SN out to redshift $z = 0.0239$ for Type Ia, $z = 0.0115$ for Type Ib/c, and $z = 0.00727$ for Type II in one of the 13 WFC GRB error boxes within 10 days is 1.1×10^{-4} (Galama *et al.* 1998a). However, if redshift is out to $z = 1$, the chance probability of the spatial-temporal coincidence with a Type Ib/c SN is $\sim 0.1\%$ (Graziani *et al.* 1998).

So, it necessary to do a systematic test by searching for correlations between SNe and GRB in catalogs of SNe and GRBs. Wang and Wheeler (1998) have performed such a study and found no correlation between Type Ia supernova and GRB and a significant (at the 10^{-5} level) correlation between Type Ib/c SNe and GRBs detected by BATSE. However, their study suffers from several deficiencies (Graziani *et al.* 1998). The analysis of both a corrected version of the Wang and Wheeler sample of Type Ib-Ic SNe and a larger sample of Type Ib-Ic SNe shows that a large evidence against the similar hypothesis between Type Ib-Ic SN and GRBs (Graziani *et al.* 1998). The limit to the fraction of Type Ib-Ic SNe that produce observable GRBs must be $< 0.70\%$ with 99.7% probability. To get more restrict limit, a much larger sample of Type Ib-Ic SNe is needed. These results are consistent with those obtained by Kippen *et al.* (1998) with all kinds of SNe (Types Ia, Ib/c, and II) and GRBs by BATSE/*Ulysses* and by Bloom *et al.* (1998). Kippen *et al.* (1998) showed no associations between GRBs with precise BATSE/*Ulysses* locations and known supernova. However, for weaker bursts with less precise locations, the possibility of a burst/supernova association cannot completely excluded with a upper limit of 1.5%. There is also no association between supernova and a subset of GRBs which have similar light curves and spectra to GRB970425 (Norris *et al.* 1998).

2.16 The Observations of Other Related Phenomena

In this section, we discuss the predicted phenomena by fireball and specific inner engine models and their observations. Let $f_{x,\gamma}$ be the ratio of energy emitted in other particles relative to γ-rays. The total fluence F_γ of a "typical" GRB observed by BATSE is 10^{-7}ergs/cm^2, and the fluence of a "strong" burst is about hundred times larger. Therefore, we should expect accompanying bursts with typical fluence of:

$$F_{x|prompt} = 10^{-3} \left(\frac{particles}{cm^2}\right) f_{x,\gamma} \left(\frac{F_\gamma}{10^{-7}\mathrm{ergs/cm^2}}\right) \left(\frac{E_x}{\mathrm{GeV}}\right)^{-1} , \qquad (2.6)$$

where E_x is the energy of these particles. This burst will be spread in time and delayed relative to the GRB if the particles do not move at the speed of light. Relativistic time delay will be significant, if the particles are not massless and their Lorentz factor is smaller than 10^8!. Similarly a deflection angle of 10^{-8} will cause a significant time delay.

In addition to the prompt burst we should expect to see a continuous background of these particles. With one 10^{51} ergs GRB per 10^6 years per galaxy, we expect $\sim 10^4$ events per galaxy in a Hubble time, provided that the event rate is constant in time. This corresponds to a background flux of

$$F_{x|bg} = 3\times 10^{-7} \times \left(\frac{particles}{cm^2 sec}\right) f_{x,\gamma} \left(\frac{E_\gamma}{10^{52}\mathrm{ergs}}\right) \left(\frac{R}{10^{-6}/galaxy}\right) \left(\frac{E_x}{\mathrm{GeV}}\right)^{-1} \quad , \quad (2.7)$$

where R is the GRB burst rate per galaxy per year.

For any specific particle that could be produced, one should calculate the ratio $f_{x,\gamma}$ and then compare the expected fluxes with fluxes from other sources and with the capabilities of current detectors.

2.16.1 Cosmic Rays

The origin of high energy cosmic rays (CR) is still a complete mystery since its discovery in 1912 (e.g. Gaisser 1990). Their power law spectrum $dn/dE \sim E^{-\alpha}$ changes slightly at the so-called "knee" around $10^{15.5}$eV and at the so-called "ankle" around $10^{18.5}$eV. There is a consensus (Gaisser 1990) that CR with energy below the knee are accelerated in Galactic supernova remnants (SNR), those with energy above the knee and below the ankle may be either Galactic or extragalactic in origin, and those with energy above the ankle are extragalactic because of their nearly isotropic sky distribution (e.g. Takeda *et al.* 1998).

Shemi & Piran (1990) pointed out that fireball model is closed related to Cosmic Rays. A "standard" fireball model involved the acceleration of $\sim 10^{-7} M_\odot$ of baryons to a typical energy of 100 GeV per baryon. Proton that leak out of the fireball will become low energy cosmic rays. However, a comparison of the GRB rate (one per 10^6 years per galaxy) with the observed low energy cosmic rays flux, suggests that even if $f_{CR,\gamma} \simeq 1$, this will amount only about 1% to 10% of the observed cosmic ray flux at these energies. Cosmic rays are believed to be produced by SNRs. Since supernova are ten thousand times more frequent than GRBs, unless GRBs are much more efficient in producing Cosmic Rays in some specific energy range, their contribution will be swamped by the SNR contribution. Dar suggested that the Galactic cosmic rays are from Galactic GRBs having jets directed mostly towards us (Dar 1998a).

Waxman (1995a) and Vietri (1995) have shown that the observed flux of Ultra High Energy Cosmic Rays (UHECRs) of energy above 10^{19}eV is consistent with the idea that these are produced by the fireball shocks provided that $f_{UHECRs,\gamma} \simeq 1$. The relativistic fireball shocks that appear in GRBs are among the few astronomical objects that satisfy the conditions for shock acceleration of UHECRs. Waxman (1995a) has shown that the spectrum of UHECRs is consistent with the one expected from Fermi acceleration within those shocks. Rachen and Mészáros (1998) argued that energetic proton cannot

be emitted directly from a GRB without great energy loss in adiabatic expansion, but that neutrons produced in charged current photo-hadronic interactions can escape the GRB and contribute to the cosmic ray proton spectrum after β-decay. Totani (1998a) also showed that UHECRs can escape from GRBs only when their energy is around the acceleration limit. This effect may compensate the Greisen-Zatsepin-Kuzmin (GZK) cutoff (Greisen 1966; Zatsepin & Kuzmin 1966), which was used by Dar (1998b, 1999) as an argument against the association between GRBs and UHECRs.

Milgrom & Usov (1995) pointed out that the error boxes of the two highest energy UHECRs contain strong GRBs – suggesting an association between the two phenomena. Waxman and Coppi (1996) argued that if deflection time of UHECRs is > 50yr, the model is consistent with observations. A possible correlation of a subset of UHECRs above 40 EeV (4×10^{19}eV among each other and with the super-galactic plane was reported by the AGASA experiment (Hayashida *et al.* 1996). This is another argument (Sigl *et al.* 1996) against the association between GRBs and UHECRs. However, we should notice that GRBs are expected to trace the luminous matter in the universe and a possible association between GRBs and Abell Cluster was reported (see section 2.15.1). Abell cluster are somewhat concentrated towards the super-galactic plane (Tully 1987)

2.16.2 Neutrinos

Waxman & Bachall (1997) suggested that collisions between protons and photons within the relativistic fireball shocks produce pions. These pions produce high energy neutrinos with $E_\nu \sim 10^{14}$eV and $f_{\nu,\gamma} > 0.1$. The flux of these neutrinos is comparable to the flux of atmospheric neutrinos but those will be correlated with the position of strong GRBs. This signal might be detected in future km^2 size neutrino detectors.

Most of the energy generated in any core collapse event and in particular in NS-NS merger is released as low energy ($\sim 5 - -10$ MeV) neutrinos (Clark & Eardley 1977). The total energy is quite large $\sim$ a few $\times 10^{53}$ ergs, leading to $f_{\nu,\gamma} \simeq 10$. However, this neutrino signal will be quite similar to a supernova signal, and at present only galactic SN neutrinos can be detected. Supernova are ten thousands times more frequent than GRBs and therefore low energy neutrinos associated with GRB constitute an insignificant contribution to the background at this energy range. Present searches for space-time correlations between GRBs in BATSE catalog and neutrinos give a null result.

2.16.3 Gravitational Waves

If GRBs are associated with NS-NS merger, then they will be associated with gravitational waves and low energy neutrinos. The spiraling in phase of a NS-NS merger

produces a clean chirping gravitational radiation signal. This signal is the prime target of LIGO (Abramovichi *et al.* 1992) and VIRGO, the two (Bardachia *et al.* 1990) large interferometers that are build now in USA and in Europe. The observational scheme of these detectors is heavily dependent on digging deeply into the noise. Kochaneck & Piran (1993) suggested that a coincidence between a chirping gravitational radiation signal from a neutron star merger and a GRB could enhance significantly the statistical significance of the detection of the gravitational radiation signal. At the same time this will also verify the NS-NS merger GRB model.

2.16.4 Black Holes

An NS-NS merger results, inevitably, in a black hole (Davies *et al.* 1994). Thus a direct implication of the NS-NS merger model is that GRBs signal to us (indirectly) that a black hole has just (with the appropriate time of flight in mind) formed.

2.16.5 Large Scale Structure and Universe

The universe and our Galaxy are transparent to MeV γ-rays (Dermer 1992). Thus, GRBs consists a unique homogeneous population of sources which does not suffer from any angular distortion due to absorption by the Galaxy or by any other object. Some authors (Cohen & Piran 1995, 1997) have explored the possibility to determine the cosmological parameters H_0, Ω, and Λ with GRBs. To do so, we need a large sample of GRBs with detected redshifts.

Direct redshift measurements would also enable us to determine the cosmological evolution of the rate of GRBs (Cohen & Piran 1997) and therefore the rates of star formation rate and of galactic evolution.

Lamb & Quashnock (1993) point out that a population of several thousand cosmological burst should show angular deviations from isotropy on a scale of a few degrees. Balázs *et al.* (1998) reported an isotropy of the distribution of short GRBs. This isotropy is not connected to the Galaxy. To estimate the cosmological parameters, the multipole moments of the GRB distribution are needed (Piran & Singh 1997).

2.16.6 Life Extinction on the Earth

BATSE on the COMPTON γ-ray observatory satellite observes one GRB per day (Fishman and Meegan 1995). That implies about a burst per million years per galaxy, if we assume that the bursts trace the mass distribution in the universe. A typical burst can release energy as high as 3×10^{54}ergs from the region probably associated with star formation in a host galaxy (see sections 2.11). If a burst explodes within 1 kpc to a planet like the Earth in a galaxy similar to Milk Way, the high energy cosmic ray jets (Dar *et al.* 1998), before its disruption by galactic magnetic field, hard X-ray

and *gamma*-ray hit the Earth atmosphere (Thorsett 1995). The burst within 1 kpc is about 1 per 10^8years (Dar *et al.* 1998). The particles and photons can produce lethal fluxes of atmospheric muons at ground level, underground and underwater, destroy the ozone layer and radio-activate the environment. They could result in a massive life extinction on the Earth every 10^8years. This is consistent with the fossil record of microbes, algae, fungi, protists, plants, and animals, which shows that the diversity of both marine and continental life increased exponentially and the diversification was interrupted by five big and a number of small mass life extinction since the end of the Precambrian about 570 million years ago (Benton 1995; Erwin 1996). Biological mutations due to ionizing radiations could cause the fast appearance of new species after the massive extinctions. MeV neutrinos may also make some contributions to the mass extinction (Collar 1996).

2.17 Conclusion

After more than thirty years of its discovery, GRBs is still a astronomical mystery. Many questions need to be answered observationally and theoretically. The main obstacle to GRB research in the past is its non-detection of counterpart in other bands, which was due to the large uncertainties in the positions given by gamma-ray detectors. The situation was changed by the instruments onboard BeppoSAX and now the counterparts, afterglow, of some bursts in X-ray, optical, and radio have been detected. While the detections have indeed brought a breakthrough to our understanding of the GRBs, the numbers of the bursts with detected afterglow in other wave-bands is still very small and cannot give a meaningful statistical results. The studies of a particular burst could lead to understanding of the phenomenon, only provided each burst event represents the all set. However, in the case of GRBs, the situation could be quite different, as there is a broad variety of individual events which have physically different, sometimes even conflicting properties. So, the statistic study in gamma-ray bands is still the most important way to understand the bulk properties of GRBs. Statistical studies fall into two categories: a search for relations between bulk properties and an examination of whether GRB bulk properties comprise a homogeneous population or are divided into distinct subclasses. We reviewed in this chapter the observational and statistical results of GRBs as a whole group or several subclasses.

We list several questions which need to be answered observationally in the near future as our conclusion to the chapter. The questions are:

1. while a cosmological distance scale seems to be generally accepted, what is the nature of the GRB host?

2. How good do we know about the history of the early cosmic evolution which is

required to understand the origin of GRBs?

3. Is there any connection between GRB and supernova which is still under debate? If yes, what is the fraction?;

4. how to explain the finding that the $\log N - \log S$ distribution of some subclass GRBs follows a $-3/2$ power law more closely than the ensemble of all bursts? Are these subclasses real or only selection effects due to, e.g., strong luminosity and spectral evolution?

5. Are there multiple populations of GRBs and, if yes, what is their typical observational properties?

6. Do short bursts have optical counterparts? Do they have different distance scale? Why afterglow accompany some GRBs and not others?

7. Are GRBs in or outside their host galaxies?

8. What is the distance distribution of GRBs?

9. What is the intrinsic luminosity function for bursts?

10. What is the multi-wavelength behaviors of GRB light curves immediately after the gamma-ray burst? Are they similar to GRB990123?

11. Is beaming important in GRBs? It seems important for some bursts but not for other. Do different subclasses have different beaming effects?

12. How large is the largest redshift z_{out} of GRBs?

13. How short is the shortest time scale of GRBs?

14. How high can the emission photon reach? Is high energy emission associated with the brightest bursts only, or with all them?

15. Can the low energy emission (afterglow) be seen before bursts?

burst	X-ray detection	O	R	γ-ray fluence in [ergs/cm^2]	redshift	Ref[1]	total energy in [ergs]
GRB970111	BeppoSAX	-	-		-	-	-
GRB970228	BeppoSAX	+	-	1×10^{-5}	≥ 1	7	-
GRB970402	BeppoSAX	-	-		-	-	-
GRB970508	BeppoSAX	+	+	2×10^{-6}	0.835	1	2×10^{51}
GRB970616	BeppoSAX	-	-	4×10^{-5}	-	-	-
GRB970815	RXTE	-	-	1×10^{-5}	-	-	-
GRB970828	RXTE	-	-	7×10^{-5}	0.33	2	1×10^{52}
GRB971024	RXTE	-	-		-	-	-
GRB971214	RXTE	+	+	1×10^{-5}	3.418	3	1×10^{53}
GRB971227	BeppoSAX	-	-	9×10^{-7}	-	-	-
GRB980109	BeppoSAX	-	-		-	-	-
GRB980326	BeppoSAX	-	-	1×10^{-6}	-	-	-
GRB980329	BeppoSAX	+	+	5×10^{-5}	~ 5	8	-
GRB980425	BeppoSAX	+	+	4×10^{-6}	0.0085	4	7×10^{47}
GRB980515	BeppoSAX	-	-	1×10^{-6}	-	-	-
GRB980519	BeppoSAX	+	+	3×10^{-5}	-	-	-
GRB980613	BeppoSAX	+	-		1.096	9	
GRB980703	RXTE	+	+	5×10^{-5}	0.966	5	1×10^{53}
GRB981226	BeppoSAX	?	-		-	-	-
GRB990123	BeppoSAX	+	+	3×10^{-4}	1.6004	6	3×10^{54}

Table 2.1: Observational data of several GRBs for which afterglow was detected. The two columns O and R indicate whether emission was detected in the optical and radio, respectively. The total energy of the burst is estimated through the observed fluence (> 20 keV and redshift, assuming spherical emission. We use the redshift of SN1998bw for GRB980425. The redshifts of the GRB970228 and GRB980329 are estimated from photometric observations.

1. Ref: (1) Metzger *et al.* (1997), (2) Yoshida *et al.* (1999), (3) Kulkarni *et al.* (1998), (4) Bloom *et al.* (1998), Bloom *et al.* (1998a), (5) Djorgovski *et al.* (1998), (6) Kelson *et al.* (1999); Hjorth *et al.* (1999), (7) van Paradijs *et al.* (1997), (8) Fruchter (1998), (9) Djorgovski *et al.* (1999)

3 Gamma-Ray Bursters – the Central Engine

3.1 Introduction

A theoretical understanding of γ-ray bursts should comprise several components: a site (where), an energy source (what), and an emission mechanism (how). Before BATSE, there was a consensus that GRBs originated from the magnetized neutron stars in the Galactic disk. The results from BATSE onboard CGRO revolutionized our idea about the burster physics. and BATSE results ruled out this possibility but did not tell us where GRBs are from. As the evidences for the distance scale were so ambiguous, we had a "great debate" on the distance scale of cosmic GRBs (Lamb 1995; Paczyński 1995) just a few couples of years ago. More than hundred models (Nemiroff 1994) have been proposed and the models were more than events, at a stage before BATSE. The number of models is still increasing.

There are three possible places for the candidate objects for GRB sources: The Oort cloud of comets, extended Galactic halo, and outside the galaxy. We will first briefly discuss all the models in four possible distances. Models at different distances require different energies for GRBs. To clarify these different models, it is necessary first to determine the distance scales to GRBs. So, we will review all the observations on the distance to GRBs. The observations of afterglow of GRBs from BeppoSAX brought us another breakthrough on the investigation of GRBs. The optical observations of the afterglow for several GRBs detect emission and/or absorption lines and determine their redshifts, which confirmed their cosmological origin. However, the number of GRBs with detected redshift is very few. So, we also discuss some indirect distance indicators in the observational characters of GRBs. These observations suggest that most GRBs, if not all, are cosmological. So, in the rest part of the chapter, we will describe, in some details, some specific models for cosmological GRBs. A couple of years ago, the simulation showed that the binary mergers could not promptly output enough energy to power GRBs, when the estimated luminosity of a typical GRB was about 10^{51}ergs. The recent simulations showed that the binary models with hyper-accretion

disk or torus could accommodate a typical GRB with that energy. Unfortunately, recent observations showed that the energy output by a typical GRB source is about $\sim 10^{52}$ergs and can be as high as $\sim 10^{53}$ ergs. So, astrophysicists who do the simulation have another march to catch this energy. Recent simulations show that the class of models based on hyper-accreting black holes has become the favorite mechanism for driving fireball. In the last section of the chapter, we discuss this hyper-accretion – black hole system.

The main impediment in all of models for cosmological GRBs and most of the extended Galactic halo models is getting the energy out as γ-rays with the observed spectral and temporal characteristics of bursts. A key idea is to separate a burst source into a fireball and its inner engine, namely the emission mechanism is separated from the energy sources. A fireball model concerned only the relativistic hydrodynamics and radiation can be studied in very detail and be easily confronted with observations without involving specific astronomical candidate objects, in the Galactic or cosmology. The recent observations have firmly confirmed the fireball model. On the contrary, the "inner engine" has to be concerned with a specific object and is very difficult to be tested as it hides from observations. So, many models on the "inner engine" are very schematic. We discuss the "inner engine" in this Chapter. The fireball model on emission mechanism will be discussed in detail in Chapters 4, 5 and 6.

3.2 Models in Four Different Distance Scales

3.2.1 The Oort Cloud of Comets

A discussion of this possibility generated a fair amount of laughter at the first BATSE conference. The Oort cloud provides the nearest possibility for a roughly isotropic, radially decreasing source density. Speculative models have been suggested, employing comet-comet collisions at typical relative velocities of ~ 1 km s^{-1} (Luchkov 1994) and comet-primordial black hole collisions (Bickert & Greiner 1993). The geometrical and physical implications of such models were examined by Horack *et al.* (1994) and Clarke *et al.* (1994).

No correlation is found between the burst angular distribution and directions towards long-period comets (Maoz 1993). The highly non-uniform radial profile of comet density ($n \propto r^{-3.5}$) also argues against the Oort cloud hypothesis because the observed V/V_{max} distribution implies a relatively uniform source distribution for the closer, brighter subset of bursts. Realizing this difficulty, White (1993) suggests a new spherical comet cloud with a shallow density profile between 35 and 10^3 AU. Li (1997) suggested, based on some statistical correlation between GRB characteristics and observational date, that GRBs are from the outer helio-sphere.

It is concluded that the Oort cloud is unlikely to be the source of γ-ray bursts.

3.2.2 Galactic Disk Objects

Before the BATSE data on the distribution of weak bursts were available, bursts were considered to emanate from within the Galactic disk, most likely from neutron stars. On physical grounds, highly magnetized neutron stars are attractive candidates for generating non-thermal γ-ray spectra with rapid temporal structure, for generating red-shifted absorption cyclotron line and annihilation line (Higdon & Lingenfelter 1990). The current data, however, now convincingly rule out such models (e.g. Mao & Paczyński 1992; Hartmann *et al.* 1995), though the possibility that some fraction of bursts are related to the disk should not be dismissed.

Two- (or multi-) component solutions could involve populations at very different distance scales or with very different luminosity function, including (1) local NS - local NS with two distinct luminosity functions (Lingenfelter & Higdon 1992), (2) a disk plus halo model (Smith & Lamb 1993) that perhaps up to 70% of all BATSE bursts could be local disk neutron stars, while the rest could reside in a dark matter halo with core radius of $R_c \sim 22$ kpc, (3) a model combining Galactic and cosmological distributions with magnetic reconnection in neutron star magneto-spheres being the source for the burst energy of the local component and fireball-environment interactions being the source for the distant bursts (Katz 1994a). One of the basic problems with these multi-component models is the general uniformity of burst appearance; there are no well established burst classes with clearly distinct spectral, temporal, flux, or angular patterns.

3.2.3 Extended Galactic Halo

The advantages of such models are the much lower energy requirements and the ability to generate cyclotron absorption lines. A satisfactory geometrical model can be constructed from burst sources in an extended Galactic halo (e.g. Brainerd 1992; Li & Dermer 1992; Podsiadlowski *et al.* 1995). Isotropy of the observed bursts requires a size $R > 100$ kpc (Hakkila *et al.* 1994). On the other hand, the failure to observe an excess of bursts from M31 places upper limits on the extent of this halo to $R \leq 300$ kpc (Briggs *et al.* 1996). This scale is larger than that of dark matter halo, which is between $\sim 10 - 100$ kpc. These models predict the emergence of a small but measurable anisotropy. One problem with this model is that there is no direct physical or other observational evidence for such an extended Galactic halo. The Galactic dark matter corona do not have the required statistical properties. It is often supposed that the bursters are neutron stars, since such compact stars are conjectured to be capable of producing intense fluxes of hard, non-thermal photons, and spectral lines. It was

argued (Brainerd 1992; Eichler & Silk 1992) that neutron stars with Population II characteristics in an extended galactic halo (EGH) could satisfy all observational constraints. High velocity neutron stars escaped from the Galactic plane can create such an extended halo (Shklovskii & Mitrofanov 1985), but the large anisotropies caused by them are unacceptable unless early bursting activity is suppressed by some special mechanism (Li & Dermer 1992; Duncan & Li 1997; Podsiadlowski, Rees & Ruderman 1995). Injection from a halo distribution of parent bodies as a result of white dwarf mergers (Hartmann 1992; Eichler & Silk 1992; Wasserman & Salpeter 1994) or from the Large Magellanic Clouds (Fabian & Podsiadlowski 1993) to form an extended Galactic halo would alleviate some of the problems associated with disk injection. The discovery (Lyne & Lorimer 1994) of high velocity radio pulsars well above the Galactic plane lends some support to the notion that neutron star formation in halo could provide the GRB sources. For NS models, one difficulty is that the full number of these stars is not large enough $N \leq 10^6 - 10^7$, which means that every NS must repeat gamma-ray bursts many times: $10^6 - 10^7$, since the total number of bursts should be $\sim 10^{13}$. The observations do not support the repetition of a large fraction of the GRBs (see the discussion in Section 2.8). It is also difficult to find the mechanism of repeating GRBs from NS and the source of energy without introducing other difficulties. Some non-NS candidates have been also proposed.

As conclusion, no known Galactic component, including the Dark Matter halo, is sufficiently isotropic to match the observational constraints. Current models postulate a new Galactic component (Briggs *et al.* 1996) or a population with very high frequency repetitions. Observations of GRBs from M31 with a ten times more sensitive detector than BATSE could help to clarify this.

3.2.4 Extragalactic Objects

Cosmological burst models were proposed shortly after the discovery of the phenomenon (Usov & Chibisov 1975);, but had been dormant until Paczyński (1986) revived these ideas prior to the launch of CGRO. Two major obstacles made them unlikely contenders; (1) Gpc distances require an enormous energy release, and (2) the small deposition volume inferred from small fluctuation time scales then implies photon densities so high that the $\gamma\gamma$ pair opacity should quenched high energy emission. Both objections are overcome when the emission is beamed, as is expected from relativistic expansion resulting from the deposition of $\sim 10^{52}$ ergs/s in a small source volume. Given that γ-ray bursts have a constant comoving density in space and time, the same intrinsic brightness and the same power-law spectrum, the most favored choice for the geometric resolution of the isotropy/non-uniformity problem is certainly that of a cosmological distribution of sources (Paczyński 1991). The observed isotropy is necessary require-

ment of models for cosmological GRBs. The apparent inhomogeneity would result from redshift effects and possibly source evolution. Cosmological distributions have been fit to the observed intensity distribution (Piran 1992; Fenimore *et al.* 1993; Woods & Loeb 1994) and satisfactory fits can be found using standard candle luminosities, standard cosmology, and no source evolution. In these models, the weakest sources are at red-shifts z_{max} of ~ 2 and at luminosities of order 10^{52} erg/s.

A number of possibilities for the energy sources have been suggested. For a list of models, one may refer to Nimeroff (1994). Any model has to answer two questions: the first is how to account for the required large and sudden energy release and the second is how this energy is transformed into a relativistically-out flowing plasma able to emit intense gamma rays with a non-thermal spectrum. All model for cosmological GRBs are associated with a compact object. For all the models, a significant difficulty is to produce a relatively clean fireball (Paczyński 1990; Shemi & Piran 1990). Suggested solutions for this problem include hypotheses that the GRB is powered by a merger of strange stars (Haensel *et al.* 1992; Usov 1998) which hold their baryons using strong interactions, the burst originates on field lines that threat an event horizon (Levinson & Eichler 1993; Iwamoto *et al.* 1998; Paczyński 1998), and a compact object together with encircling disk of matter (Eichler *et al.* 1989; Mészáros & Rees 1992; Mochkovitch *et al.* 1993), in which case the direction of the rotation axis would naturally be kept free of baryons by centrifugal force.

Binaries consisting of either two neutron stars, a neutron star and a White Dwarf, a neutron star and a black hole, a Helium star and a black hole, and a black hole and a White Dwarf (Paczyński 1986; Goodman 1986; Eichler *et al.* 1989; Narayan & Willington 1992; Mochkovitch *et al.* 1993; Narayan *et al.* 1992; Sigurdsson & Rees 1997; Ruffert & Janka 1998; Fryer & Woosley 1998a; Fryer *et al.* 1998) can release $\sim 10^{53}$ergs in $\sim$ 1ms, during their merging following the in-spiral driving by gravitational radiation taking away the orbital angular momentum. A significant of this energy is radiated as neutrinos and anti-neutrinos, a fraction of which can annihilate to produce electrons and positrons, resulting in a fireball with $\sim 10^{52}$ ergs. We will discuss in more detail the merging models of binaries in Section 3.5.

Special types of supernova (Paczyński 1998; Colgate 1974; Woosley 1993; MacFadyen & Woosley 1998; Fuller & Shi 1998; Cen 1998; Vietri & Stella 1998) have been proposed as the sources. The recent observations of a possible association between GRB980425 and SN1998bw strengthen the models (Höflich *et al.* 1998). Woosley (1993) considered the gravitational collapse of rapidly rotating Wolf-Rayet stars, a "failed type Ib supernova." In this model, a black hole with a massive accretion disk is formed, and neutrino annihilation along the rotation axis produces a fireball. The "collapsar" model refined by Woosley *et al.* (1998) and MacFadyen & Woosley (1998) is similar to "failed type Ib supernova". A "hypernovae" model by Paczyński (1998)

with geometrical similarity to "failed type Ib supernova" was proposed to explained the very high energy of GRBs which may reach $\sim 10^{53}$ ergs. In a specific speculative example of such an explosion, called micro-quasar, the inner core of a rapidly rotating massive star collapse into a $\sim 10M_{\odot}$ Kerr black hole with $\sim 10^{54} erg$ of rotational energy, while the outer core forms a massive disk/torus and a "dirty" fireball is created. A super-strong $\sim 10^{15}$G magnetic field is needed to make the object operate as a micro-quasar similar to the Blandford & Znajek (1977) model. A special "hyper-novae" model, an extremely energetic explosion of a massive carbon-oxygen (C+O) star, was proposed by Iwamoto *et al.* (1998) to reproduce the optical and the light curve of SN1998b. Vietri & Stella (1998) present another scenario, called "supranova" for the formation of GRBs. In this model, a supra-massive fast-rotating neutron star is formed during a supernova explosion of massive star and has a mass at infinity larger than the largest mass for static models. Such stars have masses in the range $M \approx 2 - 3.5 M_{\odot}$, equatorial radii $R_{eq} = 11 - 18$ km, and angular velocity $\omega \approx 8000 - 12000\,\mathrm{s}^{-1}$. Such a neutron star loses so much angular momentum through, for instance, magnetic dipole radiation that centrifugal support against self-gravity becomes impossible, becomes unstable to axisymmetric perturbations and finally implodes to a black hole. A GRB is formed during such an implosion. This may be the most baryon-clean environment proposed so far, as the surrounding medium is swept by the SN explosion in which the supra-massive neutron star forms. A mass of $\approx 0.1 M_{\odot}$ in the equatorial belt can easily reach centrifugal. Energy can be extracted via the conversion of the Poynting flux from the mini-torus (due to the large-scale magnetic field locked into the mini-torus) into a magnetic relativistic wind. Occasionally, this model will produce quickly decaying or non-detectable afterglows.

Carter (1992) suggested the tidal disruption of an ordinary star by a $10^4 - 10^6$ solar mass black hole. The tidal compression produces temperatures of order 100 keV on satisfactory time scales, but the spectrum is thermal. Dokuchaev *et al.* (1998) proposed a similar model. The models are not favored by the observations of host galaxies of transient counterparts of GRBs, as the optical transients of GRBs are observed to have a large offset to the center of host galaxy.

Usov (1992) proposed a model in which rapidly rotating neutron stars with strong magnetic fields ($\sim 10^{16}$gauss) lose rotational energy catastrophically, resulting in a mildly relativistic pair plasma. Spruit (1998) showed that such high magnetic field can be produced in X-ray binaries. Similar scenarios involving accretion-induced collapse of white dwarfs suggested by Blackman *et al.* (1996) and Yi & Blackman (1997). In these models, it is needed to clarify the way how to produce a clean fireball and to last so long time scales as the observed suggested.

Shaviv & Dar (1995) suggested that a relativistic jet with Lorentz factor of ~ 1000 is generated during a binary neutron star merger and a gamma-ray is produced by

reemission of the soft field by the atoms of the jet. This model has severe difficulties (Drozdova & Panchenko 1997). Dar (1999) suggested that when a Quark star is born due to first order phase transition in cooling NSs, highly relativistic jets are formed and γ-ray bursts are produced through inverse Compton scattering of stellar light by the the plasmoid. A Poynting dominated (or magnetically driven) baryon-poor jet from collapsed stellar objects was suggested by Mészáros & Rees (1997) to produce γ-ray bursts. In this model, a very high magnetic field ($\gtrsim 10^{15}$G) is needed. The magnetically based models are currently not sufficiently accurate to make quantitative predictions (Livio *et al.* 1998). The models involved very thin relativistic jets and high beaming effects are not favored by observations (See the discussion in Section 2.13 on beaming).

Melia & Fatuzzo (1992) suggested that γ-ray bursts are caused by extragalactic radio pulsar glitches. Ma & Xie (1996) estimated the energy and the burst rate of the explosion and found that they were in agreement with observations. However, the numerical calculation (Fryer & Woosley 1998) showed that these explosion are far too baryon-rich and have much too low an energy with $\Gamma > 40$ to explain gamma-ray bursts.

For all the models for cosmological GRBs, the overall energetics from these various progenitors do not differ by more than about one order of magnitude. The direct evidences for the cosmological origin of GRBs are the detected redshift in several optical afterglow of GRBs and probably the association between the those burst with accurate positions in the 3B catalog and Abell cluster (Kolatt & Piran 1996). We will discuss the observational evidences for cosmological distances in Section 3.3.

3.3 Distance Scale

3.3.1 Redshift

The measurements of red-shifts of several GRB optical counterparts provide the best and the only direct distance estimates for GRBs. These measurements confirm the cosmological origin of these GRBs. Unfortunately, red-shifts are available only for a few bursts, see Table 2.1. If we want to take them as the evidence for the cosmological origin of all or most of the GRBs, we have to assume that all GRBs are within one class of phenomena. Much more sample are needed.

3.3.2 The Angular Distribution

Even before these redshift measurements there was a strong evidence that GRBs originate from cosmological distances. The observed isotropically angular distribution is incompatible with a galactic disk distribution unless the sources are at distance less than 100pc. However, in this case we would expect that $\langle V/V_{max} \rangle = 0.5$ correspond-

ing to a homogeneous distribution (Schmidt *et al.* 1988) while the observations give $\langle V/V_{max}\rangle = 0.33$ (See Section 2.10).

A homogeneous angular distribution could be produced if the GRBs originate from the distant parts of the Galactic halo (extended Galactic halo). Since the solar system is located at $d = 8.5$kpc from the galactic center such a population will necessarily have a galactic dipole of order d/R, where R is a typical distance to a GRB (Katz 1992). The lack of an observed dipole strongly constrain this model. Such a distribution of source is incompatible with the distribution of dark matter in the halo. The typical distance to the GRBs must be of the order of 100kpc to comply with this constraint (Hartmann *et al.* 1995).

3.3.3 The Peak Flux Distribution

The count distribution or the peak flux distribution of the bursts observed by BATSE show a paucity of weak bursts. A homogeneous count distribution, in a Euclidean space should behave like: $N(>C) \propto C^{-3/2}$, where $N(>C)$ is the integrated number of bursts with counts more than C counts (or counts per second). The observed distribution is much flatter (see Fig. 2.10). This fact is reflected by the low $\langle V/V_{max}\rangle$ value of the BATSE data: there are fewer distant sources than expected.

The observed distribution is compatible with a cosmological distribution of sources. A homogeneous cosmological distribution displays the observed trend - a paucity of weak bursts relative to the number expected in a Euclidean distribution.

The statistics $\langle V/V_{max}\rangle$ is weighted average of the distribution $N(>f)$. In 1992, Piran (1992) compared the theoretical estimate of this statistics to the observed one and concluded that the typical redshift of the bursts observed by BATSE is $z_{max} \sim 1$. Later, Fenimore *et al.* (1993a) compared the sensitivity of PVO, which observed $N(>f) \propto f^{-3/2}$, with the sensitivity of BATSE and concluded that for BATSE bursts $z_{max} \sim 1$. This corresponds to a peak luminosity of $\sim 10^{50}$ergs/sec. Other calculations based on different statistical methods were performed by Horack & Emslie (1994), Loredo & Wassermann (1995a,b), Rutledge *et al.* (1995), Cohen & Piran (1995) and Mészáros and collaborators (Mészáros & Mészáros 1995, 1996; Horváth *et al.* 1996; Reichart & Mészáros 1997) and others. In particular, Loredo & Wassermann (1995a,b) give an extensive discussion of the statistical methodology involved.

To explain the observed peak-flux distribution properly with cosmological distribution of GRBs, we need to know the cosmic evolution of event rate and the luminosity function of GRBs. Furthermore, there are two unknown parameters related to the background cosmological models: the closure parameter, Ω, and the cosmological constant Λ. Fortunately, the likely-hood function for KS (Kolmogorov-Smirnov) test is practically independent of Ω in the range: $0.1 < \Omega < 1$. It is also insensitive to the

cosmological constant Λ (in the range $0 < \Lambda < 0.9$, in units of the critical density). This simplifies the analysis as we are left only with the luminosity function of bursts (Piran 1999).

Recall that $\langle V/V_{max} \rangle$ of the distribution of short bursts is rather close to the homogeneous Euclidean value of 0.5. This means that when analyzing the peak flux distribution one should analyze separately the long and the short bursts (Cohen & Piran 1995). For long bursts (bursts with $t_{90} > 2$sec) the likely-hood function peaks at $z_{max} = 2.1$ (see Fig. 2.10) (Cohen & Piran 1995). The allowed range at a 1% confidence level is: $1.4 < z_{max} < 3.1$. The maximal redshift $z_{max} = 2.1^{+1.1}_{-0.7}$ corresponds to 2.3×10^{-6} events per galaxy per year (for a galaxy density of $10^{-2}h^3\,\mathrm{Mpc}^{-3}$ (Kirshner *et al.* 1983)). The rate per galaxy is independent of H_0 and is only weakly dependent on Ω. For $\Omega = 1$ and $\Lambda = 0$, the typical energy of a burst with an observed fluence, F, is $7 \times 10^{50} \left(\frac{F}{10^{-7}\mathrm{ergs/cm}^2}\right)$ ergs. The distance is estimated assuming a Band's spectra with an index 1.5. These numbers vary slightly if the bursts have a wide luminosity function.

Short bursts are detected only up to a much nearer distances: $z_{max}(short) \sim 0.4$, again assuming standard candles and no source evolution. There is no significant lower limit on z_{max} for short bursts and their distribution is compatible with a homogeneous non-cosmological one. The estimate of $z_{max}(short)$ corresponds to a comparable rate of $\sim 6.3 \times 10^{-6}$ events per year per galaxy and a typical energy of $\sim 3 \times 10^{49} F_{-7}$ ergs. There are no lower limits on the energy and/or no upper limit on the rate, since there is no lower limit on $z_{max}(short)$. The fact that short bursts are detected only at nearer distances is also reflected by the higher $\langle V/V_{max} \rangle$ of the population of these bursts (Katz & Canel 1996).

The above considerations have assumed power-law functions of burst rate evolution and of the luminosity. The results should be modified if the rate of GRBs trace the star formation rate (SFR) (Totani 1997; Sahu *et al.* 1997; Wijers *et al.* 1998). The SFR has been determined recently by two independent studies (Lilly *et al.* 1996; Madau *et al.* 1996; Connolly *et al.* 1997). The SFR peaks at $z \sim 1.25$. This is a strongly evolving non-monotonic distribution, which is drastically different from the power laws considered. Sahu *et al.* (1997) find that the assumption that the GRB event rate is proportional to SFR yields a $N(> f)$ distribution compatible with the observed one (for $q_0 = 0.2$, $H_0 = 50\,\mathrm{km\,sec}^{-1}\,\mathrm{Mpc}^{-1}$) for a narrow luminosity distribution with $L_\gamma = 10^{51}$ sers/sec. Wijers *et al.* (1998) find that the implied peak luminosity is higher $L_\gamma = 8.3 \times 10^{51}$ ergs/sec and it corresponds to a situation in which the dimmest bursts observed by BATSE originate from $z \approx 6$!

The direct redshift measures of GRB970508 (Metzger *et al.* 1997) agrees well with estimated made previously using peak-flux count statistics (Fenimore *et al.* 1993a; Loredo & Wassermann 1995b; Cohen & Piran 1995). The redshift of GRB971214,

$z = 3.14$, and of GRB980703, $z = 0.996$, and the implied luminosities disagree with these estimates. A future detection of additional red-shifts for other bursts will enable us to estimate directly the luminosity function of GRBs. It will also enable us to determine the evolution of GRBs. Krumholz *et al.* (1998) and Hogg & Fruchter (1998) find that with a wide luminosity function both models of a constant GRB rate and a GRB rate following the star formation rate are consistent with the peak flux distribution and with the observed redshift of the three GRBs.

3.3.4 Time Dilation

Norris *et al.* (1995) examined 131 long bursts with a duration longer than 1.5 s and found that the dimmest bursts are longer by a factor of ≈ 2.3 compared to the bright ones. With our canonical value of $z_{max} = 2.1$ the bright bursts originate at $z_{bright} \approx 0.2$. The corresponding expected ratio due to cosmological time dilation, 2.6, is in agreement with this measurement. Fenimore and Bloom (1995) find, on the other hand, that when the fact that the burst duration decreases as a function of energy as $\Delta t \approx E^{-0.5}$ is included in the analysis, this time dilation corresponds to $z_{max} > 6$. This would require a strong negative intrinsic evolution: $\beta \approx -1.5 \pm 0.3$. Alternatively, this might agree with the model in which the GRB rate follows the star formation rate (SFR) (Totani 1997; Wijers *et al.* 1998) which gives $z_{max} \approx 6$.

3.4 Observational Constraints on the Central Engine

The observations of afterglow of GRBs have confirmed simple GRB fireball models. A fireball model is based on an "inner engine" that supplies the energy and accelerates the baryons and it is impossible to determine what it is from current observations. Unfortunately, the discovery of afterglow does not give additional direct evidence about the central engine.

Once the cosmological origin of GRBs was established, we had two direct clues on the nature of the "inner engine": the rate and the energy output. GRBs occur at a rate of about one per 10^6 years per galaxy (Piran 1992; see Section 2.1) and the total energy is $\sim 10^{52}$ergs. These estimates assume isotropic emission. Beaming with an angle θ changes these estimates by a factor of $4\pi/\theta^2$ in the rate and $\theta^2/4\pi$ in the total energy involved. These estimates are also based on the assumption that the burst rate does not vary with cosmic time. The observations that GRB hosts are star forming galaxies (Djorgovski *et al.* 1998; Bloom *et al.* 1998b; Fruchter *et al.* 1998; Hogg & Fruchter 1998) indicates that the rate of GRBs may follow the star formation rate (Totani 1997; Sahu *et al.* 1997; Wijers *et al.* 1998). In this case, the bursts are further and they take place at a lower rate and have significantly higher energy output. The output enery

may range from about $\sim 10^{51}$ ergs to about 10^{53} ergs.

The fireball model poses an additional constraint: the inner engine should be capable of accelerating $\sim 10^{-5} M_{\odot}$ to relativistic energies. One can imagine various scenarios in which 10^{52}ergs are generated within a short time. The requirement that this energy should be converted to a relativistic flow is much more difficult as it requires a "clean" system with a very low but non-zero baryonic load. This requirement suggests a preference for models based on a torus-BH system and electro-magnetic energy transfer or electro-magnetic energy generation, as these could more naturally satisfy this condition (see Thompson 1994; Mészáros & Rees 1997; Usov 1992). Paczyński (1998) has recently suggested a unique hydrodynamic model in which 10^{54}ergs are dumped into an atmosphere with a decreasing density profile. This is a cosmological variant of Colgate's (1974) galactic model. This would lead to an acceleration of fewer and fewer baryons eventually to relativistic velocities. Overall, one could say that the "baryonic load" problem is presently the most bothersome open question in the "fireball model".

The recent realization that energy conversion is most likely via internal shocks rather than via external shocks provides additional information about the inner engine: The relativistic flow must be irregular to produce the internal shocks, it must be variable on a short time scale (as this time scale is seen in the variability of the bursts), and it must be active for up to a few hundred seconds and possibly much longer (Katz *et al.* 1998) – as this determines the observed duration of the burst. The engine must be compact $\sim 10^{7}$cm to produce the observed variability and it must operate for a million light crossing times to produce a few hundred-second signals.

There are more than a hundred GRB models (Nemiroff 1994). At a certain stage, before BATSE, there were probably more models than observed bursts. Most of these models are, however, galactic and those have been ruled out if we accept the cosmological origin of GRBs. This leaves a rather modest list of viable GRB models: binary neutron star mergers or their variants – NS-BH, WD-BH, BH-He core mergers - (Eichler *et al.* 1989; Paczyński 1986; Belinikov *et al.* 1984; Goodman *et al.* 1987), failed supernova (Woosley 1993), accretion induced white dwarf collapse (Usov 1992), hypernova (Paczyński 1998) and supra-nova (Vietri & Stella 1998). All these are based on the formation of a compact objects of one type or another and the release of its binding energy. With a binding energy of $\sim 5 \times 10^{53}$ergs or higher, all these models have, in principle, enough energy to power a GRB. However, they face similar difficulties in channeling enough energy to a relativistic flow. This would be particularly difficult if indeed 10^{53}ergs are needed, as some recent burst have indicated. Paczyński's hypernova is an exception as in this model all the energy is channeled initially to a non-relativistic flow and only later a small fraction of it is converted to relativistic baryons. All these models are consistent with the possibility that GRBs are associated with star forming regions as the life time of massive stars is quite short and even the typical life time of

a neutron star binary ($\sim 10^8$ yr) is sufficiently short to allow for this coincidence.

Other models are based on an association of GRBs with massive black holes associated with Quasars or AGNs in galactic centers (e.g. Carter 1992). These are basically ruled out as all GRBs with optical afterglow are not associated with such objects. Furthermore, such objects do not appear in other small GRB error boxes searched by Schaefer *et al.* (1997). From a theoretical point of view, it is difficult to explain the observed energy and time scales with such objects.

3.5 Binary Mergers

Coalescing compact (neutron star – neutron, neutron star – black hole) binaries (Paczyński 1986, 1991; Eichler *et al.* 1989; Narayan *et al.* 1992) and collapsars (special supernova) (Woosley 1993; Paczyński 1998) (see Sections 3.2.4) are the most popular candidates for GRB sources, suggested by the recent numerical simulations (e.g Popham *et al.* (1999), Ruffert & Janka (1998a) and Eberl *et al.* (1998)).

For binary systems, the coalescence and merging of two stars into a single object is the almost inevitable end-point of close binary evolution because of the decay of the binary orbits due to gravitational radiation emission, tidal dissipation, magnetic breaking, or friction in common gaseous envelopes. A binary merger results, most likely, in a rotating black hole (Davies *et al.* 1994). The process release $\approx 5 \times 10^{53}$ergs (Clark & Eardley 1977). most of this energy escapes as neutrinos and gravitational radiation, but small fraction of this energy suffices to power a typical GRB. The discovery of the famous binary pulsar PSR1913 + 16 (Hulse & Taylor 1975) demonstrated that this decay is taking place (Taylor & Weisberg 1982, 1989). The discovery of other binary pulsars (Wolszczan 1991; Taylor & Dewey 1988; Stokes *et al.* 1991; Anderson *et al.* 1990; Nice *et al.* 1996) has shown that PSR1913 + 16 is not unique and such systems are common.

There is some disparity between various published results about the estimated rates of NS binary coalescence in the Universe. Simple statistical arguments based on the observed local population of binary radio pulsars with probable NS companions lead to an estimate $\sim 10^{-7}\,\mathrm{yr}^{-1}\,\mathrm{Mpc}^{-3}$ (Narayan *et al.* 1991; Phinney 1991). Finn & Chernoff (1993) derived a Galactic NS merger rate of $R \simeq 10^{-6}\,\mathrm{yr}^{-1}$ from radio pulsar surveys. More recently, van den Heuvel & Lorimer (1996) revised this number to $R \simeq 8 \times 10^{-6}\,\mathrm{yr}^{-1}$, using the latest galactic pulsar population model of Curran & Lorimer (1995), who gave an estimate of $3 \times 10^{-6}\,\mathrm{yr}^{-1}$. This value is consistent with the upper limit of $10^{-5}\,\mathrm{yr}^{-1}$ for the Galactic binary NS birth rate derived by Bailes (1996) on the basis of very general statistical considerations about pulsars. In contrast, theoretical models of the binary star population in our Galaxy suggest that the NS binary coalescence rate may be much higher, $\gtrsim 10^{-6}\,\mathrm{yr}^{-1}\,\mathrm{Mpc}^{-3}$ (Tutukov & Yungelson

1993; Portegies Zwart & Spreeuw 1996; Lipunov *et al.* 1998; Clark *et al.* 1979; Hills *et al.* 1991). The merger rate of BH-WD binary is very difficult to estimate. A possible range is 10^{-9} - $10^{-6}\,\mathrm{yr}^{-1}$ per Milky-Way like galaxy (Fryer *et al.* 1998). Sigurdsson & Rees (1997) predict a neutron star/white dwarf rate of $\sim 10^{-7}\,\mathrm{yr}^{-1}$ per galaxy. Low mass black holes will merger with white dwarfs roughly at the same rate (or an order of magnitude less), depending on the black hole formation rate. The Helium star/black hole merger powering a GRB is estimated to occur at a rate comparable to that of merging neutron stars and black hole/neutron star pairs (Fryer & Woosley 1998a). The estimated merger rate of binary neutron stars is a rough agreement with the simple estimate of the GRB rate (assuming no beaming and no cosmic evolution of the rate) (cf. Section 2.1). The merging rate of binary neutron stars estimated based on observations of three binary system or pulsars in the Galaxy suffered the statistical error and should be taken as lower limit, while theoretical estimate suffers from the fact that it is based on several assumptions, e.g. about the mass loss, angular momentum loss through wind, and the kick velocity obtained by a neutron as a result of supernova explosion, etc. The uncertain of assuming selection effects affects both methods,

It is not yet clear how binary neutron star systems form. The question is how the system survive the second supernova event. The binary will be disrupted if this explosion ejects more than half of its total mass. There are two competing scenarios for the formation. In one scenario, the first neutron star sink into the envelope of its giant companion and its motion within this envelope lead to a strong wind that carries away most of the secondary mass. When the secondary reaches core collapse, it has only very small envelope and the total mass ejected is rather small. in a second scenario, the second supernova explosion is asymmetric. The asymmetric explosion gives a velocity of a few hundred *km/sec* to the newborn neutron star. In a fraction of the cases, this velocity is in the right direction to keep the binary together. Such a binary system will have a comparable velocity of its mass centroid (Narayan *et al.* 1992; Lyne & Lorimer 1994; White & van Paradijs 1996; Fryer & Kalogera 1997). The second scenario has several advantages. First, it explains both the existence of binary neutron stars and the existence of high velocity pulsars (Lyne & Lorimer 1994; Cordes & Chernoff 1997,1998). Second, and more relevant to GRBs, with these kick velocities, these binaries could escape from their parent galaxy, provided that this galaxy is small enough. Such escaping binaries will travel a distance of $\sim 200\,\mathrm{kpc}(\frac{v}{200\mathrm{km/sec}})(\frac{T}{10^9})$ before they merge. The GRB will occur when the system is at a distance of the order of hundred kpc from the parent galaxy.

While a NS-NS merger has enough energy available to power a GRB, it is not clear how the GRB is produced. A central question is how a binary merger system generate the relativistic wind required to power a GRB. Most of the binding energy escapes as neutrinos (Clark & Eardley 1977), which is several $10^{53}\,\mathrm{erg\,s}^{-1}$ (Ruffert & Janka

1998a). Eichler *et al.* (1989) suggested that about one thousandth of these neutrinos annihilate and produce pairs that in turn produce gamma-ray via $\nu\bar{\nu} \rightarrow e^+e^- \rightarrow \gamma\gamma$, during the merger. This idea has bee criticized on several grounds by different authors. Jaroszynski (1996) pointed out that a large fraction of the neutrinos will be swallowed by the black hole that forms. To model the gamma-ray emission realistically, the complete hydrodynamic and nuclear evolution during the final merging of the two NS, especially in the outermost low-density regions of the merger, must be understood in detail. This is far more challenging than understanding the emission of gravitational waves, which is mostly sensitive to the bulk motion of the fluid, but is completely insensitive to nuclear processes taking place in low-density regions. Of course, the coalescence of binary neutron stars is one of the most outstanding problem in numerical relativity. Davies *et al.* (1994), Ruffert *et al.* (1996, 1997), and Ruffert & Janka (1996, 1998, 1998a) simulated neutron mergers and suggested that the neutrino emission associated with the dynamical phase of the merging is too short to provide the energy for gamma-ray bursts by neutrino-antineutrino annihilation and that the central object will not be warmed enough to produce a significant neutrino flux because the merger is nearly adiabatic (Katz 1997). Significant heating of the coalescing stars occurs only after they have plunged into each other, and the neutrino luminosities can rise to several $\times 10^{53}\,\mathrm{ergs\,sec^{-1}}$ and even exceed $10^{54}\,\mathrm{ergs\,sec^{-1}}$ in case of the more violent collisions. The energy from neutrino and anti-neutrino annihilation during the merger was proved to be too small to power a typical cosmic GRB with energy 10^{51} ergs (Janka & Ruffert 1996) by at least an order of magnitude, and more probably two or three orders of magnitude. The discrepancy become even worse, if a GRB like GRB971214 is considered.

As the tidal interaction steepens the effective interaction potential between two binary stars and destabilize the circular orbit (Lai *et al.* 1994), hydrostatic equilibrium configurations for the binary systems with sufficiently close components can become dynamically unstable (Chandrasekhar 1975), sensitively depending on the NS equation of state. This is because tidal effect are stronger for stars containing a less compressible fluid. As the dynamical stability limit is approached, the secular orbital decay driven by gravitational wave emission can be dramatically accelerated. The two stars then plunge rapidly toward each other, and merge together into a single object in just a few rotation periods (Rasio & Shapiro 1992). Including approximate general relativistic effect in a numerical simulation of a neutron merger, Wilson and collaborators (Wilson *et al.* 1996; Mathews & Wilson 1997; Marronetti *et al.* 1998; hereafter WMM) found a "binary-induced collapse instability" of neutron star, which, if correct, is that an initially stable neutron star of a very soft equation of state (EOS) close to the maximum mass could collapse to a black hole well before getting to the final phase of binary coalescence due to the tidal effect of a companion. This is a purely general relativistic effect, as in

Newtonian gravitation, the tidal field of a companion tends to *stabilize* a star against radial collapse, lowering the critical value of polytropic index for collapse below 4/3. If it is valid, there would be no GRBs. The result has been criticized by many authors (see Rasio & Shapiro (1999) for a full review) on theoretical grounds either analytically or numerically. Shibata *et al.* (1998) show, after considering approximate GR effect, that the tidal field from a binary companion stabilizes a star against gravitational collapse. However, even though there is mounting evidence against the WMM effect, the possibility that sufficiently massive, highly compact NS in coalescing binaries can collapse into black holes prior to merge will not be completely ruled out until detailed hydrodynamic simulations in full GR without approximation are carried out (Shapiro 1998).

There are two possibilities for the final fate of the compact massive remnant of the merger. One possibility is that continuing neutrino emission of the remnant would drive a dense baryonic wind off its surface which would lead to a sizable mass loss but non-relativistic expansion (Woosley & Baron 1992; Qian & Woosley 1996), a situation which is not favorable for producing gamma-ray bursts. Another possibility is that after a few milliseconds the compact massive remnant will collapse to a black hole surrounded by a hyper-accretion disk or a torus. While a numerical approach in full GR would be required in order to decide definitively between the two different outcomes, simulations with approximate treatment of relativistic effects show that the latter is most likely. The accretion power around a black hole is an alternative source of energy (Narayan *et al.* 1992; Katz 1997; Mészáros & Rees 1997). Accretion of the disk, or torus, on the central black hole may take a few dozen seconds (Katz 1997). It can produce the wind needed to produced a GRB (Popham *et al.* 1999; Mészáros *et al.* 1998) as luminous as GRB971214. If the fireball is powered by the torus binding energy, a moderate amount of beaming is required; if it is powered by Blandford-Znajek mechanism, a beaming is not necessary. The accretion disk or torus system play a key role in a binary system to produce a GRB (Mészáros *et al.* 1998; Mészáros & Rees 1992; Rees & Mészáros 1992), so we will discuss it in more detail in Section 3.6.

There are a lot of similarities among the binary systems NS-NS, NS-BH, He-BH and WD-BH, which are driven by gravitational radiation emission and result in a single black hole. Unless the mass of the black hole is rather large, the companion will be tidally disrupted before it is captured by the black hole. However, even if such a tidal disruption will take place, we still expect a collision in the binary neutron star mergers but at most a tidal disruption followed by infall of the debris on the black hole for other types of binary mergers.

It is very difficult to prove or to disprove a GRB model, as the observed radiation emerges from a distant region which is very far from the "inner engine". For binary neutron star models, one possible way is to detect the coincidence between a GRB and

a gravitational radiation signal. Binary neutron stars are one of the most promising sources of gravitational waves for near future detectors such as LIGO, VIRGO, TAMA and GEO. We can hope for some progress in the studies of GRB coming from future observations of gravitational waves.

3.6 Thick Accretion Disk or Torus – BH System

3.6.1 Energetics

Various numerical simulations of binary mergers (NS–NS or NS–BH mergers, Helium core – black hole or white dwarf – black hole mergers) and a wide category called as hypernova or collapsars (failed supernova Ib, single or binary Wolf-Rayet collapse, massive star collapse) suggest that a debris torus or a disk is formed around the black hole after the coalescence of binary stars or the collapse of a massive star in an failed supernova or hypernova. The mass of the torus is $\sim 0.1\,M_\odot$ for NS-NS (Ruffert & Janka 1998a), $\sim 0.5\,M_\odot$ for BH-NS (Eberl *et al.* 1998), $\sim 1M_\odot$ for WD-BH (Fryer *et al.* 1998a), $\sim 2M_\odot$ for BH-He core (Fryer & Woosley 1998a), $\sim 2M_\odot$ for Collapsar or hypernova (MacFadyen & Woosley 1998), respectively. In all the simulations, the magnetic field is not considered.

Two large reservoirs of energy are available in principle: the binding energy of the orbiting debris, and the spin energy of the black hole (Mészáros & Rees 1997). The first can provide up to 42% of the rest mass energy of the disk for maximally rotating black hole, while the second can provide up to 29% of the rest mass of the black hole itself. The $\nu\bar{\nu} \rightarrow e^+e^-$ process (Eichler *et al.* 1989) can tap the thermal energy of the torus produced by viscous dissipation. For this mechanism to be efficient, the neutrinos must escape before being advected into the hole; on the other hand, the efficiency of conversion is low if the neutrino production is too gradual. The black hole of ~ 2 – $10M_\odot$ accrete matter from the torus of mass $\sim 0.1M_\odot$ – $\sim 2M_\odot$ at a high rate within, at most, a typical GRB time scale of ~ 30 sec. A steady-state accretion disk at a high accretion rate $0.01M_\odot$ to $10M_\odot$ has been studied by Popham *et al.* (1999).

For a "standard model" with black hole spin parameter $a = 0$, $\alpha = 0.1$, $\dot{M} = 0.1M_\odot\,\mathrm{s}^{-1}$, and $M = 3M_\odot$,

Outside about 10^8 cm the disk is thick and advection-dominated as the densities and temperature are too small for neutrino cooling to be significant, while optical depths are too large for significant photon cooling. At $r = 10^9$ cm, the disk surface density is $\sim 10^{14}\,\mathrm{g\,g}^{-1}$. Between $\sim 10^{7.5}$ cm and $\sim 10^8$ cm, the disk becomes denser and thinner due to the energy loss via photo-integration of nuclei (alpha particles) into nucleons. At $r \sim 10^{7.5}$ cm, the cooling due to photo-integration shuts off rather abruptly as all the nuclei have bee photo-integrated. Neutrino cooling, predominantly by pair

capture, becomes important but well below the dissipation rate. As the gas continues to fall inward, the increasing temperature and density produce a rapid increase in the neutrino cooling rate, which gradually approaches the dissipation rate. As a result of these losses, the temperature is reduced, the density raised, and the disk scale height reduced compared to the advection solution. Neutrino temperatures at the last stable orbit range from 2 MeV (a BH without rotation, slow accretion) to 13 MeV (Kerr geometry, rapid accretion) and the density from 10^9 to $10^{12}\,\mathrm{gcm}^{-3}$. The sudden onset of neutrino losses owing to the high temperature dependence leads to an abrupt thinning of the disk that may provide favorable geometry for jet production. The inner disk remains optically thin to neutrino for accretion rates up to about 1 $M_\odot\,\mathrm{s}^{-1}$. The efficiency for producing a pair fireball along the rotational axis by neutrino annihilation is highly variable and very sensitive to the accretion rate. A diverse range of GRB energies seems unavoidable. Typical estimates suggest that a fireball of $\lesssim 10^{51}$, erg (Ruffert & Janka 1998a; Popham *et al.* 1999), except perhaps in the "collapsar" or failed SN Ib case, where an energy $10^{52.3}$ ergs for optimum parameters can be reached (Popham *et al.* 1999). The binary models with this mechanism may account well for the short hard bursts in BATSE sample, but not a burst like GRB971214 of energy $10^{53.5}$, ergs without invoking energy sources other than neutrino annihilation.

An alternative way to tap the torus energy is through dissipation of magnetic fields generated by the differential rotation in the torus (Paczyński 1991; Narayan *et al.* 1992; Mészáros & Rees 1997; Katz 1997). Even before the BH forms, a NS-NS merging system might lead to winding up of the fields and dissipation in the last stages before the merger (Mészáros & Rees 1992; Vietri 1996). However, a hole formed from a coalescing compact binary is guaranteed to be rapidly spinning, and, being more massive, could contain more energy than the torus. This energy can be extracted by the Blandford & Znajek (1977) effect, if magnetic fields supported by the rotating torus couple to the rotation of the hole (Mészáros & Rees 1997; Paczyński 1998). Collectively, resulting MHD outflows have been referred to as Poynting jets.

Strong magnetic fields of order 10^{15}G are needed to carry away the rotational or gravitational energy in a time scale of tens of second (Usov 1994; Thompson 1994). If the magnetic fields do not thread the black hole, then a Poynting outflow can at most carry the gravitational binding energy of the torus. For a maximally rotation and for a non-rotating black hole, this is $\epsilon_{max} = 0.42$ and 0.06 of the torus rest mass, respectively. For NS-BH, He-BH or WD-BH which have larger disks and fast rotation, the maximum energy is $\sim 8 \times 10^{53}\epsilon(M_d/M_\odot)$ ergs; for the failed supernova which is a slow rotator, it is $\sim 1.2 \times 10^{53}\epsilon(M_d/M_\odot)$ ergs, and for the fast rotating NS-NS merger it is $\sim 0.8 \times 10^{53}\epsilon(M_d/M_\odot)$ ergs, where M_d is the mass of the disk and ϵ the efficiency in converting gravitational into MHD jet energy.

If the magnetic fields generated in the torus thread the spinning black hole, the

rotational energy of the BH can be extracted via the Blandford-Znajek (Blandford & Znajek 1977) mechanism (Woosley 1993; Paczyński 1998; Mészáros & Rees 1997). Blandford-Znajek effect has been invoked as the fundamental process powering radio jets in active galactic nuclei (AGN) (e.g. Begelman *et al.* 1984). However, this mechanism has recently been criticized severely. We will discuss it in more detail in Section 3.6.2. For a rotating black hole described by the Kerr metric, a portion of the rest mass energy can in principle be extracted, while the remaining 'irreducible' mass is proportional to the area of the event horizon. The extractable energy is

$$\epsilon f(a) M_{bh} c^2$$

where ϵ is the MHD efficiency factor and $a = Jc/(GM^2)$ is the rotational parameter and J the total angular momentum of the rotating black hole. a equals to 1 for a maximally rotating black hole.

$$f(a) = 1 - \sqrt{(1 + \sqrt{1 - a^2})/2}$$

is small unless a is close to 1, where it sharply rises to its maximum value $f(1) = 0.293$, so the main requirement is a rapidly rotating black hole, $a \gtrsim 0.5$. Rapid rotation is essentially guaranteed in NS-NS merger, since the radius is close the that of a black hole and the final orbital spin period is close to the required maximal spin rotation period. For NS-NS merger, the central black hole will have a mass of about $2.5 M_\odot$ (Ruffert & Janka 1998a), the NS-NS system can thus power a jet of up to $\sim 1.3 \times 10^{54} \epsilon (M_{BH}/2.5 M_\odot)$ ergs. The scenarios less likely to produce a fast rotating BH are the NS-BH merger and the failed SNe Ib. A maximal rotation rate may also be possible in a He-BH merger, depending on what fraction of the He core gets accreted along the rotation axis as opposed to along the equator (Fryer & Woosley 1998a). For a fast rotating BH of $3 M_\odot$ (Eberl *et al.* 1998), the maximal energy carried out by the jet is then $\sim 1.6 \times 10^{54} \epsilon (M_{BH}/3 M_\odot)$ ergs.

Thus in the accretion powered jet case, even allowing for low total efficiency, say $\epsilon = 30\%$, a NS-NS merger require a modest beaming of the γ-rays by a factor $(4\pi/\Omega_j) \sim 20$ if the jet is powered by the torus binding energy, or no beaming if the jet is powered by the Blandford-Znajek mechanism, to produce a GRB of an isotropic energy of $10^{53.5}$ ergs. The beaming requirements of BH-NS and some of the other progenitor scenarios are even less constraining.

3.6.2 The Instability of a Torus-BH System and Blandford-Znajek Effect

A torus-BH system may suffer several kinds of instability, which may determine the system duration. Papaloizou & Pringle (1987) found that all *non-accreting* tori are unstable to global non-axisymmetric perturbations. However, further studies (Blaes

1987) showed that even a modest mass loss due to accretion may be sufficient to stabilize the Papaloizou-Pringle modes. The essentially local Rayleigh criterion for dynamical stability with respect to axially symmetric local perturbations demands that the specific angular momentum should increase with the distance from the axis of rotation (Seguin 1975). In the presence of a weak magnetic field, even the Rayleigh-stable tori are dynamically unstable (Balbus & Hawley 1991). The Balbus-Hawley instability does not destroy the large scale structure of tori, but instead drives a local turbulence which induces viscosity which is needed for accretion to occur.

There have been some discussions on possible "runaway instability" in relativistic tori (Abramowicz *et al.* 1983; Abramowicz, Karas & Lanza 1998; Nishida *et al.* 1996; Daigne & Mochkovitch 1997; Masuda & Eriguchi 1998). There exists a self-crossing equi-potential surface in fluid distribution around a black hole, similar to the Roche lobe in binary systems. The material inside the Roche lobe for the torus is gravitationally bound, while those configurations that just fill their Roche lobes are marginally bound and are called critical. If the relativistic Roche lobe shrinks sufficiently during the transfer of mass through its cusp-like inner edge, matter which was bound before the mass transfer will become unbound, falling catastrophically into the black hole on the dynamical time scale, producing a further shrinking of the Roche lobe. This is the runaway instability. These instabilities can be virulent in a sufficiently massive torus ($M_{disk}/M_{BH} \gtrsim 0.1$), as the self-gravity of the disk is the main source of the instability (Abramowicz *et al.* 1983; Nishida *et al.* 1996; Masuda & Eriguchi 1997). When its self gravity is neglected, a disk with a constant angular momentum is unstable to the runaway instability (Abramowicz *et al.* 1983; Daigne & Mochkovitch 1997), but is stabilized by a spread in angular momentum (Daigne & Mochkovitch 1997; Abramowicz, Karas & Lanza 1998). The strong magnetic fields may help to stabilize the disk. The disruption of a neutron star is almost certain to lead to a situation where violent instabilities redistribute mass and angular momentum within a few dynamical time scales. In a torus that was massive or thin enough to be self-gravitating, bar-mode gravitational instabilities could lead to further redistribution of angular momentum and/or to energy loss by gravitational radiation within only a few orbits. Whether a torus of given mass is dynamically unstable depends on its thickness and stratification, which in turn depends on internal viscous dissipation and neutrino cooling.

If a long burst is generated via Poynting flux – either through a relativistic wind "spun off" the torus or via the Blandford-Znajek mechanism, the required magnetic field is a few times 10^{15} G. A weaker field would extract inadequate power; on the other hand, if the large-scale field were even stronger, the energy would be damped too fast to account the longer complex bursts. Kluźniak and Ruderman (1998) point out that, starting with 10^{12} G, it only takes of order a second for simple winding to amplify the field to 10^{15} G; they argue further that magnetic stresses would then be strong enough

for flares to break out. But amplification in a newly-formed torus could well occur more rapidly, for instance via convective instabilities, as in a newly formed neutron star (Thompson 1994). Such fields can build up on very short time scales, or order of $\sim$ few ms. Convective overturning motions should stop after the disk has cooled by neutrino emission below a few MeV on a time scale of order a few second (Ruffert *et al.* 1997). But the Balbus-Hawley mechanism can also generate azimuthal magnetic fields. The nonlinear evolution and/or reconnection of such fields as they become buoyant can then lead to poloidal components at least of order $\gtrsim 10^{15}$ G. Indeed, it is not obvious why the fields cannot become even higher up to the Virial limit $B_v \sim 10^{17}$ G.

Blandford & Znajek (1977) acknowledged that if the poloidal magnetic field threading the black hole is comparable in the strength to the poloidal field threading the inner parts of the accretion disk, then the contribution from black hole to the electro-magnetic output due to the Blandford-Znajek mechanism would only be a fraction of the comparable flux from the disk. For a black hole of given mass and angular momentum, the strength of the mechanism depends crucially on that of the magnetic field threading the horizon of the hole, scaling as the square of this field strength (Blandford & Znajek 1977; Ghosh & Abramowicz 1997 for review). Rees & Mészáros (1994) argued that it only takes a residual cold disk of $10^{-3} M_{\odot}$ to confine a field of 10^{15} G. Recently, however, Ghosh & Abramowicz (1997) have called into question the mechanism as, they argued, the strength of the magnetic field threading the black-hole horizon has been widely overestimated previously for AGNs. The question is severe, if we consider that the magnetic field is generated in some way by the flux in the accretion disk. Li (1999) constructed a toy model to study this mechanism and showed that if the magnetic field is generated by a toroidal electric current residing in a thin disk around the black hole in the equatorial plane, the Blandford-Znajek mechanism may not be effective in extracting energy from a rotating black hole with disk. Livio *et al.* (1998) argue that there is no reason to suppose that the magnetic field threading the central spinning black hole is significantly stronger than the field threading the inner disk, as the currents which generate the field must be in the disk and not in the hole and the hole is in effect a very poor conductor, compared to the surrounding disk material. In order for the Blandford-Znajek mechanism to dominate the power output, it is necessary for the magnetic flux on open field lines threading the black hole to greatly exceed the magnetic flux on open field lines threading the inner regions of the disk. Therefore, they conclude that independent of the spin of the black hole, the electro-magnetic output of the disk in the Poynting flux or magnetically driven wind dominates that from the black hole. Thus the spin of the black hole is probably irrelevant to the expected electro-magnetic power output from the system.

The models proposed (e.g. Kluźniak & Ruderman 1998; Mézsáros *et al.* 1998) for the generation of threading magnetic fields in the torus-BH system for GRBs base on

the flux in the torus. Therefore the generated magnetic field by the torus to thread the black hole horizon may not be as strong as the one needed by models (10^{15} G).

Can a strong enough magnetic field threading the black hole but weakly threading the inner edge of the torus be generated during the merging of binary stars or the clapsing of a star and before the formation of a torus-BH system? No literature is found to address this question. The magnetic might be amplified before and during the merging of binary neutron stars (e.g. Mészáros & Rees 1992) due to the winding of the fields or to the nonlinear growth of MHD modes. The difficulty for my opinion is that the time for the merging or collapsing is very short, as short as hydrodynamic time. However, we have to notice that some processes such as the Balbus-Hawley mechanism (Balbus & Hawley 1991) and the convective instabilities, as in a newly formed neutron star (Duncan & Thompson 1992; Thompson 1994), are very powerful in amplifying a magnetic field. The time scale might be of order $\sim$ few millisecond. This idea will be checked more carefully.

A magnetic field configuration capable of powering the bursts is likely to have a large scale structure. Flares and instabilities occurring on the characteristic dynamical time scale of order $\sim$ millisecond would cause substantial irregularity or intermittence in the overall outflow that would manifest itself in internal shocks (Rees & Mészáros 1994). There is thus no problem in principle in accounting for sporadic large-amplitude variability, on all time scales down to a millisecond (Mészáros *et al.* 1998).

Observationally (see Section 2.3) the short burst (< 5 sec and typically ~ 0.5 sec), long (> 4 sec and typically ~ 33 sec) and intermediate (2.3 – 8 sec and typically ~ 4.4 sec) appear to represent three distinct subclasses, and one early proposal to explain this was that accretion induced collapse (AIC) of a white dwarf into a NS plus debris might be a candidate for the long bursts, while NS-NS mergers could provide the short ones (Katz & Canel 1996). As indicated by numerical simulation by Ruffert *et al.* (1997), Ruffert & Janka (1998a), and Eberl *et al.* (1998), $\nu\bar{\nu}$ annihilation will generally tend to produce short bursts $\lesssim$1 sec of of energy $\lesssim 10^{51}$ ergs in NS-NS systems, while Popham *et al.* (1999) argued that in collapsars and WD/He–BH systems longer $\nu\bar{\nu}$ bursts may be possible. Longer bursts however imply lower $e^{\pm}$ conversion efficiency, so the observed fluxes could be explained only if the jets were extremely collimated, by at least 10^{-3} – 10^{-4}.

After magnetic field have built up to some fraction of the equi-partition value with the shear motion, a magnetic viscosity develops. Assuming that $B_r B_\phi \sim B^2$, it can be characterized in the usual way by the parameter

$$\alpha \sim B^2/(4\pi\rho c_s^2) \sim 10^{-1} B_{15}^2/(\rho_{11} T_{11}) \quad , \tag{3.1}$$

where $\rho = \rho_{11} \times 10^{11}\,\mathrm{g/cm^3}$ is the torus mass density, c_s sound velocity and $T = T_{11} \times 10^{11}\,\mathrm{K}$ temperature. The typical values for temperature and mass density are taken

from the numerical simulation by Ruffert & Janka (1998a) and Popham *et al.* (1999). This viscosity continues operating also after cooling has led to the disappearance of neutrino viscosity. Assuming a value of $\alpha = 0.1$, a BH mass $3M_{\odot}$ and outer disk radius equal to the Roche lobe size, Popham *et al.* (1998) estimate life times of the torus of 0.1 sec for NS-NS or BH-NS, 10 – 20 sec for a collapsar (Failed SN Ib, hyper-nova or rotating WR), 15 – 150 sec for WD-BH and He-BH systems. However, a field of 10^{15} G may be more difficult to support in He-BH systems.

3.7 Conclusion

The origin of the GRB energy sources remains a mystery. There are more than a hundred of models but few observational constrains can be put on them. Since the detections of afterglow and redshift for several GRBs, models for cosmological GRBs have become the most favored models. The studies of afterglow and the host galaxies imply that GRBs may be associated with star formation regions. So, present popular models for cosmological GRBs are associated with the formation of a compact object, mostly a black hole, as in this process an enough energy can be released. As the energy source would generally be enshrouded by an opaque envelope of matter which would degrade any gamma rays to much softer radiation and would let the photons leak out on an un-acceptably long diffusion time scale. The solution to the problem leads the torus – BH system, in which a energetic fireball break out along the rotation axis leading to gamma-rays emission outside the envelope. Such system may exist $\gtrsim$1 sec. So, it also gives a solution to another problem that most typical estimates for the time scale of the radiation burst are of order of $\sim 10^{-4}$ sec, either from considering the free-fall time into a black hole following merger or from looking at the radiation front traversal time of a radiated-dominated fireball at the time it becomes optically thin, and this is much shorter than the observed burst time scales. It seems that a torus-BH system is a feature of all the plausible models – binary mergers or collapse of a massive star. The complex light curves and the long lasting of some GRBs might imply that a very strong magnetic field is necessary to produce a fireball.

NS-NS merger model is the most widely studied one. Numerical simulations indicate that NS-NS merger might be the source of short bursts. The longer bursts might have different sources. It was showed recently that some serious doubts have arisen about whether the Blandford-Znajek mechanism would be efficient in extracting rotational energy of a spinning black hole to power a jet. This process needs more investigation.

We cannot make any conclusion on the GRB inner engine here. A breakthrough in the study of inner engines of GRBs might result from by the detection of some other astrophysical phenomena, e.g. gravitational waves.

4 Fireball Model – Overview

4.1 Introduction

We start our discussion from the compactness problem in Sec. 4.2. The observations of large luminosity and short variability of GRBs show that the radiation density in the emission regions is so high that it should be optically thick to high energy γ-ray photons due to $\gamma\gamma \rightarrow e^{-}e^{+}$ pair productions and no high energy photon can escape from it. This is contrast with the observations that there are a large fraction of high energy γ-ray photons in the GRB spectrum and the emission regions should be optically thin to these photons. The inferred radiation should also be optically thick to GeV photons as these GeV photons interacts with low energy photons and produces electron-positron pairs. This is the compactness problem. This problem exists even if the sources are in the extended halo. To solve this problem, we have to invoke relativistic motions in the radiation sources.

Ockham's razor and the desire to limit the total energy lead us to the idea that the observed gamma-ray energy and the relativistic motion is related. The simplest model is that the source is a sphere composed of photons, electrons and positrons with relativistic motions. This kind of a sphere is called a fireball. However, a small amount of contaminant baryons in the fireball will convert the internal energy to kinetic energy of the baryons. Finally, the spectrum will be thermal. As ultra-relativistic motion is required, the baryonic load cannot be large. The hydrodynamics of the conversion process will be discussed in Sec. 4.3. This baryonic contamination problem leads us to the idea that the observed gamma-rays arise in the process of slowing down of a relativistic energy flow, at a stage that the motion of the emitting particles is still highly relativistic.

The simplest mode of relativistic energy flow is in the form of kinetic energy of relativistic particles. A variant is based on the possibility that a fraction of the energy is carried by Poynting flux (Thompson 1994; Usov 1994; Usov & Smolsky 1996; Mészáros & Rees 1997; Katz 1997), but in all models the power must be converted to kinetic energy somewhere. Overall, the different methods can be characterized by two parameters: the ratio of the kinetic energy flux to the Poynting flux and the location of

the energy conversion stage (internal or external conversion). In the following sections, we will focus on the simplest possibility, that is of a kinetic energy flux.

Within the baryonic models, the initial energy (radiation or other kinds) of the fireball is converted to kinetic energy of a shell of relativistic particles and then transported. The kinetic energy will be converted to "thermal" (random) energy of the relativistic particles via shocks. These particles then release their "thermal" energy and produce the observed radiation. The conversion processes from kinetic to random energies will be discussed in Sec. 4.6 after the Sec. 4.3 on the hydrodynamics of a fireball. There are two modes of energy conversion: (i) external shocks (Sec. 4.5.3), which are due to interaction with an external medium like the ISM; (ii) internal shocks (Sec. 4.6.4) that arise due to shocks within the flow, when fast moving particles catch up with slower ones. Similar to external and internal energy conversion occurs within other models for the energy flow.

The angular spreading problem (Section 4.4) shows that a GRB with a complex light curve is most probably produced by internal shocks. Internal shocks occur when one shell overtakes another. However, only a small fraction of the kinetic energy can be converted into radiation via internal shocks. Most of the energy has to be radiated via external shocks. The external shocks may be responsible for afterglow (Sec. 4.8) and the prompt emission of some GRBs with simple light curve.

We are led to the generic picture that GRBs are composed of a three-stage phenomenon: (i) a compact central hidden "engine" that produces a relativistic energy flow, (ii) the energy transport stage from internal energy to kinetic energy, (iii) the conversion of this energy to thermal energy and then the observed prompt γ-ray emission, and (iv) conversion of the remaining energy to radiation in other wavelengths – the "afterglow". Theoretical models should give a full description of these stages. The models for central engines have been discussed in Chapter 3. This chapter gives a general review on the theoretical works on (ii), (iii), and (iv). Our work will be presented in next chapters.

4.2 Compactness Problem and Relativistic Motion

For a GRB source emitting isotropically at a luminosity distance D_L, the observed fluence F and the rapid temporal variability of a small time scale δT imply that a total energy of $E = 4\pi D_L^2 F$ is emitted in a small region of a size $R_0 < c\,\delta T$. In these small emission regions, the high energy γ-ray photons of energy E_1 could interact with lower energy photons of energy E_2 via pair production process $\gamma\gamma \to e^+e^-$, if $\sqrt{E_1 E_2} > m_e c^2$ (up to an angular factor). The average optical depth of a spherical source for this

process is (Guilbert *et al.* 1983; Carigan & Katz 1992; Piran & Shemi 1993)

$$\begin{aligned}\tau_{\gamma\gamma} &= \frac{f_p \sigma_T F D_L^2}{R_0^2 m_e c^2} \\ &\gtrsim 10^{13} f_p \left(\frac{F}{10^{-7}\,\mathrm{ergs/cm^2}}\right)\left(\frac{D_L}{3000\,\mathrm{Mpc}}\right)^2\left(\frac{\delta T}{10\,\mathrm{msec}}\right)^{-2} ,\end{aligned} \quad (4.1)$$

where f_p is the fraction of photon pairs in the observed fluence that satisfy the pair production condition, σ_T the Thompson cross-section and $F \simeq 10^{-7}\,\mathrm{ergs/cm^2}$ is about the sensitive limit of BATSE. $D_L \sim 3000\,\mathrm{Mpc}$ and $\delta T = 10\,\mathrm{msec}$ have been taken as typical values for a GRB. Under this very large optical depth, pairs will form rapidly, if no at the initial time, and up-scatter lower energy photons via Compton scattering, resulting in a tremendous optical depth for all photons. Therefore, no photons can freely escape from the emission and no high energy photon should be observed! The observed spectrum should be thermal! However, the observed non-thermal spectrum indicates with certainty that the sources must be optically thin! This is the compactness problem. This problem was realized very early in one form by Ruderman (1975) and in another by Schmidt (1978). Both of them used it to argue against the cosmological origin of GRBs.

An alternative estimate of $\tau_{\gamma\gamma}$ is to consider the optical depth of the highest energy photons, say a GeV photon, to pair production with the lower energy photons. This estimate gives a comparable limit $\tau_{\gamma\gamma}$ with the same order of magnitude as the one in Eq. 4.1. The observation of GeV photons shows that they are able to escape freely. In other words, it means that the optical depth must be much smaller than unity (Fenimore *et al.* 1993; Woods & Loeb 1995).

To overcome compactness problem, one has to invoke special relativistic effects. For a radiation source moving towards an observer at rest with a relativistic velocity of a Lorentz factor $\Gamma = 1/\sqrt{1 - v^2/c^2} \gg 1$, photons with an observed energy $h\nu_{obs}$ have energy at the source $h\nu_{obs}/\Gamma$. Since the energy at the source is lower, fewer photons have sufficient energy to produce pairs. The fraction of photons that could produce pairs at the source is smaller by a factor $\Gamma^{-2\alpha}$ than the observed fraction f_p, where α is the high energy spectral index defined by Eq. 2.1. At the same time, the emission source radius $R_e < \Gamma^2 c\delta T$ is by a factor of Γ^2 larger than the original estimate, $R_0 < c\delta T$. Therefore, we have

$$\begin{aligned}\tau_{\gamma\gamma} &= \frac{f_p}{\Gamma^{2\alpha}} \frac{\sigma_T D_L^2 F}{R_e^2 m_e c^2} \\ &\gtrsim \frac{10^{13}}{\Gamma^{(4+2\alpha)}} f_p \left(\frac{F}{10^{-7}\,\mathrm{ergs/cm^2}}\right)\left(\frac{D_L}{3000\,\mathrm{Mpc}}\right)^2\left(\frac{\delta T}{10\,\mathrm{msec}}\right)^{-2} .\end{aligned} \quad (4.2)$$

The compactness problem can be resolved, if the emission region is moving relativistically towards us with a Lorentz factor $\Gamma \gtrsim 10^{13/(4+2\alpha)} \approx 10^2$. A more detailed discussion

(Fenimore *et al.* 1993; Woods & Loeb 1995) gives comparable limits on Γ.

Blinnikov *et al.* (1999) argued that the "non-thermal" spectrum of a GRB can be formed as a time integration of a set of thermal blackbody spectra. However, an ultra-relativistic motion is still necessary in this scenario.

In the eighties, Goodman (1986), Paczyński (1986) and Krolik and Pier (1991) realized the potential of the relativistic motion to resolve the compactness problem. There was, however, a difference between them Goodman (1986) and Paczyński (1986) considered relativistic motion in the dynamical context of fireballs, in which the relativistic motion is an integral part of the dynamics of the system, while Krolik and Pier (1991) considered a kinetic solution, in which the source moves relativistically and this motion is not necessarily related to the mechanism that produces the burst.

For a purely kinetic models, the total energy in the source frame required is at least $(Mc^2 + 4\pi FD^2/\Gamma^3 + E_{th})\Gamma$, where M is the rest mass of the source, Γ bulk Doppler factor and E_{th} the internal energy. For most scenarios that one can imagine, $\Gamma Mc^2 \gg (4\pi/\Gamma^2)FD^2$. The kinetic energy is much larger than the observed energy of the burst and the process is extremely wasteful and, generally, the energy required is so large that the model becomes infeasible.

The most energetically-economical situation is that the kinetic energy could be comparable to the observed energy and also powers the observed burst. This will be the case, if GRBs result from slowing down of ultra relativistic matter. This idea was suggested by Mészáros and Rees (1992, 1993) in the context of the slowing down of fireball accelerated materials (Shemi & Piran 1990) by the ISM and by Narayan *et al.* (1992) and independently by Rees and Mészáros (1994) and Paczyński and Xu (1994) in the context of self-interaction and internal shocks within the fireball. In both cases the introduction of energy conversion was motivated by the need to resolve the "Baryonic Contamination" problem. In the original fireball models of Goodman (1986) and of Paczyński (1986), only radiation and no baryon was considered. However, if the fireball contains even a small amount of baryons, which in reality is usually the case, all its energy will eventually be converted to kinetic energy of those baryons and no radiation directly comes out at all when the fireball becomes optically thin. This is the so called "Baryonic Contamination" problem. A mechanism was needed to recover this energy back to radiation.

Assuming that GRBs result from the slowing down of a relativistic bulk motion of massive particles, the allowed rest mass of the ultra-relativistic baryonic particles in a fireball is

$$M = \frac{\theta^2 F D_L^2}{\Gamma \epsilon_e c^2} \approx 10^{-6} M_\odot \epsilon_e^{-1} \left(\frac{\theta^2}{4\pi}\right) \left(\frac{F}{10^{-7}\,\mathrm{ergs/cm^2}}\right) \left(\frac{D_L}{3000\,\mathrm{Mpc}}\right)^2 \left(\frac{\Gamma}{100}\right)^{-1} \quad (4.3)$$

where ϵ_e is the conversion efficiency and θ is the opening angle of the emitted radiation. If the fireball is optical thick as suggested by Blinnikov *et al.* (1999), the Doppler factor

Γ could be smaller. But the allowed mass is still very small. Even though a way was found to convert back the kinetic energy of the baryons to radiation, there is still a "baryonic contamination" problem. Too much baryonic mass will slow down the flow and it will not be relativistic.

4.3 Relativistic Hydrodynamics of a Fireball under Free Expansion

4.3.1 General Description of the Hydrodynamic Evolution

Goodman (1986) and Paczyński (1986) showed that the sudden release of a large amount of gamma-ray photons into a compact region can lead to an opaque photon-electron-positron (pure radiation) fireball through the production of electron-positron pairs. In Goodman's model, a large amount of energy E_0 is suddenly release into a small volume of radius R_0. Such a situation could occur in an explosion. Paczyński (1986) considered a steady radiation and electron-positron plasma wind with rate $\dot{E}$ ejected by a compact source of size R_0 with a total energy E_0 over a time scale significantly larger than R_0/c. These two configurations display the qualitatively similar behavior. The outcome is thermal radiation. However, Shemi & Piran (1990) and Paczyński (1990) pointed out that a baryonic load will transfer all of the energy of a fireball to kinetic energy of the baryons. If the baryonic load is sufficiently small, the baryons will be accelerated to a relativistic velocity with $\Gamma \approx E_0/Mc^2$, while if it is large, the results will be a Newtonian flow with $v \simeq \sqrt{2E_0/M}$.

The evolution of a homogeneous fireball is an analogy to the Early Universe (Shemi & Piran 1990). Here and thereafter in this chapter, we will discuss the evolution of a fireball under the observer rest frame. This means that all thermal dynamic quantities (e.g. mass density, pressure, temperature and internal energy) are measured in the local fluid (comoving) frame, while coordinate related quantities (e.g. distance, time, velocity and corresponding Lorentz factor) are measured in the observer rest frame.

If the initial temperature of a pure radiation fireball is high enough, pairs will form. Because of the opacity due to pairs (Goodman 1986), the radiation cannot escape. The pairs-radiation plasma behaves like a perfect fluid with an equation of state $p = e/3$, where e is local total energy density. The fluid expands under its own pressure. As it expands, it cools with $T \propto R^{-1}$, where T is the local temperature and R the radius of the fireball measured in the observer rest frame. As the temperature drops below the pair-production threshold, the pairs annihilate. When the local temperature is around 16 keV (Goodman 1986; Paczyński 1986), the number of pairs becomes sufficiently small, the plasma becomes transparent and the photons freely escape. In the meantime, the fireball was accelerated and it is expanding relativistically outwards.

Energy conservation, as viewed from the observer frame, requires that the Lorentz factor that corresponds to this outward motion satisfies $\Gamma \propto R$. The escaping photons of local energy about 16 keV are blue shifted. An observer at rest detects them with a temperature of $T_{obs} \propto \Gamma T$. Since $T \propto R^{-1}$ and $\Gamma \propto R$, the observed temperature T_{obs} is a constant and approximately equals the initial temperature T_0. The observed spectrum is, however, almost thermal (Goodman 1986) and it is still far from the one observed in GRBs.

In addition to radiation and e^+e^- pairs, a realist astrophysical fireball should also include some baryonic matter which may be injected with the original radiation or may be present in an atmosphere surrounding the initial explosion. These baryons influence the fireball evolution in two ways. The electrons associated with this matter increase the matter opacity τ_b through Compton scattering. When the local temperature T is large, the opacity τ is dominated by e^+e^- pairs τ_p (Goodman 1986). τ_p decreases exponentially with decreasing temperature, and falls to unity, when $T = T_p \approx 0.032 m_e c^2 \approx 16\,\mathrm{keV}$ (Shemi & Piran 1990). The matter opacity τ_b, on the other hand, decreases only as R^{-2}. If at the point where $\tau_p = 1$, τ_b is still > 1, the final transition to $\tau = \tau_p + \tau_b = 1$ is delayed and occurs at a cooler temperature.

More importantly, the baryons are accelerated with the rest of the fireball and convert part of the radiation energy into bulk kinetic energy. The expanding fireball has two basic phases: a radiation dominated phase and a matter dominated phase. Initially, radiation dominates and the fluid accelerates with $\Gamma \propto R$. The fireball is roughly homogeneous in its local rest frame, but due to the Lorentz contraction its width in the observer frame Δ is $\Delta \approx R_0$, the initial size of the fireball. A transition to the matter dominated phase takes place when the mean Lorentz factor of the fireball is $\Gamma \approx E_0/Mc^2$ and the fireball has a size

$$R_\eta = R_0 \frac{E_0}{Mc^2} \approx 10^9\,\mathrm{cm}\, R_{07}\, E_{52} \left(\frac{M}{5 \times 10^{-5} M_\odot} \right)^{-1} \quad , \tag{4.4}$$

where we define $E_{52} \equiv E_0/10^{52}\,\mathrm{ergs}$ and $R_{07} \equiv R_0/10^7\,\mathrm{cm}$. After this stage, all the energy is in the kinetic energy of the matter, and the matter coasts asymptotically with a constant Lorentz factor.

The matter dominated phase is itself further divided into two sub-phases. At first, there is a frozen-costing phase in which the fireball expands as a shell with fixed radial width in its own local frame, $\sim \langle\Gamma\rangle R_0 \sim (E_0/Mc^2) R_0$. Because of Lorentz contraction the shell appears to an observer with a width $\Delta \approx R_0$. Eventually, when the size of the fireball reaches

$$R_s = \Gamma^2 \Delta \approx 10^{11}\,\mathrm{cm} \left(\frac{\Delta}{10^7\,\mathrm{cm}} \right) \left(\frac{\Gamma}{100} \right)^2 \quad , \tag{4.5}$$

variability in Γ within the fireball results in a spreading of the fireball which enters the coasting-expanding phase. In this final phase, the width of the shell grows linearly

with the size of the shell, R:

$$\Delta(R) \approx R/\Gamma^2 \approx 10^7\,\mathrm{cm}\left(\frac{R}{10^{11}\,\mathrm{cm}}\right)\left(\frac{\Gamma}{100}\right)^{-2} \quad \text{for } R > R_s \quad . \tag{4.6}$$

Initial ratio of total energy to rest mass energy, $\eta \equiv (E_0/Mc^2)$, determines the order of these transitions. Two critical values for η (Shemi & Piran 1990)

$$\eta_{pair} = \left(\frac{4^{1/6}3\sigma_T^2 E_0 \sigma T_p^4}{11^{1/6}\pi m_p^2 c^5 R_0}\right)^{1/2} \approx 2\times 10^{10}\, E_{52}^{1/2} R_{07}^{-1/2} \tag{4.7}$$

and

$$\eta_b = \left(\frac{3\sigma_T E_0}{8\pi m_p c^2 R_0^2}\right)^{1/3} \approx 2\times 10^{5} E_{52}^{1/3} R_{07}^{-2/3} \quad , \tag{4.8}$$

where $T_p \approx 16\,\mathrm{keV}$, $\sigma = 5.6705\times 10^{-5}\,\mathrm{ergs\,s^{-1}cm^{-2}K^{-4}}$ the Stefan-Boltzmann constant and m_p the proton rest mass, divide fireballs into four different types (Shemi & Piran 1990; Mészáros & Rees 1993):

- (i) **A Pair Fireball** ($\eta_{pair} < \eta$, $M < M_{pair} \approx 3\times 10^{-13} M_\odot E_{52}^{1/2} R_{07}^{1/2}$): The effect of the baryons is negligible and the evolution is of a pure photon-lepton fireball. The fireball is still dominated by its pair opacity at the time it becomes optically thin. When the temperature reaches T_p, the pair opacity τ_p drops to 1 and $\tau_b \ll 1$. At this point, the fireball is radiation dominated ($E_0 > Mc^2$) and most of the energy escapes as radiation. For the case $\eta \to \infty$ and $M_0 \to 0$, if tangled fields are present in the fireball, after the $\tau_p < 1$ there would still remains a very small fraction of pairs to providing the currents necessary to continue supporting the field and the fields continue expanding. The expansion would be stopped by the external medium.

- (ii) **Electron Dominated Opacity** ($\eta_b < \eta < \eta_{pair}$, $M_{pair} < M < M_b \approx 3\times 10^{-8} M_\odot E_{52}^{2/3} R_{07}^{2/3}$): In this case the fireball becomes optically thin before the bulk Lorentz factor saturates. In the late stages, the opacity is dominated by free electrons associated with the baryons, i.e. $\tau_p < 1$, $\tau_b > 1$. The co-moving temperature decreases far below T_p before τ reaches unity. However, the fireball continues to be radiation dominated and most of the energy still escapes as radiation.

- (iii) **Relativistic Baryonic Fireball** ($1 < \eta < \eta_b$, $M_b < M < 5\times 10^{-3} M_\odot E_{52}$): The fireball becomes matter dominated before it becomes optically thin. Most of the initial energy is converted into bulk kinetic energy of the baryons, with a final Lorentz factor $\Gamma_f \approx (E_0/Mc^2)$. We are mainly interested in this situation.

- (iv) **Newtonian Fireball** ($\eta < 1$, $M > 5\times 10^{-4} M_\odot E_{52}$): This is the Newtonian regime. The rest energy exceeds the radiation energy and the expansion never

becomes relativistic. This is the situation, for example, in supernova explosions in which the energy is deposited into a massive envelope.

This summary describes the qualitative features of a roughly homogeneously expanding fireball. Similar scaling laws exists also for inhomogeneous fireballs (Piran *et al.* 1993) as well as for relativistic winds (Paczyński 1986). For a spherical fireball with an arbitrary radial distribution of radiation and matter and at rest at the initial time, under optically thick conditions the radiation, the relativistic leptons and the matter at each radius behave like a single fluid, moving with the same velocity. The pressure p and the energy density e are related by the equation of state $p = e/3$. The relativistic conservation equations of baryon number, energy and momentum are usually studied using characteristic coordinates: r and $s \equiv t - r$ (Piran *et al.* 1993). The mass density ρ, the total energy density e and the pressure p are measured in the local frame of the fluid, but the spherical coordinate r and time t are in the observer frame, fixed with its zero point at the center of the fireball at the initial time. At the early stages of the relativistic fireball expansion, the fluid has a very rapid acceleration phase, which is usually studied numerically (Mészáros *et al.* 1993; Mészáros & Rees 1993; Piran *et al.* 1993). A self-similar solution to this stage was given by Bisnovatyi-Kogan & Murzina (1995). After a short acceleration phase, we expect that the motion of a fluid shell will become highly relativistic ($\Gamma \gg 1$). If we restrict our attention to the evolution of the fireball from this point on, we may treat Γ^{-1} as a small parameter and set $\Gamma \approx u$ accurate to order $O(\Gamma^{-2})$, where $u = \sqrt{\Gamma^2 - 1}$ is the four velocity in radial direction. Then, under a wide range of conditions, we have the following conservation laws for each fluid shell along a characteristic moving outward at the speed of light (Piran *et al.* 1993):

$$r^2 \rho \Gamma = const., \qquad r^2 e^{3/4} \Gamma = const., \qquad r^2 \left(\rho c^2 + \frac{4}{3} e \right) \Gamma^2 = const.. \tag{4.9}$$

Let t_0 be the time when the fireball first becomes ultra-relativistic, with $\Gamma \gtrsim$ a few. Eq. 4.9 gives

$$r = r_0 \frac{\Gamma_0^{1/2} D^{3/2}}{\Gamma^{1/2}}, \qquad \rho = \frac{\rho_0}{D^3}, \qquad e = \frac{e_0}{D^4}, \tag{4.10}$$

with the definition of D

$$\frac{1}{D} \equiv \frac{\Gamma_0}{\Gamma} + \frac{3\Gamma_0 \rho_0 c^2}{4 e_0 \Gamma} - \frac{3\rho_0 c^2}{4 e_0} \quad , \tag{4.11}$$

where subscript 0 denote the quantities at time t_0. The solution given by Eq. 4.10 is valid both for the radiation-dominated and matter-dominated phases of the fireball within the frozen pulse approximation.

4.3.2 The Radiation-Dominated Phase of Fireball

Initially the fireball is radiation dominated and $\Gamma \ll (e_0/\rho_0)\Gamma_0$. The first term in Eq. 4.11 dominates and we have $D \propto r$, $\Gamma \propto r$, recovering the radiation-dominated scaling (along a characteristic moving at speed of light):

$$\Gamma \propto r, \qquad \rho \propto r^{-3}, \qquad e \propto r^{-4} \quad . \tag{4.12}$$

The scalings of ρ and e given in Eq. 4.12 correspond to those of a fluid expanding uniformly in the comoving frame. Indeed, all three scalings in Eq. 4.12 can be derived for a homogeneous radiation dominated fireball by noting the analogy with an expanding universe.

Although the fluid is approximately homogeneous in its own frame, because of Lorentz contraction it appears as a narrow shell in the observer frame, with a radial width given by $\Delta r \sim r/\Gamma \sim constant \sim R_0$, the initial radius of the fireball, or the initial width of the specific shell under discussion, when we consider a continuous wind. The conservation laws given in Eq. 4.9 and Eq. 4.12 are valid as long as the radiation-dominated fireball expands ultra-relativistically with large Γ. The only possible exception is in the very outermost layers of the fireball where the pressure gradient may be extremely steep and the scaling law cannot be applied. If this deviation is ignored, the constancy of the radial width Δr in the observer frame and the scaling law means that the fireball behaves like a pulse of energy with a frozen radial profile, accelerating outward at almost the speed of light (Piran *et al.* 1993; Mészáros *et al.* 1993; Kobayashi *et al.* 1999).

4.3.3 The Matter Dominated Phase of Fireball

The radiation-dominated regime extends to a radius $r \sim (e_0/\rho_0 c^2)r_0$. At large radii, the first and last terms in Eq. 4.11 become comparable and Γ tends to its asymptotic value of $\Gamma_f = (1 + 4e_0/(3\rho_0 c^2))\Gamma_0$. This is the matter dominated regime. The transition occurs when $4e/3 = \rho c^2$, which happens when $\Gamma = \Gamma_f/2$. In this regime, $D \propto r^{2/3}$, leading to the scalings (along a characteristic moving at speed of light):

$$\Gamma \to constant, \qquad \rho \propto r^{-2}, \qquad e \propto r^{-8/3}. \tag{4.13}$$

Since $e \ll \rho c^2$, the radiation has no important dynamical effect on the motion and produces no significant radial acceleration. Therefore, Γ remains a constant on streamlines and the fluid coasts with a constant asymptotic radial velocity. Of course, since each shell moves with a velocity that is slightly less than c and that is different from one shell to the next, the frozen pulse approximation on which Eq. 4.9 is based must ultimately break down at some large radius. The modified scalings of ρ and e arise because the fireball now moves with a constant radial width in the comoving frame.

At very large late times in the matter-dominated phase, the frozen-pulse approximation begins to break down (Piran *et al.* 1993; Kobayashi *et al.* 1999). In this stage, the radiation density e is much smaller than the matter density ρ, and the Lorentz factor Γ tends to a constant value Γ_f for each shell. The flow moves strictly along the characteristic $\beta_f t - r = constant$ so that each fluid shell coasts at a constant radial speed $\beta_f = u_f/\Gamma_f$ (Piran *et al.* 1993).

Let t_c and r_c represent the time and radius at which the coasting phase begins, which corresponds essentially to the point at which the fluid makes the transition from being radiation dominated to matter dominated. The width of the pulse at time t can be given as $\Delta r(t) \approx \Delta r_c + [(\partial\beta_f/\partial R)\Delta R](t-t_c) \approx \Delta r_c + \Delta\Gamma_f(t-t_c)/\langle\Gamma_f\rangle^3 \approx R_0 + (t-t_c)/\langle\Gamma_f\rangle^2$, where $\Delta r_c \sim R_0$ is the width of the fireball when it begins coasting, $\langle\Gamma_f\rangle$ is the mean Γ_f in the pulse, and $\Delta\Gamma_f \sim \langle\Gamma_f\rangle$ is the total spread of Γ_f across the pulse. From this results, we see that within the matter dominated coasting phase there are two separate regimes. So long as $t - t_c < \langle\Gamma_f\rangle^2 R_0$, we have an approximately frozen-coasting phase in which Δr is approximately constant and the frozen pulse approximation is valid. In this regime, the scalings in Eq. 4.13 are satisfied. However, when $t - t_c > \langle\Gamma_f\rangle^2 R_0$, the fireball switches to an expanding-coasting phase where $\Delta r \propto t - t_c$ and the pulse width grows linearly with time. In this regime, the scaling of ρ reverts to $\rho \propto r^{-3}$, and, if the radiation is still coupled to the matter, $e \propto r^{-4}$.

4.3.4 Optical Depth of Fireball

Independently of the above considerations, at some point during the expansion, the fireball will become optically thin. For a pure fireball, this happens when the local temperature drops to about 16 keV at (Shemi & Piran 1990)

$$R_{pair} = \left(\frac{3E_0}{4\pi R_0^3 a T_p^4}\right)^{1/4} R_0 \approx 10^{10}\,\mathrm{cm}\, E_{52}^{1/4} R_{07}^{1/4} \quad , \tag{4.14}$$

where $R_{07} = R_0/10^7\,\mathrm{cm}$ and $a = 4\sigma/c = 7.5659 \times 10^{-15}\,\mathrm{ergs\,cm^{-3}K^{-4}}$ the radiation energy density constant.

At the "photosphere", however, $\tau \simeq 1$, energy will be radiated freely. As long as the interior temperature is high enough to support pairs, the speed of the photosphere with respect to the fluid must be less than the blackbody value. This is less than the sound speed of the radiation fluid, so that a rare-faction wave will propagate inward faster than the photosphere. The radius of the photosphere will actually increase at speeds close to c, until the temperature drops below $m_e c^2$. The fluid will thereafter become optically thin very abruptly to pair productions (Goodman 1986).

In a matter dominated fireball, the optical depth is usually determined by the ambient electrons associated with the baryons. In this case, the fireball becomes optically

thin at

$$R_e = \left(\frac{\sigma_T E_0}{4\pi m_p c^2 \eta}\right)^{1/2} \approx 6 \times 10^{13}\,\mathrm{cm}\, E_{52}^{1/2} \left(\frac{\eta}{100}\right)^{-1/2} \quad . \tag{4.15}$$

From this stage on, the radiation and the baryons are decoupled and move with different velocities. Any remaining radiation will escape freely and the radiation pressure vanishes. The baryon shells will coast with their own individual velocities. If the fireball is already in the matter dominated coasting phase, there will be no change in the propagation of the baryons. However, if the fireball is in the radiation dominated phase when it becomes optically thin, the baryons will switch immediately from accelerating phase to a coasting phase. This transition radius R_e has another crucial role in the fireball evolution. It is the minimal radius in which energy conversion and generation of the observed GRB can begin. Photons produced at $R < R_e$ cannot escape.

4.3.5 Anisotropic Fireball

It is unlikely that a realistic fireball will be spherically symmetric. The radiation will be beamed into a cone Γ^{-1}, when the matter has been accelerated along a very narrow beam with opening angle $\theta < \Gamma^{-1}$. The opening angle can also have any intermediate value, if it emerges from a beam with an opening angle $\theta > \Gamma^{-1}$, as will be the case if the source is an anisotropic fireball or an electro-magnetic accelerator with a modest beam width. The initial motion of a non-spherically symmetric fireball might be fairly complex but once $\Gamma \gg 1$ the motion of each fluid element decouple from the motion of its neighbors with angular distance larger than Γ^{-1}. This motion can be described by the same asymptotic solution, as if it is a part of a spherical shell.

4.4 Temporal Structure and Angular Spreading Problem

It turns out recently that the observed temporal structure may set a strong constraint on the energy conversion models (Fenimore *et al.* 1996a, 1999; Sari & Piran 1997). The external shock model may face difficulties in explaining short-time-scale variability in GRB light curves as a consequence of special relativistic effects resulting from the curvature of the blast wave shell. This argues that internal shocks rather than external shocks are the mechanism to convert kinetic into thermal energy and radiate it. We discuss this issue in this section then the models on external and internal shocks in next sections.

4.4.1 Time Scales

Several time-scales from Fig 4.1 are usually defined in literatures. An emphasis should be made on the difference between arrival time at the detector, T, and coordinate time,

t, (equivalent to laboratory time, the detector rest frame time, and the rest frame of the central explosion). Coordinate time is measured by clocks placed at all locations within the frame when the photons are produced. In contrast, a detector measures when the photons arrive at a single location. These two times are related as $T = (1 - \beta\cos\theta)t$, where θ is the angle between the direction of motion of the emission region and the direction to the detector. In contrast, the Lorentz transformation between the time in the rest frame of the detector (t) and time in a frame moving with the shell (t') is $t = \Gamma t'$.

- **The Radial Time Scale** T_{radial}: For an infinitely thin relativistic emitting shell with a Lorentz factor Γ_E (the subscript E for the emitting region) with a typical R_E, the observed duration between the first photon emitted at R_E and last one emitted at $R_E + \Delta R_E$ is (Ruderman 1975; Katz 1994a):

$$T_{radial} = t_C - t_A = (R_E + \Delta R_E)(\frac{1}{v} - \frac{1}{c}) - R_E(\frac{1}{v} - \frac{1}{c}) \cong R_E/2\Gamma_E^2 c \quad . \quad (4.16)$$

- **The Angular Time Scale** T_{ang}: One can receive photons only up to solid angle of Γ_E^{-1} from the line of sight due to relativistic beaming. Two photons one on the line of sight and the other at an angle of $\bar{\theta} \sim \Gamma_E^{-1}$ away emitted at the same time and radius R_E arrive with a time delay by (Ruderman 1975; Katz 1994a; Fenimore *et al.* 1996a):

$$T_{ang} = t_D - t_A = \frac{R_E}{c}(1 - \cos\bar{\theta}) \cong R_E/2\Gamma_E^2 c \quad . \quad (4.17)$$

- **Intrinsic Duration** ΔT: The duration of the flow. This is simply the time in which the source that produces the relativistic flow is active. $\Delta T = t_B - t_A = \Delta/C$, where Δ is the width of the relativistic wind, measured in the observer rest frame. The observed duration of the burst must be longer or equal to Δ/c.

- **Intrinsic variability** δT: The time scale on which the inner source varies and produces a subsequent variability with a length scale $\delta = c\delta T$ in the flow. Naturally, δT sets a lower limit to the variability time scale observed in any burst. Δ and δ must satisfy

$$\delta \leq \Delta \quad . \quad (4.18)$$

- **The Cooling Time Scale** T_{cool}: The observed time delay, while the shocked material cools. The relation to the local cooling time e/P in the fluid rest frame is

$$T_{cool} \cong [\Gamma_E(e/P) - \Gamma_E(e/P)\frac{v}{c}] \cong e/P\Gamma_E \quad , \quad (4.19)$$

where e is the internal energy density and P is the power radiated per unit volume. It differs from the usual coordinate time dilation which gives $\Gamma_E e/P$.

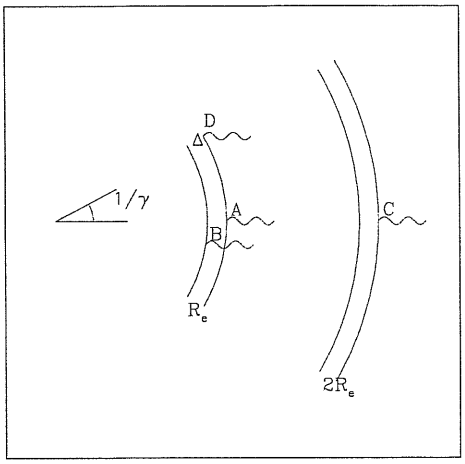

Figure 4.1: Different time scales in term of the arrival time of four photons: t_A, t_B, t_C, and t_D. $T_{radial} = t_C - t_A$; $T_{angular} = t_D - t_A$, $\Delta/c = t_B - t_A$. The time when the shell is at the center is taken as the initial time.

For synchrotron cooling, $T_{cool}(\nu) \propto \nu^{-1/2}$ (Sari *et al.* 1996) and quite generally is shorter than the hydrodynamics time scales (Sari *et al.* 1996; Katz 1997), except for the late stages of an afterglow when T_{cool} becomes the longest time scale in the system.

4.4.2 The Angular Spreading Problem and External Shocks

As long as the shell angular width is large than Γ^{-1}, any temporal structure due to irregularities in the properties of the shell or in the encountered material will be spread on a time given by T_{ang}. This means that T_{ang} is the minimal time scale for the observed temporal variability: $\delta T \geq T_{ang}$. So, we may define two types of models:

$$T = \begin{cases} T_{ang} = R_E/c\Gamma_E^2 & \text{if } \Delta < R_E/\Gamma_E^2 \quad \text{(Type-I);} \\ \Delta/c & \text{otherwise} \qquad\quad \text{(Type-II).} \end{cases} \tag{4.20}$$

In Type-I models, the duration of the burst is determined by the emission radius and the Lorentz factor. It is independent of Δ. In Type-II models, the duration of the burst is determined by the thickness of the relativistic shell Δ that is by the duration that the source is active and produces the relativistic wind. The angular spreading time is shorter and therefore irrelevant.

The majority of GRBs have a complex temporal structure (see discussion in Section 2.2) with $N \equiv T/\delta T$ of order 100. For a Type-I model, angular spreading means that at any given moment the observer sees a whole region of angular width Γ_E^{-1}. Any variability in the emission due to different conditions in different radii on time scale smaller than T_{ang} is erased unless the angular size of the emitting regions is smaller than Γ_E^{-1}. Thus, such a source can produce only a smooth single humped burst with

$N = 1$ and no temporal structure on a time scale $\delta T < T$. In other words a shell, of a Type-I model, and with an angular width larger than Γ_E^{-1} cannot produce a variable burst with $N \gg 1$. This is the angular spreading problem. Type-I models include the standard "external shock model" (Mészáros & Rees 1992a; Katz 1994a; Sari & Piran 1995) in which GRBs are produced when the relativistic shell is decelerated by the ISM, the relativistic magnetic model (Usov & Smolsky 1996) in which a magnetic Poynting flux runs into the ISM, or the scattering of star light by a relativistic shell (Shemi 1994; Shaviv & Dar 1995).

On the other hand, a Type-II model contains a thick shell $\Delta > R_E/\Gamma_E^2$ and it can produce a variable burst. The variability time scale, is again limited $\delta T > T_{ang}$ but now it can be shorter than the total duration T. The duration of the burst reflects the time that the "inner engine" operates. The variability reflects the radial inhomogeneity of the shell which was produced by the source or the cooling time if it is longer than δ/c. The observed temporal variability provides an upper limit to the scale of the radial "inner engine" varies. This is a remarkable conclusion in view of the fact that the fireball hides the "inner engine". The Type-II models include the "internal shock model" (Narayan *et al.* 1992; Rees & Mészáros 1994; Paczyński & Xu 1994), in which different parts of the shell are moving with different Lorentz factor and therefore collide with one another. A magnetic dominated version is given by Thompson (1994). Type-II models have two weakness that internal shocks are rather inefficient and the long, complex time history of a GRB must be postulated at the central site.

It is possible for an external shock to give rise a Type-II behavior, if we could set the parameters of the external shock model to satisfy $R_E \leq 2\Gamma_E c\delta T$. As shown by Sari & Piran (1995), the deceleration radius R_E (as indication c.f. Eq. 4.38 for R_N or Eq. 4.40) for a thin shell is decreasing with an increasing initial Lorentz factor Γ. However, as Γ increases above a critical value $\Gamma \geq \Gamma_c$, the shell can no longer be considered thin. In this situation, the reverse shock penetrating the shell becomes ultra-relativistic and the shocked matter moves with Lorentz factor $\Gamma_E = \Gamma_c < \Gamma$ which is independent of the initial Lorentz factor of the shell Γ. The deceleration radius is now a constant independent of the initial Lorentz factor of the shell. This emission radius R_E is always larger than Δ/Γ_E^2 – thus an external shock cannot be of Type-II (Sari & Piran 1995).

One possibility for external shocks to produce Type-II temporal structures is that the emitting regions are significantly narrower than Γ_E^{-1}. In the case, the angular size of the emitting region is limited by (Sari & Piran 1997)

$$\delta\theta \leq \frac{c\,\delta T}{R_E\bar{\theta}} = \frac{c}{R_E\bar{\theta}}\left(\frac{1}{N}\frac{R_E}{2\Gamma_E^2 c}\right) = \frac{1}{2N\Gamma_E} \quad , \tag{4.21}$$

and the corresponding size size $r_s = R_E\delta\theta$ is

$$r_s \leq \delta T c/\bar{\theta} \quad . \tag{4.22}$$

Sari & Piran (1997) takes $\bar{\theta} \sim \Gamma_E^{-1}$ as most of the emitting regions have this value. The $\bar{\theta}$ used by Fenimore *et al.* (1996a) is incorrect.

Therefore, one can produce GRBs with $T \approx T_{radial} \approx R_E/c\Gamma_E^2$ and $\delta T = T/N$, if the emission regions have angular size smaller than $1/N\Gamma_E \approx= 10^{-4}$. A narrow jet with such an extremely small opening angle would be able to produce the observed variability. Such a jet must be extremely cold in its local frame, otherwise its internal pressure will cause it to spread. It is not clear what could produce such a jet. Additionally, for the temporal variability to be produced, either a rapid modulation of the jet or inhomogeneities in the ISM are needed. This requires a very small energy $E \sim 10^{49}$ ergs (Fenimore *et al.* 1999).

The second possibility is that the local spherical symmetry of the shell is broken into many emitting regions on a scale smaller than Γ^{-1} due to some reasons (e.g. ISM dense clouds distributing in the space, star light or fragments of the shell, etc), where only a small fraction of the surface converts its energy into photons. However, whether this scenario is extremely inefficient is still under debate (Sari & Piran 1997; Fenimore *et al.* 1999; Dermer & Mitman 1999). Fenimore *et al.* (1999) explore several other ways to resolve the angular spreading problems. But none seems promising.

4.4.3 Temporal Structure in Internal Shocks

Type-II behavior arises naturally in the internal shock model (Fenimore *et al.* 1996a; Sari & Piran 1997). In this model, different shells have different Lorentz factors. Latter shell with higher Lorentz factor overtakes and collides with former shell with lower Lorentz factor. The collision convert some of their kinetic energy into internal energy and radiate. If the emission radius is sufficiently small, angular spreading will not erase the temporal variability. This requires $R_E \leq 2\Gamma_E^2\, c\,\delta T$. This condition is always satisfied as internal shocks take place at $R_E = R_s \approx \Gamma_E^2 \delta$ with $\Gamma_E \simeq \Gamma$. Since $\delta < \Delta$, we have $T = \Delta/c > T_{ang} = R_E/2\Gamma_E^2 c = \delta/c = \delta T$. Clearly multiple shells are needed to account for the observed temporal structure. Each shell produces an observed peak of duration δT and the whole complex of shells of width Δ, produces a burst that lasts $T = \Delta/c$. The angular spreading time is comparable to the temporal variability produced by the "inner engine". They determine the observed temporal structure provided that they are longer than the cooling time.

Before concluding that internal shocks can actually produce GRBs, we must address two issues. First, can internal shocks produce the observed variable structure? Second, can it be done efficiently? We will address the first question here and the second one in Section 4.6.4 where we discuss energy conversion in internal shocks.

Kobayashi *et al.* (1997) construct a simple model to calculate the temporal profile produced by an internal shocks. In this model, the relativistic flow is characterized by

multiple shells of width d and a initial separation δ. The shells are assigned random Lorentz factors of order Γ and random density or energy. The time scale depends on the cooling time, the hydrodynamic time, and the angular spreading time. As the internal energy is radiated via synchrotron emission and inverse Compton scattering and in most of cases, the electron cooling time is much shorter than the hydrodynamic time scale (Sari *et al.* 1996; Katz 1994a), only the hydrodynamic time and the angular spreading time have been considered.

The hydrodynamic time scale is determined by the time that the shock crosses the shell, whose width is d. In fact, there are two shocks: a forward shock and a reverse shock. Detailed calculations (Kobayashi *et al.* 1997) reveals that this time scale (in the observer rest frame) is of order of the light crossing time of the shell d/c.

A collision takes place at $\Gamma^2\delta$ and the angular spreading yields an observed pulse of width $\sim \delta/c$. If $\delta > d$, the overall pulse width δT is determined by angular spreading. The shape of the pulse become asymmetric with a fast rise and a slower decline which GRBs typically show (see Section 2.2).

Typical synthetic temporal profile shows that internal shocks can produce the observed highly variable temporal profiles, provided that the source of the relativistic particles, the "inner engine", produces a sufficiently variable wind (Kobayashi *et al.* 1997). Long bursts require relativistic winds that last hundred seconds, with a rapid variability on a time scale of a fraction of a second. Thus, internal shock models provide direct information on the "inner engine" of GRBs: It produces the observed complicated temporal structure. This severely constrains numerous models.

4.5 Slowing Down of Relativistic Motion

4.5.1 Shock Conditions

The cross section for a direct electro-magnetic or nuclear interaction between the relativistic particles and the ISM particles is too small to convert efficiently the kinetic energy to radiation. GRBs are the relativistic analogues of SNRs. Like in supernova remnants (SNRs),one can expect that collisionless shocks will form in GRBs (Mészáros & Rees 1992a; Katz 1994a) and the particles in a fireball can effectively be slowed down via the collisionless shocks.

Simply considering the collision and merger of a rapid point object of rest mass m_r and a slower one of rest mass m_s into a single mass denoted by m, from energy and momentum conservations we can obtain two interesting limits. The first is for external shocks or a shock between relativistic ejecta and a non-relativistic material which is ejected from the source before it exploded. In this case, m_s is at rest $\Gamma_s = 1$ and a

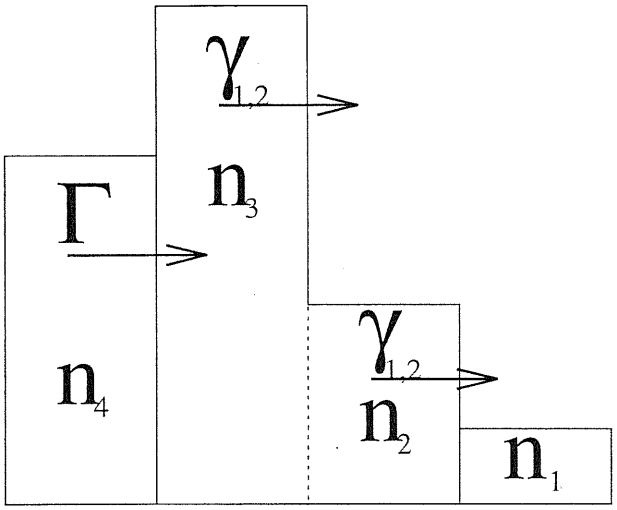

Figure 4.2: The definition of different regions in a interaction via a shock

mass:

$$m_s \approx m_r/\Gamma_r \ll m_r \tag{4.23}$$

is needed to yield $\Gamma_m \approx \Gamma_r/2$ and $e \approx m_r c^2/2$. The external mass needed to converted half of the kinetic energy is smaller than the original mass by a factor of Γ_r (Mészáros & Rees 1992a; Narayan *et al.* 1992).

The second case corresponds to an internal collision between shells that are moving at different relativistic velocities: $\Gamma_r \gtrsim \Gamma_s \gg 1$. The conversion efficiency of kinetic energy into internal energy in internal collision is (Kobayashi *et al.* 1997; Katz 1997):

$$\epsilon = 1 - \frac{(m_r + m_s)\Gamma_m}{(m_r\Gamma_r + m_s\Gamma_s)} \quad . \tag{4.24}$$

A conversion of a significant fraction of the initial kinetic energy to internal energy requires that the difference in velocities between the shells will be significant: $\Gamma_r \gg \Gamma_s$ and that the two masses will be comparable $m_r \approx m_s$.

As the interaction time-scale of the shell with the ISM may be relatively long, one has to turn to the hydrodynamics of the interaction and takes into account the shocks formed during the collision.

When a cold shell collides with another cold shell or moves into the cold ISM, two shocks form: an forward shock propagating into the ISM or into the external shell and a reverse shock propagating into the inner shell, with a contact discontinuity between the shocked material (Kobayashi *et al.* 1999). The shock structure is determined by the Lorentz factor Γ of the motion of the inner shell (denoted by 4) relative to the outer one - or the ISM (denoted by 1) and the ratio of the particle number densities between these regions, $f \equiv n_4/n_1$ (see Fig. 4.2).

Three cases are interesting:

(i) Ultra-relativistic shock ($\Gamma \gg 1$) **with** $f > \Gamma^2$. This configuration is called "Newtonian", as the reverse shock is non-relativistic (or mildly relativistic). In this case, the energy conversion takes place in the forward shock. Let Γ_2 be the Lorentz factor of the motion of the shocked fluid relative to the rest frame of the fluid at 1 (an external observer for interaction with ISM and the outer shell in case of internal collision) and let $\bar{\Gamma}_3$ be the Lorentz factor of the motion of this fluid relative to the rest frame of the relativistic shell 4. Then the Lorentz factors, particle and energy density (n, e) in the shocked regions are

$$\Gamma_2 \approx \Gamma; \qquad \bar{\Gamma}_3 \approx 1, \tag{4.25}$$

$$n_2 \approx 4\Gamma n_1; \qquad e \equiv e_2 = 4\Gamma^2 n_1 m_p c^2; \qquad n_3 = 7n_4; \qquad e_3 = e. \tag{4.26}$$

(ii) Late stage with $f < \Gamma^2$. During the propagation of the shell, the density ratio decreases. Both the forward and the reverse shocks are relativistic. The shock conditions give for regions 1 and 2 (Blandford & McKee 1976)

$$\Gamma_2 = f^{1/4}\Gamma^{1/2}/\sqrt{2}; \qquad n_2 = 4\Gamma_2 n_1; \qquad e \equiv e_2 = 4\Gamma^2 n_1 m_p c^2 \tag{4.27}$$

and for the reverse shock

$$\bar{\Gamma}_3 = f^{-1/4}\Gamma^{1/2}/\sqrt{2}; \qquad n_3 = 4\bar{\Gamma}_3 n_4. \tag{4.28}$$

In addition, we have $e_3 = e$ and $\bar{\Gamma} \cong (\Gamma/\Gamma_2 + \Gamma_2/\Gamma)/2$ which follows from the equality of pressures and velocity on the contact discontinuity. Comparable amounts of energy are converted to thermal energy in both shocks when both are relativistic, while only a negligible amount of energy is converted to thermal energy in the reverse shock if it is Newtonian (Sari & Piran 1995).

(iii) Internal shocks with $f \approx 1$ **and a relative Lorentz factor of order of a few.** Both shells have similar densities. In this case, for an adiabatic index of 4/3 we have

$$\Gamma_2 = \sqrt{(\Gamma^2 + 1)/2}; \qquad n_2 = (4\Gamma_2 + 3)n_1; \qquad e_2 = \Gamma_2 n_2 m_p c^2 \tag{4.29}$$

Both shocks are mildly relativistic and their strength depends on the relative Lorentz factors of the two shells.

The relativistic shocks are characterized by Γ_{sh} that describes the shock velocity as well as the "thermal" motion of the shocked particles. It is measured relative to a rest frame in which the unshocked material is at rest. The Lorentz factor of the forward shock is usually different from the Lorentz factor of the reverse shock. The emitting region – the shocked material – moves relativistically relative to an observer at rest at infinity. This is characterized by a Lorentz factor Γ_E.

4.5.2 Internal Shock

The arguments presented in Section 4.4 will show external shocks cannot explain the complex observed temporal structures in some GRBs, while internal shocks probably provide the best way.

If an inner shell with a Lorentz factor $\Gamma_r \sim \Gamma$ is faster than an outer shell with Lorentz factor $\Gamma_s \sim \Gamma$, in the observer rest frame, the inner shell will overtake the outer one at

$$R_\delta \approx \Gamma^2 \delta \approx 10^{14}\,\mathrm{cm}\,\delta_{10}(\Gamma_{100})^2 \quad , \tag{4.30}$$

where $\delta = \delta_{10} \times 10^{10}\,\mathrm{cm}$ is the initial separation between the shells in the observer rest frame and we have assumed that $\Gamma_r - \Gamma_s \sim \Gamma = 100\Gamma_{100}$. Internal shocks are relevant only if they appear before the external shock. From the discussion in section 4.6.5, this requires

$$\xi^{3/2} > \zeta \quad , \tag{4.31}$$

where the dimensionless parameter xi (ξ) is given by Eq. 4.37 in Sec. 4.5.3.1 and the parameter zeta (ζ) is defined to characterize the variability of the flow and

$$\zeta \equiv \frac{\delta}{\Delta} \leq 1 \quad . \tag{4.32}$$

From the discussion in Sec. 4.4.3, for internal shocks the duration of the burst $T \approx \Delta/c$ and the duration of individual spikes $\delta T \approx \delta/c$. Atypical value for ζ is $\zeta = \delta/\Delta \approx \delta T/T \approx 0.01$.

The condition of Eq. 4.31 can be turned into a condition for Γ

$$\Gamma \leq 2.8 \times 10^3 \left(\frac{\zeta}{0.01}\right)^{-1/2} \left(\frac{T}{10\,\mathrm{sec}}\right)^{-3/8} \left(\frac{l}{10^{18}\,\mathrm{cm}}\right)^{3/8} \quad , \tag{4.33}$$

where the Sedov length l is defined in Eq. 4.36 and relation $T = \Delta/c$ is used. It follows that internal shocks take place in relatively "low" Γ regime.

The optical thinness required to Compton scattering of the photons on the shell implies that from Eq. 4.15, Eq. 4.30 and the condition $R_e \leq R_\delta$

$$\Gamma \geq \left(\frac{E_0 \sigma_T}{4\pi c^2 \delta^2 m_p}\right)^{1/5} = 130\, T_{10}^{-2/5} \zeta_{0.01}^{-2/5} E_{52}^{1/5} \quad . \tag{4.34}$$

Requiring the emission region optically thin ($\tau_{\gamma\gamma} < 1$) to pair productions $\gamma\gamma \to e^+e^-$ leads (Fenimore *et al.* 1993; Woods & Leob 1995)

$$\Gamma > 570\,(\zeta_{0.01} T_{10})^{-1/4} \quad , \tag{4.35}$$

where $\zeta_{0.01} = \zeta/0.01$, $T_{10} = T/10\,\mathrm{sec}$ and we have used Eq. 4.30 for R_δ. This constraint due to the $\gamma\gamma$ interaction is generally more important than the constraint due to Compton scattering as generally $\tau_{\gamma\gamma} > \tau_e$.

Three conclusions can be drawn. First, if the spectrum of the observed photons extends beyond 100 MeV (as was the case in the bursts detected by EGRET) and if those high energy photons are emitted in the same region as the low energy ones, the condition on the pair production, $\tau_{\gamma\gamma}$, Eq. 4.35 is stronger than the condition on Compton scattering Eq. 4.34. This increases the required Lorentz factors. Second, the Compton scattering limit poses also a lower limit on Γ. However, this is usually less restrictive than the limit from pair production. Finally, in the (δ, Γ) plane, optically thin internal shocks are produced only in a narrow region (Sari & Piran 1997a). The region is so small that there is no single value of Γ that produces peaks with all durations. The internal shock scenario suggests that bursts with very narrow peaks should not have very high energy tails and that very short burst may have a softer spectrum.

4.5.3 External Shocks with the ISM

4.5.3.1 Newtonian *vs* Relativistic Reverse Shocks

A relativistic shell will interact with the ISM and produce external shocks. The interaction between a relativistic flow and an external medium depends, like in SNRs, on the Sedov length within which the ISM rest mass energy equals the energy E_0 of the GRB

$$l \equiv \left(\frac{3E_0}{4\pi n_{ism} m_p c^2}\right)^{1/3} \simeq 10^{18}\,\mathrm{cm}\left(\frac{E_{52}}{n_{1ism}}\right)^{1/3} , \tag{4.36}$$

where $n_{1ism} \equiv n_{ism}/1\,\mathrm{particle/cm^3}$.

There are two possible types of external shocks (Sari & Piran 1995; Kobayashi *et al.* 1999), characterized according to the nature of the reverse shock: Newtonian Reverse Shock (NRS) vs Relativistic Reverse Shock (RRS). If the reverse shock is relativistic (RRS), then it reduces significantly the kinetic energy of each layer that it crosses. Each layer within the shell loses its energy independently from the rest of the shock. The energy conversion process is over once the reverse shock crosses the shell (the *left panel* of Fig. 4.3). A Newtonian or even mildly relativistic reverse shock (NRS) is comparatively weak. Such a shock reduces the energy of the layer that it crosses by a relatively small amount. Significant energy conversion takes place only after the shock has crossed the shell several times (the *right panel* of Fig. 4.3). The shell behaves practically like a single object and it loses its energy only by the time that it accumulates an external mass equal to M/Γ.

The question which scenario is taking place depends on the parameters of the shell relative to the parameters of the ISM, actually on a combined dimensionless parameter ξ constructed from l, Δ and Γ (Sari & Piran 1995)

$$\xi \equiv (l/\Delta)^{1/2}\Gamma^{-4/3} . \tag{4.37}$$

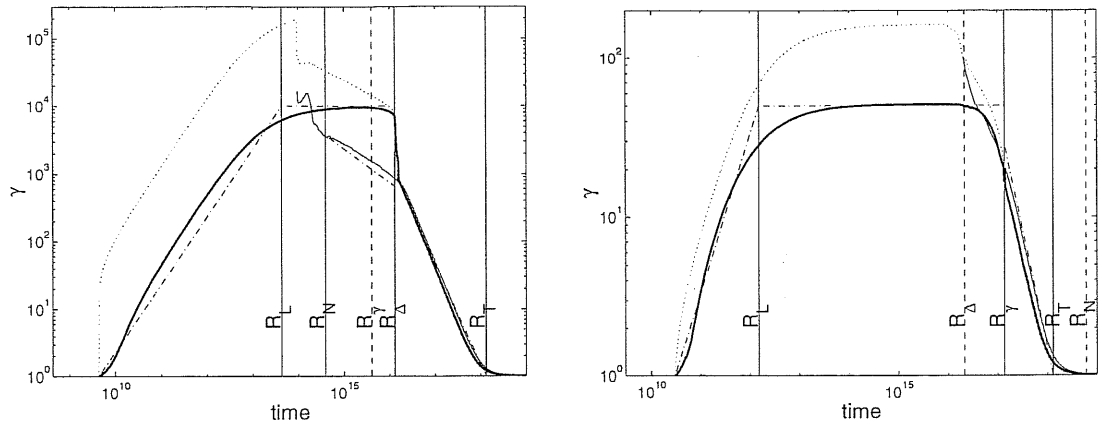

Figure 4.3: Fireball evolution from its initial formation to the final Newtonian Sedov solution. The energy is extracted due to the interaction with the ISM via a forward shock and a reverse shock, if relativistic. The thick solid line is for the average value of the Lorentz factor, thin solid line for the value in forward shock, dotted line for the maximal value, dashed dotted line for an analytic estimate. (*left*) for a relativistic reverse shock. The used parameters are: $\xi = 0.1$, $E_0 = 10^{52}$ ergs, $\gamma_0 = 10^4$, $R_0 = 4.3 \times 10^9$ cm. (*right*) for a Newtonian reverse shock. The used parameter are $\xi = 43$, $E_0 = 10^{52}$ erg, $\Gamma_0 = 50$ and $R_0 = 3 \times 10^{10}$ cm. From Kobayashi *et al.* (1999).

As the shell propagates outwards, it is initially very dense and the density ratio between the shell and the ISM $f \equiv n_4/n_1$ is extremely large (more specifically $f > \Gamma^2$) (see Sec. 4.5.1 for the definition of different regions). The reverse shock is initially Newtonian. Such a shock converts only a small fraction of the kinetic energy to thermal energy. As the shell propagates, the density ratio, f, decreases (like R^{-2} if the width of the shell is constant and like R^{-3} if the shell is spreading). Eventually, the reverse shock becomes relativistic at R_N where $f = 1$. The question where the kinetic energy is converted depends on whether the reverse shock reaches the inner edge of the shell before or after it becomes relativistic.

There are four different radii that should be considered. The following estimates assume a spherically symmetric shell of initial energy E_0 and rest mass M. The reverse shock becomes relativistic at R_N, where $f = n_4/n_1 = 1$:

$$R_N = l^{3/2}/\Delta^{1/2}\Gamma^2 \quad . \tag{4.38}$$

Using the expression for the velocity of the reverse shock into the shell (Eq. 4.28), we find that the reverse shock crosses the shell at R_Δ (Sari & Piran 1995)

$$R_\Delta = l^{3/4}\Delta^{1/4} \quad . \tag{4.39}$$

A third radius is R_γ, where the shell collects ISM mass of M/Γ (Katz 1994a; Mészáros

R_0	Initial Radius	$c\,\delta t$	$\approx 10^7 - 10^8$cm
R_η	Transition to MD	ηR_0	$\approx 10^9$cm
R_{pair}	Optically thin to pairs	$(3E_0/4\pi R_0^3 a T_p^4)^{1/4} R_0$	$\approx 10^{10}$cm
R_s	Becomes spreading	$\Gamma^2 R_0$	$\approx 10^{12} - 10^{11}$cm
R_e	Optically thin to electrons	$(\sigma_T E_0/4\pi\eta m_p c^2)^{1/2}$	$\approx 10^{13}$cm
R_δ	Internal collisions	$\Gamma^2\delta$	$\approx 10^{12} - 10^{14}$cm
R_γ	ES becomes important	$\Gamma^{-2/3} l$	$\approx 10^{16}$cm
R_Δ	RS cross shell	$l^{3/4}\Delta^{1/4}$	$\approx 10^{16}$cm
R_N	RS becomes relativistic	$l^{3/2}/\Delta^{1/2}\Gamma^2$	$\approx 10^{17}$cm
l or L	Non relativistic ES	l $^{(a)}$ or $l\Gamma^{-1/3}$ $^{(b)}$	$\approx 10^{17} - 10^{18}$cm
l	Sedov Length	$l = (3E_0/4\pi n_{ism} m_p c^2)^{1/3}$	$\approx 10^{18}$cm

Table 4.1: Critical Radii. MD: Matter-Dominated phase, ES: External Shock, RS: Reverse Shock. (a) for adiabatic fireball; (b) for radiative fireball.

& Rees 1992a). For NRS, this is where an effective energy release occurs

$$R_\gamma = \frac{l}{\Gamma^{2/3}} = \left(\frac{3E_0}{4\pi n_{ism} m_p c^2 \Gamma^2}\right)^{1/3} = 5.4 \times 10^{16}\,\mathrm{cm}\, E_{52}^{1/3} n_{1ism}^{-1/3} \Gamma_{100}^{-2/3} \quad , \tag{4.40}$$

Finally, we have $R_\delta = \Gamma^2\delta$ (see Eq. 4.30). All the critical radii arising in the fireball evolution are summarized in Table 4.1.

The different radii are related by the dimensionless parameter ξ, and this determines the character of the shock

$$R_\delta/\zeta = R_\Delta/\xi^{3/2} = R_\gamma/\xi^2 = R_N/\xi^3 \quad . \tag{4.41}$$

If $\xi > 1$, then

$$R_\delta < R_\Delta < R_\gamma < R_N \quad . \tag{4.42}$$

The reverse shock remains Newtonian or at best mildly relativistic during the whole energy extraction process (Sari & Piran 1995). The reverse shock reaches the inner edge of the shell at R_Δ while it is still Newtonian. At this stage, a reflected rarefaction wave begins to move forwards. This wave is, in turn, reflected from the contact discontinuity, between the shell material and the ISM material, and another reverse shock begins. The overall outcome of these waves is that in this case the shell acts as single fluid element of mass $M \approx E_0/\Gamma c^2$ that is interacting collectively with the ISM. It follows from Eq. 4.23 that an external mass $m = M/\Gamma$ is required to reduce Γ to $\Gamma/2$ and to convert half of the kinetic energy to thermal energy. Energy conversion takes place at R_γ. Comparison of R_γ with R_e (equation 4.15) shows that the optical depth is much smaller than unity.

If the shell propagates with a constant width, then $R_N/\xi = R_\gamma = \sqrt{\xi}R_\Delta$ and for $\xi > 1$ the reverse shock remains Newtonian during the energy extraction period. If there are significant variations in the particle velocities within the shell, it will spread during the expansion. If the typical variation in Γ is of the same order as Γ, then the shell width increases like R/Γ^2. Thus Δ changes with time in such a manner that at each moment the current width, $\Delta(t)$, satisfies $\Delta(t) \sim \max[\Delta(0), R/\Gamma^2]$. This delays the time that the reverse shock reaches the inner edge of the shell and increases R_Δ. It also reduces the mass density in the shell which, in turn, reduces f and leads to a decrease in R_N. The overall result is a triple coincidence $R_N \approx R_\gamma \approx R_\Delta$ with a mildly relativistic reverse shock and a significant energy conversion in the reverse shock as well. This means that due to spreading a shell which begins with a value of $\xi > 1$ adjusts itself so as to satisfy $\xi = 1$ (Sari & Piran 1995; Kobayashi *et al.* 1999).

For $\xi \geq 1$, we find that $T_{radial} \sim T_{ang} \sim R_\gamma/\Gamma^2 > \Delta$. Therefore, NRS can produce only smooth bursts. The burst duration is determined by the slowing down time of the shell. Given the typical radius of energy conversion R_γ, this time scale is

$$\delta T \approx T_{obs} \approx R_\gamma/(\Gamma_E^2 c) \approx R_\gamma/(\Gamma^2 c) \approx 170\,\mathrm{sec}\, E_{52}^{1/3} n_{1ism}^{-1/3} \Gamma_{100}^{-8/3} \quad . \tag{4.43}$$

If Γ or Δ are larger, then $\xi < 1$. In this case, the order is reversed

$$R_N < R_\gamma < R_\Delta \quad . \tag{4.44}$$

The reverse shock becomes relativistic very early. Since $\Gamma_{sh} = \Gamma_2 \ll \Gamma$, the relativistic reverse shock converts very efficiently the kinetic energy of the shell to thermal energy. Each layer of the shell that is shocked loses effectively all its kinetic energy at once and the time scale of converting the shell kinetic energy to thermal energy is the shell crossing time. The kinetic energy is consumed at R_Δ, where the reverse shock reaches the inner edge of the shell. Using Eq. 4.39 for R_Δ and Eq. 4.27, we find that at R_Δ

$$\Gamma_E = \Gamma_2 = (l/\Delta)^{3/8} \tag{4.45}$$

Note that Γ_E is independent of Γ. The observed radial or angular time scales are

$$T_{radial} \approx T_{ang} \approx R_\Delta/\Gamma_E^2 c \approx \Delta/c \simeq 30\,\mathrm{sec}\,\Delta_{12} \quad . \tag{4.46}$$

Thus even for RRS, we find that $\delta T \sim T$ and there is only one time scale. This time scale depends only on Δ and it is independent of Γ! Spreading does not affect this estimate, since for $\xi < 1$ spreading does not occur before the energy extraction.

In the following discussion, we focus on the RRS case and we express all results in terms of the parameter ξ. By setting $\xi < 1$ in the expressions, we obtain results corresponding to RRS, and by choosing $\xi = 1$ in the same expressions we obtain the spreading NRS limit. We shall not discuss the case of non-spreading NRS ($\xi \gg 1$),

since spreading will always bring these shells to the mildly relativistic limit ($\xi \sim 1$). Therefore, in this way, the same formulae are valid for both the RRS and NRS limits.

If $\xi > 1$, it follows from Eq. 4.42 that internal shocks will take place before external shocks. If $\xi < 1$, then the condition for internal shocks $R_\delta < R_\Delta$ becomes Eq. 4.31: $\xi^{3/2} > \zeta$. As we have seen earlier in section 4.5.2, this sets an upper limit on Γ for internal shocks.

4.5.3.2 The Hydrodynamics

Within the external shock models, there are several possible physical assumptions that one can make. The "standard" models assumes adiabatic hydrodynamics (energy losses are negligible and do not influence the hydrodynamics), slow cooling (the electrons radiate a small fraction of the energy that is generated by the shock) and synchrotron emission (Paczyński & Rhoads 1993; Katz 1994a; Mészáros & Rees 1997a; Wijers *et al.* 1997; Waxman 1997, 1997a; Mészáros *et al.* 1998b). However, there are also other possibilities. First, the electron energy might be radiated rapidly. In this case, the radiation processes is fast and the observed flux is determined by the rate of energy generation by the shock. If the electrons carry a significant fraction of the total internal energy, fast cooling will influence the hydrodynamics which will not be adiabatic any more. In this case, we have a radiation (Vietri 1997; Katz & Piran 1997) which differs in its basic scaling laws from the adiabatic one.

We discuss first a simple model for the slowing down of the shell. In this model, the slowing down is described by a series of infinitesimal inelastic collisions between the shell and the infinitesimal external masses. We assume a homogeneous shell described by its rest frame energy M (rest mass and thermal energy) and its Lorentz factor Γ. Initially, $E_0 = \Gamma_0 M_0 c^2$. The shell collides with the surrounding matter. We denote the mass of the ISM that has already collided with the shell by $m(R)$. As the shell propagates, it sweeps up more ISM mass. Additional ISM mass elements, which are at rest, collides inelastically with the shell.

Assuming a fixed fraction of the shocked thermal energy (relative to the observer frame), ε, is radiated, energy and momentum conservation yield relations between the Lorentz factor and the total mass of the shell and between $m(R)$ (and therefore R) and Γ as function of Γ (Piran 1999).

The latter relation gives us the well known result that to significantly decelerate the ultra-relativistic shell to $\Gamma_0/2$ ($\Gamma_0 \gg 1$), a mass $m \cong M_0/(2\Gamma_0)$ is required. This result is independent of the cooling parameter ε.

From the relation between $m(R)$ and Γ, a second simple result can be obtained in the limit that $\Gamma_0 \gg \Gamma \gg 1$

$$m(R) = \frac{M_0}{(2-\varepsilon)\Gamma_0}\left(\frac{\Gamma}{\Gamma_0}\right)^{-2+\varepsilon} \quad , \tag{4.47}$$

so that $\Gamma \propto R^{-3/(2-\varepsilon)}$. For $\varepsilon = 0$, this yields the well known adiabatic result

$$\frac{4\pi}{3} R^3 n m_p c^2 \Gamma^2 = E_0 \quad , \tag{4.48}$$

and $\Gamma \propto R^{-3/2}$ (Blandford & McKee 1976; Katz 1994a; Wijers *et al.* 1997; Waxman 1997). For $\varepsilon = 1$, this yields the completely radiative result

$$\frac{4\pi}{3} R^3 n m_p c^2 \Gamma \Gamma_0 = E_0 \quad , \tag{4.49}$$

and $\Gamma \propto R^{-3}$ (Blandford & McKee 1976; Katz 1994a; Katz & Piran 1997).

For comparison with observations, we have to calculate the observed time that corresponds to different radii and Lorentz factors. The well-known formula

$$t_{obs} = \frac{R}{2\Gamma^2 c} \tag{4.50}$$

is valid only for emission along the line of sight from a shell that propagates with a constant velocity. As the shell decelerates this formula should be used only in a differential sense (Sari 1997). The situation is complicated further if we recall that the emission reaches the observe from an angle of order Γ^{-1} around the line of sight. Averaging on all angles yields another numerical factor (Waxman 1997b; Sari 1998; Panaitescu & Mészáros 1998) and altogether we get

$$t_{obs} = \frac{R}{c_\gamma \Gamma^2 c} \quad , \tag{4.51}$$

where the value of the numerical factor, c_γ, depends on the details of the solution and it varies between ~ 3 and ~ 7. Using Eqs. 4.51 and 4.48 or 4.49, we obtain the following relations between R, Γ, and t

$$R(t) \cong \begin{cases} (3Et/\pi n m_p c)^{1/4}, & \text{adiabatic} \\ (4ct/L)^{1/7} L, & \text{Radiative} \end{cases} \tag{4.52}$$

$$\Gamma(t) \cong \begin{cases} (3E/256\pi n m_p c^5 t^3)^{1/8}, & \text{adiabatic} \\ (4ct/L)^{-3/7}, & \text{Radiative} \end{cases} \tag{4.53}$$

where $L \equiv (3E/4\pi n m_p c^2 \Gamma)^{1/3}$ is the radius where the external mass equals the mass of the shell.

One can proceed and use the relation between R and Γ and t_{obs} (4.52 and 4.53) to estimate the physical conditions at the shocked material using Eqs. 4.26. Then one can estimate the emitted radiation from this shock using Eqs. 4.64 and 4.63. However, we want to explore the Blandford-McKee self-similar solution (Blandford & McKee 1976), which describes more precisely the adiabatic expansion. This solution is inhomogeneous with a well determine radial profile. The matter at the front of the shell moves faster than the average speed. This influences the estimates of the radiation emitted from the shell.

4.5.3.3 The Blandford-McKee Self-Similar Solution

Blandford and McKee (1976) discovered a self-similar solution that describes the adiabatic slowing down of an extremely relativistic shell propagating into the ISM. Using several simplifications and some algebraic manipulations, we rewrite the Blandford-McKee solution (Sari 1997)

$$\begin{aligned} n(r,t) &= 4n\Gamma(t)\left[1+16\Gamma(t)^2(1-r/R)\right]^{-5/4}, \\ \Gamma(r,t) &= \Gamma(t)\left[1+16\Gamma(t)^2(1-r/R)\right]^{-1/2}, \\ e(r,t) &= 4nm_pc^2\Gamma(t)^2\left[1+16\Gamma(t)^2(1-r/R)\right]^{-17/12}, \end{aligned} \tag{4.54}$$

where $n(r,t)$, $e(r,t)$ and $\Gamma(r,t)$ are, respectively, the density, energy density and Lorentz factor of the matter behind the shock and $\Gamma(t) = \Gamma(R(t))$ is the Lorentz factor of material just behind the shock. $n(r,t)$ and $e(r,t)$ are measured in the fluid rest frame, while $\Gamma(r,t)$ is relative to an observer at rest. The total energy in this adiabatic flow equals $E = E_0$, the initial energy. The scaling laws of $R(t)$ and $\Gamma(t)$ that follow from these profiles and from the condition that the total energy in the flow equals E is

$$\begin{aligned} R(t) &= \left(\tfrac{17Et}{\pi m_p nc}\right)^{1/4} = 3.2\times10^{16}\,\mathrm{cm}\,E_{52}^{1/4}n_{1ism}^{-1/4}t_s^{1/4}, \\ \Gamma(t) &= \tfrac{1}{4}\left(\tfrac{17E}{\pi nm_pc^5t^3}\right)^{1/8} = 260E_{52}^{1/8}n_{1ism}^{-1/8}t_s^{-3/8}. \end{aligned} \tag{4.55}$$

The scalings Eq. 4.55 are consistent with the scalings Eqs. 4.52 and 4.53 which were derived using conservation of energy and momentum. Eq. 4.55 provide the exact numerical factor that cannot be calculated by the simple analysis. These equations can serve as a starting point for a detailed radiation emission calculation and a comparison with observations.

The Blandford-Mckee solution is adiabatic and as such it does not allow for any energy losses. With some simplifying assumptions, it is possible to derive a self-similar radiative solution in which an arbitrary fraction of the energy generated by the shock is radiated away (Cohen *et al.* 1999).

4.6 The Emission

4.6.1 Synchrotron Emission from Relativistic Shocks

The most likely radiation process in GRB itself and in afterglow is synchrotron emission (Mészáros *et al.* 1993; Katz 1994a; Sari *et al.* 1996; Cohen *et al.* 1997; Vietri 1997; Liang *et al.* 1997; Schaefer *et al.* 1998), although at some extent thermal emission or thermalization has to be considered (e.g. Liang *et al.* 1997; Ghisellini & Celotti 1998; Blinnikov *et al.* 1999; Eichler & Levinson 1999). Synchrotron emission is determined by the magnetic field B and the electron energy power-law distribution with a minimal energy Lorentz factor $\gamma_{e,min}$ and an index p of the expected power-law distribution.

These parameters should be determined from microscopic physical processes in the shocks. However, they are difficult to be estimated from first principles and usually characterized with two dimensionless parameters ε_B and ε_e (Sari *et al.* 1996).

ε_B is defined as the ratio of the magnetic field energy density to the total thermal energy e:

$$\varepsilon_B \equiv \frac{U_B}{e} = \frac{B^2}{8\pi e} \quad , \tag{4.56}$$

so that, after substituting the shock conditions we have

$$B = \sqrt{32\pi}\varepsilon_B^{1/2}\Gamma_{sh}m_p^{1/2}n_1^{1/2}c \quad . \tag{4.57}$$

ε_B is the measure of the deviation of a magnetic field from its thermal equipartition. There have been some attempts to estimate ε_B in the literatures (Mészáros *et al.* 1993; Thompson 1994; Usov & Smolsky 1996). Additionally, the magnetic field is usually assumed to be randomly oriented in space.

The parameter ε_e is defined to measures the fraction of the total thermal energy e which goes into random motions of the electrons:

$$\varepsilon_e \equiv \frac{U_e}{e} \quad . \tag{4.58}$$

ε_e is the measure of the deviation of the thermal energy of electrons from their thermal equipartition.

As the electrons is heated by shocks, the distribution of their Lorentz factors is most likely a power-law

$$N(\gamma_e) \sim \gamma_e^{-p} \quad \text{for } \gamma_e > \gamma_{e,min} \quad . \tag{4.59}$$

The minimum Lorentz factor $\gamma_{e,min}$ of the distribution is related to ε_e and to the total energy $e \sim \Gamma_{sh}nm_pc^2$ with (Sari *et al.* 1996)

$$\gamma_{e,min} = \frac{m_p}{m_e}\frac{p-2}{p-1}\varepsilon_e\Gamma_{sh} = \frac{p-2}{p-1}\langle\gamma_e\rangle \quad , \tag{4.60}$$

where $\langle\gamma_e\rangle$ is the average γ_e of the electrons given as

$$\langle\gamma_e\rangle = \frac{m_p}{m_e}\varepsilon_e\Gamma_{sh} \quad . \tag{4.61}$$

A electron is called "typical", if its γ_e equals to $\langle\gamma_e\rangle$.

The energy spectral index p is required to be > 2 so that the energy does not diverge at large γ_e. It is usually fixed by fitting the model to the observed high energy spectra of GRBs. The mean spectral index of GRBs at high photon energies (Band *et al.* 1993) corresponds a typical value $p \approx 2.5$, which agrees with the value inferred from afterglow observations (see Section 4.8).

For the shock-heated maximal electron energy $\gamma_{e,max}$, the acceleration time equals to the cooling time. Equating the acceleration time, which is determined by the Larmor radius R_L and the Alfvén velocity v_A (Hillas 1984), to the synchrotron cooling time in the local frame, we can estimate (Piran 1999)

$$\gamma_{e,max} = \sqrt{\frac{24\pi\varepsilon_B^2 q_e}{B\sigma_T\Gamma_{sh}}} = 3.7\times10^8 \frac{\varepsilon_B^{3/4}}{\Gamma_{sh}^{1/2} n_1^{1/4}} \quad , \tag{4.62}$$

where q_e is the electron charge. This value is quite large and generally it does not affect the observed spectrum in the soft gamma-ray range.

The power emitted by a single electron due to synchrotron radiation in the local frame is given by

$$P_{syn} = \frac{4}{3}\sigma_T c U_B \gamma_e^2 \quad , \tag{4.63}$$

where σ_T is the Thomson cross section. The characteristic synchrotron photon has energy in the observer frame

$$(h\nu_{syn})_{obs} = \frac{\hbar q_e B}{m_e c}\gamma_e^2\Gamma_E \quad . \tag{4.64}$$

The cooling time of the electron in the fluid frame is then $\gamma_e m_e c^2/P$ and in the observer frame

$$t_{syn}(\nu) = \frac{3m_e c}{4\sigma_T U_B \gamma_e \Gamma_E} = \frac{3}{\sigma_T}\sqrt{\frac{2\pi q_e m_e c}{B^3\nu\Gamma_E}} \quad , \tag{4.65}$$

where ν is the frequency in the observer frame. Eq. 4.65 implies that t_{syn} at a given observed frequency is independent of the electron energy distribution within the shock, provided there are electrons with required γ_e emitting in the considered frequency. The scaling is not very different from the observed relation $\delta T \propto \nu^{-0.4}$ (Fenimore *et al.* 1995). Although it is not clear if the cooling time not others determines the temporal profile, the cooling time scale above sets a lower limit to the variability time scales of a GRB as the burst cannot possibly contain spikes that are shorter than its cooling time.

The instantaneous synchrotron spectrum of a single electron with an initial energy $\gamma_e m_e c^2$ is a power law with $F_\nu \propto \nu^{1/3}$ up to $\nu_{syn}(\gamma_e)$ and an exponential decay above it (Rybicki & Lightman 1979). If the electron is energetic, it will cool rapidly until it will reach $\gamma_{e,c}$, the Lorentz factor of an electron that cools on a hydrodynamic time scale. For a rapidly cooling electron, the time integrated spectrum (over the hydrodynamic time scale) above $h\nu_{syn}(\gamma_{e,c})$ is given as $F_\nu \propto \nu^{-1/2}$ from $\nu_{syn}(\gamma_{e,c})$ up to $\nu_{syn}(\gamma_e)$ and $F_\nu \propto \nu^{-p/2}$ for the higher frequency. For slow cooling, we have the instantaneous spectrum for the lower part of the spectrum and $F_\nu \propto \nu^{-(p-1)/2}$ for the upper part.

$\gamma_{e,c}$ is estimated by comparing t_{syn} (Eq. 4.65) with the hydrodynamic time scale (in the observer rest frame) t_{hyd}

$$\gamma_{e,c} = \frac{3m_e c}{4\sigma_T U_B \Gamma_E t_{hyd}} \quad . \tag{4.66}$$

fast cooling occurs if $\gamma_{e,c} < \gamma_{e,min}$. All the electrons cool rapidly and the electron distribution effectively extends down to $\gamma_{e,c}$. If $\gamma_{e,c} > \gamma_{e,min}$, only the high energy tail of the distribution (those electrons above $\gamma_{e,c}$) cool and the system is in the **slow cooling** regime.

For the GRB itself, the relativistic shocks must emit their energy effectively, otherwise there will be a serious inefficiency problem. We will not be able to explain the variability, if the cooling time is too long. The electrons must cool rapidly and release all their energy and $\gamma_{e,min} > \gamma_{e,c}$ (Sari *et al.* 1996) and all the electrons cool down roughly to $\gamma_{e,c}$. The observed flux F_ν at gamma-ray is given by (Sari *et al.* 1998)

$$F_\nu \propto \begin{cases} (\nu/\nu_c)^{1/3} F_{\nu,max} & \nu_c > \nu \quad , \\ (\nu/\nu_c)^{-1/2} F_{\nu,max} & \nu_m > \nu > \nu_c \quad , \\ (\nu_m/\nu_c)^{-1/2} (\nu/\nu_m)^{-p/2} F_{\nu,max} & \nu > \nu_m \quad , \end{cases} \tag{4.67}$$

where $\nu_m \equiv \nu_{syn}(\gamma_{e,min})$, $\nu_c \equiv \nu_{syn}(\gamma_{e,c})$ and $F_{\nu,max}$ is the observed peak flux at ν_c. The energy (νF_ν) peak is at ν_m. The emitted power is

$$P_{fast} = \varepsilon_e \frac{dE}{dt} \quad , \tag{4.68}$$

where $\frac{dE}{dt}$ is the power generated by the shock.

It is argued that during the latter stages of an external shock, there will likely be transition from fast to slow cooling (Mészáros & Rees 1997a; Waxman 1997; Katz & Piran 1997). When $\gamma_{e,c} > \gamma_{e,min}$, only those electrons with $\gamma_e > \gamma_{e,c}$ can cool and is called slow cooling, because the electrons with $\gamma_e \sim \gamma_{e,min}$ forming the bulk of the population, do not cool. Integration over the electron distribution gives (Sari *et al.* 1998)

$$F_\nu \propto \begin{cases} (\nu/\nu_m)^{1/3} F_{\nu,max} & \nu_m > \nu \quad , \\ (\nu/\nu_m)^{-(p-1)/2} F_{\nu,max} & \nu_c > \nu > \nu_m \quad , \\ (\nu_c/\nu_m)^{-(p-1)/2} (\nu/\nu_c)^{-p/2} F_{\nu,max} & \nu > \nu_c \quad . \end{cases} \tag{4.69}$$

The peak flux $F_{\nu,max}$ is at ν_m, while the energy peak is at ν_c. The emitted power is

$$P_{slow} = N_e P_{syn}(\gamma_{e,min}) \quad , \tag{4.70}$$

where N_e is the number of electrons in the emitting region and $P_{syn}(\gamma_{e,min})$ the synchrotron power of an electron with $\gamma_{e,min}$.

A typical spectrum is shown in Fig. 4.4.

The radiative efficiency depends on three factors: (i) the available electron energy density (ε_e), (ii) the available magnetic energy density ($\min\left[1, \sqrt{\varepsilon_B/\varepsilon_e}\right]$) if $\varepsilon_B < \varepsilon_e$ and the inverse Compton radiation is not in the observed 100 keV band, and (iii) the specific Lorentz factor $\hat{\gamma}_e$, with which an electron emitted synchrotron photon at 110 keV band. Only radiation of electrons with $\gamma \geq \hat{\gamma}_e$ can contribute significantly to the

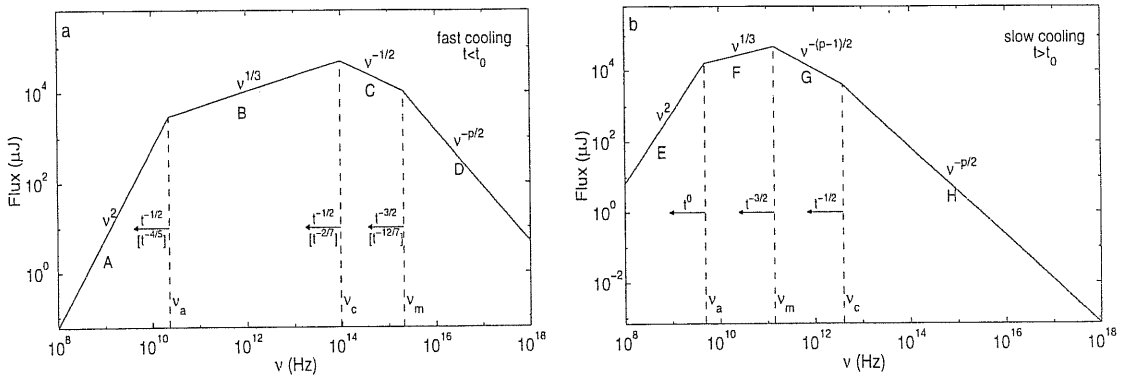

Figure 4.4: Synchrotron spectrum of a relativistic shock with a power-law electron distribution. **(a)** fast cooling, which is expected at early times. The spectrum consists of four segments, identified as A, B, C and D. Self-absorption is important below ν_a. The frequencies, ν_m, ν_c, ν_a, decrease with time as indicated; the scalings above the arrows correspond to an adiabatic evolution, and the scalings below to a fully radiative evolution. **(b)** Slow cooling, which is expected at late times. The evolution is always adiabatic. The four segments are identified as E, F, G and H. From Sari *et al.* (1998).

observed soft gamma rays. If the distribution of electron energy is a power law of an index $p = 2.5$ and $\hat{\gamma}_e > \gamma_{e,min}$, the total efficiency is given by (Sari *et al.* 1996)

$$\varepsilon = \varepsilon_e \min\left[1, \sqrt{\varepsilon_B/\varepsilon_e}\right] (\gamma_{e,min}/\hat{\gamma}_e)^{1/2} \quad . \tag{4.71}$$

So, in order to be efficient, both electron and magnetic energy densities should be close to equipartition. Additionally, $\hat{\gamma}_e$ should be larger but not very much than $\gamma_{e,min}$. This estimate is different if the observed γ-ray are produced by inverse Compton scattering.

4.6.2 Synchrotron Self-absorption

While it is not clear if the self-absorption present in the GRB itself, it may appear at late time and typically in radio emission (Katz 1994a; Katz & Piran 1997; Waxman 1997a; Wijers & Galama 1998; Granot *et al.* 1998). When it appears, the energy spectrum below the absorption frequency ν_a will have a steep cutoff.

At the self-absorption frequency, the optical depth τ_{ν_a} along the line of sight is unit. From Rybicki & Lightman (1979), $\tau_\nu \simeq \alpha'_{\nu'} R/\Gamma$, where $\alpha'_{\nu'}$ is is the absorption coefficient

$$\alpha'_{\nu'} = \frac{p+2}{8\pi m_e \nu'^2} \int_{\gamma_{e,min}}^{\infty} P'_{\nu',e}(\gamma_e) \frac{N(\gamma_e)}{\gamma_e} d\gamma_e \quad . \tag{4.72}$$

The self-absorption frequency ν_a can be estimated only when we have a model for the hydrodynamics (Wijers & Galama 1998; Granot *et al.* 1998).

In GRB afterglow, the expected self-absorption spectrum (Katz 1994) is $\approx \nu^2$, as in relativistic shocks nearly all the low-frequency radiation is emitted by electrons with a single, and much higher, characteristic synchrotron frequency and the radiation in

self-absorption range is due to the low energy tail of electrons radiating effectively at higher energies. This is different from the the well known spectrum (Rybicki & Lightman 1979) $\approx \nu^{5/2}$ which is for the case that the synchrotron frequency of the electron emitting the self-absorbed radiation is inside the self-absorption range. The observed flux is given as (Katz 1994; Katz & Piran 1997)

$$F_\nu \approx \nu^2 (k_B T_e / (\Gamma m_p c^2)) R^2 \quad , \tag{4.73}$$

where R is the radius of the radiating shell, k_B Boltzmann constant, and T_e the temperature in the comoving frame of the shocked radiating matter moving with Lorentz factor Γ. The factor $k_B T_e/(\Gamma m_p c^2)$ describes the equi-partition degree of electrons in the plasma of an internal energy per particle $\Gamma m_p c^2$.

4.6.3 Inverse Compton Emission

Inverse Compton (IC) scattering can influence the spectrum even if the system is optically thin to Compton scattering (e.g. Rybicki & Lightman 1979). In view of the high energies involved, it is usually assumed that only one IC scattering takes place, while saturated (repeated) repeatedly IC scattering (Liang *et al.* 1997, 1999) is possible as in the emitting regions the Thomson depths may initially be very large ($\tau \gg 1$) (Zdziarski *et al.* 1990; Fenimore *et al.* 1982). The impulsively accelerated, non-thermal, relativistic leptons (electrons and pairs) repeatedly Compton up-scatter self-emitted (radio-IR) synchrotron photons into gamma-ray energies.

In generally, after one upper scattering, the photon energy is so high that in the electron rest frame it is above the Klein-Nishina energy and in this case the decrease in the Compton cross section makes this scattering unlikely. The importance of IC depends on the Compton parameter $Y = \Gamma^2 \tau_e$ which equals to the ratio of energy rates via IC and via synchrotron radiation. For fast cooling (Sari *et al.* 1996)

$$\begin{array}{ll} Y = \varepsilon_e/\varepsilon_B & \text{if } \varepsilon_e \ll \varepsilon_B \\ Y = \sqrt{\varepsilon_e/\varepsilon_B} & \text{if } \varepsilon_e \gg \varepsilon_B \end{array} \quad . \tag{4.74}$$

IC is not important and can be neglected when $Y < 1$.

If $Y > 1$, a large fraction of the low energy will be emitted via IC process. If IC up-scattered photons are in the observed energy band, the observed radiation will be IC instead of synchrotron photons. If IC photons are too energetic and beyond the observed energy range, IC will not influence the observed spectra directly. However, as IC will take a significant fraction of the energy of the cooling electrons, it will influence the observations in two ways: shortening the cooling time (the emitting electrons will be cooled by both synchrotron and IC process) and influencing the overall energy budget (e.g. scattering a fraction of the synchrotron emission in the observed range)

and therefore reducing the production efficiency of the observed synchrotron radiation by a factor of Y.

Synchrotron photons emitted by electrons of $\gamma_{e,1}$ and inverse Compton scattered by electrons of $\gamma_{e,2}$ will be observed at the energy (Sari *et al.* 1996)

$$(h\nu_{IC})_{obs} = \frac{\hbar q_e B}{m_e c}\gamma_{e,1}^2\gamma_{e,2}^2\Gamma_E = 13\,\mathrm{MeV}\left(\frac{B}{1\,\mathrm{G}}\right)\left(\frac{\Gamma_E}{100}\right)\left(\frac{\gamma_{e,1}}{(m_p/m_e)}\right)^2\left(\frac{\gamma_{e,2}}{(m_p/m_e)}\right)^2 . \tag{4.75}$$

The photon energy is boosted by a factor of $\gamma_{e,2}^2$.

To produce an observed photon at frequency ν, the electron Lorentz factor γ_e required by self IC scattering mechanism ($\gamma_{e,2} = \gamma_{e,1}$) is the square root of that required by synchrotron radiation mechanism. The required value for γ_e is rather low relative to the typical value $\langle\gamma_{e,ext}\rangle \sim \varepsilon_e(m_p/m_e)\Gamma_{sh}$ (see Eq. 4.61) in an external shock. In internal shocks, $\langle\gamma_{e,int}\rangle \sim \varepsilon_e(m_p/m_e)$ is lower but in this case the equi-partition magnetic field is much stronger and of the order of few thousand Gauss, or higher. Thus, IC might produce the observed photons in internal shocks if ε_B is rather small (of order10^{-5}).

Electrons are cooled both by synchrotron and by IC. The latter is more efficient if $Y > 1$ and the cooling is enhanced by the Compton parameter Y. The cooling time scale is given

$$t_{IC} = \frac{6\pi}{\sigma_T}\left(\frac{\varepsilon_B^2\hbar m_e^3 c^3 q_e}{\varepsilon_e^2 B^7\Gamma_E^3 h\nu}\right)^{1/4} = 8\times10^4\,\mathrm{sec}\sqrt{\frac{\varepsilon_B}{\epsilon_e}}\left(\frac{B}{1\,\mathrm{G}}\right)^{-7/4}\left(\frac{\Gamma_E}{100}\right)^{-3/4}\left(\frac{h\nu}{100\,\mathrm{keV}}\right)^{-1/4} . \tag{4.76}$$

4.6.4 Emission from Internal Shock

Provided that the different parts of the shell have comparable Lorentz factors differing by factor of ~ 2, the internal shocks are mildly relativistic. The thermal Lorentz factor of protons will be of order unity, and the shocked regions will still move highly relativistically towards the observer with approximately the initial Lorentz factor Γ. In front of the shocks, the particle density of the shell is given by the total number of baryons $E_0/\Gamma m_p c^2$ divided by the co-moving volume of the shell at the radius R_δ which is $4\pi R_\delta^2\Gamma\Delta$. Using Eq. 4.29 for the particle density n and the thermal energy density e behind the shocks, we obtain the number density of shocked regions

$$n_{int} \approx \frac{3E(\Gamma_{int}/2)^{1/2}}{4\pi\Gamma^6 c^2 m_p\delta^2\Delta} = 2\times10^{10}\,\mathrm{cm}^{-3}\,E_{52}(\Gamma_{int}/2)^{1/2}\Gamma_{100}^{-6}\Delta_{12}^{-1}\delta_{10}^{-2} , \tag{4.77}$$

$$e_{int} = (\Gamma_{int}/2)^{1/2} n_{int} m_p c^2 , \tag{4.78}$$

where $\Gamma_{int} \sim$ afew is the Lorentz factor of the internal shock and $\Gamma_{100} = \Gamma/100$, $\Delta_{12} = \Delta/10^{12}$ cm, $\delta_{10} = \delta/10^{10}$ cm.

From Eq. 4.57,

$$B_{int} = 6 \times 10^5 \,\mathrm{G}\, \varepsilon_B^{1/2} E_{52}^{1/2} \Gamma_{100}^{-3} \Delta_{12}^{-1/2} \delta_{10}^{-1} (\Gamma_{int}/2)^{1/2} \quad . \tag{4.79}$$

From Eqs. 4.57, 4.61, 4.64 and 4.77, the typical synchrotron frequency from an internal shock emitted by a "typical" electron is

$$(h\nu_{syn})_{obs}|_{\langle\gamma_e\rangle} = \frac{\hbar q_e B}{m_e c} \langle\gamma_e\rangle^2 \Gamma = 220\,\mathrm{keV}\, E_{52}^{1/2} \varepsilon_B^{1/2} \zeta_{0.01} (\Gamma_{int}/2)^{1/2} \Gamma_{100}^{-2} \Delta_{12}^{-3/2} (\gamma_e/(m_p/m_e))^2 \quad . \tag{4.80}$$

and the corresponding observed synchrotron cooling time is

$$t_{syn}|_{\langle\gamma_e\rangle} = 1.3 \times 10^{-6}\,\mathrm{sec}\, \varepsilon_B^{-1} \delta_{10}^2 \Delta_{12} \Gamma_{100}^5 E_{52}^{-1} (\Gamma_{int}/2)^{-1} \quad . \tag{4.81}$$

We can estimate the minimal synchrotron frequency

$$(h\nu_{syn})_{obs}|_{\gamma_{e,min}} = 24\,\mathrm{keV}\, E_{52}^{1/2} \varepsilon_B^{1/2} \varepsilon_e^2 \zeta_{0.01}^{-1} (\Gamma_{int}/2)^{5/2} \Gamma_{100}^{-2} \Delta_{12}^{-3/2} \quad , \tag{4.82}$$

where we have used Eq. 4.60 to obtain $\gamma_{e,min}$.

The energy emitted by a "typical electron" is around 220 keV. The energy emitted by a "minimal energy" electron is about one order of magnitude lower than the typical observed energy of $\sim$ 100 keV. This should correspond to the break energy of the spectrum. These results seem in a good agreement with the observations. However, both ε_B and ε_e might be significantly lower than unity. A lower value of ε_B or ε_e might be compensated by a higher value of Γ_{int}. Shocks with higher Γ_{int} are more efficient (see section 4.6.4).

The synchrotron cooling time at a given frequency (in the observer frame) is

$$t_{syn}(h\nu) = 2 \times 10^{-6}\,\mathrm{sec}\, \varepsilon_B^{-3/4} \left(\frac{h\nu_{obs}}{100\,\mathrm{keV}}\right)^{-1/2} \delta_{10}^{3/2} \Delta_{12}^{3/4} \Gamma_{100}^4 E_{52}^{3/4} (\Gamma_{int}/2)^{-3/4} \quad . \tag{4.83}$$

We recover the general trend $t_{syn} \propto (h\nu)^{-1/2}$ of synchrotron emission. However, if this cooling time is much shorter than T_{ang}, it does not determine the width of the observed peaks. It will correspond to the observed time scales if, for example, ε_B is small. But then "typical" photon energy will be far below the observed range. Therefore, it is not clear this relation can explain the observed dependence of the width of the bursts on the observed energy.

The calculations of section 4.6.3 suggest that the typical Inverse Compton (IC) radiation from internal shocks will be at energy higher by a factor γ_e^2 than the typical synchrotron frequency and in the GeV or even TeV range. This radiation might contribute to the prompt very high energy emission that accompanies some GRBs (Pilla & Loeb 1998).

However, if the magnetic field is extremely low: $\varepsilon_B \sim 10^{-12}$, then we would expect the IC photons to be in the observed $\sim$ 100 keV region: $h\nu_{IC-int} \sim 800$ keV. Using

Eqs. 4.76 and 4.79, we find that the cooling time for synchrotron self Compton in this case is $\tau_{IC-int} \sim 1$ sec. It is too large for some bursts and possibly adequate for others. It is more likely that if inverse Compton is important it contributes to the very high (GeV or even TeV) signal (Pilla & Loeb 1998).

The efficiency of a single collision between two shells was calculated in Sec. 4.5.1. For multiple collisions, the efficiency of conversion of kinetic energy into thermal energy depends on the nature of the random distribution. It is highest if the energy is distributed equally among the different shells. Numerical calculations (Mochkovitch *et al.* 1995) show that the efficiency of this process is low (less than 2%) if the initial spread in Γ is only a factor of two, but could be much higher (Kobayashi *et al.* 1997) if the initial spread is larger. The most efficient case is when the shells have a comparable energy but very different Lorentz factors. In this case (shell mass $m_i \propto \Gamma_i^{-1}$, and spread of Lorentz factor $\Gamma_{max}/\Gamma_{min} > 10^3$), the efficiency is as high as 40%. For a moderate spread of Lorentz factor $\Gamma_{max}/\Gamma_{min} = 10$, with $m_i \propto \Gamma_i^{-1}$, the efficiency is 20%.

The overall efficiency of conversion of kinetic energy into radiation should be calculated by multiplying this by the radiative efficiency, discussed in Sec. 4.6.1 (Eq. 4.71). The resulting values may be rather small (Mochkovitch *et al.* 1995; Kobayashi *et al.* 1997; Katz 1997). It seems that some sort of beaming is required to overcome an overall energy crisis. Highly relativistic flow with a kinetic energy and a Lorentz factor comparable to the original one remains after the internal shocks. The low efficiency of conversion of kinetic energy to γ-rays is the main concern with any internal shock model.

4.6.5 Radiation from External Shocks

The interaction between the outward moving shell and the ISM takes place in the form of two shocks: a forward shock that propagates into the ISM and a reverse shock that propagates into the relativistic shell. This results in four distinct regions (cf Fig. (4.2): the ISM at rest (denoted by the subscript 1), the shocked material which has passed through the forward shock (subscript 2 or f), the shocked shell material which has passed through the reverse shock (3 or r), and the unshocked material in the shell (4) (Kobayashi *et al.* 1999). The nature of the emitted radiation and the efficiency of the cooling processes depend on the conditions in the shocked regions 2 and 3. Both regions have the same energy density ε. The particle densities n_2 and n_3 are, however, different and hence the effective "temperatures", i.e. the mean Lorentz factors of the random motions of the shocked protons and electron, are different.

The bulk of the kinetic energy of the shell is converted to thermal energy via the two shocks at around the time the shell has expanded to the radius R_Δ. At this radius,

the conditions at the forward shock are as follows,

$$\Gamma_2 = \Gamma\xi^{3/4}, \qquad n_2 = 4\Gamma_2 n_1, \qquad e_2 = 4\Gamma_2^2 n_1 m_p c^2, \tag{4.84}$$

while at the reverse shock we have

$$\bar{\Gamma}_3 = \xi^{-3/4}, \qquad \Gamma_3 = \Gamma\xi^{3/4}, \qquad n_3 = 4\xi^{9/4}\Gamma^2 n_1, \qquad e_3 = e_2. \tag{4.85}$$

Substitution of $\Gamma_{sh} = \Gamma_2 = \Gamma\xi^{3/4}$ in Eq. 4.57 yields

$$B = \sqrt{32\pi}\varepsilon_B^{1/2}\Gamma\xi^{3/4}m_p^{1/2}n_{1ism}^{1/2}c = 40\,\mathrm{G}\,\varepsilon_B^{1/2}\xi^{3/4}\Gamma_{100}n_{1ism}^{1/2} \quad . \tag{4.86}$$

If the magnetic field in region 2 behind the forward shock is obtained purely by shock compression of the ISM field, the field would be very weak with $\varepsilon_B \ll 1$. Such low fields are incompatible with observations of GRBs. There may be some kind of a turbulent instability which may bring the magnetic field to approximate equipartition. In the case of reverse shock, magnetic fields of considerable strength might be present in the pre-shock shell material if the original exploding fireball was magnetic. The exact nature of magnetic field evolution during fireball expansion depends on several assumptions. Thompson (1994) found that the magnetic field will remain in equipartition if it started off originally in equipartition. Mészáros, Laguna & Rees (1993), on the other hand, estimated that if the magnetic field was initially in equipartition it would be below equipartition by a factor of 10^{-5} by the time the shell expands to R_Δ. It is uncertain which, if either, is right. As in the forward shock, an instability could boost the field back to equipartition. Thus, while both shocks may have $\varepsilon \ll 1$ with pure flux freezing, both could achieve $\varepsilon \to 1$ in the presence of instabilities. In principle, ε_B could be different for the two shocks, but people take them the same in both shocks.

In both regions 2 and 3, the electrons have a power law distribution with a minimal Lorentz factor $\gamma_{e,min}$ given by Eq. 4.60 with the corresponding Lorentz factors for the forward and the reverse shock.

From Eq. 4.60 for $\gamma_{e,min}$, the characteristic synchrotron energy for the forward shock is

$$(h\nu_{syn})_{obs}|_{\gamma_{e,min}} = 160\,\mathrm{keV}\,\varepsilon_B^{1/2}\varepsilon_e^2(\Gamma_2/100)^4 n_{1ism}^{1/2} \quad , \tag{4.87}$$

and cooling time is

$$t_{syn}|_{\gamma_{e,min}} = 0.085\,\mathrm{sec}\,\varepsilon_B^{-1}\varepsilon_e^{-1}(\Gamma_2/100)^{-4}n_{1ism}^{-1} \quad . \tag{4.88}$$

The characteristic frequency and the corresponding cooling time for the "typical" electron (c.f. Eq. 4.60) are larger by a factor of $[(p-1)/(p-2)]^2$ and shorter by a factor of $[(p-2)/(p-1)]^2$, respectively.

These photons seem to be right in the observed soft gamma-ray range. However, one should recall that the frequency calculated in Eq. 4.87 depends on the fourth power

of Γ_2. An increase of the canonical Γ_2 by a factor of 3 (that is $\Gamma_2 = 300$ instead of $\Gamma_2 = 100$) will yield a "typical" synchrotron emission at the 16 MeV instead of 160 keV. The Lorentz factor of a "typical electron" in the reverse shock is lower by a factor ξ^3, while the cooling time scale is longer by a factor $\xi^{-3/4}$.

Using Eq. 4.64, we can check the conditions in order that there are electrons with a Lorentz factor $\hat{\gamma}_e$ that is emitting soft gamma-rays with energies $\sim 100\,\mathrm{keV}$. The results show that the reverse shock always has electrons with the right Lorentz factors to produce soft gamma-ray synchrotron photons. However, the situation is marginal in the case of the forward shock. If $\Gamma > 100$ and if the heating of the electrons is efficient, i.e. $\varepsilon_e|_f \sim 1$, then most of the electrons may be too energetic. Of course, as an electron cools, it radiates at softer energies. Therefore, even if $\gamma_{e,min}$ is initially too large for the synchrotron radiation to be in soft gamma-rays, the same electrons would at a later time have $\gamma_e \sim \hat{\gamma}_e$ and become visible. However, the energy remaining in the electrons at the later time will also be lower (by a factor of $\hat{\gamma}/\gamma_{e,min}$), which means that the burst will be inefficient. For simplicity, we ignore this radiation.

From Eq. 4.64 and Eq. 4.65, we obtain the cooling time scale as a function of the observed photon energy to be

$$t_{syn}(h\nu) = 1.4 \times 10^{-2}\,\mathrm{sec}\,\varepsilon_B^{-3/4}\left(\frac{h\nu_{obs}}{100\,\mathrm{keV}}\right)^{-1/2}\Gamma_{100}^{-2}n_{1ism}^{-3/4} \quad . \tag{4.89}$$

Eq. 4.89 is valid for both the forward and reverse shock, and is moreover independent of whether the reverse shock is relativistic or Newtonian.

The cooling time calculated above sets a lower limit to the variability time scale of a GRB, since the burst cannot possibly contain spikes that are shorter than its cooling time. However, it is unlikely that this cooling time actually determines the observed time scales (Piran 1999).

4.6.6 Phases in a Relativistic Decelerating Shell

There are several phases in the deceleration of a relativistic shell (Paczyński & Rhoads 1993; Katz 1994a; Mészáros & Rees 1997a; Wijers *et al.* 1997; Waxman 1997, 1997a; Mészáros *et al.* 1998b; Vietri 1997; Katz & Piran 1997): fast cooling with either radiative ($\varepsilon_e \sim 1$) or adiabatic ($\varepsilon_e \ll 1$) hydrodynamics is followed by slow cooling with adiabatic hydrodynamics. Then if the shell is non spherical its evolution changes and a phase of sideways expansion and much faster slow down begins when the Lorentz factor reaches θ^{-1} (Rhoads 1997). Finally, the shell becomes Newtonian when enough mass is collected and $\Gamma \approx 1$. In the following, we discuss the time scale for the different transitions. $\gamma_{e,min} \equiv c_\gamma \varepsilon_e (m_p/m_e)\Gamma$ and $t_{obs} = (1+z)R/4c_t c\Gamma^2$ are defined such that the tractor c_γ and c_t reflects some of the uncertainties in the model. The canonical values of these factors are $c_\gamma \approx 0.5$ and $c_t \approx 1$.

The deceleration begins in a fast cooling phase. If ε_e is close to unity, this cooling phase will also be radiative. The first transition is from fast to slow cooling. There are several different ways to estimate this transition. One can compare the cooling time scale to the hydrodynamic time scale; alternatively one can calculate the fast cooling rate (given by the rate of energy generation by the shell) and compare it to the slow cooling rate (given by the emissivity of the relativistic electrons). If we calculate this time as the time when the "typical electron" cools ($\nu_c = \nu_m$, we have

$$t_{fs} = \begin{cases} 210\,\mathrm{days}\,\varepsilon_B^2\varepsilon_e^2 E_{52} n_{1ism} & \text{adiabatic} \\ 4.6\,\mathrm{days}\,\varepsilon_B^{7/5}\varepsilon_e^{7/5} E_{52}^{4/5}\Gamma_{100}^{-4/5} n_{1ism}^{3/5} & \text{radiative} \end{cases} . \quad (4.90)$$

All methods of estimating t_{fs} give the same dependence on the parameters, but the numerical factors is quite sensitive to the definition of this transition.

If the solution is initially radiative, the transition from fast to slow cooling implies a transition from a radiative hydrodynamics to adiabatic hydrodynamics at

$$t_{rad-ad} = 1.3\,\mathrm{days}\, E_{52}^{4/5} n_{1ism}^{3/5}\varepsilon_e^{7/5}\varepsilon_B^{7/5}[(1+z)/2]^{12/5}(c_\gamma/0.5)^{14/5} c_t^{-12/5}(\Gamma_0/100)^{-4/5}. \quad (4.91)$$

During a radiative evolution, the energy in the shock decreases with time. The energy that appears in Eq. 4.91 in the radiative scalings is the initial energy. When a radiative shock switches to adiabatic evolution, it is necessary to use the reduced energy to calculate the subsequent adiabatic evolution. The energy $E_{f,52}$ which one should use in the adiabatic regime is related to the initial $E_{i,52}$ of the fireball by

$$E_{f,52} = 0.022\,\varepsilon_B^{-3/5}\varepsilon_e^{-3/5} E_{i,52}^{4/5}\Gamma_{100}^{-4/5} n_{1ism}^{-2/5} . \quad (4.92)$$

If the shell is not spherical and it has an opening angle θ, the evolution will change when $\Gamma \sim \theta^{-1}$ (Rhoads 1997). Earlier on, the jet expands too rapidly to expands to sideways and it evolves as if it is a part of a spherical shell. After this stage, the jet expands sideways and it accumulates much more mass and slows down much faster. This transition will take place, quite generally, during the adiabatic phase

$$t_\theta \approx 0.5\,\mathrm{days}\, E_{52}^{1/3} n_{1ism}^{-1/3}\left(\frac{1+z}{2}\right)\left(\frac{\theta}{0.1}\right)^{8/3} c_t^{-1} . \quad (4.93)$$

The shell eventually becomes non-relativistic, when $R \approx l = (3E_0/4\pi n_{ism} m_p c^2)^{1/3}$ for an adiabatic solution. This corresponds to a transition at

$$t_{NR,ad} \approx l/c \approx 300\,\mathrm{days}\, E_{52}^{1/3} n_{1ism}^{-1/3} . \quad (4.94)$$

A radiative shell loses energy faster and it becomes non-relativistic at $R = L = l/\Gamma_0^{1/3} = (3E_0/4\pi n_{ism} m_p c^2 \Gamma_0)^{1/3}$. This will take place at

$$t_{NR,rad} \approx 65\,\mathrm{days}\, E_{52}^{1/3} n_{1ism}^{-1/3}(\Gamma_0/100)^{-1/3} . \quad (4.95)$$

However, the earlier estimate of the transition from fast to slow cooling suggests that the shell cannot remain radiative for such a long time.

4.6.7 The Expected Instantaneous Spectrum and Light Curve

We discuss the expected instantaneous spectrum and light curve from a relativistic decelerating shell. The discussion is limited to a spherical shock propagating into a homogeneous external matter.

At any given time, a detector receives photons which were emitted at different times in the observer frame, at different distances behind the shock front and at different angles from the line of sight (LOS) to the center of the GRB. The properties of the matter are different at each of these points, and so are the emissivity and the frequency of the emitted radiation.

In the early calculations, Mészáros & Rees (1997), Waxman (1997), Sari *et al.* (1998) considered emission from a single representative point. Later works included more detailed calculations. Synchrotron emission was considered from the shock front (Sari 1998; Panaitescu & Mészáros 1998) and monochromatic emission was considered from a uniform shell (Waxman 1997a).

Actually to obtain the observed spectrum, we should integrate over the shell profile and over different angles relative to the line of sight. A full calculation of the integrated spectrum over a Blandford-McKee self similar solution and with realistic emission has been done by Granot *et al.* 1999. They show that different radial points from which the radiation reaches the observers simultaneously conspire to have practically the same synchrotron frequency and therefore they emit the same spectrum. Hence the radial integration over a Blandford-McKee profile does not change the observed spectrum. This result considerably differs from the full calculation for a homogeneous shell (Waxman 1997b). On the other hand, the contribution from angles away from the line of sight is important and it shapes the observed spectrum, the light curve and the shape of the afterglow (Fig. 4.5). Both the spectra and light curve are flat near the peak. This causes difficulty in explaining the shape of the optical peak of GRB970508.

Sari *et al.* (1998) made an analysis of the spectrum over a wider range of the frequencies. We discuss some more about the results. We consider two extreme limits for the hydrodynamic evolution: fully radiative and fully adiabatic. If ε_e is somewhat less than unity during the fast cooling phases ($t < t_{fs}$), only a fraction of the shock energy is lost to radiation. The scalings will be intermediate between the two limits of fully radiative and fully adiabatic discussed here. For simplicity, we assume that all the observed radiation reaches the observer from the front of the shell and along the line of sight.

The instantaneous synchrotron spectra from a relativistic shock were described in section 4.6.1. They do not depends on the hydrodynamic evolution but rather on the instantaneous conditions at the shock front, which determines the break energies ν_c and ν_m. The only assumption made is that the shock properties are fairly constant

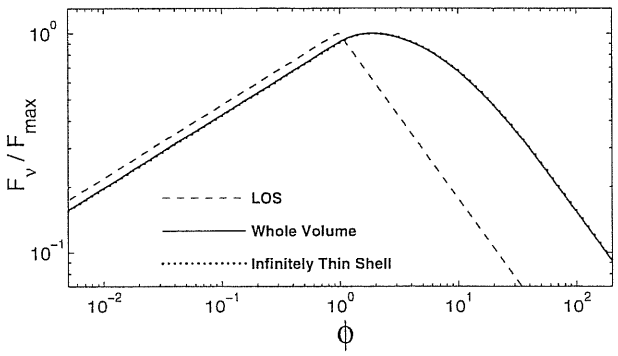

Figure 4.5: Calculated flux from a Blandford-McKee solution. $F_\nu(t)$ is plotted as a function of $\phi \equiv Constant \times \nu t^{3/2}$. Thus if we consider a constant time this figure yields the spectrum, while if we consider a fixed frequency it yields the light curve. The solid curve depicts emission from the whole fireball. The dashed line depicts the spectrum resulting from emission only along the line of sight. The dotted line is the emission from an infinity thin shell, which is very similar to that obtained for the full fireball due to a coincidence that rises from the choice for the magnetic field model. From Granot *et al.* (1999).

over a time scale comparable to the observation time t.

Using the adiabatic shell conditions (Eqs. 4.52 and 4.53), Eq. 4.26 for the shock conditions, Eq. 4.64 for the synchrotron energy and Eq. 4.66 for the "cooling energy", we find

$$\begin{aligned} \nu_c &= 2.7 \times 10^{12}\,\mathrm{Hz}\,\varepsilon_B^{-3/2} E_{52}^{-1/2} n_{1ism}^{-1} t_d^{-1/2} \quad , \\ \nu_m &= 5.7 \times 10^{14}\,\mathrm{Hz}\,\varepsilon_B^{1/2} \varepsilon_e^2 E_{52}^{1/2} t_d^{-3/2} \quad , \\ F_{\nu,max} &= 1.1 \times 10^5 \mu\mathrm{J}\,\varepsilon_B^{1/2} E_{52} n_{1ism}^{1/2} D_{28}^{-2} \quad , \end{aligned} \tag{4.96}$$

where $t_d = t/1\,\mathrm{day}$ and $D_{28} = D_L/10^{28}\,\mathrm{cm}$ and we have ignored cosmological redshift effects. Fig. 4.4 depicts the instantaneous spectrum in this case.

For a fully radiative evolution, we find

$$\begin{aligned} \nu_c &= 1.3 \times 10^{13}\,\mathrm{Hz}\,\varepsilon_B^{-3/2} E_{52}^{-4/7} \Gamma_{100}^{4/7} n_{1ism}^{-13/14} t_d^{-2/7} \quad , \\ \nu_m &= 1.2 \times 10^{14}\,\mathrm{Hz}\,\varepsilon_B^{1/2} \varepsilon_e^2 E_{52}^{4/7} \Gamma_{100}^{-4/7} n_{1ism}^{-1/14} t_d^{-12/7} \quad , \\ F_{\nu,max} &= 4.5 \times 10^3 \mu\ J\,\varepsilon_B^{1/2} E_{52}^{8/7} \Gamma_{100}^{-8/7} n_{1ism}^{5/14} D_{L,28}^{-2} t_d^{-3/7} \quad . \end{aligned} \tag{4.97}$$

These instantaneous spectra are also shown in Fig. 4.4.

The light curves at a given frequency depend on the temporal evolution of the break frequencies ν_m and ν_c and the peak power $N_e P_{syn}(\gamma_{e,min})$ (see Eq. 4.70). These depend, in turn, on how Γ and N_e scale as a function of t.

The spectra presented in Fig. 4.4 show the positions of ν_c and ν_m for typical parameters. In both the adiabatic and radiative cases, ν_c decreases more slowly with time than ν_m. Therefore, at sufficiently early times we have $\nu_c < \nu_m$, i.e. fast cooling. At

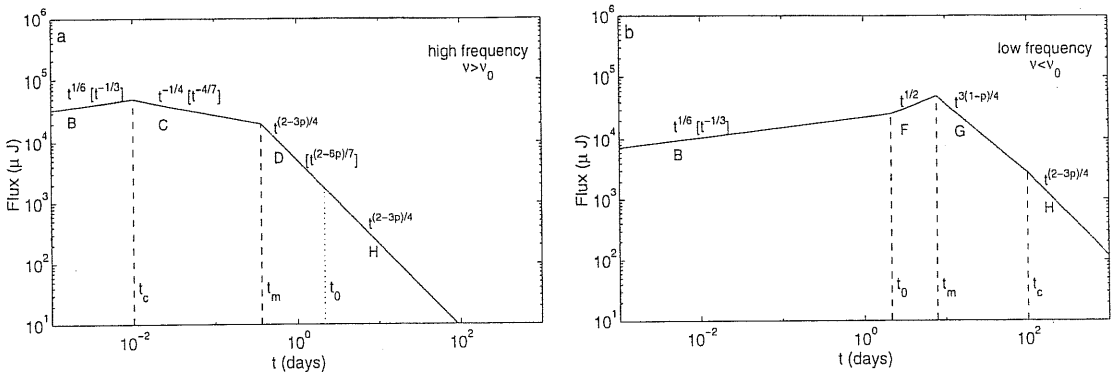

Figure 4.6: Light curve due to synchrotron radiation from a spherical relativistic external shock, ignoring the effect of self-absorption. (a) The high frequency case ($\nu > \nu_0$). The light curve has four segments, separated by the critical times, t_c, t_m, t_0. The labels, B, C, D and H indicate the correspondence with spectral segments in Fig.??. The observed flux varies with time as indicated; the scalings within square brackets are for radiative evolution (which is restricted to $t < t_0$) and the other scalings are for adiabatic evolution. (b) The low frequency case ($\nu < \nu_0$). From Sari *et al.* (1998).

later times, we have $\nu_c > \nu_m$, i.e. slow cooling. The transition between the two occurs when $\nu_c = \nu_m$ at t_{fs} (see Eq. 4.90). At $t = t_{fs}$, the spectrum changes from fast cooling (Fig. 4.4a) to slow cooling (Fig. 4.4b). In addition, if $\varepsilon_e \approx 1$, the hydrodynamic evolution changes from radiative to adiabatic. However, if $\varepsilon_e \ll 1$, the evolution remains adiabatic throughout.

Once we know how the break frequencies, ν_c, ν_m, and the peak flux $F_{\nu,max}$ vary with time, we can calculate the light curve. For a fixed frequency (e.g. $\nu = 10^{15}\nu_{15}\,\mathrm{Hz}$), from the first two equations in Eqs. 4.96 and 4.97 we see that there are two critical times, t_c and t_m, when the break frequencies ν_c and ν_m cross the observed frequency ν

$$t_c = \begin{cases} 7.3\times10^{-6}\,\mathrm{days}\,\varepsilon_B^{-3}E_{52}^{-1}n_{1ism}^{-2}\nu_{15}^{-2} & \text{adiabatic} \\ 2.7\times10^{-7}\,\mathrm{days}\,\varepsilon_B^{-21/4}E_{52}^{-2}\Gamma_{100}^{2}n_{1ism}^{-13/4}\nu_{15}^{-7/2} & \text{radiative} \end{cases}, \tag{4.98}$$

$$t_m = \begin{cases} 0.69\,\mathrm{days}\,\varepsilon_B^{1/3}\varepsilon_e^{4/3}E_{52}^{1/3}\nu_{15}^{-2/3} & \text{adiabatic} \\ 0.29\,\mathrm{days}\,\varepsilon_B^{7/24}\varepsilon_e^{7/6}E_{52}^{1/3}\Gamma_{100}^{-1/3}n_{1ism}^{-1/24}\nu_{15}^{-7/12} & \text{radiative} \end{cases}. \tag{4.99}$$

There are only two possible orders for the three critical time, t_c, t_m, t_{fs}, namely $t_{fs} > t_m > t_c$ and $t_{fs} < t_m < t_c$. We define the critical frequency, $\nu_0 = \nu_c(t_{fs}) = \nu_m(t_{fs})$

$$\nu_0 = \begin{cases} 1.8\times10^{11}\,\mathrm{Hz}\,\varepsilon_B^{-5/2}\varepsilon_e^{-1}E_{52}^{-1}n_{1ism}^{-3/2} & \text{adiabatic} \\ 8.5\times10^{12}\,\mathrm{Hz}\,\varepsilon_B^{-19/10}\varepsilon_e^{-2/5}E_{52}^{-4/5}\Gamma_{100}^{4/5}n_{1ism}^{-11/10} & \text{radiative} \end{cases}. \tag{4.100}$$

When $\nu > \nu_0$, we have $t_{fs} > t_m > t_c$ and we refer to the corresponding light curve as the **high frequency light curve**. Similarly, when $\nu < \nu_0$, we have $t_{fs} < t_m < t_c$, and we obtain the **low frequency light curve**.

Fig. 4.6a depicts a typical high frequency light curve. At early time, the electrons cool fast and $\nu < \nu_m$ and $\nu < \nu_c$. Ignoring self-absorption, the situation corresponds to segment B in Fig. 4.4, and the flux varies as $F_\nu \sim F_{\nu,max}(\nu/\nu_c)^{1/3}$. If the evolution is adiabatic, $F_{\nu,max}$ is constant, and $F_\nu \propto \nu^{1/6}$. In the radiative case, $F_{\nu,max} \propto t^{-3/7}$ and $F_\nu \propto \nu^{-1/3}$. The scaling in the other segments, which correspond to C, D, H in Fig. 4.4, can be derived in a similar fashion and are shown in Fig. 4.6a.

Fig. 4.6b shows the low frequency light curve, corresponding to $\nu < \nu_0$. In this case, there are four phases in the light curve, corresponding to to segments B, F, G, and H. The time dependences of the flux are indicated on the plot for both the adiabatic and the radiative cases.

For a relativistic electron distribution with a power distribution γ_e^{-p}, the uppermost spectral part behaves like $\nu^{-p/2}$. The corresponding temporal index (for adiabatic hydrodynamics) is $-3p/4$. In terms of the spectral index α, this yields the relation $F_\nu \propto t^{(1-3\alpha)/2}$. Alternatively for slow cooling, there is also another frequency range (between ν_m and ν_c) for which the spectrum is given by $\nu^{-(p-1)/2}$ and the temporal decay is $-3(p-1)/4$. Now we have $F_\nu \propto t^{-3\alpha/2}$. Note that in both cases there is a specific relation between the spectral index and the temporal index which could be tested by observations.

4.7 The Internal - External Scenario

We have seen in Sec. 4.4.3 that external shocks cannot produce bursts with a complicated temporal structure without the cost of low efficiency. However, some smooth bursts might be produced in this ways and even in some bursts emission is probably from both internal and external shocks (Sari 1997). In addition, it is general believed that the observed afterglow is produced by external shocks (see Sec. 4.8).

Internal shocks can convert only a fraction of the total energy to radiation (Mochkovitch *et al.* 1995; Katz 1997; Kobayashi *et al.* 1999). After the flow has produced a GRB via internal shocks, it will interact via an external shock with the surrounding medium (Sari & Piran 1997). This will produce the afterglow - a signal that will follow the GRB. The idea of an afterglow in other wavelengths was suggested earlier (Paczyński & Rhoads 1993; Katz 1994a; Mészáros & Rees 1997) but it was suggested as a follow up of the external shock scenario. In this case, the afterglow would have been a direct continuation of the GRB activity and its properties would have scaled directly to the properties of the GRB.

According to internal-external models, i.e. internal shocks for the GRB and external shocks for the afterglow, different mechanism produce the GRB and the afterglow. Therefore the afterglow should not be scaled directly to the properties of the GRB. This was in fact seen in the recent afterglow observations (Katz & Piran 1997). In

all models of external shocks, the observed time satisfy $t \cdot \propto R/\Gamma_e^2$ and the typical frequency satisfies $\nu \propto \Gamma_e^4$. Since most of the emission takes place at practically the same radius and all that we see is the variation of the Lorentz factor we expect quite generally (Katz & Piran 1997): $\nu \propto t^{2\pm\chi}$. The small parameter χ reflects the variation of the radius and it depends on the specific assumptions made in the model. We would expect that $t_x/t_\gamma \sim 5$ and $t_{opt}/t_\gamma \sim 300$. The observations of GRB970508 show that $(t_{opt}/t_\gamma)_{obs} \approx 10^4$. This is in a clear disagreement with the single external shock model for both the GRB and afterglow.

Under quite general conditions, the initial typical synchrotron energy for either the forward or the reverse external shock may fall in the soft GRB band. In this case, the initial stage of the afterglow might overlap the γ-ray emission from the internal shock (Sari 1997). The result will be superposition of a rapidly varying signal on top of a long smooth and softening pulse. This possibility should be explored in greater detail.

However, we have to notice that internal shock models suffer from several difficulties. Internal shocks require a variable flow. The situation in which an inner shell is faster than an outer shell is Rayleigh-Taylor unstable (Waxman & Piran 1994). The instability develops when the fireball becomes matter dominated before the shocks form and it may affect the energy conversion process. The full implications of this instability have to be understood. The second is that internal shock models require long-lasting central engines. And, the third is its inefficiency which leads to the necessity of beaming effects.

4.8 Confronting the Observations and the Problem of Present Models

4.8.1 The observations of Afterglow

It is generally believed that the observed afterglow results from slowing down of a relativistic shell on the external ISM. The afterglow is produced, in this case, by an external shock. A second alternative is that the central engine that powers the GRB continues to emit energy for much longer duration with a lower amplitude (Katz *et al.* 1998) and may produce the earlier part (first day or two in GRB970228 and GRB970508) of the afterglow. It is most likely that both processes take place to some extent. We compare the observations of afterglow with the prediction of theoretical models.

Shortly after the observations of GRB970228, Wijers *et al.* (1997) showed that the decline in the intensity in X-ray and several visual bands (from B to K) fit the afterglow model well. The previous discussion indicates that this agreement shows that the high energy tail (or late time behavior) is produced by a synchrotron emission

from a power-law energy distribution of electrons

There are much more data on the afterglow of GRB970508. The light curves in the different optical bands generally peak around two days. There is a rather steep rise before the peak which is followed by a long power law decay (see Figs. 2.12 and 2.13). In the optical band, the observed power law decay for GRB970508 is -1.141 ± 0.014 (Galama *et al.* 1998b). This implies a spectral index $\alpha = -0.761 \pm 0.009$ for an adiabatic slow cooling model, which was used by Wijers *et al.* (1997). However, the observed spectral index is $\alpha = -1.12 \pm +0.04$ (Galama *et al.* 1998). A fast cooling model, for which the spectral index is $p/2$ and the temporal behavior is $3p/4 - 1/2$, fits the data better as the temporal law implies that $p = 2.188 \pm 0.019$ while the spectral index implies, consistently, $p = 2.24 \pm 0.08$ (Sari *et al.* 1998; Galama *et al.* 1998c). Unfortunately, at present this fit does not tell us much about the nature of the hydrodynamic processes and the slowing down . The prompt afterglow of the bright GRB990123 was observed (see Sec. 2.11). The observations of optical prompt emission and its time decay is consistent reverse shock model with $p = 2.5$ (Sari & Piran 1999). The observation after about 0.1 days gives a evolution of $\propto t^{-1.1}$ (Fruchter *et al.* 1999; Kulkarni *et al.* 1999; Sagar *et al.* 1999; Galama *et al.* 1999) and a spectrum $\propto \nu^{-0.671}$ (Galama *et al.* 1999). The difficulties for all the present fireball models is that the none can explain the initial rise of the optical flux in GRB980508 and GRB990123. Some physic process must be present at the initial stage of the afterglow but we do not know.

Using both the optical and radio data, one can try to fit the whole spectrum and to obtain the unknown parameters that determine the fireball evolution (Wijers & Galama 1998; Granot *et al.* 1998). Wijers and Galama (1998) have attempted to do so, using the spectrum of GRB970508. They have obtained a reasonable set parameters. However, more detailed analysis (Granot *et al.* 1998) reveals that the solution is very sensitive to assumptions made on how to fit the observational data to the theoretical curve. Moreover, the initial phase of the light curve of GRB970508 does not fit any of the theoretical curves. This suggests that at least initially an additional process might be taking place. Because of the inability to obtain a good fit for this initial phase, there is a large uncertainty in parameters obtained in this way. The possible values of ε_e and ε_B for GRB980508 is ~ 0.6 and ~ 0.01, respectively. These value can also be used for GRB990123 (Sari & Piran 1999).

No deviation in the observed decaying light curve from a single power law was observed for GRB970508, until it faded below the level of the surrounding nebula. This suggests that there are was no significant beaming in this case. If the outflow is in the form of a jet, the temporal behavior will change drastically when the opening angle of the jet equals to $1/\Gamma$ (Rhoads 1997). However, the observation of the afterglow of GRB990123 and several others indicate that beaming effect does exist in some bursts

(see the discussion in Sec. 2.13). The beaming factor Γ may be in the range of 1 to ~ 300. This may indicate that a significant fraction of the spread in the luminosity is given by the beaming effect (Sari *et al.* 1999).

4.8.2 The Observations of GRBs

Fireball models are originally devised to explain the observations of GRB itself but until now it is not as successful as in the explanation of the observations of afterglow. The complex light curve seems to be produced via internal shocks and internal shocks can indeed produce such a light curve. However, no explanation has been given to a specific burst. As for spectrum, there are more question needed answering than those have been. Optically thin synchrotron shock model are general accepted as the standard model. However, studies show that the gamma-ray burst photons near the peak of the spectrum at several hundred keV may be thermal (Eichler & Levison 1999) which suggested a more compact fireball ($R \lesssim 10^{12}$ cm) than expected by the standard optically thin shock models, or even whole the emission may be from blackbody radiation (Blinnikov *et al.* 1999).

One prediction of the synchrotron shock models is that the asymptotic low-energy (below the spectral break) photon slope α ($n_\nu \propto \nu^{-\alpha}$) should be between $-2/3$, in the case of an instantaneous synchrotron spectrum, and $-3/2$, when the spectrum is integrated over the radiative decay of electron energies (Cohen *et al.* 1997). These predictions are tested by Cohen *et al.* (1997) with 11 times-integrated GRBs and are consistent with the observations. However, time-resolved spectra (Crider *et al.* 1997; Preece *et al.* 1998) clearly yield that a significant number of bursts have maximum α values beyond the "line-of-death": $\alpha > -2/3$ predicted by the synchrotron models. One way to circumvent this problem (Papathanassiou 1999) may be that the flow has a high compactness and produces a large density of pairs that live long enough to contribute to the synchrotron emission and turn it optically thick in the BATSE range. A compact fireball model has also been suggested by Eichler & Levinson (1999).

Saturated Comptonization model, propose by Liang (1997) and examined by Liang *et al.* (1997, 1999) with observations of GRBs, require a large optical depth ($\tau \gg 1$) at the initial stages of GRBs. Indeed, the observed time variation of spectral index α in the time-resolved spectra of some bursts (Crider *et al.* 1997) and more clearly the increase of the optical depth (~ 4) in GRB950818 during its early evolution (Crider *et al.* 1997) show that some bursts are produced when the fireball transit from optically thick to optically thin. Therefore, it is necessary dynamically and radiatively to investigate fireball from completely optically thick case through the transition to completely optically thin. That is what we can address in our models, which will presented in next chapters.

5 General Relativistic Hydrodynamics for Fireball

5.1 Introduction

We discussed the present situation of the research on fireball models in Chapter (4) on the aspects of hydrodynamics, emission mechanism and the predication confronting the observations of GRB itself and the afterglow. The fireball model is very successful in the explanation of the light curve of afterglow but still over-simplified for GRB itself. The basic picture might be that the internal shocks produces the complex gamma-ray emission and the external shocks are in charge of the afterglow. While the latter picture has been confirmed, there are still many questions on the physical situations in the shocked regions and in the external medium.

The situation for GRB itself is very complicated. The possible observation of prompt reverse shock associated external shock in GRB990123 imply that the external shock is important during the gamma-bursts. Generally, it is assumed that the internal shocks happen after the fireball becomes optically thin. However, the observations show that internal shocks may happen when the fireball is during its early stages and still optically thick or during the transition from optically thick to optically thin stages. It is still not clear about the hydrodynamics and the spectrum property in the transition. So, it is necessary to investigate the fireball and its radiation mechanism from the initial acceleration through the transition stages to optically thin and finally the stages stopped by the external medium. Another complication is that GRBs is probably associated the formation of a compact object, most possibly a black hole. A fireball will expand immediately after it is produced along the rotational axis of the system. As it is produced very close to a compact object, general relativity is important during its initial stage but not at later stage.

To simulate the whole process, a numerical scheme which can treat relativistic shocks has to run for many orders of magnitude increase in the radius and therefore should not be too expensive. Generally, methods based on a Riemann solver can give highly accurate results for relativistic shocks but are expensive. Methods using

artificial viscosity to treat shocks are less expensive but give a much less accurate treatment of the shock than those based on Riemann solvers unless one uses adaptive-mesh-refinement techniques or other special treatments in this region. However, if one's main concern is for the hydrodynamics of the whole region and for the radiation instead of for the detailed structure of the shock, the straightforward artificial viscosity technique is still satisfactory.

The final goal of our project is to use general relativistic hydrodynamics supplemented by artificial viscosity for shocks to investigate the relativistic radiation hydrodynamics of fireballs. General relativity may be important at the initial stage of the expansion of a fireball but not in its late stage, as a fireball is produced at a position near a black hole in present favorite center-engine models (see Chapter (3)). Note also that while this can, for the most part, be satisfactorily treated within special rather than general relativity, it is nevertheless convenient to five a general relativistic formulation since we will be dealing with accelerated frame which can be handled very naturally in this way and, in any case, GR does not introduce any serious further complications for as long as one remains within spherically symmetry as we will here. As it is a large project, we divide it into three stages. The first is to study the hydrodynamic expansion of the fireball into vacuum. The second is to study the interaction of fireball and external medium. The final step is to investigate the radiation hydrodynamics of the fireball interacting with the external medium.

We present the physical background and derive the equations which we will use in the numerical simulation. In the next chapter, we will discuss the numerical technique used and some results which have been obtained. The project is still going.

We first review the relativistic hydrodynamic equations for a spherically symmetric system in Sec. (5.2) and Sec. (5.3) and the characteristic forms of the hydrodynamic equations in Sec. (5.4). As we will treat the interaction of the fireball with the external medium and possibly the internal shocks, depending on the initial conditions, we derive the junction conditions at the shock and contact discontinuities (in Sec. (5.5)), following the argument of May & White (1967). The interaction of fireball with the external medium belongs to the Riemann problem for the initial configuration, which we will discuss a bit more in Sec. (5.6). This discussion will be helpful in the studies of the boundary in the expansion into vacuum and of the shocks and the contact surface in the interaction of the fireball with the external medium.

Although we have not yet started our numerical simulation on the relativistic radiation hydrodynamics, we derive in Sec. (5.7) the relativistic radiation hydrodynamic equations which will be used in further work.

5.2 The Metric in Comoving Frame

The most general metric exhibiting spherical symmetry is

$$ds^2 = a^2(dx^0)^2 - b^2(dx^1)^2 - R^2 d\Omega^2 + \chi dx^0 dx^1 \quad , \tag{5.1}$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta\, d\phi^2$ and the metric coefficients a, b, χ and R are functions of x^0 and x^1 only. One can always define a new time coordinate so that the metric is made diagonal and we make this choice putting $\chi = 0$.

Let $x^0 \equiv ct$ and $x^1 \equiv \mu$ be the time and the radial coordinates respectively, we re-write the metric as

$$ds^2 = a^2c^2dt^2 - b^2 d\mu^2 - R^2 d\Omega^2 \quad . \tag{5.2}$$

Since the proper circumference of a shell characterized by the radial coordinate μ at time t is

$$\int ds = \int g_{\theta\theta}{}^{1/2}\, d\theta = 2\pi R(t,\mu) \quad , \tag{5.3}$$

the metric component $R(t,\mu)$ is often referred to as the "Schwarzschild circumference coordinate". $g_{\theta\theta}$ is the $\theta\theta$ component of the metric tensor $g_{\alpha\beta}$. In Newtonian and special relativistic limits, it represents the Eulerian radial coordinate of a sphere of fluid with the comoving coordinate μ and at time t.

Thus, the metric together with the four-velocity of the fluid in the comoving frame

$$u^\alpha|_{comoving} \equiv cdx^\alpha/ds|_{streamlines} = [a^{-1}c, 0, 0, 0] \tag{5.4}$$

uniquely determines the comoving spherically symmetric frame except for an arbitrary scaling of the radial coordinate μ and an arbitrary scaling of the coordinate time t. We have used the condition

$$c^2 = u^\alpha u_\alpha = g_{\alpha\lambda}u^\lambda u^\alpha \tag{5.5}$$

to normalize Eq. 5.4. We have adopted Einstein notation: repeated Greek indices are summed from 0 to 3 and Latin indices from 1 to 3.

In order to investigate shocks in the fluid, it is convenient to introduce a relativistic analog of the classical Eulerian coordinate frame. The closest frame is referred to as the Schwarzschild frame using R as the radial coordinate. The metric in this frame is

$$ds^2 = A^2c^2dT^2 - B^2dR^2 - R^2d\Omega^2 \quad , \tag{5.6}$$

where A and B are functions of R and T only. Denoting the radial component of the fluid four-velocity in the Schwarzschild frame by u and using the normalization equation Eq. (5.5), we find

$$u^\alpha = [A^{-1}(c^2 + B^2u^2)^{1/2}, u, 0, 0] \quad . \tag{5.7}$$

The transformation law for the contravariant metric coefficients between the comoving and the Schwarzschild frames gives following relations

$$1/A^2c^2 = (T_t^2/a^2c^2) - (T_\mu^2/b^2) \quad , \tag{5.8}$$

$$-1/B^2 = (R_t^2/a^2c^2) - (R_\mu^2/b^2) \quad , \tag{5.9}$$

$$0 = (R_tT_t/a^2c^2) - (R_\mu T_\mu/b^2) \quad , \tag{5.10}$$

where the subscripts indicate conventional partial derivatives, e.g. $T_t \equiv \partial T/\partial t$ and $T_\mu \equiv \partial T/\partial \mu$. The transformation of the contravariant four-velocity yields

$$u = R_t/a \quad , \tag{5.11}$$

$$(c^2 + B^2u^2)^{1/2}/A = cT_t/a \quad . \tag{5.12}$$

We introduce another quantity Γ

$$\Gamma \equiv R_\mu/b \quad , \tag{5.13}$$

and together with equation (5.11), we write Eq. (5.9) as

$$1/B^2 = \Gamma^2 - u^2/c^2 \quad . \tag{5.14}$$

From Eqs. (5.8), (5.10), (5.11), (5.13) and (5.14), we obtain

$$T_t/a = B\Gamma/A \quad , \tag{5.15}$$

and

$$T_\mu/b = Bu/Ac^2 \quad . \tag{5.16}$$

5.3 Perfect Fluid Hydrodynamics with Spherical Symmetry

For a perfect fluid, the energy moment tensor takes the form

$$T_\alpha{}^\beta = \rho\omega g_{\alpha\lambda}u^\lambda u^\beta - g_\alpha{}^\beta p \quad , \tag{5.17}$$

where ρ is the proper rest mass density, p proper pressure and ω the relativistic enthalpy given by

$$\omega = 1 + [\epsilon + (p/\rho)]/c^2 \quad , \tag{5.18}$$

where ϵ is the proper specific internal energy (per unit mass).

The metric in Eq. (5.2) is determined from the Einstein equation

$$-\frac{8\pi G}{c^4}T_\alpha{}^\beta = R_\alpha{}^\beta - \frac{1}{2}g_\alpha{}^\beta R_\lambda{}^\lambda \quad , \tag{5.19}$$

where $R_\alpha{}^\beta$ is the Ricci tensor defined in term of the Riemann curvature tensor (Misner *et al.* 1973). Using Eqs. (5.2), (5.4) and (5.17), the Einstein equations reduce to

$$4\pi G\rho\left(1+\frac{\epsilon}{c^2}\right)R^2R_\mu = \frac{c^2}{2}\left[R+\frac{RR_t^2}{a^2c^2}-\frac{RR_\mu^2}{b^2}\right]_\mu \tag{5.20}$$

$$\frac{4\pi G}{c^2}pR^2R_t = -\frac{c^2}{2}\left[R+\frac{RR_t^2}{a^2c^2}-\frac{RR_\mu^2}{b^2}\right]_t \tag{5.21}$$

$$4\pi G\rho\omega R^3 = c^2\left[R+\frac{RR_t^2}{a^2c^2}-\frac{RR_\mu^2}{b^2}\right]+\frac{R^3c}{ab}\left[\left(\frac{ca_\mu}{b}\right)_\mu-\left(\frac{b_t}{ac}\right)_t\right] \tag{5.22}$$

$$0 = \frac{a_\mu}{a}R_t+\frac{b_t}{b}R_\mu-R_{\mu t} \quad . \tag{5.23}$$

Eqs. (5.20) to (5.23) together with the equation of state

$$p = p(\rho,\epsilon) \tag{5.24}$$

provide us only five equations for six unknowns a, b, R, ρ, ϵ and p. One more equation is needed. Actually, the radial coordinate scaling factor b can be fixed by identifying μ with the proper rest mass m_0 within a sphere of a "radius" μ; that is

$$m_0 = \int_0^\mu 4\pi bR^2\rho\, d\mu \quad , \tag{5.25}$$

and $\mu = m_0$.

Differentiating it, we have

$$d\mu = 4\pi bR^2\rho\, d\mu \quad . \tag{5.26}$$

It follows that

$$b = \frac{1}{4\pi R^2\rho} \quad . \tag{5.27}$$

For a system with spherical symmetry, this choice of radial coordinate is almost universally used for classical Lagrangian hydrodynamics in spherical symmetry since carrying the mass as a coordinate automatically guarantees mass conservation.

From Einstein equation (5.19), we have local conservation law of energy and momentum

$$T_\alpha{}^\beta{}_{;\beta} = 0 \quad , \tag{5.28}$$

where the semicolon denotes covariant derivatives associated with the metric $g_{\alpha\beta}$. In the comoving frame, this equation can be written explicitly as

$$(T_0{}^\alpha{}_{;\alpha}=0) \qquad [\rho(1+\epsilon/c^2)]_t+[(b_t/b)+2(R_t/R)]\rho\omega = 0 \quad , \tag{5.29}$$

$$(T_1{}^\alpha{}_{;\alpha}=0) \qquad (a_\mu/a)+(p_\mu/\rho\omega c^2) = 0 \quad . \tag{5.30}$$

Substituting Eq. (5.27) into Eq. (5.29), we find

$$\epsilon_t = (p/\rho^2)\rho_t \quad , \tag{5.31}$$

which is just a restatement of the first law of thermodynamics when there is no heat transfer ($\epsilon_t = -p(1/\rho)_t$. From Eqs. (5.11), (5.27) and (5.23), we have the mass conservation equation

$$\frac{(\rho R^2)_t}{\rho R^2} = -a\frac{u_\mu}{R_\mu} \quad . \tag{5.32}$$

Eq. (5.30) with Eq. (5.18) can be written as

$$\frac{(a\omega)_\mu}{a\omega} = \frac{\epsilon_\mu - (p/\rho)(\rho_\mu/\rho)}{\omega c^2} \quad . \tag{5.33}$$

From Eqs. (5.20) and (5.21), If we define a mass function $m(\mu, t)$ such that

$$m_\mu = 4\pi\rho(1+\epsilon/c^2)R^2 R_\mu \quad , \tag{5.34}$$

$$m_t = -\frac{4\pi}{c^2}pR^2 R_t \quad , \tag{5.35}$$

we obtain the general relativistic analogue of the Lorentz contraction factor

$$\Gamma = (1 + u^2/c^2 - 2Gm/Rc^2)^{1/2} \quad , \tag{5.36}$$

where we have used Eqs. (5.11) and (5.13) and the gravitational mass is given by

$$m(\mu, t) = 4\pi \int_0^\mu \rho(1+\epsilon/c^2)R^2 R_\mu \, d\mu \quad . \tag{5.37}$$

Using Eqs. (5.36), (5.21), (5.13), (5.11), (5.23) and (5.30), we get the "equation of motion"

$$u_t = -a\left[4\pi R^2 \frac{\Gamma}{\omega}p_\mu + \frac{Gm}{R^2} + \frac{4\pi G}{c^2}pR\right] \quad . \tag{5.38}$$

Equations (5.11), (5.31), (5.32), (5.18), (5.33), (5.36) (or (5.13)), (5.37), (5.38), and the equation of state, comprise a complete set of equations for spherically symmetric general-relativistic hydrodynamics. These equations together with initial and boundary conditions describe the evolution of our system. The boundary conditions at the center $\mu = 0$ are $u = 0$, $R = 0$ and $m = 0$. We also assume that the system is not singular at the center in which case we must have $\Gamma = 1$ there. The initial and outer boundary conditions will be discussed in detail in Chapter 6.

The Newtonian limit can be obtained by letting $c^2 \to \infty$ and taking $a = 1$ everywhere. To get special-relativistic limit, one just sets $G = 0$. If we change in the equations $4\pi R^2 \to \rho$ and put $G = 0$ and $b = 1/\rho$, we get the equations for plane symmetric geometry.

5.4 The Characteristic Form of the Hydrodynamic Equations

The system we obtained is first order quasi-linear hyperbolic in one space dimension and time. For such a system, it is often more convenient to rewrite it in characteristic forms, giving a set of ordinary differential equations along particular directions.

Miller & Pantano (1989) gave the characteristic equations of our system. Here we follow their argument. To get the characteristic equations of the system, we use Eqs. (5.11) and (5.13) to rewrite Eq. (5.32) as

$$\rho_t + (4\pi\rho^2 R^2 a/\Gamma)u_\mu + 2a\rho u/R = 0 \quad . \tag{5.39}$$

Substituting Eq. (5.39) into Eq. (5.31), we obtain

$$p_t + \omega {c_s}^2((4\pi\rho^2 R^2 a/\Gamma)u_\mu + 2a\rho u/R) = 0 \quad , \tag{5.40}$$

where c_s is the local sound speed defined by $c_s \equiv (\partial p/\partial e)_s^{1/2}$ with partial derivative at constant specific entropy, where $e = \rho(1 + \epsilon/c^2)$. Eqs. (5.39), (5.40) and (5.38) can be written in a compact form

$$\frac{\partial Y}{\partial t} + A\frac{\partial Y}{\partial \mu} + D = 0 \quad , \tag{5.41}$$

where

$$Y = \begin{bmatrix} u \\ p \\ \rho \end{bmatrix}, \quad A = \begin{bmatrix} 0 & a\Gamma/b\rho\omega & 0 \\ a\rho\omega c_s^2/b\Gamma & 0 & 0 \\ a\rho/b\Gamma & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} \alpha \\ \beta\omega c_s^2 \\ \beta \end{bmatrix}, \tag{5.42}$$

where $\alpha = aG[(m/R^2) + 4\pi R/c^2]$ and $\beta = 2au\rho/R$. Since our system is hyperbolic and A is a 3×3 matrix, A has 3 different real eigenvalues λ_i with $i = 1, 2, 3$ (Jeffrey 1978). λ_i is given by

$$l_i A = \lambda_i l_i \qquad \text{for } i = 1, 2, 3 \quad , \tag{5.43}$$

where l_i is the ith left eigenvector of A. The eigenvalue equation is then

$$\lambda(\lambda^2 - a^2 c_s^2/b^2) = 0 \quad , \tag{5.44}$$

which gives the eigenvalues

$$\lambda_0 = 0 \quad , \tag{5.45}$$

$$\lambda_\pm = \pm 4\pi a\rho R^2 c_s \quad . \tag{5.46}$$

The corresponding eigenvectors (up to a multiplicative constant) are

$$l_0 = [0, -1/\omega c_s^2, 1] \quad , \tag{5.47}$$

$$l_\pm = [\pm 1, \Gamma/\rho\omega c_s, 0] \quad . \tag{5.48}$$

Then, we have the characteristic equation of (5.41)

$$l_i\frac{\partial Y}{\partial t} + \lambda_i l_i\frac{\partial Y}{\partial \mu} + l_i D = 0 \quad , \tag{5.49}$$

or

$$l_i dY + l_i D dt = 0 \quad , \tag{5.50}$$

along

$$d\mu/dt = \lambda_i \quad . \tag{5.51}$$

For $i = 0$, from Eqs. (5.47), (5.49) and (5.51) we obtain the advective characteristic equation

$$dp = \omega c_s^2 d\rho \quad \text{along } d\mu = 0 \quad . \tag{5.52}$$

For $i = \pm$, we get the forward (+) and backward (-) characteristic equations

$$\begin{aligned} du \pm \frac{\Gamma}{\rho\omega c_s} dp + \left[G\left(\frac{m}{R^2} + \frac{4\pi p R}{c^2}\right) \pm \frac{2u\Gamma c_s}{R} \right] a\, dt &= 0 \\ \text{along} \quad d\mu &= \pm 4\pi R^2 a \rho c_s\, dt \quad . \end{aligned} \tag{5.53}$$

From Eqs. (5.11) and (5.35), we get the advective characteristic equations for R and m

$$dR = a\, u\, dt \quad , \tag{5.54}$$

$$dm = -\frac{4\pi}{c^2} a p u R^2\, dt \quad , \tag{5.55}$$

along

$$d\mu = 0 \quad . \tag{5.56}$$

For ω, we have only a constraint equation from Eq. (5.33)

$$d\ln(a\omega) = (\ln(\epsilon) + (p/\rho\epsilon)\ln(\rho))(\epsilon/\omega c^2) \quad , \tag{5.57}$$

along $dt = 0$.

The characteristic form of the hydrodynamics equations has been widely used in numerical calculations in newtonian theory, special and general relativity. For instance, Vitello & Salvati (1976) used special relativistic characteristic methods to compute the hydrodynamic evolution of a fireball under free expansion. Miller & Pantano (1990) use general relativistic characteristic methods in their studies on the growth of hadronic bubbles during the cosmological quark-hadron transition.

5.5 Junction Conditions for Shocks and Contact Discontinuities

For a nonlinear hyperbolic system, the solution of the equations may become discontinuous even though the initial values and boundary values are smooth (for more discussions, see e.g. Kröner 1997).

In this section, we will discuss the junction conditions for shock and contact discontinuities. We will focus on shocks but note that contact discontinuities represent a special case of this treatment where there is no fluid flow across the discontinuity surface. In the treatment of non-relativistic shocks, one obtains three independent junction conditions resting on the existence of integral forms of the conservation laws of energy, momentum, and mass (Courant & Friedrich 1976). However, in the general relativistic treatment of shocks, the situation becomes more complicated, as the fluid equations are not derived from integral conservation laws and the metric coefficients $g_{\alpha\beta}$ may be discontinuous at a shock.

The relativistic junction conditions have been discussed by many authors (e.g. Synge 1960; Israel 1966). Israel (1966; see also Maeda 1986) used a general way to treat two types of discontinuities: surface layers characterized by a δ-discontinuity in the density (e.g. two different phases) and boundary surfaces (shock waves and contact discontinuities) characterized by a finite jump in some quantities. Synge's method is for a particular types of coordinate frames, called admissible coordinates. Schwarzschild frame belongs to the "particular class" of frame, while, unfortunately, comoving frame does not. May & White (1967) follow Synge to get the junction conditions for Schwarzschild frame and then transform them to the comoving frame. We will follow their arguments to introduce the junction conditions.

5.5.1 Junction Condition for the Metric

We first explore the metric junction conditions. As the space-time interval between two events **on the world surface of the shock** measured by two observers one just ahead of the shock and the other just behind the shock is the same, we have

$$ds^2 = a_+^2 c^2 dt^2 - b_+^2 d\mu^2 - R_+^2 d\Omega^2 = a_-^2 c^2 dt^2 - b_-^2 d\mu^2 - R_-^2 d\Omega^2 \quad , \tag{5.58}$$

where the subscripts $\pm$ refer to "ahead of " and "behind" the shock, respectively. We shall adopt the standard notation

$$[\xi]_\pm \equiv \xi_+ - \xi_- \quad . \tag{5.59}$$

As we are interested in spherically symmetric shocks, we can select $dt = d\mu = 0$ and get relation

$$[R]_\pm = 0 \quad , \tag{5.60}$$

which means that the proper "circumferential" length $R\,d\Omega$ on the shock front measured by observers + and - at a fixed time is the same. From Eqs. (5.58) and (5.27), we have

$$0 = [a^2 c^2 - (\dot{\mu}_s / 4\pi\rho R^2)^2]_\pm \quad , \tag{5.61}$$

where $\mu_s(t)$ is the location of shock front and $\dot{\mu}_s \equiv d\mu_s/dt$ is the shock speed. From the continuity of R at all time, it follows that (dR_s/dt) is also continuous on the surface. As

$$dR = (\partial R/\partial t)dt + (\partial R/\partial \mu)d\mu = au\,dt + \Gamma b\,d\mu \quad , \tag{5.62}$$

we find another junction condition equation

$$[au + b\Gamma\dot{\mu}_s]_\pm = 0 \quad . \tag{5.63}$$

Following an argument identical to that above, from Eq. (5.6) we obtain the junction conditions for the metric in Schwarzschild frame

$$[R]_\pm = 0 \quad , \tag{5.64}$$

$$\left[A^2c^2 - B^2c^2S^2\right]_\pm = 0 \quad , \tag{5.65}$$

where the S is the shock speed divided by c

$$S \equiv (dR_s/dT)/c \quad . \tag{5.66}$$

Using the transformation relations (5.8) and (5.10), and Eqs. (5.11), (5.13), (5.15) and (5.16), we obtain

$$cS = \frac{R_t + \dot{\mu}_s R_\mu}{T_t + \dot{\mu}_s T_\mu} = A\sqrt{\Gamma^2 - u^2/c^2}\,\frac{au + b\dot{\mu}_s\Gamma}{a\Gamma + ub\dot{\mu}_s/c^2} \quad . \tag{5.67}$$

As in the case we are interested in no source contributes on the shock front to the mass function (mass conservation), we have

$$[m]_\pm = 0 \quad . \tag{5.68}$$

Therefore, from Eqs. (5.14), (5.36) and (5.65), we have

$$[A]_\pm = 0, \qquad [B]_\pm = 0 \quad . \tag{5.69}$$

5.5.2 Junction Conditions for Energy Momentum Tensors

As the Schwarzschild frame belongs to the particular class of coordinates, called admissible coordinates, from Synge (1960) the junction conditions are

$$[T_\alpha{}^\beta \partial f/\partial x^\beta]_\pm = 0 \quad , \tag{5.70}$$

where $f(x^0, x^1, x^2, x^3) = 0$ is the equation for the world surface of the shock and $\partial f/\partial x^\beta$ is a vector normal to the shock. Eq. (5.70) states that the flux of energy and momentum into one side of the shock surface equals the flux out of the other side.

Since at a shock surface,

$$f = R_s(T) - R = 0 \quad , \tag{5.71}$$

then

$$\partial f/\partial x^{\beta} = [S, -1, 0, 0] \quad . \tag{5.72}$$

Substituting Eqs. (5.6), (5.17), (5.72), and then (5.7) into (5.70), we get

$$\left[(\rho\omega(c^2 + B^2u^2) - p)S - \rho\omega A(c^2 + B^2u^2)^{1/2}u\right]_{\pm} = 0 \quad , \tag{5.73}$$

$$\left[-\rho\omega B^2 u A^{-1}(c^2 + B^2u^2)^{1/2}S + \rho\omega B^2u^2 + p\right]_{\pm} = 0 \quad . \tag{5.74}$$

From Eqs. (5.14) and (5.69), we obtain

$$[\Gamma^2c^2 - u^2]_{\pm} = 0 \quad . \tag{5.75}$$

From Eqs. (5.7), (5.14), (5.61), (5.63), (5.67), (5.69), (5.73), (5.74), and (5.75), we get the final form of our junction conditions as following

$$\left[\frac{\dot{\mu}_s}{4\pi R^2}u\left(1 + \frac{\epsilon}{c^2}\right) - ap\Gamma\right]_{\pm} = 0 \tag{5.76}$$

$$\left[\frac{\dot{\mu}_s}{4\pi R^2}\Gamma c^2\left(1 + \frac{\epsilon}{c^2}\right) - apu\right]_{\pm} = 0 \tag{5.77}$$

$$\left[\frac{\dot{\mu}_s}{4\pi R^2}\frac{\Gamma}{\rho} + au\right]_{\pm} = 0 \tag{5.78}$$

where we have used relation $b = 1/4\pi\rho R^2$. These three equations, together with Eqs. (5.60), (5.61) and (5.68), give the required junction conditions for spherically symmetric general-relativistic shocks.

From the energy equations (5.76) and (5.77) and (5.61), we get the notable equation

$$[a\omega]_{\pm} = 0 \quad . \tag{5.79}$$

Equation (5.78) and (5.33) together imply that if an initially isentropic medium is penetrated by a shock of uniform strength, $a\omega = constant$ throughout the medium.

The junction conditions for slab can easily be obtained from Eqs. (5.61) and (5.75) - (5.78) by replacing $\dot{\mu}_s/4\pi R^2$ by $\dot{M}$, the flux of rest mass per unit area crossing the shock surface. To study special relativistic, just taking $G = 0$ to neglect the effects of gravitation.

5.5.3 Contact Discontinuities

The discontinuity conditions (5.60), (5.61), (5.68), (5.75), (5.76), (5.77) and (5.78) admit a trivial or degenerate solution. If the flux $\dot{\mu}_s$ through the surface of discontinuity is zero, i.e. no fluid crosses it, we call it a contact discontinuity. From Eqs. (5.60), (5.61), and (5.78), we obtain the junction conditions for a contact continuity

$$[R]_{\pm} = 0 \tag{5.80}$$

$$[a]_{\pm} = 0 \tag{5.81}$$

$$[u]_{\pm} = 0 \quad . \tag{5.82}$$

From Eq. (5.75) and (5.82), we get

$$[\Gamma]_{\pm} = 0 \quad . \tag{5.83}$$

Eq. (5.76) (or (5.77)), (5.82) and (5.83) altogether give

$$[p]_{\pm} = 0 \quad . \tag{5.84}$$

Of course, we have

$$[m]_{\pm} = 0 \quad . \tag{5.85}$$

We cannot use Eq. (5.79) to get $[a\omega]_{\pm} = 0$, as we have used condition $\dot{\mu}_s \neq 0$ to derive (5.79). Therefore, $a\omega$ could have a jump across a contact discontinuity.

The junction conditions for the metric and from the conservation of energy-momentum do not require the temperature and mass density to be continuous across the discontinuity surface. A contact surface moves with the fluid and separates two zones of different (and temperature); but the pressure p, velocity u, general-relativistic Doppler contraction factor Γ and time scaling factor a being the same on both sides. A contact discontinuity may separate not only parts of the fluid but also two different fluids.

It is obvious that in reality such a contact surface cannot be maintained for an appreciable length of time; heat conduction between permanently adjacent particles on either side of the discontinuity would soon make our idealized assumption unrealistic. A contact layer will gradually fade out, depending on the ratio of the hydrodynamic time scale and the heat conductive time scale. In the case we are interested in, the assumption is reasonable.

5.6 Initial Discontinuity and Relativistic Riemann Problem

When a spherical-symmetric fireball with very high pressure and high density is set within an external medium having very low pressure and density, the fireball will interact with the external medium strongly. This kind of interaction is a special case of a more general type of phenomenon, called the Riemann problem, in three dimensions with spherical symmetry.

The Riemann problem concerns finding the fluid flow resulting from an initial state in which the fluid on the right-hand side of some boundary is in an arbitrary constant state (r) and on the left-hand side in an another (different) constant state (l). The surface of their contact is a surface of discontinuities of pressure, mass density, and velocity. The discontinuities in the initial conditions are called **initial discontinuities**. The discontinuities of the various quantities can be of any values and it is not necessary to have any relation between them.

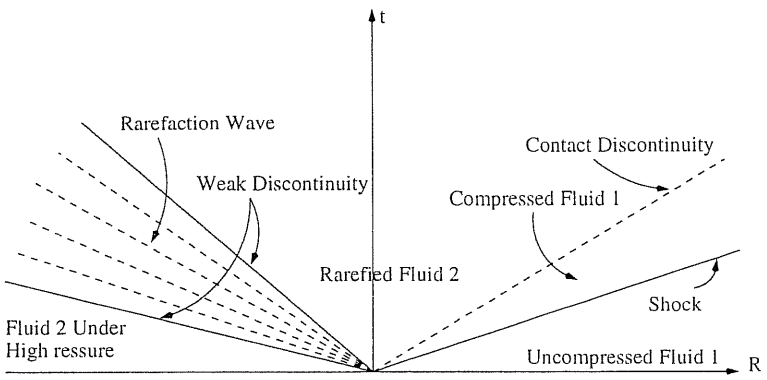

Figure 5.1: Wave motion produced in the Riemann problem in R–t plane. The rarefaction wave is centered.

Since to be stable surfaces of discontinuity in a flow, certain conditions must hold, initial discontinuities which do not satisfy the conditions cannot continue to exist subsequently but instead in general split into several possible discontinuities, shock wave, contact discontinuity, weak discontinuity and a centered rare-faction wave (Fig. 5.1 shows the situation for 1D shock tube with high pressure fluid2 expanding out into uncompressed fluid 1). The discontinuities must move away from the position of the initial discontinuity, in both directions depending on inequalities prevailing for the initial quantities. The two given states r and l must be connected by the discontinuities formed from the initial discontinuity. As the state of the flow is determined by three independent quantities, e.g. p, ρ, and u, it is necessary to have three free parameters in order to go from an arbitrary state l to another arbitrary state r. A shock wave propagating into a given state of matter is completely determined by one parameter (see the discussion of the junction condition in Sec. 5.5). The state of a flow leaving a rarefaction wave is also completely determined only by one parameter, when the state of the flow entering it is given. At the contact (tangential) discontinuity stationary with respect to the flow on both sides of it, the pressure and velocity are continuous, but the mass density (and thus temperature) is not (see Sec. 5.5).

Many modern numerical schemes, including the well-known PPM scheme, devised to simulate fluid dynamics on computers are based on the solution of the Riemann problem (for review see, e.g. Woodward & Colella 1984; Colella & Woodward 1984; Anile 1989). The non-relativistic Riemann problem has been discussed in detail by Courant & Friedrich (1976), by Landau & Lifshitz (1979) and also by Zel'dovich & Raizer (1967). A solution to the Riemann problem in Newtonian hydrodynamics was given by Godunov (1959). The analytic solution to the Riemann problem in relativistic hydrodynamics was derived by Thompson (1986) for the particular case, called *shock-tube problem*, when the initial states are at rest and by Smoller & Temple (1993) and by Martí & Müller (1994) for the general case. The application of the methods to the

studies of GRBs utilizing the solution of the relativistic Riemann problem have been discussed by many authors (e.g. Wen *et al.* 1997; Kobayashi *et al.* 1999; Aloy *et al.* 1999).

We will be interested in a simple case, but not only, that a fireball expands into vacuum. If the specific internal energy $\epsilon = \eta c^2$ and the expansion is adiabatic with adiabatic index γ, Martí & Müller (1994) show that the maximum Doppler factor of the rarefaction wave in shock-tube problem is given by

$$\Gamma_{max} = \frac{1}{2}\left[2\gamma\eta\left(1+\sqrt{1+\frac{1}{\gamma\eta}}+\frac{1}{2\gamma\eta}\right)\right]^{1/\sqrt{\gamma-1}} , \tag{5.86}$$

where γ is adiabatic index.

5.7 General Relativistic Radiation Hydrodynamics

5.7.1 Relativistic Radiative Transfer and PSTF Tensor Formalism

To treat properly the coupling between radiation and matter in general relativity, one has to seek the relativistic theory of radiation hydrodynamics. However, the general relativistic form of the radiative transfer equation describing the variation of the radiation field as it propagates through a moving fluid,

$$\frac{d\mathcal{N}}{dl} = \mathcal{S} \tag{5.87}$$

is too complicated to handle in full detail at present time. In Eq. (5.87), $\mathcal{N}$ is the photon occupation number which is a relativistic invariant, l is a non-affine parameter along the photon trajectory in phase-space, and $\mathcal{S}$ is a source function which describe the effects of the interaction between matter and radiation, its actual form depending on the radiative processes which are considered. The total derivative is taken in the phase space, as $\mathcal{N} = \mathcal{N}(x^\alpha, p^\alpha)$, where p^α is the photon four-momentum.

A number of different approaches (Rezzolla & Miller 1994) have to be applied to obtain its approximate solution. Among the methods, we are particularly interested in the covariant moment formalism introduced by Thorne (1981), as this approach can give a quite accurate solution similar to other methods and has successfully been used in a number of calculations (Nobili *et al.* 1993; Zampieri *et al.* 1993). In this approach (Thorne 1981; see also Turolla & Nobili 1988; Rezzolla & Miller 1994), the fundamental idea is to replace Eq. (5.87) by a hierarchy of moment equations written in terms of **projected symmetric and trace-free (PSTF)** tensors which are properly defined at each point in the projected tangent space to the fluid four-velocity u^α.

If $I = I(x^\alpha, p^\alpha)$ is the specific intensity of the radiation field, the relation $\mathcal{N}_\nu = (c^2/2h)(I_\nu/\nu^3)$ holds, where $\nu = u^\alpha p_\alpha/h$ is the specific photon frequency measured in

the comoving frame. Then, the kth Projected Symmetric Trace-Free (PSTF) moments of $\mathcal{N}$ and $\mathcal{S}$ in the local rest frame of the standard fluid are given as

$$\begin{aligned}
\mathcal{M}_\nu^{\alpha_1\cdots\alpha_k} &\equiv \left(\int \frac{\mathcal{N}\delta(2\pi\nu+\vec{p}\cdot\vec{u})}{(-\vec{p}\cdot\vec{u})^{k-2}} p^{\alpha_1}\cdots p^{\alpha_k} d\vec{V}_p\right)^{PSTF} \\
&= \frac{1}{c}\left(\int I n^{\alpha_1}\cdots n^{\alpha_k} d\Omega\right)^{PSTF} \qquad (5.88) \\
\mathcal{S}_\nu^{\alpha_1\cdots\alpha_k} &\equiv \left(\int \frac{\mathcal{S}\,\delta(2\pi\nu+\vec{p}\cdot\vec{u})}{(-\vec{p}\cdot\vec{u})^{k-2}} p^{\alpha_1}\cdots p^{\alpha_k} d\vec{V}_p\right)^{PSTF} \\
&= \frac{h}{c^3}\left(\int \nu^3 \mathcal{S} n^{\alpha_1}\cdots n^{\alpha_k} d\Omega\right)^{PSTF} , \qquad (5.89)
\end{aligned}$$

where $d\vec{V}_p$ is the invariant momentum-space volume element on the light cone, $\delta(y)$ is the Dirac delta function and n^α is the unit vector which gives the direction of propagation of a photon as seen in the rest frame. Integration with respect to Ω is over solid angle in the projected space and "Trace-Free" denotes the consequence of the usual tensor operation.

In spherical symmetry, we can define a local orthonormal tetrad carried by an observer comoving with the standard fluid as $\{\vec{e}_{\hat{0}}, \vec{e}_{\hat{r}}, \vec{e}_{\hat{\theta}}, \vec{e}_{\hat{\varphi}}\}$, with $\vec{e}_{\hat{0}} = \vec{u}$, $\vec{e}_{\hat{r}}$ being in the radial direction and $\vec{e}_{\hat{\theta}}$, $\vec{e}_{\hat{\varphi}}$ being orthogonal to each other and to $\vec{e}_{\hat{r}}$. Since in this case, I and $\mathcal{S}$ are invariant under rotations of the photon direction n^α about $\vec{e}_{\hat{r}}$, it is possible to show that all of the components of each PSTF moment of rank k can be evaluated in the comoving frame as functions of the radial one

$$W_k \equiv \mathcal{M}^{\hat{r}_1\cdots\hat{r}_k} = 2\pi\frac{k!(2k+1)}{(2k+1)!!}\frac{1}{c}\int I P^k(\mu)\, d\mu \quad , \qquad (5.90)$$

$$S_k \equiv \mathcal{S}^{\hat{r}_1\cdots\hat{r}_k} = 2\pi\frac{k!(2k+1)}{(2k+1)!!}\frac{h}{c^3}\int \nu^3 \mathcal{S} P^k(\mu)\, d\mu \quad , \qquad (5.91)$$

where $P^k(\mu)$ is the Legendre polynomial of order k, μ is the $\vec{e}_{\hat{\alpha}}$ component of photon direction $n^{\hat{r}}$. In particular

$$\begin{aligned}
\mathcal{M} &= W_0 \quad , \qquad (5.92) \\
\mathcal{M}^\alpha &= W_1 e^\alpha_{\hat{r}} \quad , \qquad (5.93) \\
\mathcal{M}^{\alpha\beta} &= W_2\left(e^\alpha_{\hat{r}} e^\beta_{\hat{r}} - \frac{1}{2}e^\alpha_{\hat{\theta}} e^\beta_{\hat{\theta}} - \frac{1}{2}e^\alpha_{\hat{\varphi}} e^\beta_{\hat{\varphi}}\right) \quad . \qquad (5.94)
\end{aligned}$$

The frequency integrated radial moments of the specific intensity, w_k, and the source function, s_k, are thus given by

$$\begin{aligned}
w_k(r) &= \int_0^\infty W_k(r,\nu)\, d\nu \quad , \qquad (5.95) \\
s_k &= \int_0^\infty S_k(r,\nu)\, d\nu \quad . \qquad (5.96)
\end{aligned}$$

Moment equations can be obtained by taking the PSTF moments of the relativistic transfer equation (5.87) and this gives rise to a recursive hierarchy of differential equations, which is not closed since at any given order k_{max} it contains moments up to order k_{max+1} in the frequency-integrated case (or k_{max+2} if frequency-dependent). This means that it is necessary to make some "ad hoc" assumption to close the system and this is usually done on physical grounds by introducing suitable closure functions which relate $W_{k_{max+1}}$ or $w_{k_{max+1}}$ to moments of lower order. Since the behavior of all of the moments is known in the asymptotic limits, it is sufficient to prescribe a reasonable smooth function that connects these two limits (e.g. Rezzolla & Miller 1994).

The frequency integrated moment equations (5.95) and (5.96) give a clear physical interpretation for the first three integrated moments of the hierarchy: the zeroth moment w_0 is the energy density of the radiation, the first moment cw_1 is the radiative energy flux, and the second moment w_2 is the shear stress tensor of the radiation fluid. The scalar source moments s_0 and s_1 represent the transfer of energy and momentum between the two fluids.

The stress-energy tensor for the radiation $T_R^{\alpha\beta}$ is completely defined in terms of the first three moments and higher-order moments do not enter into this definition:

$$T_R^{\alpha\beta} = w_0 u^\alpha u^\beta + 2w_1 e_{\hat{r}}^{(\alpha} u^{\beta)} + w_2 e_{\hat{r}}^\alpha e_{\hat{r}}^\beta + \frac{1}{3} w_0 P^{\alpha\beta} \quad , \tag{5.97}$$

where $P^{\alpha\beta} \equiv g^{\alpha\beta} + u^\alpha u^\beta$ is the projection operator orthogonal to $\vec{u}$. This equation implies that if the hierarchy of moments is truncated at the second order, it is possible to derive the equations governing the hydrodynamics of the radiation by starting from the conservation laws of energy and momentum of radiation. It was shown that if we truncate the moments at the second order, the error in the calculation values of the radiation variables (lower order moments) is typically not larger than $\sim 15\%$ for the range of parameter values typical for real astrophysical fluids (Turolla & Nobili 1988). In our calculation, we will restrict ourselves to the use of the first three scalar moments and the first two scalar source moments.

5.7.2 The Equations of Radiation Hydrodynamics

Following the discussion in Sec. (5.7.1), the radiation hydrodynamic equations can be derived from the standard conservation laws for the energy and momentum of the "radiation fluid". The conservation laws give (Rezzolla & Miller 1994)

$$-u_\alpha T_R^{\alpha\beta}{}_{;\beta} = s_0 \tag{5.98}$$

$$n_j P^j{}_\alpha T_R^{\alpha\beta}{}_{;\beta} = s_1/b \quad , \tag{5.99}$$

where $\vec{n}$ is a radial space-like unit vector normal to $\vec{u}$. Making use of the stress-energy tensor (5.97) and of the metric relation (5.2), Eq. (5.98) and (5.99) can be written

explicitly as

$$(w_0)_t = ac\left[s_0 - 2\frac{\Gamma}{R}w_1 - \frac{u}{Rc}\left(\frac{8}{3}w_0 - w_2\right) - 8\pi\rho R^2 w_1 a_\mu/a - 4\pi\rho R^2 (w_1)_\mu\right] + \frac{(\rho R^2)_t}{\rho R^2}\left(\frac{4}{3}w_0 + w_2\right) \quad , \tag{5.100}$$

$$(w_1)_t = ac\left(s_1 - \frac{2u}{Rc}w_1 - 3w_2\frac{\Gamma}{R}\right) - 4\pi\rho R^2 ac\left(\frac{1}{3}w_0 + w_2\right)_\mu - 4\pi\rho R^2 ac\left(\frac{4}{3}w_0 + w_2\right)\frac{a_\mu}{a} + 2w_1\frac{(\rho R^2)_t}{\rho R^2} \quad , \tag{5.101}$$

$$w_2 = f_E w_0 \quad , \tag{5.102}$$

where we have used the definition (5.11) for R_t, (5.13) for R_μ and $b = 1/4\pi\rho R^2$ (see also below). In equations (5.100) and (5.101), w_0, w_1 and w_2 have the dimensions of energy density and cs_0 an cs_1 are in units of erg cm^{-3} s^{-1}. Eq. (5.102) is the closure equation, in which f_E is a variable Eddington factor and indicates the degree of anisotropy of the radiation. It can take values ranging from 0 for complete isotropy (which could, for instance, be caused by the medium being extremely optically thick) to 2/3 for complete anisotropy. A key point is that an expression for f_E has to be supplied, constructed on the basis of physical considerations and how this is done is "ad hoc". We will be discussing it when we use it in the numerical computations.

For the combined fluids (the "standard" fluid and the "radiation fluid"), the stress-energy tensor in Einstein Eq. (5.19) is now ${T_\alpha}^\beta = {T_{R\alpha}}^\beta + {T_{M\alpha}}^\beta$ with ${T_{R\alpha}}^\beta$ given by Eq. (5.97) and ${T_{M\alpha}}^\beta$ by Eq. (5.17), if we treat the standard fluid as perfect apart from the radiative transfer. Following a similar procedure as in Sec. (5.3), we can explicitly write the hydrodynamic equations with the metric Eq. (5.2), arising from the Einstein equation (5.19), and from the conservation equations for energy and momentum of the combined fluids and the continuity equation for the standard fluid

$$u_\alpha {T^{\alpha\beta}}_{;\beta} = 0 \quad , \tag{5.103}$$

$$n_j {P^j}_\alpha {T^{\alpha\beta}}_{;\beta} = 0 \quad , \tag{5.104}$$

$$(\rho u^\alpha)_{;\alpha} = 0 \quad . \tag{5.105}$$

The equations can be expressed as (Rezzolla & Miller 1994)

$$\epsilon_t = \frac{p}{\rho^2}\rho_t - acs_0 \tag{5.106}$$

$$u_t = -a\left[4\pi R^2\frac{\Gamma}{\omega}p_\mu + \frac{\Gamma}{\rho\omega}s_1 + \frac{Gm}{R^2} + \frac{4\pi GR}{c^2}(p + \frac{1}{3}w_0 + w_2)\right] \tag{5.107}$$

$$\frac{(\rho R^2)_t}{\rho R^2} = -a\frac{u_\mu}{R_\mu} + \frac{aG}{\rho Rc^3}\frac{w_1}{R_\mu} \tag{5.108}$$

$$\frac{(a\omega)_\mu}{a\omega} = \frac{\epsilon_\mu - (p/\rho)(\rho_\mu/\rho)}{\omega c^2} - \frac{s_1}{4\pi c^2\omega\rho^2 R^2} \tag{5.109}$$

$$m_\mu = \frac{4\pi R^2 R_\mu}{c^2}\left[\rho(c^2+\epsilon) + w_0 + \frac{u}{\Gamma c}w_1\right] d\mu \tag{5.110}$$

$$R_t = au \tag{5.111}$$

$$\Gamma = 4\pi\rho R^2 R_\mu = (1 + u^2/c^2 - 2Gm/Rc^2)^{1/2} \tag{5.112}$$

$$\omega = 1 + [\epsilon + (p/\rho)]/c^2 \tag{5.113}$$

$$b = 1/4\pi\rho R^2 \quad . \tag{5.114}$$

All of the quantities have the same meaning as defined before and the coordinate continue to be co-moving with respect to the standard fluid. The effective gravitational mass m (black hole + gas + radiation) internal to any shell with label μ can also be calculated using the alternative equation

$$m_t = -\frac{4\pi R^2 R_t}{c^2}\left(p + \frac{1}{3}w_0 + \frac{\Gamma c}{u}w_1 + w_2\right) \quad . \tag{5.115}$$

The equations (5.100) – (5.101) and (5.106) – (5.114) need to be supplemented by the equation of state

$$p = p(\epsilon, \rho) \quad , \tag{5.116}$$

and by the equations for the source scalar moments s_0 and s_1, which we will be discussing together with our further numerical computations.

5.7.3 Junction Conditions for Discontinuity Surfaces

The junction conditions for metric and for the energy and momentum of the standard fluids are the same relations as given by Eqs. (5.60) – (5.61), (5.68), and (5.76) – (5.78) and we repeat them here

$$[R]_\pm = 0 \tag{5.117}$$

$$\left[a^2c^2 - \left(\frac{\dot{\mu}_s}{4\pi\rho R^2}\right)^2\right]_\pm = 0 \tag{5.118}$$

$$[m]_\pm = 0 \tag{5.119}$$

$$[a\omega]_\pm = 0 \tag{5.120}$$

$$\left[\frac{\Gamma\dot{\mu}_s}{4\pi\rho R^2}\left\{\rho\left(c^2+\epsilon\right) + w_0 + \frac{uw_1}{\Gamma c}\right\} - au\left\{p + \left(\frac{1}{3} + f_E\right)w_0 + \frac{\Gamma c}{u}w_1\right\}\right]_\pm = 0 \tag{5.121}$$

$$\left[\frac{\dot{\mu}_s}{4\pi R^2}\frac{\Gamma}{\rho} + au\right]_\pm = 0 \quad . \tag{5.122}$$

It is reasonable to neglect any interaction between the radiation and the discontinuity region so that the energy and momentum junction conditions for the radiation fluid

simply reduce to the corresponding continuity equations

$$\left[a\left(\frac{\dot{\mu}_s}{4\pi\rho R^2}\right)\left(\frac{4}{3}+f_E\right)w_0-\left\{a^2+\left(\frac{\dot{\mu}_s}{4\pi\rho R^2}\right)^2\right\}w_1\right]_{\pm} = 0 \qquad (5.123)$$

$$\left[\left\{a^2\left(\frac{1}{3}+f_E\right)+\left(\frac{\dot{\mu}_s}{4\pi\rho R^2}\right)^2\right\}w_0-2a\left(\frac{\dot{\mu}_s}{4\pi\rho R^2}\right)w_1\right]_{\pm} = 0 \ . \qquad (5.124)$$

At a contact discontinuity, $\dot{\mu}_s \equiv 0$. From Eqs. (5.117) to (5.124), we have the junction conditions

$$[R]_{\pm} = 0 \qquad (5.125)$$

$$[a]_{\pm} = 0 \qquad (5.126)$$

$$[m]_{\pm} = 0 \qquad (5.127)$$

$$[\Gamma]_{\pm} = 0 \qquad (5.128)$$

$$[p]_{\pm} = 0 \qquad (5.129)$$

$$[u]_{\pm} = 0 \ . \qquad (5.130)$$

$$[w_1]_{\pm} = 0 \qquad (5.131)$$

$$\left[\left(\frac{1}{3}+f_E\right)w_0\right]_{\pm} = 0 \ . \qquad (5.132)$$

As in the case of a pure standard fluid, we cannot use Eq. (5.120) to argue that $[\omega]_{\pm} = 0$ for the combined fluid at a contact discontinuity.

6 Numerical Computation of Dynamical Fireball Expansion

6.1 Introduction

In the last chapter, we introduced the equations for general relativity hydrodynamics and radiation hydrodynamics. As it has been mentioned, the radiation hydrodynamics will be studied in our further work. In this chapter, the equations without radiation transfer will be used for a numerical study of dynamical fireball expansion. For convenience, we repeat the hydrodynamic equations in sec. (6.2) and discuss the variants of some equations. One of the important features of a non-linear hyperbolic system is the possible appearance of shocks during the evolution of the system. To properly treat shocks whenever they appear inside the domain of interest, we introduce the numerical artificial viscosity technique in Sec. (6.3).

Our numerical difference scheme will be introduced by studying the free expansion of a fireball in Sec. (6.4). As the system is radiation dominated with a very high initial ratio of the total energy to rest mass energy ($\eta \equiv E_0/M_0c^2 \geq 100$), numerically it is not easy to handle such a system. We will study this system very carefully. A more realistic system of a fireball in an external medium is studied in Sec. 6.5. The fireball interacts with the external medium via an ultra-relativistic external shock, which needs special treatment (Sec. 6.5.2). The fireball and the shock wave are connected via a contact discontinuity, where the quantities on both sides have to satisfy junction conditions. We trace the contact discontinuity with forward, backward and advective characteristic equations together with the junction conditions in Sec. (6.5.3). Between the contact discontinuity and the shock front is the shock wave, which is growing very rapidly over many orders of magnitude in the change of radius and in the rest mass ever since the system is set. The required resolution for this region rapidly changes with evolutionary time, with an extremely high resolution requirement at the initial stage and a relatively low requirement at a late stage. The special character requires a re-gridding technique (Sec. 6.5.4). We will discuss our results in Sec. (6.5.5).

To resolve relativistic shock and contact discontinuity, high resolution scheme is

required. To make the finite difference–viscosity method competitive to PPM, adaptive mesh refinement (AMR) technique is essential. In Sec. (6.6), we present a one-dimension AMR which we are trying to combine with our specific system. However, as the technique is complicated and we have not yet completely worked it out, we just give a brief introduction to it.

6.2 Hydrodynamic Equations

For convenience, we list here the basic hydrodynamic equations derived in chapter (5):

$$u_t = -a\left[4\pi R^2\frac{\Gamma}{\omega}p_\mu + \frac{Gm}{R^2} + \frac{4\pi G}{c^2}pR\right] \tag{6.1}$$

$$R_t = au \tag{6.2}$$

$$\frac{(\rho R^2)_t}{\rho R^2} = -a\frac{u_\mu}{R_\mu} \tag{6.3}$$

$$[\ln\epsilon]_t = \frac{p}{\rho\epsilon}[\ln\rho]_t \tag{6.4}$$

$$p = p(\epsilon,\rho) \tag{6.5}$$

$$\frac{(a\omega)_\mu}{a\omega} = \left\{[\ln\epsilon]_\mu - \frac{p}{\rho\epsilon}[\ln\rho]_\mu\right\}\frac{\epsilon}{\omega c^2} \tag{6.6}$$

$$m(\mu,t) = 4\pi\int_0^\mu \rho(1+\epsilon/c^2)R^2R_\mu\,d\mu \tag{6.7}$$

$$\Gamma = \left(1+\frac{u^2}{c^2}-2\frac{Gm}{Rc^2}\right)^{1/2} \tag{6.8}$$

$$\omega = 1+[\epsilon+(p/\rho)]/c^2 \quad . \tag{6.9}$$

Eqs. (6.1)-(6.9) give a complete set of equations for the unknowns: u, R, ρ, ϵ, a, m, Γ, ω and p. Our numerical computation basically follows the order of the equation in the list. The forms of Eqs. (6.4) and (6.6) have been changed to those we use in the code, as the factor $p/\rho\epsilon$ is usually more slowly varying than p/ρ^2 as in Eq. (5.31). We have an alternative way to calculate Γ

$$\Gamma = 4\pi\rho R^2R_\mu \quad . \tag{6.10}$$

In principle, the values of Γ calculated from Eqs. (6.10) and (6.8) should be the same. However, we will find that it is very useful to distinguish them. We will call the one obtained from Eq. (6.10) Γ_ρ and the other from Eq. (6.8) Γ_u. We also use Eq. (6.3) in another form. From Eq. (6.2), we have

$$(R_t)_\mu = a_\mu u + au_\mu \quad . \tag{6.11}$$

Since $(R_t)_\mu = (R_\mu)_t$, using Eq. (6.11), we can write Eq. (6.3) as

$$\frac{(\rho R^2R_\mu)_t}{\rho R^2R_\mu} = u\frac{a_\mu}{R_\mu} \quad . \tag{6.12}$$

This equation and Eq. (6.10) together give

$$\frac{\Gamma_t}{\Gamma} = u\frac{a_\mu}{R_\mu} \quad . \tag{6.13}$$

When we investigate free expansion of a fireball, we will also use this relation to calculate Γ and Eq.(6.10) for density ρ.

The gravitational mass will also be computed with an alternative

$$m_t = -\frac{4\pi}{c^2} p R^2 R_t \quad , \tag{6.14}$$

When we study a fireball expanding into ISM via external shocks, we will use the above equations in both the fireball and and the shock wave (or shocked) regions. We assume that the ISM ahead of the external shocks is at rest and not influenced by the fireball. Therefore, the dynamic equations will not be applied in ISM. As far as the equation of state, we take the gamma-law

$$p = (\gamma - 1)\rho\epsilon \quad . \tag{6.15}$$

For relativistic plasma, $\gamma = 4/3$. More generally, γ is taken between 1 and 5/3.

6.3 Artificial Viscosities

When we derive the equations, a perfect fluid is assumed. However, In a non-linear hyperbolic system, the kinetic energy associated with acoustic disturbances is channeled progressively to higher and higher wave number modes. In nature, sound waves evolve to become shock waves and the kinetic energy ultimately appears as thermodynamic internal energy.

Owing to the finite mesh size in numerical computations with a finite difference scheme, the energy can be channeled only to an upper-limited wavenumber of two zone widths. If no dissipation mechanism is included, the energy will accumulate and oscillate in this wavenumber, and may finally destroy the solution. This energy can be dissipated into internal energy by introducing a diffusion within the difference scheme or an artificial viscosity.

In our case, reverse shocks may occur in the inner region of the fireball, when the fireball expands into ISM and interacts with the ISM via an external shock. In addition, when the initial Lorentz factor distribution in the fireball is not constant and an inner shell with higher Lorentz factor catches up an outer shell with lower Lorentz factor, internal shocks happens inside the fireball.

von Neumann & Richtmyer (1950) first introduced artificial viscosities to treat shocks by adding a scalar stress to the pressure. This scalar stress, which with wavelength-dependent coefficient has negligible effect for long-wavelength modes and

becomes important only for short wavelengths, is contrived to spread shocks over several zones of the mesh and to automatically ensure that junction conditions for shocks are satisfied.

There are numerous possible expressions for the artificial viscosity which have the desired effect. In our calculation, we adopt the one devised by May & White (1967) for the use with spherically symmetric relativistic problems:

$$Q = \begin{cases} k^2 \rho \left(\Delta\mu / R^2\right)^2 \left[\left(R^2 u\right)_\mu\right]^2 / \Gamma & \text{if} \quad \rho_t > 0 \\ 0 & \text{if} \quad \rho_t < 0 \end{cases}, \tag{6.16}$$

where $\Delta\mu$ is one zone width and $k^2 = 2$ which spreads the shock over ~ 3 or 4 zones.

6.4 Free Expansion of a Relativistic Fireball

6.4.1 Initial and Boundary Conditions

Now, we consider a fireball of total energy $E_0 = 10^{52}$ ergs and a baryonic contamination of total rest mass M_0 confined initially in a sphere of radius $R_0 = 10^7$ cm. This fireball is set in a region with zero mass density and pressure. In the computation, *cgs* units are used. We will be mainly interested in the case that $\eta = 10^2$ because in the fireball models for GRBs the favorite value is $\eta \sim 300$. Such a fireball consists of $e^{\pm}$ and photons and is radiation-dominated at its initial stage. It is initially at rest and expected to expand freely under the radiation pressure. For simplicity, the initial fireball configurations are that the mass at $t = 0$ is uniformly distributed within $R \leq R_0$ and that the total radiation energy $E_0 \gg M_0 c^2$ is deposited uniformly throughout the fireball. This approximation is not very restrictive, since free expansion behaves essentially ballistic after a few expansion time scales. The fireball is optically thick to its own electrons via pair-productions and Compton-scattering and the resulting photon-pair-electron-baryon fluid is essentially an isentropic fluid which can be modeled as a gas of adiabatic index $\gamma = 4/3$, as long as the photon are coupled to the matter by radiation drag. The same applies to the case where the energy is mainly in the form of random magnetic fields. After decoupling, the adiabatic index becomes $\gamma = 5/3$.

Two coordinate systems: comoving frame and Schwarzschild frame (or laboratory frame), are introduced in the computation. Both have their coordinate zero point at the fireball center at the initial stage and the latter is fixed relative to an observer at distance. All the thermodynamic quantities (density ρ, specific internal energy ϵ, pressure p, specific enthalpy ω and temperature T) and the time t and radial coordinate μ are measured in the comoving frame, while the radial four velocity u, radius R, m and effective Lorentz factor Γ are relative to the Schwarzschild frame. Therefore, we

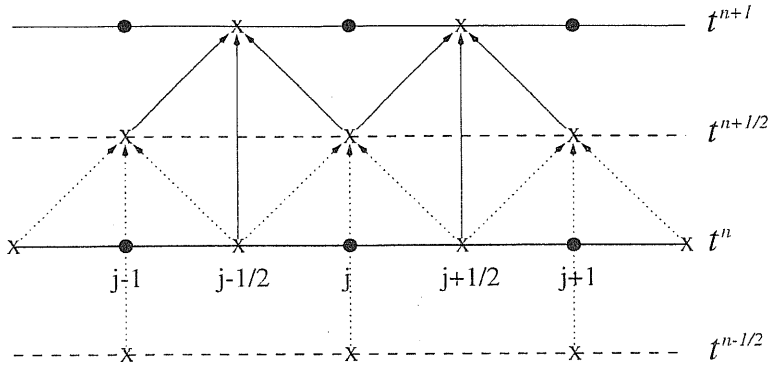

Figure 6.1: The schematic presentation of the two-step finite difference scheme. Zone boundaries are at j and $j-1$ and the region in between is called zone $j-1/2$ with a constant $\Delta\mu_{j-1/2}$. Time-step size $\Delta t^{t+1/2} = t^{n+1} - t^n$ is adjusted for each cycle, according to the computational requirements.

have the initial conditions

$$\rho_0 = \frac{3M_0}{4\pi R_0^3} = const. \quad \epsilon_0 = \eta c^2 = const. \quad u_0 = 0 \quad . \tag{6.17}$$

The distribution of pressure p and ω are set with Eqs. (6.15) and (6.9), respectively. Gravitational mass m is integrated from Eq. (6.7). We take $\Gamma_0 = \Gamma_\rho = \Gamma_u$ and calculate the value from the equation (6.8). $\Gamma_0 \simeq 1$ as the gravitational effect is negligible small. As we have known the distribution functions of Γ and ρ, the radial function $R(\mu, 0)$ can be set from Eq. (6.10). Before we can uniquely fix the time-scaling factor a with the constraint function Eq. (6.6), we have first to set the boundary conditions. In numerical computation, the consistent values of Γ, $R(\mu, 0)$ and m are set iteratively.

As discussed at the end of the section (5.3), the inner boundary conditions at $\mu = 0$ are $u|_{\mu=0} = 0$, $R|_{\mu=0} = 0$, $m|_{\mu=0} = 0$ and $\Gamma|_{\mu=0} = 1$, which means that we have assumed that the system is not singular at the center ($m/R \to 0$ when $R \to 0$). At outer boundary $\mu = M_0$, initially $R = R_0$ and $u = 0$. During the evolution of the expansion of the fireball, we set outer boundary conditions $\rho|_{\mu=M_0} = 0$ and $\epsilon|_{\mu=M_0} = 0$, and then we have $\omega|_{\mu=M_0} = 1$ and $p_{\mu=M_0} = 0$. a is set equal to unity at outer boundary to synchronize the coordinate time t with clock time as measured by an observer at infinity.

6.4.2 The Standard Two-Step Scheme

In the computational code, we use the standard explicit two-step finite-difference scheme (May & White 1967; Potter 1973; Kröner 1997) to convert the set of the hydrodynamic differential equations (6.1)–(6.9) into difference equations on a finite mesh grid of 400 zones. Each zone is a spherical shell of matter with the shell boundaries being identified with a time-independent value of μ and indexed with a subscript j (Fig. (6.1)). The calculation proceeds in coordinate time by time steps $\Delta t^{n+1/2}$ from t^n

to $t^{n+1} = t^n + \Delta t^{n+1/2}$. The truncation error in space direction is minimized by using constant $\Delta\mu$ which is chosen such that the fireball surface is at a zone boundary. If $J-1$ is the total number of the zone regions, then $\Delta\mu = M_0/(J-1)$. The index j runs from 1 to J.

Among the unknowns,

- (i) the velocity u and Γ_u is defined at the zone boundaries and the half time step with $u_j^{n+1/2}$ and $\Gamma_j^{n+1/2}$, respectively;
- (ii) radius R and gravitational mass m are defined at zone boundaries and the full time step and with R_j^n and m_j^n, respectively;
- (iii) Γ_ρ, ρ, ϵ, p, ω, and a are defined at mid-zones and the full time-step with $\Gamma_{j-1/2}^n$, $\rho_{j-1/2}^n$, $\epsilon_{j-1/2}^n$, $p_{j-1/2}^n$, $\omega_{j-1/2}^n$ and $a_{j-1/2}^n$, respectively;
- (iv) the artificial viscosity Q is defined at mid-zones and the half time-step with $Q_{j-1/2}^{n+1/2}$.

Necessary temporary auxiliary values of the various quantities are obtained by interpolation, if possible. Although it is necessary to limit the necessity of auxiliary steps, we should also try to minimize the error introducing by interpolation. For instance, when $Q = 0$ the equation of state (6.15) gives $p/\rho\epsilon = \gamma - 1$. If we interpolate p, ρ and ϵ separately, then usually $p/\rho\epsilon \neq \gamma - 1$ and the difference is sometimes very large. In the computation, we linearly interpolate mass density ρ and specific internal energy ϵ and compute p with the equation of state. To correctly center the equations at time direction, extrapolations are needed for some quantities. However, any kind of extrapolation is doubtful and would be dangerous. So, we proceed one time-step by extrapolating the necessary temporary auxiliary values and, after we get the solution at the time instant, we replace the extrapolated values with interpolated ones and repeat the integration cycle once more.

As an example, we given the difference equation of the differential equation (6.1) for velocity u_t

$$u_j^{n+1/2} = u_j^{n-1/2} - \Delta t^n a_j^n \left\{ 4\pi (R_j^n)^2 \frac{\Gamma_j^n}{\omega_j^n} \frac{p_{j+1/2}^n + Q_{j+1/2}^n - p_{j-1/2}^n - Q_{j-1/2}^n}{\Delta\mu_j} + \frac{G m_j^n}{(R_j^n)^2} + \frac{4\pi G}{c^2}(p+Q)_j^n R_j^n \right\} , \tag{6.18}$$

where the subscript j with $j = 1, J$ indicates the zone boundary with $R_1 = 0$ and $R_J = R_0$ at initial time and the superscript n is the time index. For the difference equations for other differential equations, May & White (1967) are referred. For the outer boundary half zone, we cannot center it. One side or upstream scheme is use to calculate the velocity at the surface. One side scheme is stable if the time-step has been

confined by Courant condition (LeVeque 1992), as the coefficient of p_μ in Eq. (6.1) is always negative.

The truncation error in time direction would also be minimize by using constant time-step $\Delta t^{n+1/2}$. However, the time step has to be adjusted by the relativistic generalization of the Courant-Friedrich-Lewy (CFL) stability condition

$$\Delta t \leq \lambda_c \Delta\mu / 4\pi\rho R^2 a \Delta t \quad , \tag{6.19}$$

with $\lambda_c \leq 1$, which means that the sound velocity must be smaller than the mesh velocity. This restriction prevents sound signals from propagating through more than one zone in a time step. In our computation, we take $\lambda_c = 0.01$. The time step is also controlled by limiting the change rates of ρ, ϵ, R and Γ per cycle to some extent (1% or 0.1% in our computation). In order to prevent the time step from varying too large from one cycle to another, we require that

$$\Delta t^{n+1/2} \leq (1 + \lambda_t)\Delta t^{n-1/2} \quad , \tag{6.20}$$

with $\lambda_t \leq 1$. $\lambda_t = 0.01$ is used in the computation. This requirement can restrict the truncation error to very small due to the non-constancy of the time step. This is very important for the initial cycles, as in order to smooth the initial outer boundary $\Delta t^{1/2}_{init}$ is set with Courant condition by taking $\lambda_c = 10^{-8}$ and therefore very small.

The results about Γ_u, Γ_ρ and R with the standard scheme for the case $\eta = 10^2$ are shown in Fig. (6.2). Computed points (diamond or triangle) are connected with lines to show the relative positions of neighbor points in the plots. We will give more discussion about the results in Sections (6.4.3) and (6.4.4).

6.4.3 Analysis of the Discontinuity and Treatment of Boundary

The initial configuration of the fireball described in Sec. (6.4.1) gives a initial discontinuity at the outer boundary: a finite jump of mass density, internal energy, and pressure from the fireball to vacuum (we call a region vacuum if its mass density and pressure are zero). It is an analogue of a piston withdrawn from a quiet gas accelerating from initial rest state to a constant terminal velocity greater than the escape speed of the gas in an infinitely small time interval (i.e. instantaneously). It is a special Riemann problem. A centered rarefaction wave will formed at the outer boundary and the discontinuity will be immediately smoothed out in the subsequent motion. The rarefaction wave thins the gas just behind of the surface down to density zero; pressure and sound speed are likewise decreased to zero there. The rarefaction wave is a complete rarefaction wave and ends in a vacuum.

From the junction conditions Eqs. (5.80) – (5.84) for contact discontinuity, we obtain the conditions inside the fireball but just behind the surface

$$R_- = R_0 \ , \qquad a_- = 1 \ , \qquad p_- = 0 \ , \qquad \omega_- = 1 \quad . \tag{6.21}$$

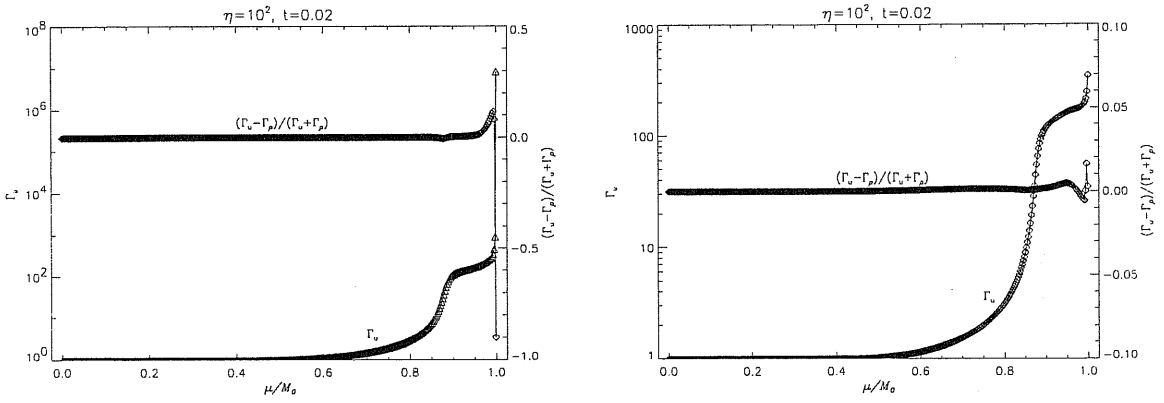

Figure 6.2: Lorentz factor Γ and the relative error $(\Gamma_u - \Gamma_\rho)/(\Gamma_\rho + \Gamma_\rho)$ *vs* normalized Lagrangian mass coordinate μ/M_0 at comoving frame coordinate time $t = 0.02$ sec. (*left*) the results obtained with standard two-step scheme for the inner region and one-side upstream method at the outer boundary half zone. (*right*) the results after the modification (see text). At outer boundary, $a = 1, p = 0, \omega = 1$, $Q = 0$.

The centered rarefaction wave can be simulated by integrating the advective and forward characteristics with setting dt to 0 (Eqs. (5.52), (5.53), (6.8), (6.9), (6.15) and $c_s = c_s(\rho, \epsilon)$) with respect to p from the initial value inside the fireball down to zero. The result gives a velocity at the outer boundary $\Gamma_R = 2.66 \times 10^4$ immediately after the system starts expanding, which is consistent with that computed from Eq. (5.86). This value should be the upper limit for a numerical computation with difference scheme, as numerical results give the average Lorentz factor over one zone. For $\eta = 10^4$, the maximum Γ_R is 7.698×10^8, which is different from the one given by Mészáros *et al.* (1993) and Kobayashi *et al.* (1999) as their numerical value is the average of one zone and resolution-dependent.

As the centered rarefaction wave region has zero width (cf Fig. (5.1) and Sec. (5.6) for the discussion on Riemann problem) at the initial time, it is impossible to resolve it out with any finite grid. A very high resolution mesh can only give an approximate solution with very slow convergence even if a second order method is used. To get faster convergence and higher resolution, a large number of mesh points are needed. The needed number is so high, it is technically forbidden for method of constant mesh width with present computation resources. In this case, Adaptive Mesh Refinement technique is necessary (cf Sec. (6.6)). A first order of numerical method can gives a very smeared solution and a second order method give a solution with better resolution but with some oscillations at the position where initial discontinuities were. The Lorentz factor at the boundary in Fig. (6.2) (*left*) is much larger than the value Γ_R and that of its neighbor. This is because the jump boundary conditions behave as a wall and waves with any wavenumber cannot pass through it and are reflected. The wave help to

accelerate the surface layer. The relative difference between Γ_u and Γ_ρ is unacceptably large.

Numerically, as fluid passes through the rarefaction wave with width much less than the boundary mesh zone size, the internal energy, mass density, pressure, specific enthalpy ω and time scaling factor a vary several magnitudes and therefore the values of quantities at a specific point (e.g. zone boundary or mid-zone) are no longer the typical values of the zone, unless the rarefaction has completely passed through it. However, when we use difference equations to study the differential system, we are studying the average properties of a small zone and assuming that the quantities in difference equation should be representative. We will try to use the averages instead of values only at some points. When the zone width $\Delta\mu \to 0$, the method should approach the differential equations.

At boundary, $p_\mu \to \infty$ in (6.18). However, $p_\mu dt$ should be finite when $\Delta t \to 0$ and the Lorentz factor at the surface should approach the finite value Γ_R. From the coefficient of p_μ in the Eq. (6.1), we have a function

$$F(\mu, t) \equiv 4\pi a R^2/\omega \quad . \tag{6.22}$$

If J is the index of the outer boundary of the fireball, we calculate F at J and at the zone center $J-1/2$ to get to F_J and $F_{J-1/2}$. In the numerical computation, we use their average $\langle F\rangle = (F_J F_{J-1/2})^{1/2}$. The Lorentz factor Γ at the boundary is computed with equation (6.8). To do so, the difficulty is that $R_{J-1/2}$ is not defined. We interpolate it from R_J and R_{J-1} by

$$R_{J-1/2} = (R_J R_{J-1})^{1/2} \quad . \tag{6.23}$$

The difference equation for Lorentz factor Γ is basing on Eq. (6.13) and the mass density is calculated with Eq. (6.10). The results is the same if we use Eq. (6.13) to compute the mass density and Eq. (6.10) for the Lorentz factor. But the choice here gives slightly smoother solutions The reason probably is that both in Eq. (6.13) and Eq. (6.3) at outer boundary zone $u\frac{a_\mu}{R_\mu}$ and $a\frac{u_\mu}{R_\mu}$ are singular when $d\mu \to 0$. However, u equals to zero but a has a finite value at initial time and the change of Γ is thus smaller than that of ρR^2. a used in Eqs. (6.2) and (6.13) should be an average over a finite region and smooth. We use a interpolated value

$$a^* = 1/\omega^* \tag{6.24}$$

with

$$\omega^* = \omega_{J-1/2}(1-\alpha) + \alpha\omega_- \quad , \tag{6.25}$$

where

$$\alpha = \begin{cases} t/\Delta t_c & \text{if } t < \Delta t_c, \\ 1 & \text{if } t > \Delta t_c. \end{cases} \tag{6.26}$$

Δt_c is the measure of smoothness, calculated as a fraction λ of the time during which a sound wave travel from the boundary to the center of the last zone and

$$\Delta t_c = \lambda \frac{(\Delta\mu/2)}{4\pi\rho R^2 a \Delta t} \quad , \tag{6.27}$$

where $\Delta\mu$ is the size of last zone. The solution in the last few zones is sensitive to the choice of λ. We use $\lambda = 3/4$ to make Γ_ρ and Γ_u consistent with each other. As Γ is essentially energy per unit mass, in principle, Γ_ρ and Γ_u should have no difference within calculation error.

The constraint equation (6.6) is applied to the half zone from the boundary to the center of the last zone. We interpolate ρ and ϵ and compute p with the equation of state. Thus, this equation has a regular behavior at the boundary.

After these special treatment of the boundary zone, we obtained a regular solution which is also shown in Fig. (6.2). The relative difference between Γ_ρ and Γ_u at the boundary is now less than 1%. When we compare Γ_u with Γ_ρ, we interpolate Γ_u to zone-center. The relative errors shown in Fig. (6.2) have included those introduced by the interpolation.

6.4.4 Computational Accuracy and Method beyond the Standard Scheme

When the rarefaction wave propagates toward the center of the fireball, it slowly spreads and finally at somewhere inside the fireball can be fairly resolved by a suitable number meshes. Fig. (6.2) (*left*) shows that the discrepancy between Γ_u and Γ_ρ is larger at the region near boundary but decreases quickly towards the inner region. When the rarefaction wave cannot resolved completely, we have similar difficulties in the inner region as at boundary. Difference equations in conservative form or high resolution may help to solve the problem. However, a uniform mesh with a currently reasonable number of grid points is not suitable for all of the domain interest, while, in general, a non-uniform suffers serious truncation error problems. An uniform adaptive mesh refinement technique may be the option, which we will discuss more in Sec. (6.6).

Here, as we are mainly interested in the hydrodynamics of the whole region, we may modify the scheme beyond the standard technique to obtain satisfying results. The modification is not straightforward. Experimentally, we find that the main difficulties are associated with the coefficient of p_μ in the equation of motion (6.18) or (6.1), which varies by several orders of magnitudes during the computation. The idea therefore is to make the coefficient of the differential quantities to be represented. For the equation of motion (6.18) or (6.1), we use the average $\bar{R}$ and $\bar{\Gamma}$. $\bar{R}_j$ is computed from R_j, R_{j-1} and R_{j+1} by

$$\bar{R}_j = \left\{ [R_j(R_{j+1} - \delta R_j)]^{1/2} [R_j(R_{j-1} + \delta R_j)]^{1/2} \right\}^{1/2} \qquad \text{for } j = 2, J-1 \quad , \tag{6.28}$$

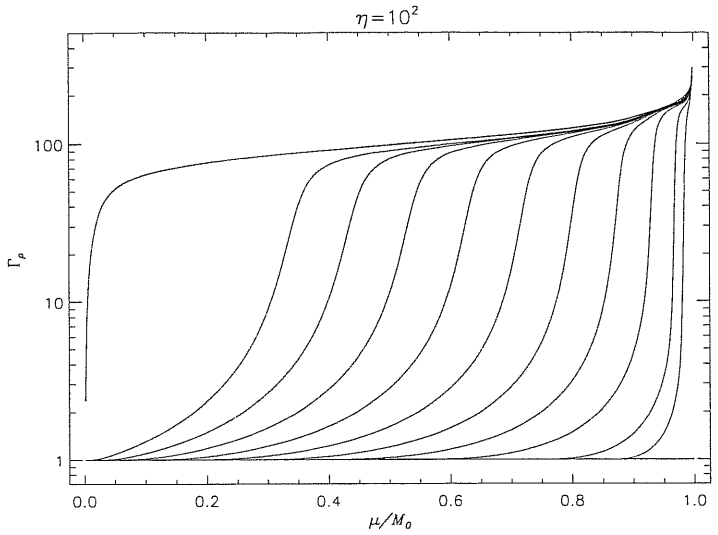

Figure 6.3: Lorentz factor Γ_ρ as a function of normalized rest mass coordinate μ/M_0. The curves are corresponding to expansion factor $(ct + R_0)/R_0$ (from right to left) 0, 0.1, 0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6, 1.8 and 10 $\times\eta$. The fireball starts expanding at $t = 0$.

where δR_j is a small quantity and, important only for the points with j close to 2 in order that $\bar{R}_2 \neq 0$ ($R_1 = 0$). We take $\delta R_j = R_2(R_2/R_j)$, where $R_2 = \Delta R_2$ is the size of the innermost zone. Then the average of R_j used in $F(\mu, t)$ is given by

$$\langle R\rangle_j = (R_j + \bar{R}_j)/2 \quad . \tag{6.29}$$

For the Lorentz factor Γ in Eq. (6.18), we use the interpolation value at j with

$$\bar{\Gamma}_j = \left(\Gamma_{j-1/2}\Gamma_{j+1/2}\right)^{1/2} \quad \text{for } j = 2, J-1 \quad , \tag{6.30}$$

where $\Gamma_{j\pm1/2}$ is calculated with Eq. (6.10).

Another modification to the numerical scheme presented by May & White (1967) is that when we need a at half time-step, we interpolate $a\omega$ and ω instead of directly interpolating a, as we do in the space direction.

6.4.5 Results and Discussion

We do the numerical computation with 400 grid zones. To compare with the standard two-step scheme, we also show in Fig. (6.2) Γ_u and its relative difference with Γ_ρ at coordinate time $t = 0.02\,\text{sec}$ for $\eta = 10^2$. The relative difference is now $< 1\%$ over all the grid point except for one with $< 2\%$. We have checked the relative errors for different time: from the initial stage to when all matter has been accelerated to its saturating velocity. Since Γ is the specific energy, we may say that the energy is conserved during our computation and the numerical calculation is accurate. As we

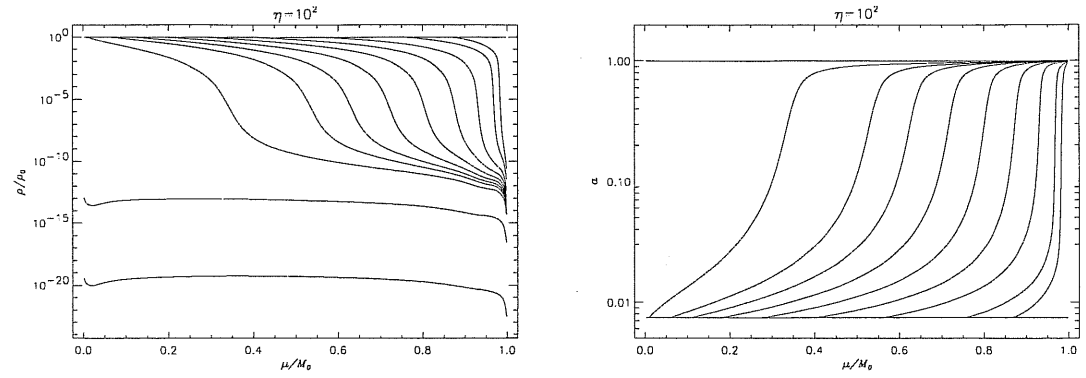

Figure 6.4: Normalized rest mass density ρ/ρ_0 (*left*) and time scaling factor (*right*) *vs* normalized rest mass coordinate μ/M_0. M_0 and ρ_0 are the initial total rest mass and rest mass density. Different curves are for different time t: 0.00, 0.0033, 0.0063, 0.013, 0.020, 0.027, 0.033, 0.040, 0.047, 0.060, 0.33 and 33 sec from top to bottom for mass density ρ and from bottom to top for time scaling factor a. For a, the last two curves overlaid each other.

chose mass as the radial coordinate in the comoving frame, the conservation of mass is automatically guaranteed.

The Lorentz factor Γ_ρ as the function of radial coordinate μ and time t is shown in Fig. (6.3). When the fireball start expanding, the width of the rarefaction wave is very small and the matter in the boundary layer pass through this wave and almost jump to the saturating value of velocity, which is higher than the random Lorentz factor $\eta = 10^2$ due to the initial discontinuity at outer boundary. When the rarefaction wave propagate toward the center of the fireball, its width spreads slowly and matter passes through this region and is accelerated to the asymptotic value of velocity about $\sim \eta$. In the innermost small region with mass about 10 % of the total matter, the Lorentz factor drops quite steeply to unity.

The maximum Lorentz factor Γ_ρ and the average Lorentz factor Γ_ρ grow initially almost linearly with radius, and saturates to a value $\Gamma_\rho \sim \eta$ after the radius expanding a factor of η (Fig. (6.5).

The propagation of this rarefaction wave and the weak discontinuity are very clear from the changes of the mass density ρ and time scaling factor a with time in Fig. (6.4). At the outer boundary, the density is thinning to very low value. When the rarefaction wave completely pass through the fireball, all the matter is accelerated to the saturated velocity except for those in a small region near outer and inner boundaries. Then, mass density is thinning to a nearly constant value over the inner $\sim$ 80% region away from the outer and the inner edges of the fireball and the mass distributes quite uniformly. However, the mass density uniformly drops in time due to the spherical geometry (cf Eq. (6.10)). This is the coasting stage. The time scaling factor a reflects the property of the co-moving frame in general relativity. It is not given by other authors, as they

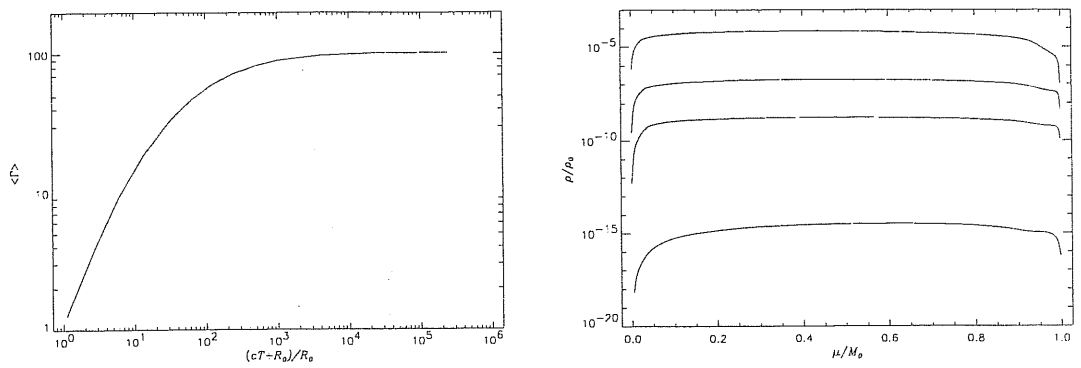

Figure 6.5: The results at instants with laboratory frame time slice T = *constant*. (*left*) Average bulk Lorentz factor Γ as a function of the expansion factor $(cT + R_0)/R_0$. (*right*) Comoving frame mass density against the Lagrangian mass coordinate, when the expansion factor $(cT + R_0)/R_0$ equals (from top to bottom) to 13.5, 126, 1001 and 240001.

used comoving frame only in special relativity. We also show it as a function of mass coordinate and time in Fig. (6.4). At each matter point in the fireball, whenever the matter is accelerated to its maximum velocity, the time scaling factor a rapidly approaches its asymptotic value which is unity. At this moment the internal energy is almost completely converted into bulk kinetic energy and the region becomes matter dominated and the enthalpy approaches unity.

In our computation, we used different initial conditions from those of Mészáros *et al.* (1993) and present our results in comoving frame. Although they also gave their results in a comoving frame, there are some differences between theirs and the results presented here. They presented the results at time instants in laboratory frame and the results here are at time instants in comoving frame. Eq. (5.15) shows that they are in general different. When the acceleration has saturated, u/c and Γ are $\sim \eta$ and nearly time-independent. From Eqs. (5.14) and (5.10), $B = \left(1 - \frac{2Gm}{Rc^2}\right)^{1/2} \simeq 1$. The time scaling factor A at Schwarzschild frame can be set to unity. From Fig. (6.4), $a \cong 1$ every where inside the fireball. From Eq. (5.15), we have $T \sim \Gamma t \sim \eta t \propto t$ except for the points close to both boundaries. An observer in the Schwarzschild frame would observe a uniformly fireball shell. As the Lorentz factor Γ_{out} at outer boundary is much larger than the Lorentz factor Γ_{in} in the center, which is very close to unity, we have $T_{out} \gg T_{in}$. So, in the Fig. 4b of Mészáros *et al.* (1993) the mass density in the center region decreases toward the center of the fireball, while the mass density presented in Fig. (6.4) increases slightly. For convenience, we give the results on the average bulk Lorentz factor and comoving mass density, obtained by Mészáros *et al.* (1993) in Fig. (6.7).

To see more clearly the consistency between our results and those obtained by

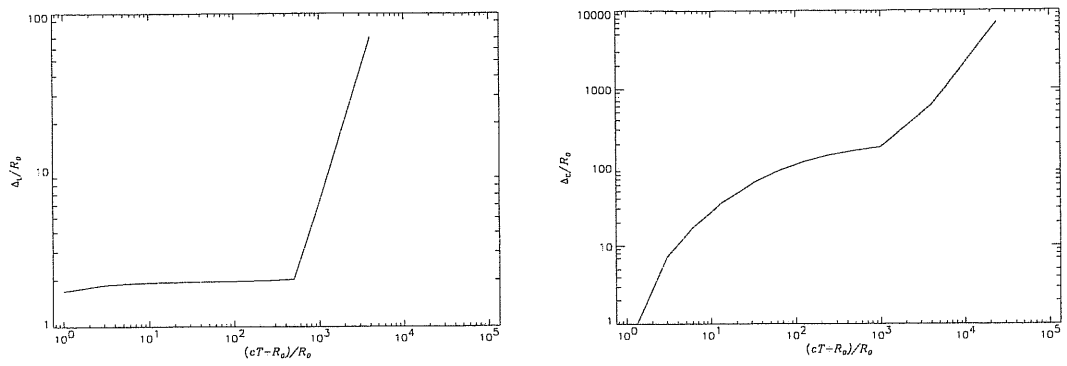

Figure 6.6: (*left*) Lab frame width of the shell defining as within it most (80%) of the mass is contained, as a function of the expansion factor $(cT + R_0)/R_0$ for $\eta = 10^2$. (*right*) Comoving frame width with the same definition, as a function of the expansion factor $(cT + R_0)/R_0$ for $\eta = 10^2$.

Mészáros *et al.* (1993), we present the average Lorentz factor Γ as a function of the expansion factor $(cT+R_0)/R_0$ in the Schwarzschild frame ($A = 1$ and $B = \left(1 - \frac{2Gm}{Rc^2}\right)^{-1/2}$) in Fig. (6.5)(*left*) and the comoving frame mass density against the Lagrangian mass coordinate in Fig. (6.5)(*right*), measured in laboratory time instants T.

Although the initial conditions are different, both results are comparable. The average Lorentz factor (defined by respecting to the rest mass) in Fig. (6.5) initially nearly linearly increases with radius until an expansion factor about η and saturates to a value $\Gamma \simeq \eta$. It is consistent with that in Fig. 3 in Mészáros *et al.* (1993) and the results which we described in Chapter (4). Fig. (6.5)(*right*) is consistent with the Fig .4*b* in Mészáros *et al.* (1993), although we use different initial conditions. The difference is significant between Fig. (6.4)(*left*) and Fig. (6.5)(*right*), although both figures depicts the variation of the comoving frame mass density as a function of Lagrangian mass coordinate. The mass is more uniform in Fig. (6.5).

We can check our results by comparing them with the analytically expected values. Since the expansion is adiabatic, the comoving density is connected to the comoving volume by

$$(\rho/\rho_0)^{1/3} = (V_0/V)^{1/3} \quad . \tag{6.31}$$

During the initial stage, $\Gamma \propto R$. The comoving volume in this phase is $V \propto R^2\Delta r \propto R^2\Gamma\Delta R$, where r is the radial in comoving frame and $\Delta r \propto \Gamma\Delta R$ and the dimensions transverse to the motion will be the same in the lab frame and comoving frame. From our results on the shell width in laboratory frame shown in Fig. (6.6)(*left*) and the shell width in comoving frame shown in Fig. (6.6)(*right*), $\Delta r \propto \Gamma R_0$.

Thus, we have $V \propto R^3$ and

$$\rho/\rho_0 \sim (R_0/R)^3 \quad . \tag{6.32}$$

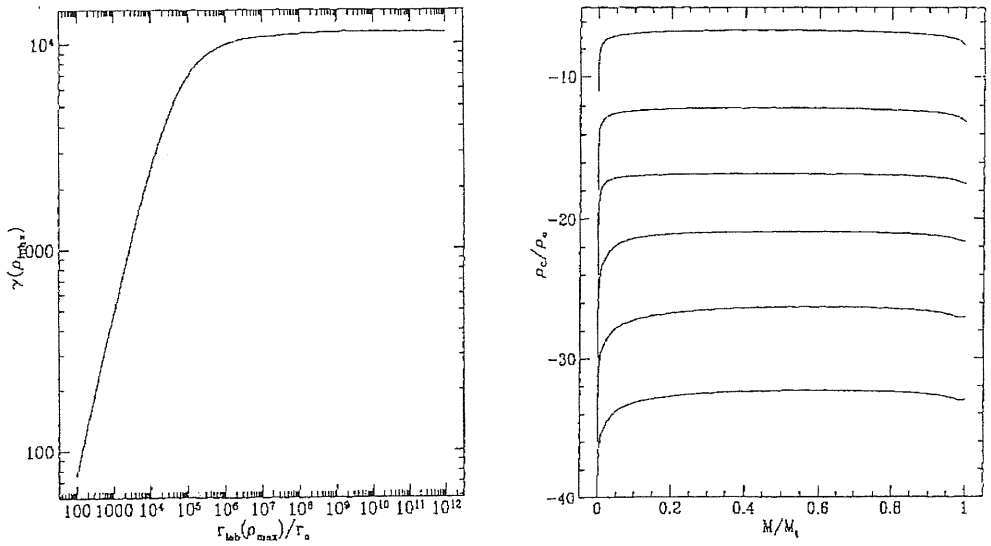

Figure 6.7: Results obtained by Mészáros *et al.* (1993). (*left*) The average bulk Lorentz factor Γ as a function of the expansion factor $R(\rho_{max})/R_0$ for $\eta = 10^4$. $R(\rho_{max})$ is the radius corresponding maximum mass density in lab (or Schwarzschild) frame and equivalent to $(cT + R_0)/R_0$. (*right*) Comoving frame mass density against the Lagrangian mass coordinate (M), at the times (from top to bottom) when the expansion factor $R(\rho_{max})/R_0$ is $\eta^{1/2}$, η, $\eta^{3/2}$, η^2, $\eta^{5/2}$, η^3, for $\eta = 10^4$. M_t is the total rest mass of fireball.

So, for an expansion factor of ~ 1000, we have $\rho/\rho_0 \sim 10^{-9}$, which is consistent with the value given by the third curve in Fig. (6.5)(*right*).

Although the curves given in Figs. (6.5) and (6.6) for Schwarzschild frame is not very smooth due to the use of a few points which are translated from the comoving frame, the difference between Fig. (6.5) and Fig. (6.7) is clear. The average bulk Lorentz factor and comoving mass density in Fig. (6.5) does not follows the simply linear scaling law so strictly as those given in Fig. (6.7) and the transition regions are relatively broader. The reason is that the initial random Lorentz factor η used by us is much smaller than that used by Mészáros *et al.* (1993) and the rest mass energy becomes important soon after the system starts expanding. The difference is also clear in Fig. (6.6) about the width of the fireball shell in comoving frame. The width first fast increases but not linearly (linearly in Mészáros *et al.*) and then pass a slowly but non-constant (nearly constant in Mészáros *et al.*) phase. After this slowly increase period, the shell nearly linearly broadens again. For $\eta = 10^4$, the width have a relatively-long linearly increasing period and following an almost constant phase. This difference is important when the theoretical results confront the observations of GRBs itself as the favorite Lorentz factor now is ~ 300 which is far below 10^4. Therefore, to understand GRBs, fully numerical computation is necessary, after simple scaling-law analysis.

6.5 Expanding into ISM

6.5.1 Initial and Boundary Conditions

We now study a fireball expanding into ISM. We are interested from the momentum when a fireball is set within the ISM. The initial conditions for fireball are the same as presented in Sec (6.4.1): the fireball is at rest $u_0 = 0$, mass density is constant through the fireball $\rho_0 = constant = \frac{3(E_0/\eta c^2)}{4\pi R_0^3}$, specific internal energy has a uniform distribution $\epsilon_0 = \eta c^2$. the distributions of R, m, Γ, p an ω can be consistently set with the equations given in Sec. (6.2) (similar to the case expanding into vacuum, see Sec. (6.4.1)). The time scaling factor a is a coordinate quantity. Although it can be scaled arbitrarily for a system, we should set it consistently for the whole system consisting of different physical properties. So, before discussing its initial setting, we should have the initial distributions of various quantities in ISM.

We consider a simple ISM which consist of atomic hydrogen and has time-independent uniform distributions of number density, pressure and temperature in the region which we are interested. The atomic hydrogen assumption is not important, as we are interested in the dynamic effect of the ISM. If it is molecular instead of atomics, we just scale its number density. Uniformly-distributing number density (thus mass density) $n_{ism} = 1\text{particle/cm}^3$ and temperature (proportional to internal energy) of $T = 100\,\text{K}$ are taken. We assume that the equation of state of ISM is Eq. (6.15) with $\gamma = 5/3$. The ISM is at rest within the time interval we are interested and therefore $u_{ism} = 0$. The quantities p, ω, R, m and Γ are set in the same ways as inside the fireball. A non-uniform ISM will be considered in our further work.

At far distance $\omega = 1$, we set $a = 1$ to synchronizes coordinate time and the clock time of an observer at far distance. The distribution of a in the ISM is computed with the constraint equation (6.6). The time coordinate t inside the fireball and therefore its scaling factor a are determined uniquely. However, we do not have junction for initial discontinuity but indeed have for both contact and shock discontinuities. We have difficulty to correctly set the initial distribution of a inside the fireball. We may set it by separately synchronizing the coordinate time in both regions to the same observer at far distance. Fortunately, with the numerical technique we used here, the splitting of the initial discontinuity is not affected by the initial setting of a and the centered rarefaction wave and the junction conditions will immediately and correctly set the distribution inside the fireball.

As the external medium has very low density, it is dynamically not important at early stage. However, we expect that the existence of ISM will affect the numerical computations in several ways. From the analysis in Sec. (5.6), the initial discontinuity will instantaneously split into three kinds of discontinuities, which shrink to one point at the initial time but connect each other by the jump conditions. As the existence of

external medium, the centered rarefaction wave no longer terminates to vacuum but instead at the contact discontinuity with finite pressure and mass density.

In the technique used here, we will take the instant when the initial discontinuity has just properly split as our initial time and follow the evolution of the system thereafter. We use the following strategy to split the initial discontinuity. We integrate the forward and advective characteristic equations (5.52), (5.53), (6.8), (6.9) and (6.15) (with $dt = 0$) just behind the contact discontinuity, setting $dt = 0$ and integrating with respect to pressure p away from the initial fireball values. As $dt = 0$, it does not matter what value of a we take. From the junction conditions Eqs. (5.80)–(5.85), the velocity and pressure should be continuous across the contact discontinuity. Thus, when the rarefaction wave accelerates the matter just behind the contact discontinuity to some velocity, the fluid just ahead of the contact discontinuity will have the same velocity and a shock of zero width will be formed there. This shock should satisfy the shock junction conditions Eqs. (5.60), (5.61), (5.68), (5.76), (5.78) and (5.79), and produces a higher pressure in the shocked fluid. The integration will be terminated when the variable pressure just behind the contact discontinuity is equal to the pressure in the shocked fluid. At this moment, the quantities satisfy all the junction conditions for the contact discontinuity and for shocks and we take this time phase as the initial stage of the system and the conditions inside the shock as the internal outer boundary conditions of the fireball regions. We will integrate the two region separately. The shock propagates into the ISM while the rarefaction wave propagates into the fireball. After the split is successfully finished, the time scaling factor a inside the fireball just behind the contact discontinuity will take the value at the shock and it will become the outer boundary value of a inside the fireball. The distribution is thus calculated with Eq. (6.6). The resulting state obtained by this technique is consistent with those obtained with the analytic equations given by Martí & Müller (1994). In the case described here, the splitting velocity corresponds to $\Gamma_R = 2.5 \times 10^4$, which is slightly less than the value for free expansion in Sec. (6.4.3).

After the initial stage, we follow the evolution of the discontinuities with the technique developed in Sec. (6.5.2) and take the shock front as the outer boundary of the whole solution domain.

For the region within the contact discontinuity and the shock wave between the continuity and the shock front, we will use the technique introduced in Sec. (6.4). The results presented in Sec. (6.4.5) show that the weak discontinuity and rarefaction wave inside the fireball can be treated with the general method.

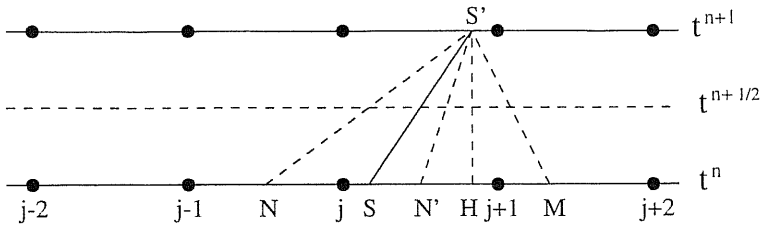

Figure 6.8: Motion of the shock front across the computational grid. SS' represents the shock front path and characteristics are shown with dashed lines. Shock front is always kept between points j and $j+1$.

6.5.2 The Treatment of Shocks

As soon as the system is set, a relativistic shock immediately splits away from the initial discontinuity. The shock front is so thin that we may consider it as a layer without thickness and apply the shock junction conditions directly on it. We know the exact state of the external medium just ahead of the shock front. To uniquely determine the states of the fluid just behind of the shock front and the front properties, the junction condition equations for shocks (Eqs. (5.60), (5.61), (5.68), (5.76), (5.78) and (5.79)) are not complete and we need one more condition.

The shock junction conditions need to be solved simultaneously with the fluid equations for the immediate neighborhood of the interface and, in doing so, it is important that causal structure is represented correctly. As it has been discussed in Sec. (5.4), the system of the hydrodynamic equations is hyperbolic and can be written in characteristic form. Physically, disturbances inside the fluid are propagating with sound velocity along the characteristics during the motion of the fluid. When shocks propagate into the external medium, the shock front moves supersonically relatively to the fluid just ahead of it and subsonically relatively to the fluid just behind it. Fig. (6.8) shows the characteristic curves which can be drawn from time t^{n+1} back to the base time level t^n. Only the forward characteristic curve can be drawn back from the shock front position at time t^{n+1} to t^n just behind the shock front, while all the backward, forward and advective characteristic curves can be drawn just ahead of the shock. The state of the fluid ahead of shock front is completely determined by the initial data and the boundary conditions in the external medium and thus is not affected by the presence of the shock.

The two states just behind and just ahead of the shock front are connected by the shock junction conditions. In order to make numerical computation efficient, we explicitly write the junction conditions in different forms. From the outer boundary condition for time scaling factor a, we have

$$a_+\omega_+ \equiv 1 \quad . \tag{6.33}$$

Eq. (6.33) and the junction condition equation (5.79) give

$$a_- = 1/\omega_- \quad , \tag{6.34}$$

where the enthalpy just behind and ahead of the shock front are given by

$$\omega_\pm = 1 + (\epsilon_\pm + p_\pm/\rho_\pm)/c^2 = 1 + \frac{\epsilon_\pm}{c^2}\left(1 + \frac{p_\pm}{\epsilon_\pm \rho_\pm}\right) \quad . \tag{6.35}$$

For ISM, $\omega_+ \simeq 1$.

From Eqs. (5.61) and (5.60), we obtain

$$f_s = \left(\frac{a_+^2 - a_-^2}{v_+^2 - v_-^2}\right)^{1/2} \quad , \tag{6.36}$$

where f_s is defined by

$$f_s \equiv \frac{\dot{\mu}_s}{4\pi c (R_s)^2} \tag{6.37}$$

where $R_s = R_- = R_+$ is the radial position of shock front in the Schwarzschild frame, and $v_\pm$ given by

$$v_\pm \equiv 1/\rho_\pm \quad , \tag{6.38}$$

are the specific volume per rest mass in the comoving frame. For ultra-relativistic shocks, $f_s \simeq \rho_+$ as $\rho_- \gg \rho_+$.

Equation (5.76) together with Eq. (5.60) can be written explicitly as

$$f_s \left(\frac{u_-}{\Gamma_- c}\right)\left(1 + \frac{\epsilon_-}{c^2}\right)\left[1 + \frac{\epsilon_-}{c^2}\left(1 + \frac{p_-}{\rho_- \epsilon_-}\right)\right] - \left(\frac{p_-}{\rho_- \epsilon_-}\right)\rho_- \frac{\epsilon_-}{c^2} = -a_+ \Gamma_+ \left(\frac{p_+}{c^2}\right)\left(\frac{\omega_-}{\Gamma_-}\right) \quad , \tag{6.39}$$

where we have used the condition for ISM that $u_+ = 0$ and the Lorentz factor $\Gamma_\pm$ just ahead of and behind the shock front are computed with

$$\Gamma_\pm = \left(1 + \frac{u_\pm^2}{c^2} - 2\frac{G m_s}{R_s c^2}\right)^{1/2} \quad , \tag{6.40}$$

where m_s is the effective gravitational mass contained inside the sphere denoted by μ_s.

After some algebraic operations, we obtain the equation for ϵ_-

$$\left[\left(\frac{f_s}{\rho_-}\right)\left(\frac{u_-}{\Gamma_- c}\right)\left(\frac{\rho_- \epsilon_-}{p_-} + 1\right)\right]\left(\frac{\epsilon_-}{c^2}\right)^2 + \left[\left(\frac{f_s}{\rho_-}\right)\left(\frac{u_-}{\Gamma_- c}\right)\left(2\frac{\rho_- \epsilon_-}{p_-} + 1\right) - 1\right]\left(\frac{\epsilon_-}{c^2}\right) + \left(\frac{f_s}{\rho_-}\right)\left(\frac{u_-}{\Gamma_- c}\right)\left(\frac{\rho_- \epsilon_-}{p_-}\right) + a_+ \Gamma_+ \left(\frac{p_+}{\rho_- c^2}\right)\left(\frac{\omega_-}{\Gamma_-}\right)\left(\frac{\rho_- \epsilon_-}{p_-}\right) = 0 \quad . \tag{6.41}$$

The solution of Eq. (6.41) is

$$\left(\frac{\epsilon_-}{c^2}\right) = \left(\frac{\Gamma_- c}{u_-}\right)\left(\frac{\rho_-}{f_s}\right)\frac{f_\epsilon}{1+C_-} \quad , \tag{6.42}$$

where the condition $\epsilon_- \geq 0$ has been used to select the solution and f_ϵ is defined by

$$f_\epsilon \equiv \frac{1}{2}\left\{1 - A_- + \sqrt{1 - 2A_- + B_-^2 - 4B_-(C_-+1)C_-\left(\frac{\omega_-}{\Gamma_-}\right)a_+\Gamma_+\frac{p_+}{\rho_- c^2}}\right\} , \tag{6.43}$$

and

$$C_- \equiv \frac{\rho_-\epsilon_-}{p_-} \quad , \tag{6.44}$$

$$B_- \equiv \left(\frac{f_s}{\rho_-}\right)\left(\frac{u_-}{\Gamma_- c}\right) \quad , \tag{6.45}$$

$$A_- \equiv B_-(2C_-+1) \quad . \tag{6.46}$$

For ultra-relativistic shocks and the equation of state (6.15), $f_\epsilon \simeq 1$ (as $p_+ \ll \rho_+ c^2$ and $\rho_- \gg f_s \simeq \rho_+$) and therefore $\epsilon_- \propto \rho_-$.

From Eq. (5.78) together with Eqs. (6.42) and (6.44), we can obtain the equation for ρ_-

$$\left(\frac{\Gamma_+ v_+}{\Gamma_-}\right)\left(\frac{\Gamma_- c}{u_-}\right)\frac{f_\epsilon}{C_-}\rho_-^2 - \left[\frac{u_-}{\Gamma_- c} + \left(\frac{\Gamma_- c}{u_-}\right)\frac{f_\epsilon}{C_-} - \left(\frac{\Gamma_+ v_+}{\Gamma_-}\right)f_s\right]\rho_- - f_s = 0. \tag{6.47}$$

The solution of equation (6.47) is

$$\rho_- = \left(\frac{\Gamma_-}{\Gamma_+}\right)\left(\frac{f_\rho}{f_\epsilon}\right)\rho_+ \quad , \tag{6.48}$$

where

$$f_\rho \equiv \frac{1}{2}\left\{\left[C_-\left(\frac{u_-}{\Gamma_- c}\right)^2 + f_\epsilon\right]\left[1+\sqrt{1-f_{\rho\rho}}\right] - \left(\frac{\Gamma_+}{\Gamma_-}\right)\left(\frac{f_s}{\rho_+}\right)\left(\frac{u_-}{\Gamma_- c}\right)C_-\left[1-\sqrt{1-f_{\rho\rho}}\right]\right\} \tag{6.49}$$

$$f_{\rho\rho} \equiv \frac{4(u_-/c\Gamma_-)^3(\Gamma_+ v_+/\Gamma_-)f_s}{[(u_-/c\Gamma_-)^2 + (f_\epsilon/C_-) + (\Gamma_+ v_+/\Gamma_-)f_s(u_-/c\Gamma_-)]^2} \quad . \tag{6.50}$$

For ultra-relativistic shocks and the equation of state (6.15), $f_\rho \simeq \frac{\gamma}{\gamma-1}$ and

$$\rho_- \simeq \frac{\gamma}{\gamma-1}\Gamma_-\rho_+ \quad . \tag{6.51}$$

With Eq. (6.48), the Eq. (6.42) can be written as

$$\left(\frac{\epsilon_-}{c^2}\right) = \left(\frac{\Gamma_- c}{u_-}\right)\frac{f_\rho}{1+C_-}\left(\frac{\rho_+}{f_s}\right)\frac{\Gamma_-}{\Gamma_+} \quad . \tag{6.52}$$

Therefore, $\epsilon_- \simeq \Gamma_- c^2$ for ultra-relativistic shocks.

From Eqs. (6.48) and (6.52), equation of state $p = p(\rho, \epsilon)$ can be re-written as

$$p_- = \left(\frac{p_-}{\rho_- \epsilon_-}\right) \rho_- \epsilon_- = \Gamma_-^2 f_p \quad , \tag{6.53}$$

where

$$f_p \equiv \frac{f_\rho^2}{C_-(1+C_-)f_\epsilon} \left(\frac{\rho_+}{f_s}\right) \left(\frac{\Gamma_- c}{u_-}\right) \frac{\rho_+ c^2}{\Gamma_+^2} \quad . \tag{6.54}$$

For ultra-relativistic shocks and the equation of state (6.15), $f_p \simeq \gamma \rho_+ c^2$ and $p_- \simeq \gamma \Gamma_-^2 \rho_+ c^2$.

Junction condition Eq. (5.60) cannot be directly used to trace the radial position R_s of shock front. From Eq. (5.62), we obtain along the world line just ahead of the shock front

$$\left.\frac{dR}{dt}\right|_s = b_+ \Gamma_+ \dot{\mu}_s + a_+ u_+ \quad . \tag{6.55}$$

This equation together with $u_+ = 0$, Eq. (5.27) and (6.37) gives

$$\dot{R}_s = \Gamma_+ \frac{f_s}{\rho_+} c \quad . \tag{6.56}$$

Since for relativistic shocks $f_s \simeq \rho_+$ and $\Gamma_+ \simeq 1$, we have $\dot{R}_s \simeq c$.

The situation for the gravitational effective mass m_s is similar to that for R_s. From the general form of mass

$$dm = \frac{\partial m}{\partial \mu} d\mu + \frac{\partial m}{\partial t} dt \tag{6.57}$$

together with Eqs. (5.34), (5.35), (5.13), (5.27) and (6.2), we have, along the world line just ahead of the shock front

$$\dot{m}_s = \Gamma_+ \left(1 + \frac{\epsilon_+}{c^2}\right) \dot{\mu}_s \quad , \tag{6.58}$$

where we have use the condition $u_+ = 0$.

We have expressed all the quantities just behind the shock front with the Lorentz factor Γ_-, which should be given with high accuracy. However, the finite difference scheme can give values only at grid points or half zones and the moving shock front also prevents us from correctly centering the zone just behind the shock front. Its value obtained with any extrapolation making use of the values computed in the neighbor zones are doubtful. Actually, as discussed before the forward characteristic equation (5.53) can be used to accurately calculate Γ_-.

Fig. (6.8) shows the shock front path SS' and the characteristic curves from the shock front position at t^{n+1} back to the base time level t^n. In the fluid just behind of the shock, along the forward characteristic (the curve NS') given by

$$d\mu = +4\pi R_{1/2}^2 a_{1/2} \rho_{1/2} c_{s1/2} \, dt \tag{6.59}$$

we have the characteristic equation

$$du + \frac{\Gamma_{1/2}}{\rho_{1/2}\omega_{1/2}c_{s1/2}}dp + \left[G\left(\frac{m_{1/2}}{R_{1/2}^2} + \frac{4\pi p_{1/2}R_{1/2}}{c^2}\right) + \frac{2u_{1/2}\Gamma_{1/2}c_{s1/2}}{R_{1/2}}\right]a_{1/2}\,dt = 0 \quad , \tag{6.60}$$

where the subscript index 1/2 indicates that the quantities take the values at the point at which NS′ curve interacts with the time line $t^{n+1/2}$ (the dashed line) in Fig. (6.8) and

$$du = u_{S'} - u_N = u_- - u_N \tag{6.61}$$

$$dp = p_{S'} - p_N = p_- - p_N \quad . \tag{6.62}$$

Using Eq. (6.40), we write Eq. (6.60) as

$$u_- - u_N + A_{1/2}^f\left(\frac{p_-}{\Gamma_-^2}\right)\left[1 + \left(\frac{u_-}{c}\right)^2 - 2\frac{Gm_s}{R_s c^2}\right] - A_{1/2}^f p_N + B_{1/2}^f \Delta t = 0 \quad , \tag{6.63}$$

where

$$A_{1/2}^f \equiv \frac{\Gamma_{1/2}}{\rho_{1/2}\omega_{1/2}c_{s1/2}} \tag{6.64}$$

$$B_{1/2}^f \equiv \left[G\left(\frac{m_{1/2}}{R_{1/2}^2} + \frac{4\pi p_{1/2}R_{1/2}}{c^2}\right) + \frac{2u_{1/2}\Gamma_{1/2}c_{s1/2}}{R_{1/2}}\right]a_{1/2} \quad . \tag{6.65}$$

The solution of Eq. (6.63) is

$$\frac{u_-}{c} = -\frac{c\Gamma_-^2}{2A_{1/2}^f p_-} + \sqrt{\left(\frac{c\Gamma_-^2}{2A_{1/2}^f p_-}\right)^2 + \frac{\Gamma_-^2}{A_{1/2}^f p_-}u_N + \frac{\Gamma_-^2 p_N}{p_-} - 1 + 2\frac{Gm_s}{R_s c^2} - \frac{B_{1/2}^f}{A_{1/2}^f}\frac{\Gamma_-^2}{p_-}\Delta t}\,, \tag{6.66}$$

where the condition $u_- > 0$ is used to choose the solution. With Eq. (6.53), Eq. (6.66) is written as

$$\frac{u_-}{c} = -\frac{c}{2A_{1/2}^f f_p} + \sqrt{\left(\frac{c}{2A_{1/2}^f f_p}\right)^2 + \frac{c}{A_{1/2}^f f_p}\left(\frac{u_N}{c}\right) + \frac{p_N}{f_p} - 1 + 2\frac{Gm_s}{R_s c^2} - \frac{B_{1/2}^f}{A_{1/2}^f}\frac{\Delta t}{f_p}}\,. \tag{6.67}$$

To compute R_s, μ_s and m_s at S′, we do standard interpolation to calculate $\dot{R}_s^{n+1/2}$, $f_s^{n+1/2}$, and using equations to compute $\dot{\mu}_s^{n+1/2}$ and $\dot{m}_s^{n+1/2}$. To ensure mass conserving, we use

$$(R_s^2)^{n+1/2} = \frac{1}{3}[(R_s^{n+1})^2 + R_s^{n+1}R_s^n + (R_s^n)^2] \tag{6.68}$$

when we compute $\dot{\mu}_s^{n+1/2}$.

The Eqs. (6.34)–(6.38), (6.40), (6.43)–(6.46), (6.48)–(6.50), (6.52)–(6.54), (6.56), (6.58)–(6.59), (6.64)–(6.65) and (6.67) consists of the set of equations treating shock

fronts. With these equations, we calculate the unknowns at the shocks: (1) the position μ_s of the shock front, (2) its velocity $\dot{\mu}_s$, (3) the radial coordinate R_s, (4) the time scaling factor a_-, (5) the rest mass density ρ_-, (6) specific internal energy ϵ_-, (7) pressure p_-, (8) specific enthalpy ω_-, (9) effective gravitational mass m_s, (10) radial four velocity u_-, and (10) Lorentz factor Γ_-.

In principle, the equations should be solved simultaneously. Here we try to find their solutions iteratively. To compute the quantities at time t^{n+1}, our strategy is:

- (1) linear extrapolating f_s^{n+1}, $\dot{R}_s^{n+1/2}$, ϵ_-^{n+1}, $\rho_- n+1$, and u_-^{n+1}.
- (2) calculating R_s^{n+1} with $\dot{R}_s^{n+1/2}$, $\dot{\mu}_s^{n+1}$ with Eqs. (6.37) and then μ_s^{n+1}, $\dot{m}_s^{n+1}$ with (6.58) and then m_s^{n+1}, v_-^{n+1} with (6.38), p_-^{n+1} with (6.15), ω_-^{n+1} with (6.35), a_-^{n+1} with (6.34), and Γ_-^{n+1} with (6.36).
- (3) The position of the foot of the forward characteristic is determined by (6.59) and u_N, ρ_N and ϵ_N are interpolated with second order interpolation.
- (4) computing Γ_+ with (6.40).
- (5) computing f_s with (6.36)
- (6) estimating $\dot{R}_s$ with (6.56) and R_s.
- (7) Computing $\dot{\mu}_s$ with (6.37), $\dot{m}_s$ with (6.58), and m_s.
- (8) Computing Γ_+ again.
- (9) Computing f_ρ with (6.49), f_ϵ with (6.43), and f_p with (6.54).
- (10) calculating u_- with (6.67) and Γ_- with (6.40).
- (11) ϵ_- and ρ_- are calculated with (6.52) and (6.51), respectively. p_- is computed with the equation of state (6.15). v_- is computed.

If necessary, repeat (5) to (11) to obtain improved values. When we need value at t^{n+1} and $t^{n+1/2}$, we extrapolate the values first time and then interpolate them after first cycle if the accuracy is not enough. The values of variables at the foot of the forward characteristic curve (point N in Fig. (6.8)) are computed with second order interpolation with the three points S, $j-1$ and $j-2$ or S, $j-1/2$ and $j-3/2$ for zone boundary or mid-zone quantities, respectively. To determine the foot with Eq. (6.59), first we use the values at half way along SS$'$ with extrapolation and then use the values interpolated between points S$'$ and N.

The convergence is very fast. At the initial stage, repetition is not necessary. After the system evolve to late stage, the iteration number is a few times.

6.5.3 The Treatment of Contact Discontinuities

After the splitting of the initial discontinuity immediately when the system is set, only contact discontinuity (CD) is left to the position in the comoving frame where the initial discontinuity was. For convenience, we will call the part of the system inside the contact discontinuity as fireball (it is the initial fireball) while the region between contact discontinuity and the shock front as the shock wave or shocked region. The states at the both sides of the contact discontinuity are no longer independent and must be connected by the junction conditions given in Sec. (5.5.3). Across CD, u, R, m, Γ, p and a are continuous but ρ, ϵ and ω are not. For simplicity, we set the grid so that the CD is at a zone boundary. Thus, u, R, m and Γ can be calculated with the scheme for outer boundary, described for free expansion. The value of p at CD is needed and can be computed with interpolation as its effects on the calculation results is not great. ρ and ϵ are discontinuous, but their values at CD are not used explicitly. a and ω have definition at zone centers. a is calculated with the constraint equation (6.6) at each time slice from the outer boundary inward to the inner boundary. As entropy and ω are not continuous at CD, (6.6) cannot be used across the CD but only separately on both sides. In the shocked region, it can be computed with the outer boundary provided by the shock front. For the fireball, however, we have to compute a and ω at CD and take them as the outer boundary for this region. Since a is continuous across the CD, we simply set a_- inside the fireball just behind the CD equal to a_+ in the shocked region just ahead of the CD. So, the key point is to accurately calculate $\omega_\pm$.

Since $[\omega]_\pm \neq 0$ at CD, any value of ω obtained with interpolation based on the values of points at two sides or with extrapolation based the values of points only at one side looks doubtful. One possible way is properly to interpolate those quantities which are continuous across the CD, and then to compute $\omega_\pm$ with related equations. Among all the thermodynamic quantities of the fluid, only pressure is continuous. So, we calculate pressure p_{CD} at the CD with suitable interpolation (for example logarithmic interpolation). It should be noted that p changes by many orders of magnitudes behind the contact discontinuity up to the stage at which the rarefaction wave has completely passed through the boundary zones. Fortunately, during this period the zone width just ahead of the CD is by many orders of magnitude smaller than the zone size just behind the CD, as the ISM has very low density and the period is very short. We use the equation of state (6.15) and the energy equation (6.4), including artificial viscosity term, to get $\rho_\pm$ and $\epsilon_\pm$. We then use Eq. (6.9) to compute $\omega_\pm$. In this method, we need to compute pressure at CD properly. The advantage is that (1) we may use the technique developed in Sec. (6.4) for treating the zone just behind the CD and (2) the calculated velocity is the average of a half zone instead a value for the special point

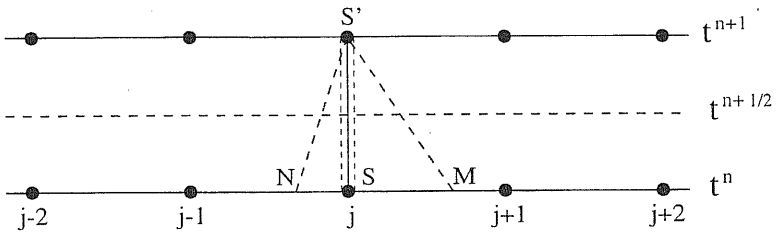

Figure 6.9: The characteristic structure of a contact discontinuity (CD). NS′ is the forward characteristic curve and MS′ is the backward characteristic curve. The advective characteristic curves are drawn on both sides of the CD (solid SS′ line).

CD. The difference between them are very large, as the latter would be the solution of Riemann problem and provide a singular boundary for the region inside the CD and is not easy to treat. For the former, the initial velocity is not very large (we use $\Gamma \sim 3$), its effects on the final solution is not significant and it is also easy to smooth the solution at the CD.

We also try another way is to use characteristic equations and the junction conditions to compute $\omega_{\pm}$. As CD is at rest relative to the fluid just behind and just ahead of it. Disturbances at both sides can propagate to it along forward and backward characteristics. Fig. (6.9) shows that there four characteristic curves can be drawn from the time t^{n+1} to the base level t^n: forward, backward, and two advective characteristic curves on both sides of CD. The two advective characteristic curves are just ahead of and just behind the CD. At the CD, we need to calculate R, a, u, Γ, p, m, $\rho_{\pm}$, $\epsilon_{\pm}$ and $\omega_{\pm}$. The forward and backward characteristic equations (5.53) first give p and u. With advective characteristic equations (5.52), we compute for $\rho_{\pm}$. (5.54) and (5.55) are used to computed R and m, respectively. Lorentz factor Γ is calculated with Eq. (6.8). After $\omega_{\pm}$ is given with Eq. (6.9), we can compute a with the constraint equation (5.57). Here, only discontinuous quantities are denoted with $\pm$. The feet of the forward and backward characteristic curves (N and M points in Fig.(6.9)) are determined by Eq. (5.53).

The computation is carried out iteratively. During the calculation, if we need quantities at time $t^{n+1/2}$ or t^{n+1} which are not known, we first use the extrapolation and refine them after we finish one cycle of iteration. All the middle values of the quantities in the characteristic equations are the values of the points where the $t^{n+1/2}$ line intersects with the the characteristic curves in Fig.(6.9).

The values at feet of the characteristic curves at base time level t^n are computed with interpolation. On the shock wave side, it is always safe to use second order interpolation, while on the fireball side linear interpolation is more stable and used in the code.

There are difficulties with this technique. Firstly, it is difficult to calculate the

values at the foot of the forward characteristic curves. The second is that the difference equations will automatically take the value at CD as the average over a small region and the final solution will give very high Lorentz factor at the CD.

6.5.4 Regridding Procedure for the Shock Wave Region

One of the very important feature of our system is that a very strong shock forms just ahead of the CD once the fireball is set within the external medium. Since the great contrast in mass density and internal energy density, a grid with a resolution similar to the one used in the fireball cannot be used in the shock wave, especially at the initial stage. The required resolution in the shock wave is many orders of magnitude higher than that in the fireball region. So, we devise a regrid procedure to handle this problem. For the shock wave, we start with a grid with only one special zone ($J = 1$ in Fig. (6.8)), whose size changes, when the shock sweeps out into the ISM. We define a limit size

$$\Delta\mu_{max} = 10^{-N_{sh}} \frac{M_0}{J_{fb}} \quad , \tag{6.69}$$

where J_{fb} is the total zone number of the grids inside the fireball and N_{sh} is a positive number of any value, depending on the requirement on the initial resolution of the region. If the size of the special zone exceeds the limit, its size is fixed and it becomes the first zone of the grid inside the shocked region. The grid number increases by one. The zone just behind the shock front becomes the new special zone and it immediately starts growing. We will always keep the zone just behind the shock front as the special zone. Now, $\Delta\mu_{max}$ is set to a new value, e.g. the size of the first regular zone next to the special zone. When the new special zone grows exceeding the new limit again, we regrid the region between the CD and the shock front with two equal zones through a general regridding procedure. All quantities are computed in the way described below. The grid number increases by one and the next special zone grows again. Each time when we regrid the shocked regions with a uniform grid, we re-set $\Delta\mu_{max}$ to a new value. The regridding procedure is repeated. As it is impossible and also not necessary to use a very large grid number, we set a upper limit number J_{sh}. If the zone number $j + 1$ (or the index of the special zone) exceeds this maximum number, j will be set to a number J_{min} with $1 < J_{min} < J_{sh}$ and regrid the whole shocked region. At this stage, the zone $j + 1$ is the special zone. Whenever the size of the special zone becomes larger than $\Delta\mu_{max}$, we start the regridding procedure for the whole shock wave region (including the special zone) with a uniform grid with zone number $j + 1$. The regridding is always done with a general regridding procedure. In this case, the regridding procedure is flexible.

In the numerical calculation, we have used $N_{sh} = 30$, $J_{sh} = 11$ and $J_{min} = 10$ which turns to be good for our special interest.

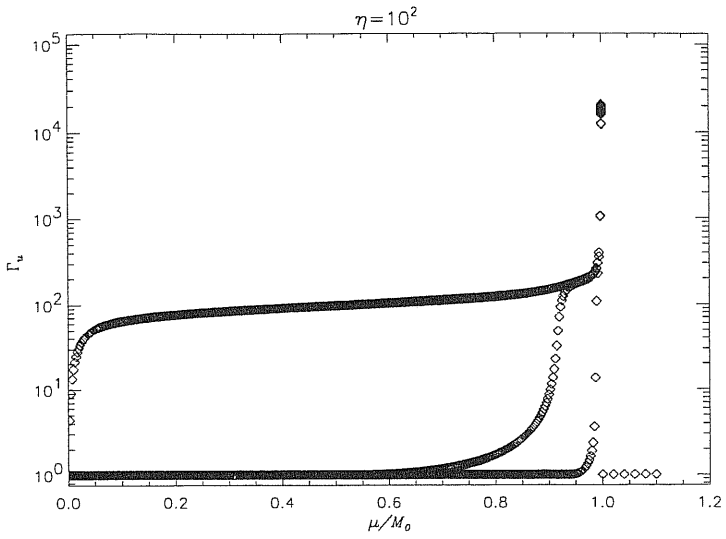

Figure 6.10: Lorentz factor Γ as a function of normalized rest mass coordinate μ/M_0. The curves correspond to expansion factor $(ct + R_0)/R_0$ values (from right to left) 10^3, 10^4 and $10^5 \times \eta$. The fireball starts expanding at $t = 0$. The contact discontinuity is at $\mu/M_0 = 1$. The Lorentz factor in the ISM is nearly unity.

It is important to correctly calculate the quantities when we regrid the shock region. Experimentally, we find that the quantities defined at grid boundary: R, m, u, and $a\omega$ should be interpolated by either linearly or logarithmically. If we take $J_{sh} = 21$ and $J_{min} = 7$, namely, an expansion factor of 3, these quantities will directly take the values at old grid boundary. ϵ, Γ and ρ are iteratively computed from the constraint equations (6.6), (6.7) and (6.10), respectively. Pressure is calculated from the equation of state (6.15) and ω from (6.9). As we have Γ_u and Γ_ρ, we use their average values.

6.5.5 Results and Discussion

We discuss the results obtained with the technique described above. For the fireball region we use 400 zones, while in the shock wave region J_{sh} is set equal to 10 and the parameter J_{min} is set to 9. The initial discontinuity is completely split at the initial time and the CD is treated with characteristic methods.

6.5.5.1 Free Acceleration Phase

With this special treatment of the initial discontinuity, an ultra-relativistic shock is formed just ahead of the contact discontinuity and propagates into the external medium with a initial velocity corresponding to a value of $\Gamma \sim 2.5 \times 10^5$. At the mean time, a rarefaction wave forms and propagates inward, into the fireball and accelerates the matter inside the fireball. As the external medium has very low density, its effects

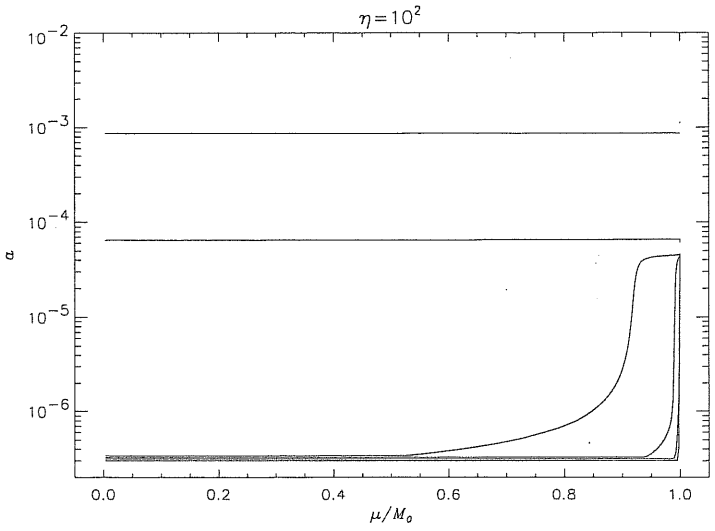

Figure 6.11: The time scaling factor a as a function of normalized rest mass coordinate μ/M_0. The curves are corresponding to expansion factor $(ct + R_0)/R_0$ (from bottom to top): 10, 10^2, 10^3, 10^4, 10^5, and $10^6 \times \eta$. The fireball starts expanding at $t = 0$. The contact discontinuity is at $\mu/M_0 = 1$.

on the expansion of the fireball at the initial stage can be neglected. The fireball will then expand freely, as was described in Sec. (6.4.5). We will present our results in the comoving frame. If we say at some instant, we mean t = constant instead of T = constant. One of the advantage to present the results in the comoving frame is that the whole process of the acceleration of matter in different layers is clearly disclosed (see Sec. (6.4.5)). This is one of the reasons, but not the most important, why we chose comoving frame. We do compare our results obtain here with those in Sec. (6.4.5), which are consistent with those found by other authors.

To compare with the results presented in Sec. (6.4.5), we give the evolution of the Lorentz factor Γ as a function of the mass coordinate and the comoving coordinate time t in Fig. (6.10), which we can compare with Fig. (6.3). Except for a very thin layer just behind the contact discontinuity, the matter will be accelerated to the saturated value of $\Gamma \simeq \eta = 10^2$, as expected. The evolution inside the fireball shown in Figs. (6.10) is nearly identical to that in Fig. (6.3). However, we have to note that the difference of the acceleration time between the two figures is several orders of magnitude. This is not the case in Lab frame. The reason is that although we have synchronized our coordinate time t to the same observer at distance, the time scaling factor a just behind the surface of the free fireball is unity due to the complete centered rarefaction wave while in the case of fireball with interaction with ISM a just behind the contact discontinuity is scaled to the value in the shock wave just ahead of the CD and is much smaller than unity (see Fig. (6.11)).

Fig. (6.11) shows that when the matter inside the fireball is accelerated to the

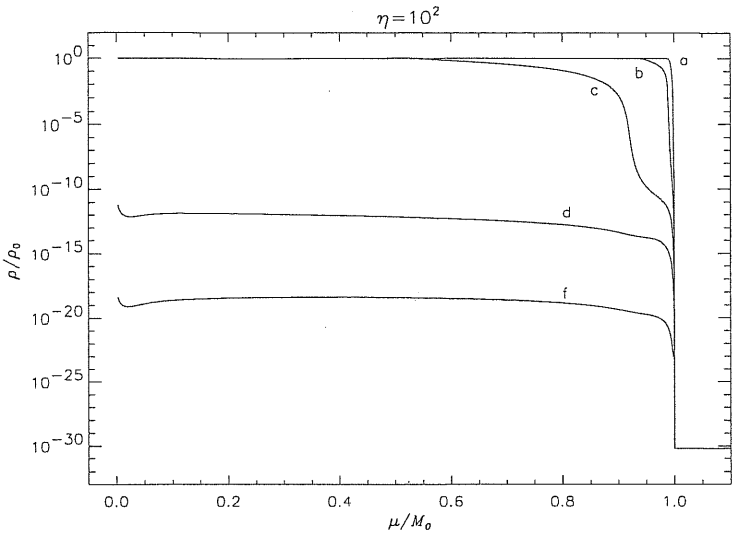

Figure 6.12: The rest mass normalized with the initial mass density ρ_0 *vs* rest mass coordinate divided by the initial total mass. The CD is at $\mu/M_0 = 1$. The curves from a to f are corresponding to the expansion factor $(ct + R_0)/R_0$: 10^2, 10^3, 10^4, 10^5, and $10^6 \times \eta$. When it starts, $t = 0$.

saturating velocity, a increases to maximum at some instant as in the case of the free expansion. The behavior of a is different from the case for a free expansion. During the stage that rarefaction has not yet completely passed through the shell, a shift up slightly. While, after it crosses the shell, a increases significantly. This change connects to the variation of the strength of the external shock, which we will discuss in Sec. (6.5.5.2). This is clear in Fig. (6.11) from the slow increase of the value of a in the un-rarefacted part of the fireball. The last two curves in Fig. (6.11) are quite far apart. This indicates a dramatical change in strength of the external shocks.

The rest mass density as a function of rest mass coordinate and coordinate time is shown in Fig. (6.12). The distributions of mass density inside the fireball is similar to those in Fig. (6.4). A very narrow rarefaction wave is sent out and propagates from the CD toward the center of the fireball. The matter is expanding very significantly by the rarefaction wave. Before and after the rarefaction wave, the mass density is nearly constant. After the rarefaction wave completely crosses the fireball, the distribution of the mass inside the fireball is nearly constant except in the center and the region close to the CD. However, the same as for the Lorentz factor, similar density distributions occur but at different times and the rarefaction wave passes the fireball with a different speed from a free fireball. The curve d in Fig. (6.12) is corresponding to the relative density $\rho/\rho_0 \sim 10^{-12}$ and time $t = 3.3 \times 10^3$ sec, while in Fig. (6.4) a curve with a similar relative mass density $\rho/\rho_0 \sim 10^{-13}$ corresponds to a time $t = 0.33$ sec. These two curves are almost identical: an almost equal constant density distribution in the inner region, a slight rise near the inner boundary and a dramatical decrease near

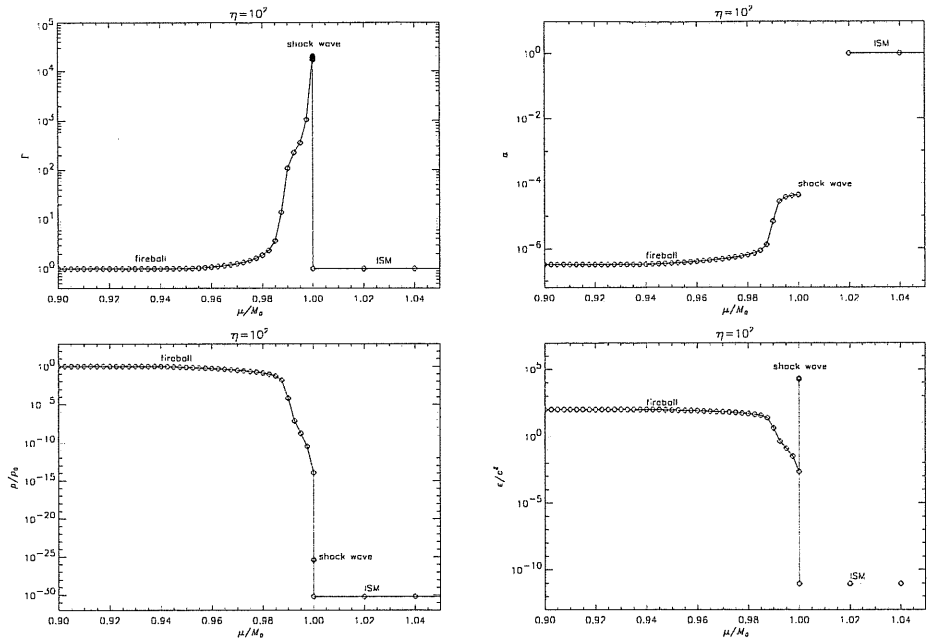

Figure 6.13: The fine structures of Γ, a, ρ and ϵ in the regions near the contact discontinuity and shock wave at $(ct + R_0)/R_0 = 10^3\eta$. The shock wave cannot be resolved in this figures. Rest mass coordinate has been divided by the initial total rest mass. The CD is exactly at $\mu/M_0 = 1$, while the shock front is at $\mu/M_0 \simeq 1$.

the CD or surface. These two curves are roughly corresponding to the time when the rarefaction wave completely passes through the fireball region. Therefore the ratio of the time intervals when the rarefaction completely passes through the fireball is 10^4.

From the relations (5.15), (5.14) and (6.8), we have

$$T \simeq \int \Gamma a \, dt \quad , \tag{6.70}$$

which gives

$$T \sim \Gamma_{max} a_{max} t \quad . \tag{6.71}$$

From Figs. (6.11) and (6.10), $\Gamma \sim 10^2$ and $a \sim 7 \times 10^{-5}$. Thus, Eq. (6.71) gives for $T_{ism} \sim 33$ sec. While it is read that from Figs. (6.3) and (6.3) for the cure at time $t = 0.33$ sec, $a \simeq 1$ and $\Gamma \sim 10^2$. We have the corresponding Lab time $T_{free} \sim 33 \sim sec$ and thus $T_{ism} \sim T_{free}$. The two curves are roughly identical on the Lab time level. Therefore, an observer in the Lab frame would measure the expansions to be the same, while another observer in the comoving frame would have different results for the expansions of the two fireballs. This is the reason why other authors did give some results for the initial stages of the expansion of two fireballs: one into vacuum and other interacting with ISM.

The results for enthalpy ω and internal energy can be explained in the same way.

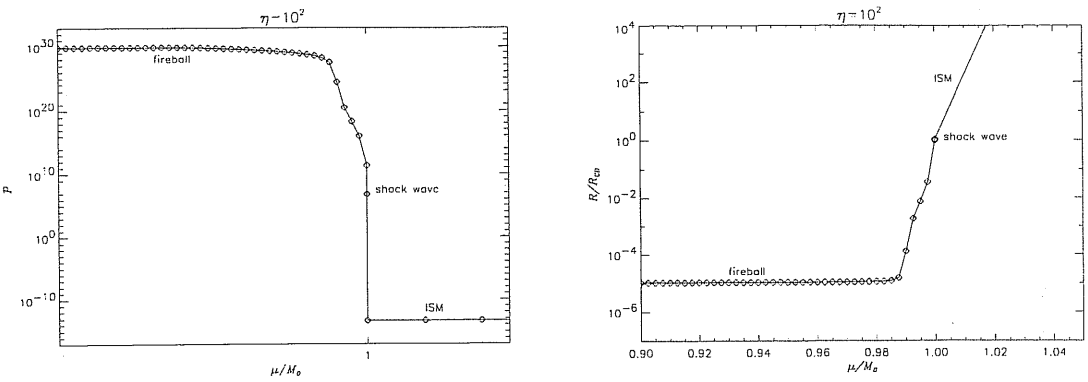

Figure 6.14: The fine structures of pressure p and R in the regions near the contact discontinuity and shock wave at $\mu/M_0 \simeq 1$ at $(ct + R_0)/R_0 = 10^3\eta$. R has been divided by the radial position of the contact discontinuity (CD).

6.5.5.2 The Interaction with ISM

At the beginning, a very thin layer is accelerated by the centered rarefaction wave to a relativistic velocity and the layer produces a very thin ultra-relativistic shock. When the shock moves out, it sweeps up more and more external matter and accelerates this matter to a relativistic motion. As the total energy contained in the shock wave region would be very small, its effect is very small and the fireball expands freely. However, as all the quantities have to satisfy the junction conditions when they cross the contact discontinuity and the shock front, we have to solve the whole system consistently. Fig. (6.13) shows the structures of the Lorentz factor Γ, time scaling factor a, mass density ρ and specific internal energy ϵ at the early stage $t = 33$ sec. At this stage, the shock wave region is too small and cannot be resolved in the diagrams with rest mass as space coordinate. The state of ISM is constant and independent of time. All the quantities are discontinuous at the shock front, while at the contact discontinuity Γ and a are continuous and ρ and ϵ are discontinuous. The pressure p is continuous at the contact discontinuity but discontinuous at the shock front (Fig. (6.14)). As R is continuous both at contact discontinuity and at shock front (Fig. (6.14), we may take it as the space coordinate of the plots.

The features are very clear in the plots if R is taken as the space coordinate, though the curves are no longer smooth at the region near contact discontinuity. As the very low density external medium is accelerated by the very thin surface layer of the fireball with extremely high Lorentz factor $\Gamma \sim 2.5 \times 10^5$, the relativistic shock propagates outward into the ISM and accelerates more and more matter. However as the energy contained within the thin surface layer is very small, this shock cannot keep so high velocity for long time when it sweeps out the external medium. Fig. (6.15) shows that the the Lorentz factor Γ at the shock wave is almost constant when the fireball is at its

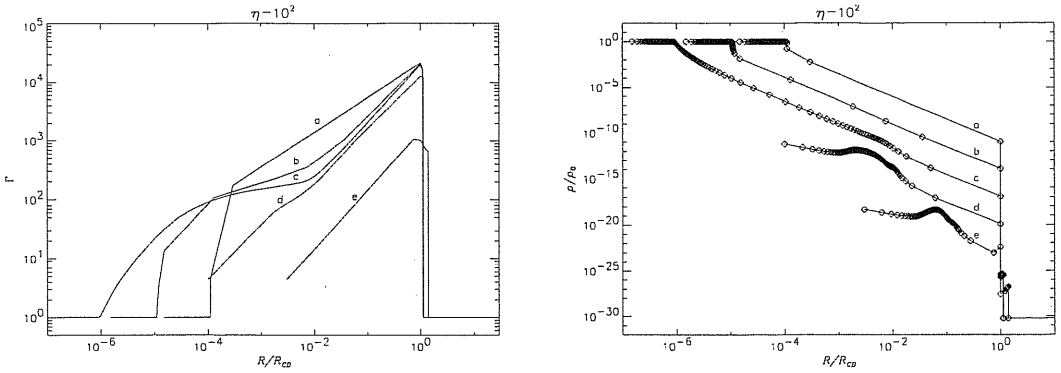

Figure 6.15: The Lorentz factor Γ (*left*) and normalized rest mass density ρ/ρ_0 by initial rest mass density (*right*) as a function of the normalized radius R/R_{CD} with R_{CD} the radial position of the contact discontinuity. The time series for the curves from a to e are $t = 3.3$, 33, 3.3×10^2, 3.3×10^3 and 3.3×10^4 sec.

acceleration phase (initial free expansion) and the shock wave region does not spread out relative to the contact discontinuity. At this phase, the fireball behaves like a piston pushing the small shock region and continuously provides energy to it. However this phase is very short and the Lorentz factor of the shock wave decrease very rapidly from a value much higher than the average value of η, as the energy is consumed due to the adiabatic expansion. Meanwhile, the velocity of the very thin layer just behind the contact discontinuity decreases accordingly, due to the reverse shock. However, the reverse shock in other part of the fireball is still Newtonian at this moment and does not significantly reduce the Lorentz factor of the whole fireball region and converts a significant part of the bulk kinetic energy into thermal energy. The Lorentz factor of inner part of the fireball after the initial acceleration keeps as a constant until very late time and linearly increases with the relative radius. Kobayashi *et al.* (1999) presented their results with an absolute relative radial coordinate $R - R_{CD}$. We prefer to use $\lg(R/R_{CD})$ as R varies by many orders of magnitude in the solution domain and a plot with logarithmic radial coordinate is preferred. As R_{CD} is related to the Lorentz factor, the ratio is insensitive to the Lorentz factor and the broadening of the the shock wave region is due to the dispersion of the Lorentz factor.

Fig. (6.15) also shows the rest mass density as a function of the relative radius R/R_{CD}. The mass density just behind the CD is higher than the value just ahead of the CD. It decreases with time very rapidly. When the reverse shock becomes relativistic just at the small layer behind the CD, the decrease becomes slow and the mass density develops a local maximum. Most of the fireball mass concentrates in a fairly thin shell. This region slowly catch up the shock wave region. When the two regions become very close, compression happens and the reverse shock becomes

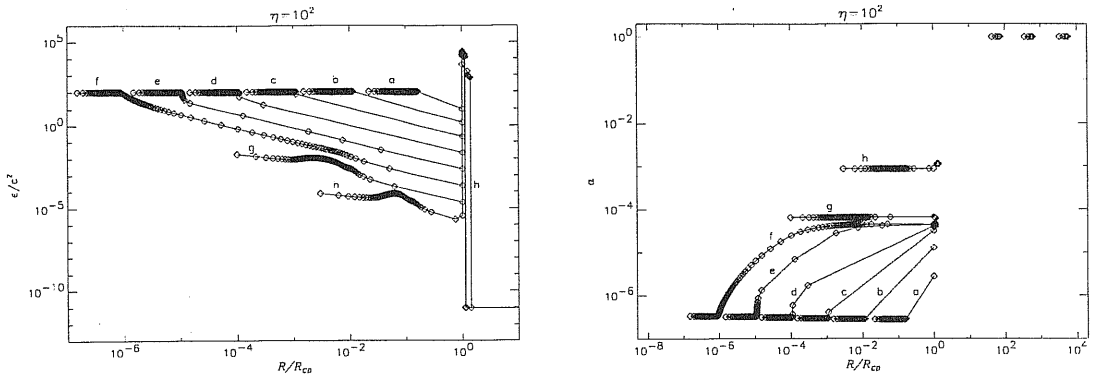

Figure 6.16: The specific internal energy ϵ/c^2 (*left*) and the time scaling factor a (*right*) as a function of the normalized radius R/R_{CD} with R_{CD} the radial position of the contact discontinuity. The time slices for the curves from a to h are $(ct + R_0)/R_0 = 0.11$, 1.0, 10, 10^2, 10^3, 10^4, 10^5, and $10^6 \times \eta$.

important. From equation (4.38), the reverse shock would become relativistic at

$$R_N = l^{3/2}/\Delta^{1/2}\Gamma^2 \sim 10^{18} \text{ cm} \quad , \tag{6.72}$$

where l is the Sedov length given by Eq. (4.36) and $l \simeq 10^{18}$ cm. While for our system the external shock becomes important when it sweeps up rest mass $\sim M_0/\Gamma$ which will happen at $R \sim 10^{16}$ cm.

Fig. (6.16) gives the specific internal energy and the time scaling factor a s a function of relative radius R/R_{CD}.

Figure (6.17) shows an example for the various quantities as function of radius R divided by the radius where the contact discontinuity is.

6.6 Adaptive Mesh Refinement

6.6.1 Introduction

When we study the fireball expanding into the external medium, we have to handle the sharp discontinuities at shocks and contact discontinuities, which is one of the challenges presented by ultra-relativistic hydrodynamic problems for compressible fluids. The piecewise parabolic method (PPM) (Colella & Woodward 1984; Martí & Müller 1996; Aloy *et al.* 1999) provided a powerful and accurate method to treating strong shocks, However, PPM is expected to be computationally too expensive for our problem. Wen *et al.* (1997) presented a hybrid method that may be competitive and faster. Finite difference methods with artificial viscosity have been a popular choice when dealing with shocks, but at the cost of smearing shocks and contact discontinuities over several zones. To make this method competitive to the PPM, adaptive-mesh refinement techniques have to be used (Woodward & Colella 1984).

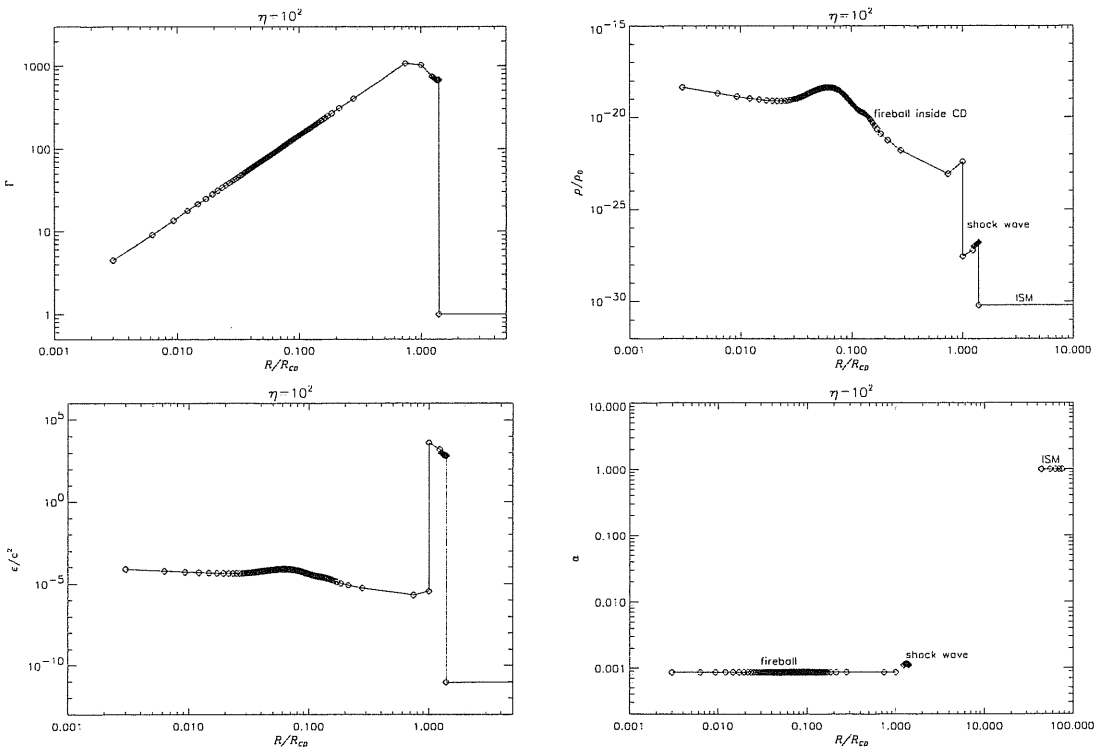

Figure 6.17: The Lorentz factor Γ (*upper-left*), rest mass density ρ (*upper-right*), specific internal energy (*lower-left*) and the time scaling factor a (*lower-right*) as a function of the relative radius R/R_{CD} at comoving time instant $(ct + R_0)/R_0 = 10^6 \times \eta$. The different region is indicated in the figures both for the mass density ρ and for the time scaling factor a.

Adaptive mesh refinement (AMR) methods are devised to treat a category of problems known as "two-scale problems", the solutions of which exhibit a characteristic feature - the spatial and/or temporal scales of variation in the solution vary dramatically from one region to another. The inability to efficiently resolve two-scale problems using conventional uniform approaches has long been recognized and AMR have been discussed widely (Brandt 1977; Winkler 1977; Berger & Oliger 1984; Woodward & Colella 1984; Choptuik 1989; Babuska & Bieteman 1985; Masso *et al.* 1995; Hamade & Stewart 1996; Wild 1996). The AMR method devised by Berger & Oliger (1984) to solve hyperbolic partial differential equations is very successful and has been applied in different physical systems (e.g. Skamarock 1989; Choptuik 1993), particularly in numerical relativity. Choptuik applied Berger-Oliger in his numerical relativity computation and reduce the computational load by a factor of 10^8. Although Berger & Oliger try to make their method applicable for multi-dimension systems, it becomes very complicated when it is applied in systems with dimensions higher than one. Recently, Wild (1996) extended Berger-Oliger's method to multi-dimensions in a simpler way, which is very useful in the study of the coalescence of two compact objects, particularly binary black holes.

In this section, we only briefly discuss the technique for one dimensional AMR, as we have not yet completely implemented it. This technique basically follows the Berger-Oliger algorithm but with some modifications. For higher dimensional AMR, Wild's method (1996) is suggested.

6.6.2 Grid Structure and Integration Algorithm

At the start of computation, only a uniform base grid G_0 is constructed and will remain fixed for the duration of the computation. During a computation, refined locally-uniform sub-grids will be created adaptively in response to some feature in the transient solution, such as the appearance of shock fronts. The fine sub-grids may contain even finer sub-grids within their boundaries, i.e. sub-grids can be recursively generated. If a grid is contained within l coarser grids, it is called at l level. Refined grids G_{l+1} within the grid G_l at l level are at level $l+1$, and so on. Therefore, base grid G_0 is at 0 level, as it is not contained in any coarser grid. In this way, a nested sequence of grids with finer and finer discretizations may be created over some portion of the spatial domain. All grids at level l in the hierarchy have mesh width h_l.

In practice, a set of possible mesh discretization $\{h_0, h_1, h_2, \cdots, h_{max}\}$ are specified in advance with $h_l = r_{l+1}h_{l+1}$, where the refinement factor r_{l+1} is a small integral number. Typically, the same refinement factor will be used for all levels and refinement will be allowed up to some maximum level l_{max}. A good choice for this refinement ratio will depends on how much of the domain needs what amount of refinement. In the

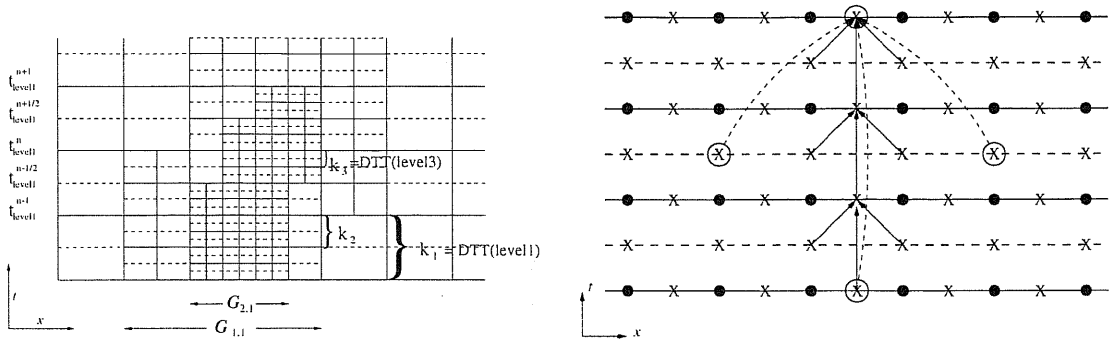

Figure 6.18: (*left*) One dimensional space-time mesh refinement. (*right*) Rishardson truncation error estimation procedure.

two-step integration scheme, a refinement ratio of 3 is preferred. In special cases, where it is expected that all areas needing refinement will need a lot of it, higher values of $r = 5, 7, 9, \cdots$, can be used efficiently.

Sub-grids are not patched into the coarse grid and instead overlay a coarser grid. Each grid is defined as an independent computational entity (independently of other grids), with its own solution vector, storage, etc. In this way, each sub-grid can be integrated in time independently of the other grids except for the determination of its boundary values. A grid with mesh width h_l is required to be contained in a grid at level $l-1$, which in turn is required to be contained in a level $l-2$ grid, and so on to the coarsest grid. All intermediate grids between the finest and the coarsest are maintained. Each point in a fine grid at level l must be in the interior, not just on the boundary, of a grid at the next coarser level $l-1$, unless it is on the physical boundary of the domain. This is important for the determination of values at grid boundary.

When a grid at level l is refined with a refinement factor r in space, the same factor r is used to set the time step on the level l grids, $k_l = k_{l-1}/r$. Figure (6.18) illustrates the refinement procedure in one space dimension and time. In this way, the mesh ratio $\lambda = k_l/h_l$ of the time step to space step is kept constant on all grids and thus Courant condition will be satisfied on all grids once it is satisfied on the base grid G_0.

The grid time integration steps are interleaved so that before advancing a grid, all its sub-grids are integrated to the same time. At every coarse grid step, all grids should be at the same time. In order to provide the boundary conditions for a fine grid (see Sec.(6.6.3)), we integrate first the coarser grid one time step and then r steps for the fine grid. This procedure starts from the base level to the highest level. After the finest grid advances to the time level with the coarser grid, the coarser advances one time step again.

This refinement technique makes it easy to solve different equations on different grids or to use different integrators in different regions of the solution, e.g. first order integrator on a refined grid around a shock and a higher order method elsewhere and

a more expensive method (e.g. Glimm scheme used by Wen *et al.* (1997)) only on refined grids and less expensive integrator on the coarse grid.

6.6.3 Boundary Conditions, Grid Initialization and Updating

There are two kinds of boundaries for AMR: (1) physical boundaries provided by partial differential equations (PDEs) or external boundaries and (2) internal boundaries, which are in the interior of the problem domain. Wild (1996) investigated the internal boundary properties in detail. The boundary values are computed using values from underlying coarse grids. We use either second order interpolation or second order difference scheme to calculate the boundary values. As before the integration of fine grid we advance one time step on coarse grid, it is possible to interpolate the quantities with high order interpolation. If necessary, u, ρ, ϵ and $a\omega$ are computed with either way. p is calculated with the equation of state. ω (then a from $a\omega$), R, m and Γ are computed with Eqs. (6.9), (6.2), (6.14) and (6.8), respectively.

When a new fine grid is created, its initial values are interpolated using the finest grids from the already existing grid structure. Nothing need be done to remove a grid that is no longer needed except reclaim its storage space. If there is no finer grid but a part of the new grid points are overlaid by a part of a old grid at the same fine level, the overlaying part is directly mapped to the new grid before the old fine grid is removed. For the part of a new grid that does not overlap with a grid at the same level or finer grids, its initial values are calculated from the values of its base coarser grid with piecewise cubic interpolations. The procedure of the computation of the quantities are here similar to that use in Sec. (6.5.4). The following quantities are interpolated: u, $a\omega$, ρ, ϵ, viscosity. ω (then a from $a\omega$), R, m and Γ are consistently computed with Eqs. (6.9), (6.10), (6.7) and (6.8), respectively. It should be noted that u is defined at half time step and we have to be careful when we do the interpolation. The initialization of grids at time $t = 0$ employs a slightly different strategy. Only at this time we can initialize the grids with the initial conditions rather than interpolation.

Whenever a fine grid is nested in a coarse grid and they have been integrated to the same point in time, the coarse grid values are updated. As u is defined at half time step, u is always updated before the fine grid advances to the time of the coarse grid. The criterion used for updating the various quantities on the coarse grid is: (i) quantities defined at the grid points (such as R, m, u) are updated by injecting the fine grid solution values onto the coarse grid points. If grid points do not match up at regular intervals, interpolation is used to find the values from the fine grid, which replace the coarse grid values. (ii) quantities defined in the mid-zones are then specified consistently with the constraint equations (6.6), (6.7), (6.8), (6.9), (6.10) and (6.15)

6.6.4 Truncation Error Estimation

After every several time steps, the error at all grid points are calculated and the grid structure is adjusted according to it. Unfortunately, the solution error is very difficult, if not impossible, to determine during an integration. Examination of the relationship between solution error and truncation error show that truncation error is one of the main sources of solution error (Skamarock 1989). Therefore, local truncation error is used to decide where to refine the grid. The method described here is problem-independent.

The local error compute is based on Richardson extrapolation (1910). Given a set of hyperbolic partial differential equations

$$u_t = \mathcal{L}[u] \quad , \tag{6.73}$$

where $\mathcal{L}$ is a spatial differential operator and u may be a vector. Let Q be a two-level explicit difference operator and if the solution is smooth enough, the local truncation error is given by

$$\begin{aligned} \tau &= u(x,t+k) - Q_{h,k}[u(x,t)] \\ &= k[k^{q_1}a(x,t) + h^{q_2}b(x,t)] + kO(k^{q_1+1} + h^{q_2+1}) \\ &= \tau_L + kO(k^{q_1+1} + h^{q_2+1}) \quad , \end{aligned} \tag{6.74}$$

where the leading term is denoted by τ_L, q_1 and q_2 are the orders of accuracy in time and space, respectively, and $a(x,t)$ and $b(x,t)$ are usually higher-order derivatives of u. If u is smooth enough, then if we integrate three time steps on a grid of step sizes h and k, and assuming that the order of accuracy in time and space are equal with $q = q_1 = q_2$

$$u(x,t+3k) - Q^3_{h,k}[u(x,t)] = 3\tau_L + kO(k^{q+1} + h^{q+1}) \quad . \tag{6.75}$$

Next, a single time step of $3k$ is taken on a grid with size $3h$. Then

$$\begin{aligned} \tau &= u(x,t+3k) - Q_{3h,3k}[u(x,t)] \\ &= (3k)[(3k)^q a(x,t) + (3h)^q b(x,t)] + kO(k^{q+1} + h^{q+1}) \\ &= 3^{q+1}\tau_L + kO(k^{q+1} + h^{q+1}) \quad . \end{aligned} \tag{6.76}$$

From Eqs. (6.75) and (6.76), we obtain the equation estimating the local truncation error

$$\tau_L = \frac{Q^3_{h,k}[u(x,t)] - Q_{3h,3k}[u(x,t)]}{3^{q+1} - 3} + kO(k^{q+1} + h^{q+1}) \quad . \tag{6.77}$$

The technique requires taking three time steps on the base grid and a single giant time step on a grid with widths of $3h$ and $3k$. Subtraction and appropriate normalization

yield the truncation error calculate. This is illustrated in Fig. (6.18) for two step scheme, which has $q = 2$.

The advantage to use Eq. (6.77) is that (1) the exact form of the truncation error need not be known and (2) the exact same method of integration can be used for the error estimation. As in our case, a small time step is needed to satisfy the controlling condition, the truncation error resulting from the time discretization are small and the spatial discretization is responsible for most of the error. In this case, one may first keep k constant and take a step based on $3h$ differences and use Eq. (6.77) to compute the error.

Since there are several variables, we may use only the velocity error calculates, as the motion equation is the most complicated and velocity may have the largest calculation error and also it is easiest to be computed. One possible scheme for using error calculates from all the dependent variables would be to scale and non-dimensionalize the error computes and add them together.

6.6.5 Clustering and Grid Generation

In the regridding algorithm, we first calculate the local truncation error at all grid points in level l, using the way describing in Sec. (6.6.4). We flag those points where the error e exceeds a tolerance ϵ and all those grid points in level l which are interior to grids at level $l + 2$, even if $e < \epsilon$ based on the level l grid. It is possible that finer grid has larger error, as in finer grid small scale structure may appear. Since each flagged point at level l will be in a level $l + 1$ grid, proper level-nesting is assured. In this way, it automatically makes sure that when two sub-grid move to each other and close enough, they are merged and form one single grid.

When the flagging procedure has finished for all grids at level l, we start separating the flagged points into distinct cluster. In one dimension, the leftmost and rightmost flagged points of the coarser grid form the left and right boundary of the new sub-grid. The cluster in this case would consist of all the flagged points between and including the leftmost and rightmost flagged points. Possibly, if a long enough gap of un-flagged points is found, two or more separate sub-grids may be formed instead. After a grid is created, it will be enlarged to include a safety zone around all its flagged points. This buffer zone determines how often grids must be examined verse how large they are. We use a buffer of width two coarse grid points. Flagged points which are closer together than twice the size of the buffer zone are put in the same grid refinement.

After the grids are clustered, a special storage is allocated to each grid and the initialization procedure is used to initialize all the new grids. If the current grid structure G has l levels, we start regridding procedure from the finest level l_h, if $l_h < l_{max}$, the maximum permitted level. Based on the error computes of level l, we might create new

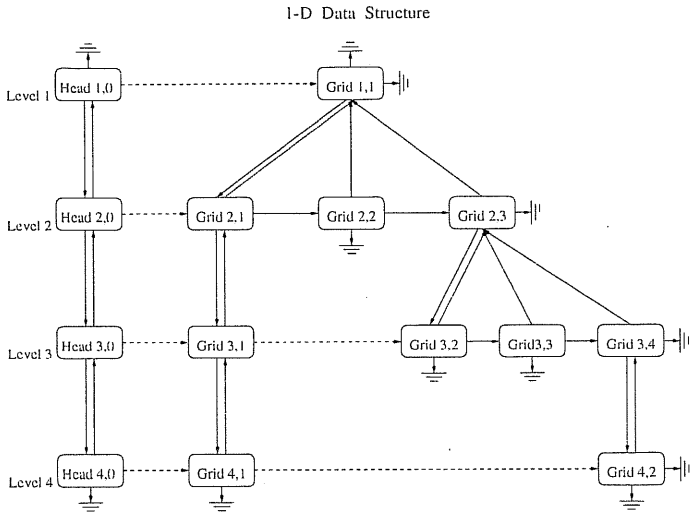

Figure 6.19: A four-level tree structure of grids. Solid lines indicate the links between parents and children and between brothers. Dotted line indicate cousin or neighbors. The arrow indicates the pointer direction in the nodes. The terminal arrow pointers to a special node "-1".

grids at level $l+1$. Next, based on calculates from the level $l-1$ grids, we recreate a new level l, making sure it includes the new level $l+1$. Continuing, the error computes on level $l-2$ are used to generate level $l-1$, making sure level l is properly nested, and so on to the coarsest level.

6.6.6 Data Structures

To keep track of the relationships between the grids as well as the solution storage for each grid, a structure is needed. From the nesting requirement for one dimensional mesh refinement that each fine grid must be entirely contained in a coarser grid at the next level, we define a tree data structure, where each node represents a grid, and make a correspondence between a parent node and a coarser parent grid (See Fig. (6.19) for the illustration). Sub-grids are considered the sons of their parent grid. Brothers are subgrids within the same coarser grid. If fine grids are at the same level of refinement with different parents, we call them neighbors.

We define two arrays: one for the record of the information of the grid and the other for the storage of the solution. The information of a grid which will be stored in each node of the tree is (1) grid location in space coordinates, (2) number of grid points, (3) level in tree, (4) offspring pointer, (5) sibling pointer, (6) parent pointer, (7) pointer to the next grid at the same level, (8) time to which this grid has been integrated, (9) index into main storage array where solution values are stored. The block we define with Fortran 77 is shown in Fig. (6.20).

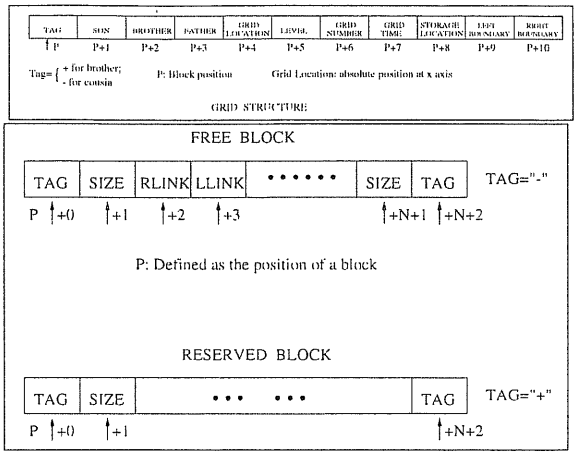

Figure 6.20: The definition of a block in an array, (*upper*) for a grid block which is use to keep grids information and the pointers, (*lower*) for the storage of solution. P is the position of first element of a block. When a block is reserved, $TAG = +$, while if it is available, $TAG = -$. SIZE records the total length of the block (or the total number of elements in a block). RLINK is the right pointer to the position of next free block in the linked list and LLINK is the left pointer to the position of the free block before it in the linked list. The number after P is the relative position of an element in the block.

Because the tree can grow or shrink dynamically, some form of dynamic storage allocation is needed, both for grid information in each node and the solution storage for each grid. Since Fortran 77 does not provide such a facility, the storage management is explicitly provided by the mesh refinement program. A linked list of free nodes (Knuth 1973) is kept and nodes are assigned to newly create grids and reclaimed when a grid is removed.

The storage area for the solution values on all grids is managed as a linked list of used and available blocks of storage. An array is divided into block and linked together. The definition of a block is shown in Fig. (6.20).

When a grid is created, storage space is reserved from the sorted list of free blocks using a first-fit algorithm (Knuth 1973). In this algorithm, the list of free blocks of storage is scanned from a random position until a large enough block is found. The requested space is allocated and the unused portion (larger than some amount) is returned to the list. When a grid no longer exists, its storage space is inserted back into the linked list of free space. When no large enough block can be found for a newly created grid, we restart our program with a larger memory area.

7 Gamma Ray Bursts and Active Galactic Nuclei: Jets and Accretion Discs

7.1 Introduction

In last Chapter, we discussed our work on a fireball model with spherical symmetry. Such a model has successfully explained the observations of afterglows in some GRBs. Present popular models on the central engine of GRBs suggest that a fireball may be produced along the axis of rotation of a black hole surrounded by a torus or thick disk. Such a fireball may move mildly-relativistically. Indeed, beaming effects are observed in some GRBs. There are many theoretical discussions in the literatures on jet model in GRBs. The geometry and the physical process of the central sources are similar to those in AGNs (Dermer & Chiang 1999), although AGNs involve continuous accretion processes and GRBs originate from one-time catastrophic accretion events. Additionally, it is still possible that GRBs are created inside a pre-existing jet, like the one in SS433, from a neutron star which collapses to a black hole due to massive accretion (Pugliese *et al.* 1999). The jets in Active Galactic Nuclei (AGNs) can be seen as extra-galactic analogues of the jets in SS433.

In this chapter, we will report our work on the jets and the accretion disks in AGNs. In Sec. (7.2), we will study the statistics of radio jets in AGNs, basing on a list of the detected radio jets in AGNs (Liu 1999; Liu & Xie 1992; Liu & Zhang 1999). As the catalogue is about several tens of pages and in order not to make the thesis too long, we will not put the catalogue in the thesis. Interested readers are referred the references.

A long-lived jet is believed to come from a jet-disk symbiosis system. If the jet in an AGN points very closely toward us, the object will be classified as a Blazar. Blazars have special characters different from other AGNs. The spectra are dominated by non-thermal radiation and are in general believed to originate from synchrotron-self Compton (SSC) radiation in blast wave, in which sense Blazars can be physically compared with GRBs. The jet-disk system in both Blazars and GRBs is essential. In

Blazars, beaming effects are dominated. It is easier to study the jet effect than the effect of disk.

The short term variabilities observed in GRBs and Blazars may be due to the activities inside jets, while the long time scale variabilities may be partly due to the activity of the disk. It was generally suggested that Blandford & Znajek mechanism extracts the rotational energy of a black hole and power the jet. However, this mechanism has been criticized. The effect of the accretion disk on the production and activities of jets and fireballs may be more important than what we expected before. The effects of the accretion disk may be detected not only in the observations of spectrum but also of the variabilities.

An important difference between the accreting disks in Blazars and GRBs concerns the accretion rates, which in GRBs are of the order of about one solar mass per second while in Blazars are about one solar mass per year. In standard accretion disk models on AGNs, accretion rate is low and the time scale of the dwarf nova type limit cycle instability is too long to be observed directly for a central black hole of mass of $10^6 M_\odot$ or higher (Lin & Shields 1986; Meyer-Hofmeister 1993). The observed AGNs may represent the outburst state of the disk, although some or all quasars could constitute a steady population having markedly higher fueling rates than other AGN (Burderi *et al.* 1998). Accretion disks around super-massive black holes in Active Galactic Nuclei may be treated as slim disks (Abramowicz *et al.* 1988; Honma *et al.* 1991; Szuszkiewicz & Miller 1997, 1998). For a slim disk with a low viscosity parameter $\alpha \sim 0.001$ limit-cycle behavior is not expected (Szuszkiewicz & Miller 1997), while a slim disk with a high $\alpha \sim 0.1$ may have limit-cycle behavior connected with thermal instability (Honma *et al.* 1991; Szuszkiewicz & Miller 1998). The time scale of the thermal limit cycle in such a slim accretion disk around a central mass of $10^6 M_\odot$ to $10^{10} M_\odot$ varies from a few years to several hundred thousand years.

The activities of the central part of an accretion disk in normal AGNs may not be observable due to its weakness comparing to the radiation from other parts in the AGNs, while the activities may be amplified by jets which are believed to originate from the central part of the accretion disks surrounding a black hole. The activities of the central part of the disks may raise the jets to an active phase. The rapid variabilities are from the jets, while the long term variabilities may be the signals of the activity in disks.

BL Lac objects are the special class of AGNs whose jets are directed almost toward us, characterized by high optical linear polarization, no or weak spectral lines and large amplitude variability in all of the electromagnetic spectrum. They are one of the main monitoring targets for a long time. If the thermal limit cycle behavior of accretion disks exists, periodic instability may be observed in some QSOs and BL Lac objects with available observational data for around 100 years.

We will discuss the periodicity in Blazars in Sec. (7.3), which is probably the signature of slim accretion disks with high α in Blazars (Liu *et al.* 1995, 1997; Liu 1996).

7.2 The Statistics of Extragalactic Radio jets

7.2.1 Description of the New Catalogue

Here we present a description of the new catalogue of jets that has been prepared (Liu 1999; Liu & Xie 1992; Liu & Zhang 1999).

With the use of many new instruments with high sensitivity and large dynamic range during the past years, the structure and dynamics of various types of extragalactic radio sources have observationally been investigated in very great detail (e.g. Parma *et al.* 1991, Leahy *et al.* 1997 and references therein; Bridle *et al.* 1994; Garrington *et al.* 1991). Great improvement has been made on the knowledge of galaxy structures and the understanding of the physics. Many more jets in radio, IR, optical, UV, and X-ray wave-bands have been detected, since the publication of the last list of jets. A recent review of the progress in the study of pc scale jets was given by Zensus (1997).

Whether a jet can be detected depends not only on the properties of the radio source itself but also on the instrument used. With presently available instruments, radio jets have been detected in various types of radio sources: FR I and FR II radio galaxies, Seyfert galaxies, radio quasars, and BL Lac objects, though the detected rates of jets are different in different categories of radio sources with the same instruments under similar working conditions. The different detected rates of jets can be reconciled with the unification model of AGNs (Bridle & Perley 1984; Antonucci 1993).

Observationally to distinguish a jet from other narrow features of a galaxy, we adopt the morphological definition of a jet, which was first given by Bridle (c.f. Bridle *et al.* 1994), that a "jet" is a narrow feature that must be

- (a) at least four times as long as it is wide,
- (b) separable at high resolution from other extended structures (if any) either by brightness contrast or spatially, and
- (c) aligned with the radio nucleus of the parent object where it is closest to it.

There were 542 sources of jets satisfying the definition known prior to October in 1997 (Liu & Zhang 1999). In addition to the IAU name of a source and references to observations in radio, IR, optical, UV, X-ray and γ-ray wave-bands, the information of identification, jet direction and morphology, redshift if available, source, core and jet luminosities are also given. The ratio of luminosities between the brighter and fainter jets is used to define the sidedness of jets. We say that a source is one-sided or two-sided, if the ratio is always larger than 4 or always less than 4, respectively. Otherwise,

Table 7.1: General statistical result of jets. $H_0 = 100Km \cdot s^{-1}Mpc^{-1}$ and $q_0 = 0.5$

height 1pt		Z		$\lg P_j^{1.4}$		$\lg P_t^{1.4}$		$\lg P_c^5$	
Type	N	Min	Max	Min	Max	Min	Max	Min	Max
QSOs	193	0.104	3.886	24.74	27.94	25.44	28.76	23.96	28.48
G	260	0.0012	3.395	20.99	27.84	21.68	27.94	20.38	26.87
S	40	0.0033	0.518	21.75	25.42	21.71	27.73	20.47	25.54
BL	25	0.0177	1.404	23.00	24.84	23.68	28.15	23.41	28.20
U	24	—	—	—	—	—	—	—	—

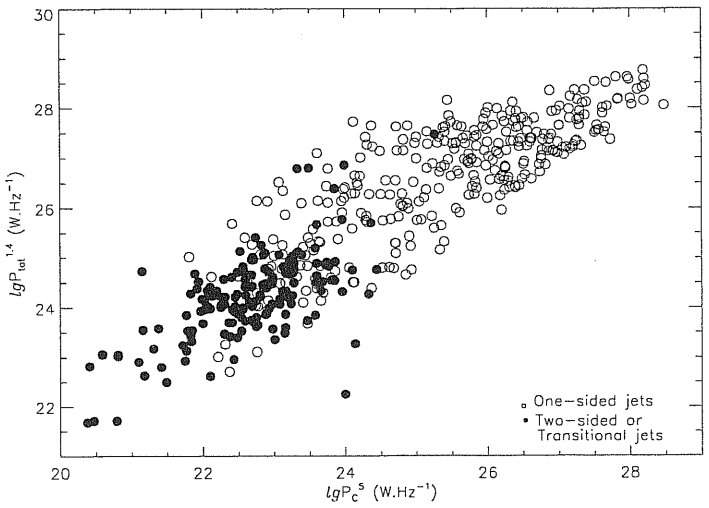

Figure 7.1: The change of the sidedness of jets with source and core powers.

we say that it is intermediate. We mark those sources in which superluminal motion is observed.

The general statistical result of the catalogue is given in Table (7.1). In column one is the source type. Column two gives the number of sources. The rest of the columns present the minimum and maximum of redshift, jet, total and core luminosities, respectively. $LgP_j^{1.4}$, $\lg P_t^{1.4}$, and $\lg P_c^5$ give the power in units of $W \cdot Hz^{-1}$ of the jet at 1.4 GHz, the source at 1.4 GHz and the core at 5 GHz, respectively. We try to use integrated flux densities for jets with a correction to background. "Typical" values are used for variable total and core emission. Some $S_j^{1.4}$, $S_t^{1.4}$, and S_c^5 values are estimated from neighboring frequencies, assuming power-law spectra with spectral index of $\nu^{-0.65}$ for the jets, $\nu^{-0.7}$ for the total emission and ν^0 for the cores if no spectral information is available, extrapolating or interpolating.

Fig. (7.1) shows the change of sidedness of jets with total and core luminosities. All BL Lac objects and QSOs are one-sided. High luminosity radio galaxies are also

one-sided.

7.2.2 Discussion

Among the 542 sources, 193 are QSOs, 260 radio galaxies, 40 Seyfert galaxies and 25 BL Lac objects. 24 sources have no identifications. We do not try to distinguish type-one and type-two Seyferts. Jets are detected in the sources with redshift from $Z_{min} = 0.0012$ to $Z_{max} = 3.886$. The closest radio source with detected jets is radio galaxy Cen A. The minimum redshift $Z_{min} = 0.104$ for QSOs is very close to the definition redshift limit 0.1 of QSOs. Owing to the importance of high redshift radio galaxies in the study of formation and evolution of galaxies, many very powerful instruments have been dedicated to the observations and consequently brought about the detection of a large number of jets in galaxies with very high redshift. Parsec scale jets have been detected in many Compact Flat-spectrum (CFS) and CSS sources now. 69 sources in this catalogue have only jets of pc scale length.

We classify radio sources as radio galaxies, QSRs, BL Lac objects, and Seyfert galaxies, based on the optical properties of the sources. Detection of jets in these four types of source are common. The sources can be classified, based on morphology in radio waveband, from extended to compact sizes as FR I and FR II, triples, Doubles, CFSs and CSSs. Jets are detected in all types of source with different detection rates. Large scale jets are detected in a large fraction of weak radio galaxies and extended QSRs, although the fraction becomes lower in the lower range of power of core radio emission (Parma *et al.* 1991). This may result from the selection effect as a consequence of the correlation between opening angle and radio power (Bridle 1986; Parma *et al.* 1991). Jets were detected only in a small fraction of distant radio galaxies with similar radio power to QSRs with instruments operating under similar parameters. However, the detection rate has significantly increased with instruments with higher power, sensitivity and resolution (Carilli *et al.* 1997). Only pc scale jets in CSSs are detected.

7.3 The Possible Signatures of Slim Accretion Disks in Blazars

7.3.1 The Observational Data and Variability Analyses

Here we present a description of our work on AGN variability which has been publishe in several papers (Liu *et al.* 1995, 1997; Liu 1996).

7.3.1.1 ON231

We have compiled all available B-band optical observation data on ON231 (W COM) from Wolf (1916), Kurochkin (1971), Romano (1972), Craine and Warner (1973), Wing

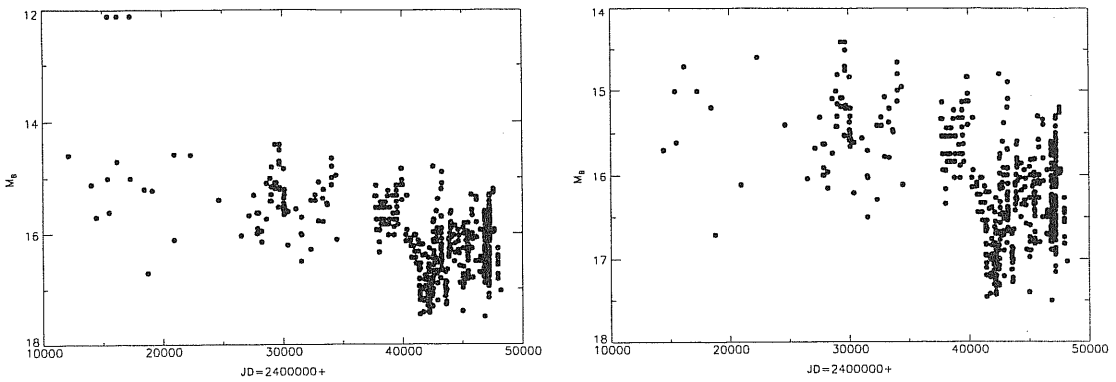

Figure 7.2: The historical light curve of ON231 from 1892 to 1992, collected from the references listed in the text. (*left*) the whole data and (*right*) after excluding Wolf's data.

(1973), Strittmatter *et al.* (1972), Pollock *et al.* (1974), Battistini and Braccesi (1974), McGimsey *et al.* (1975), Véron and Véron (1975), Tapia *et al.* (1976), Owen *et al.* (1978), Pollock *et al.* (1979), Schaefer (1980), Barbieri and Romano (1981), Zekl *et al.* (1981), Guz-Gonzalez and Huchra (1984), Moles *et al.* (1985), Weistrop *et al.* (1985), Xie *et al.* (1990, 1991, 1992, 1987, 1988), Webb *et al.* (1988), Barbieri *et al.* (1988). The B-band data for ON231 consist of 560 observations, dating back to 1892. Because of our purpose to study the large-amplitude variations, we show no error bars for individual observations. The errors are small compared with the variations being studied and including error bars in the figure would be confusing. We consider no observations published with large observational date uncertainty (about a month), for example, those given by O'Dell *et al.* (1978), and Zekl *et al.* (1981), as being inconsistent with purposes of searching for short (about a week) and medium (several months) periods. The reason that we use the B-band observations rather than those of V or other bands for ON231 is that much more published observations in B band exist, especially in the early century. We translate the photographic magnitude m_{pg} for a few observations into standard blue magnitude m_B by the approximate relation m_B=m_{pg}+0.11. The long term light curve is shown in Fig. (7.2).

From Fig. (7.2), it can be seen that ON231 is a very active object. The source may be as bright as m_{pg} = 11.5 on the plate of Wolf (1916) and as faint as m_{pg}=17.5 on the plate of Zekl *et al.* (1981), and this gives a variation range of 6 magnitudes. However, Wolf's observations are quite uncertain. On the other hand, if the entry in the General Catalog of Variable stars (Kukarkin *et al.* 1969; Birand 1971) is to be believed, ON231 is only of m_{pg}=13 at maximum, which gives a total amplitude of about 4.5 magnitudes.

Owing to the too-large uncertainties, we do not include Wolf's observations in Fig. (7.2) (*right*) and in the following discussions and analyses, unless we note it particularly. Before 1972, when the source reached its minimum, neglecting the oldest (1898-1916) quite uncertain observations (Wolf 1916), we can see that two major out-

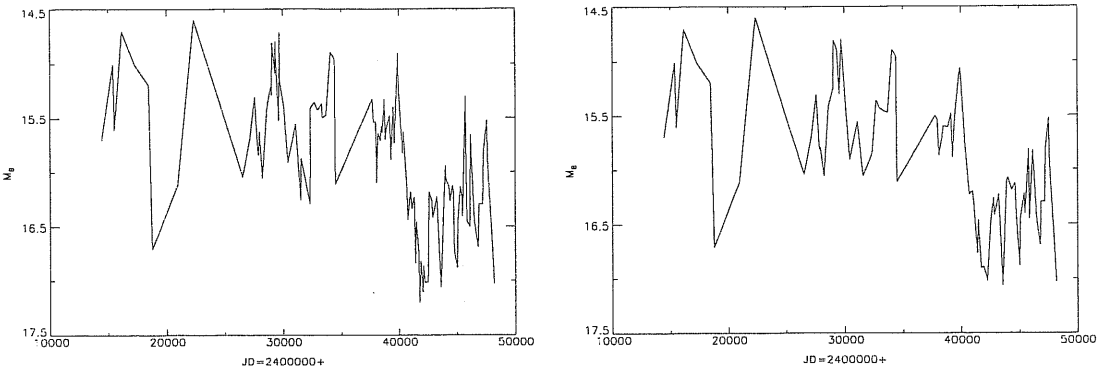

Figure 7.3: The mean light curve for ON231 in the form of 100-day means (*left*) and 200-day means (*right*).

bursts were completely observed, one in 1940, the other in 1968 (Pollock *et al.* 1974). These two outbursts have similar structures with one magnitude pulse of half-width about 3 years. Another possibly similar active period may exist in 1951 – 1953 but was not completely covered by observations. In addition to these, a number of other outbursts in the long-term light curves also exist and can clearly be seen in the light curve averaged over 100 and 200 days (Fig. (7.3)). The light curve in Fig. (7.3) has a straight line in the region of JD=2435000 to JD=2437700 because of the lack of observations between 1953 and 1961. The observations are very few before 1930, but the curve still shows violent activity of the source even excluding the variation given by Wolf(1916). Since 1962 (JD=2437700), many more observations have been made for ON231, so the complexity of the behavior of its intensities can clearly be seen in Fig. (7.3). After peaking at about JD=2440000 of the 1968 outburst, the light curve decreases very sharply (about two magnitudes in amplitude), going through a small amplitude outburst superimposed on it. Xie *et al.* (1991) observations show three similar outbursts, one on 1987 April 19, the second on 1988 March 13, and the third on 1989 April 9. This may imply a quasi-periodic variation in ON231.

To show the long-term tendency clearly, we average the light curve over 6000 days. From the result and previous discussions, we can conclude that the light curve of ON231 consists of a slow secular decrease of the intensity and rapid outbursts overlapping on it. Fig. (7.3) also shows that the variability structure of ON231 is not sinusoidal and may consist of, if any, many periodic outbursts with different periods. Therefore, Fourier analysis (e.g. Deeming analysis) is not very suitable for it. That Barbieri *et al.* (1988) fail to find any clear periodicity using this method is not surprising.

In Sec. (7.3.2), we use another powerful method, Jurkevich's V_m^2 to analyze the data of ON231, because this method is insensitive to the complicated shape of the light curve and long term tendency of intensity, and is capable of finding multiple periods.

7.3.1.2 Mkn421

The X-ray selected BL Lac object Mkn421, at $z = 0.0308$ has attracted much attention since it was identified as a BL Lac object. It is one of the objects simultaneously observed at all electro-magnetic frequencies and one of the few objects radiating strong gamma-rays. Many observations to search for its variability in the optical wave-band have been performed.

All observational data in B band on Mkn421 are available in the following studies: Miller (1975), Ulrich *et al.* (1975), Véron & Véron (1975, 1976), Miller *et al.* (1977), O'Dell *et al.* (1978), Mufson *et al.* (1980), Puschell & Stein (1980), Zekl *et al.* (1981), Gagen-Torn *et al.* (1983), Sitko *et al.* (1983), Cruz-Gonzalez & Huchra (1984), Moles *et al.* (1985), Sitko *et al.* (1985), Makino *et al.* (1987), Xie *et al.* (1987, 1988), Sillanpää *et al.* (1988, 1991), Sitko & Sitko (1991), Takalo (1991), and Takalo *et al.* (1992). Data for Mkn421 consists of 565 observations, dating back to 1900. Since we are searching for the long time variability, we include those data estimated from Miller's figures with relatively large date uncertainties (less than one month). The B-band observations are used here because there are more data available in B-band than in other bands. We translate the photographic magnitude m_{pg} by the approximate relation $m_B = m_{pg}+0.11$ and the flux density, f_B, by $m_B = [\log(4.49\times10^6)-\log f_B(mJy)]/0.4$ (Sitko *et al.* 1985).

The error caused by the conversion from photographic to photo-electric values is not larger than 0.2 magnitudes. The object does not produce a stellar image in deep photographic exposures, so photometric data obtained with different entrance sizes are different. About half of our data are taken from Miller's paper (Miller 1975), where the uncertainty of the iris photometry measures are typically 0.1–0.2 magnitudes. A large fraction of the remaining data are obtained with a diaphragm of about 26″, within which the contribution of the host galaxy is less than 0.2 magnitude. Thus the difference between magnitudes derived through different entrance diaphragm sizes is less than 0.2 magnitude. Therefore, the photometric and photo-electric data are consistent within 0.2 magnitude, a very small value compared with the large range of variation of the object, $\Delta B \geq 4.7$ mag. The magnitude uncertainties introduce noise and introduce an uncertainty on the parameters of the temporal features possibly detected on the Jurkevich plot (see Sec. (7.3.2.3).

The long-term light curve is shown in Fig. (7.4). Because of our purpose to investigate large-amplitude variations, we do not show individual error-bars, as explained previously. The effect of errors on the periodicity analysis will be discussed in Sect. (7.3.2.3). Mkn421 is very active, with a range of variation of $\Delta B \geq 4.7$ mag. The source reached a maximal brightness of 11.6 mag in 1934 January and was brighter than 12.5 mag on three occasions in 1901, 1916 and 1936 (Miller 1975). After reaching a maximum $B = 12.75$ in 1982 April, Mkn421 faded out until 1986 . There are fewer

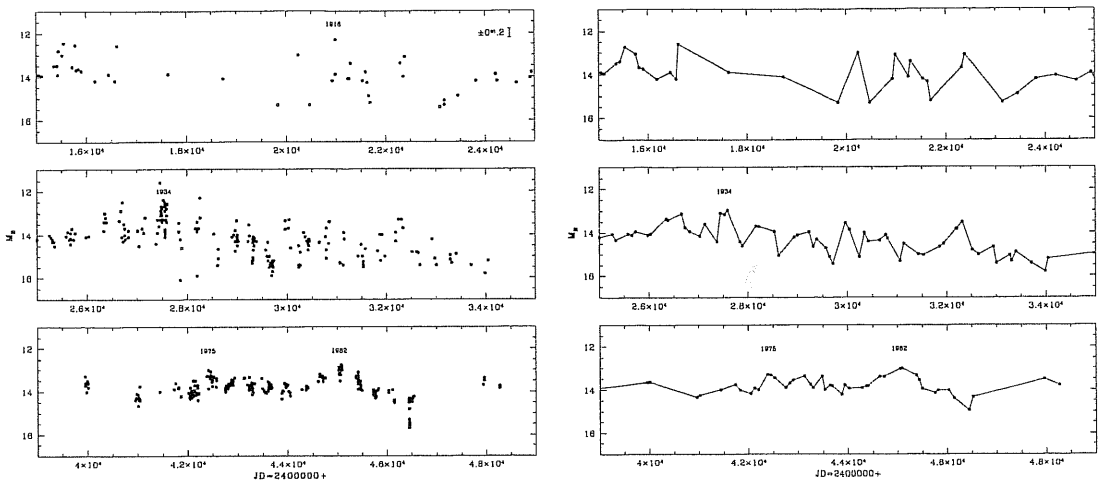

Figure 7.4: (*left*) The long-term light curve of Mkn421 from 1900 to 1991. The discontinuity of the light curve between 2435000 and 2439000 is due to lack of observations. (*right*) The mean light curve of data in the left figure over a 100-day mean.

observations available for Mkn421 after 1986, in B band. The observations, however, still show that the source brightened again (Takolo 1991; Takolo *et al.* 1992).

To reduce small amplitude intra-day fluctuations , we averaged the light curve over 1 day. No significant difference has been found. In order to probe the long-term behavior of the variations, we averaged the light curve over 100-days (Fig. (7.4) *right*). Because of the different quality of our data at different epochs, the impact of flickering in recent data is washed out, while still being significant in the early data. As the object varies in intensity by about 0.5 mag on a time scale of several hours (Xie *et al.* 1988), the largest difference between early epoch data point and the mean value can be estimated only within an uncertainty of 0.5 mag. We have averaged the light curve. The resulting light curves are similar. The peaks in 1934, 1975 and 1982 remain unchanged. The difference between the left panel and right panel in Figs. (7.4) is quite significant. It indicates that Mkn421 has large intensity variations on a time scale of a few months. The variability of Mkn421 shows two modes: a short-term one with a time-scale of a few months to several years and a longer-term one with a time scale of the order of ten years. We will analyze the repetition of the bursts in the light curve using the Jurkevich method (1971) in Sec. (7.3.2.3).

7.3.2 The Periodicity in BL Lac Objects

7.3.2.1 The Jurkevich Method

The outbursts in BL Lac objects are very complicated. At least two types of variability, comprising a superposition of short-term and long-term variations, may be simultaneously present. The one on shorter time scales usually with large amplitude of variations

is probably non-periodic, whereas some indications of periodic or quasi-periodic fluctuations on longer time scale exist. The best candidate for true periodicity (Trimble & McFadden 1998) is the BL Lac object OJ287, whose 1994-1995 December outbursts occurred as expected from the 12 yr period (Sillanpää *et al.* 1996). Another good candidate is ON231 of $\sim$ 13 yr periodicity (Liu *et al.* 1995). An outburst was observed in 1994-1997 (Tosti *et al.* 1998) as predicated with this periodicity.

To establish the reality of periods, two prerequisites must be satisfied. First, the data sample must be at least six times the length of the periodicity being claimed. Second, the possible period should be of large amplitude and thus directly visible in the light curve. Where the amplitude is low, many more than six cycles of the period will be necessary to demonstrate its reality. For a period order of ten years, which we are interested in, only a few objects are capable of satisfying both conditions.

The variation of brightness of OJ287 shows that the periodic oscillation amplitude of the flux is much smaller than that of flux due to sporadic variations. It is difficult to find a suitable technique to treat such complicated data and reliably determine the possible periodicity. Because of the existence of sporadic variation with large amplitude, the traditional fourier analysis will be dominated by this "noise". The method is also very sensitive to the mean shape of the periodicity structures and uneven sampling. So it is not a suitable method for period search for a complicated, non-sinusoidal outburst structure such as that in many BL Lac objects. Here, we use an alternative and powerful method, the Jurkevich V_m^2 test which has no such weakness. It is also less inclined to generate spurious periodicities than Fourier analysis.

The Jurkevich method is based on the expected mean square deviation. It tests a run of trial periods around which the data are folded. All data are assigned to m groups according to their phases around each trial period and the variance V_i^2 for each group (bin), and the sum V_m^2 of all groups are computed. For a trial period equal to the true one, if any, V_m^2 reaches its minimum, and a "good" period will give a much reduced variance relative to those given by other false trial periods and with almost constant values. No firm rule exists for assessing the significance of a minimum in the V_m^2 plot.

After analyzing the definition of the statistical F-test, we have a good guide in the fractional reduction of the variance, as in Kidger *et al.* (1992) and Liu *et al.* (1995, 1997),

$$f = \frac{(1 - V_m^2)}{V_m^2} \quad , \tag{7.1}$$

where V_m^2 is the normalized value. In the normalized plot, a value of $V_m^2 = 1.0$ implies that f=0 and hence there is no periodicity. The best periods can be identified from the plot. In general, a value $f \geq 0.5$ generally implies that there is a very strong periodicity in the data, whilst $f < 0.25$ usually implies that the periodicity, if genuine,

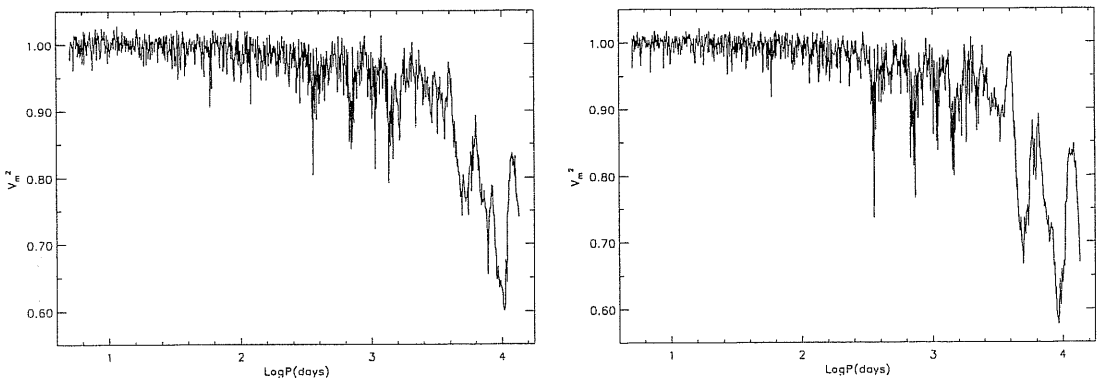

Figure 7.5: (*left*) The normalized Jurkevich test results for the period search for ON231, using the collected data but excluding Wolf's. The deep minimum corresponds to a period of 28.0 yr. (*right*) Similar to the left panel but for the 1-day means. The minima corresponding to periods of 28.0 yr, and 13.6 yr are very significant.

is a weak one. A further test is the relationship between the depth of the minimum and the noise in the "flat" section of the V_m^2 curve close to the adopted period. If the absolute value of the relative change of the minimum to the "flat" section is larger than ten times the standard error of this "flat" section, the periodicity in the data can also be considered as significant, and the minimum as highly reliable. In the Jurkevich test, the parameter m can be modified: more groups give higher sensitivity, but fewer data points per group create more noise in the plot. So we analyze the data sample mainly using $m = 10$, which gives us over 50 points per group for all the sample objects. For comparing the results, we also use $m = 20$ in some cases. During our analysis, we choose a very small interval between two successive trial periods, to search for short periods. The results of the run of values of V_m^2 is plotted against period in the same way that a fourier analysis plots power against frequency.

7.3.2.2 ON231

The result of the analysis with the Jurkevich method search with $m = 10$ is shown in the left panel of Fig. (7.5). A relatively broad minimum of $V_m^2 = 0.599$ in this V_m^2 plot clearly occurs at a trial period of around 28 years, corresponding to a value of $f = 0.669$. This period is consistent with the separation between the two well-observed outbursts in 1940 and 1968. A similar analysis with $m = 20$ shows that the V_m^2 plot with relative larger noise also has a minimum of $V_m^2 = 0.461$ at about 28 years. Within the feature, we clearly see the fluctuations around the minimum which may be caused by the large amplitude variability of intensity as short as several hours, which is definitely non-periodic. In order to reduce the noise in the V_m^2 plot and to search for moderately long periods, e.g. a period of one year, we make a analysis for the data

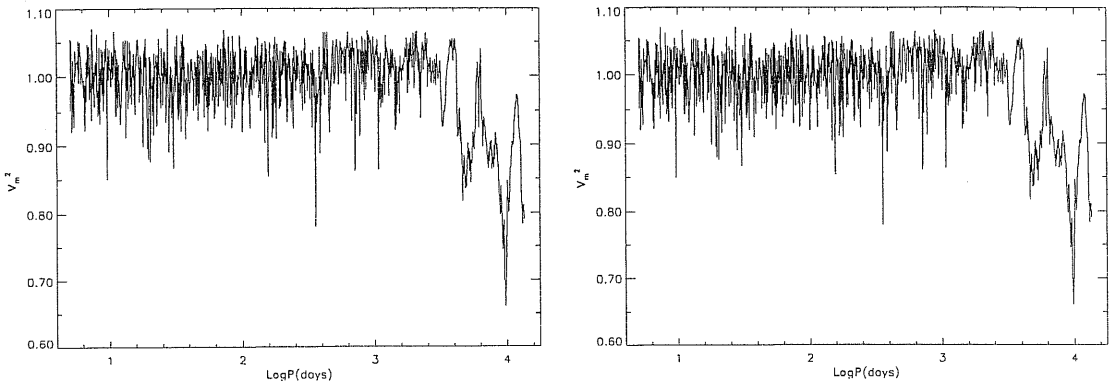

Figure 7.6: (*left*) The normalized Jurkevick test results for the period search for ON231, using the averaged data over 100-day mean. The deepest minimum corresponds to a period of 25.7 yr. (*right*) As left figure but for the 1-d means of the data including Wolf's (1918).

sample averaged in single days. The result is shown in the right panel of Fig. (7.5). The noise in this figure is indeed less than that in the left panel in the same figure. Although the width of the minimum is also broad, it is much less so than that in the left figure, and its depth is larger. The analysis with $m = 20$ shows the same effects.

In Fig. (7.5), there is also another minimum at a period of 13.6 ± 1.3 years. It is not the deepest one, but it is so significant with $V_m^2 = 0.668$ (in Fig. (7.5)) that we do not doubt its reality. This period is half of the first one, so we consider that they are related and the longer is just a harmonic of the shorter. We have calculated the half width at half minimum as a formal error (also see Kidger *et al.* 1992). The error is considered (c.f. Jurkevich 1971) to contain all effects on the precision, including random variations in the exact interval between outbursts, poor coverage of some of the early outbursts and large error in some of the early photographic photometry, the large and changing width of the outburst structure, the uncertainty of observed data estimated from the figures in literature (e.g. Pollock *et al.* 1974) for a few observations, and random variations in intensity.

Apart from the periods mentioned above, in our detailed V_m^2 plots (Fig. (7.5)) a series of very interesting and significant minima indicate the possible existence of other periods of $P1 = 0.981 \pm 0.005$ years (358 days), $P2 = 2.02$ years ($\sim 2P_1$), $P3 = 3.01$ years ($\sim 3P_1$), and so on. Of all the periods, the one of $P = 0.981$ years is the shortest and most significant and is the fundamental one. It has a very deep structure of V_m^2=0.736 in Fig. (7.5) ($f = 0.359$, larger than 0.25 and less than 0.5) which is 22 times larger than the noise. However, before concluding that it is real, we have to note that the astronomical year is one solar year and the discussion in Sec. (7.3.2.6) shows that the Jurkevich method would give a spurious period of one year, if there exists periodicity in the data.

To counteract the heavy weighting of recent data, we have calculated 100 day means

of the light curve. This interval is long enough to smooth the light curve considerably, but short enough compared to the possible 13.6 yr and $\sim$ 28 yr periods and thus unlikely to distort the long-term variations too considerably. The result is shown in Fig. (7.3). This light curve consists of 113 points. When it is tested with the Jurkevich method, the noise in the resultant plot (Fig. (7.6) *left*) is far increased, but a well defined minimum is seen in the V_m^2 statistic at a period of 26.7+0.8 years. The minima at periods of 0.981 years and 13.6 years are also clear.

All the results discussed above are based on the data sample excluding Wolf's observations, because of their large uncertainties in magnitude. However, the fact that Wolf's data are in the early region, where only a few observations exist, and shows possible but surprisingly large outbursts indicates that we should check its effects on the previous results. To do so, we have tested the light curve for the one-day means and shown the result in Fig. (7.6) (*right*). In this V_m^2 plot, the positions at which the minima occur shift slightly and the lowest minimum is at a period of 25.7 years. The significance of this minimum decreases somewhat. The possible one-year period has almost no change after considering Wolf's observations.

Many special efforts have been made to search for short (several days) periods. We find it very difficult to do this. In the studies, as long as the trial period deviates from exact one by as little as 1×10^{-4} days, the V_m^2 will change a lot. So, to test it, we have worked with many different intervals as small as 0.0005 days and m=20. Occasionally, we find a minimum at a period of 8.000 ± 0.001 days in the V_m^2 plot with a interval of trial periods of $\Delta(\lg(p)) = 0.01$ days and $m = 20$. This period has a significance of $V_m^2 = 0.771$ ($f = 0.297$) and the depth is about 12 times noise intensity. A further analysis shows that when we do the test from a trial period of 2 days to 100 days the V_m^2 plot reaches its minima at periods of $2, 3, 4, \cdots 10$ days. However, these minima have not so large significance that we can determine their reality.

7.3.2.3 Mkn421

The result of the analysis with $m = 10$ is shown in Fig. (7.7) (*upper*). A minimum of $V_m^2 = 0.653$ ($f = 0.532$) is significant at a trial period of 23.1 ± 1.1 years. A similar analysis with $m = 20$ shows that $V_m^2 = 0.615$ ($f = 0.627$) at the period of 23.5-year. In addition to the period of 23.1 ± 1.1 years, the broad minimum at $P = 15.5$ years is also significant with $V_m^2 = 0.701$ and $f = 0.427$ but not as certain as the one obtained at $P = 23.1$ years. We have considered the half width at half minimum as the "formal" error (c.f. Jurkevich 1971) to derive all effects on the precision, including random variations in the exact interval between outbursts, poor coverage of some of the early outbursts and the larger error in some of the early photographic photometry, the uncertainty of observed data estimated from the figures in the literature, random variations in

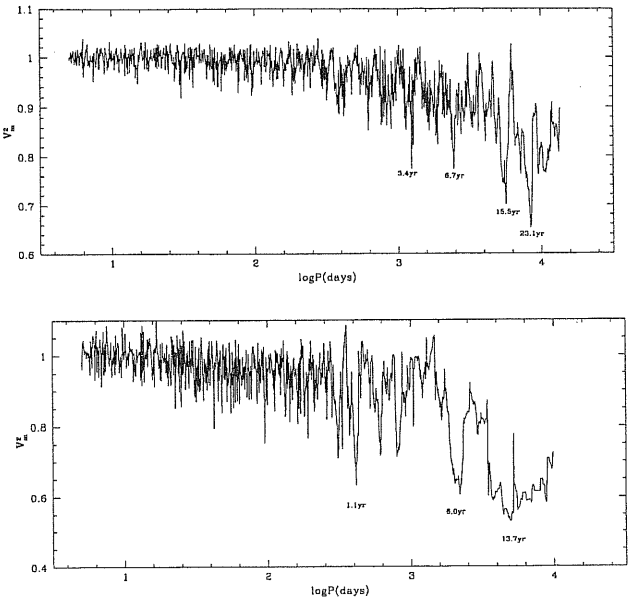

Figure 7.7: (*upper*) Results of the normalized Jurkevich test for the period search, in Mkn421. The deepest minimum corresponds to a period of 23.1 year. The minima corresponding to periods of 3.4 years, 6.7 years, and 15.5 years are also conspicuous. (*lower*) Results as for the left panel but excluding Miller's data and with m=5. The minima corresponding to periods of 1.1, 6.0, and 13.7 years are significant.

intensity, and the changing width of the outburst structure. The errors caused by the conversion from photographic to photo-electric values and by the measurement with different diaphragms are considered in this analysis as random variations in the intensity. They would reduce the depth of minima and therefore the significance of the obtained periodicity. These errors also increase the "formal" error, and this effect has been taken into account. The fluctuations seen around the minimum may also be caused by flickering, which is definitely non-periodic. The broad width of the minima may also result from the broad structures of the bursts, the drift of the real period and the effect of adjacent periods, if present.

To compensate for the heavy weighting of recent data, we use the 100-day averaged light curve. This interval is short enough compared to the possible periods of 15.3 years and 22.8 years and unlikely to prevent finding the long term periodicity. The result of the Jurkevich test shows a larger noise due to fewer points in every group and flickering effects in the early epochs. However, a minimum of V_m^2 near the possible period of 22.8 years is still seen.

Considering the redshift of 0.0308, the period of 23.1 years (P_{obs}) corresponds to 22.4 years in the rest frame of the source (P_{res}).

The Jurkevich analyses of the observational data with Miller's and without Miller's data (for which the errors are rather large) for Mkn421 provide similar results (See

Sec. (7.3.2.6)). These results are independent of the parameter m. Our analysis shows that probably two periods exist in the light curve of the BL Lac object Mkn421: one of around 15 years and another of around 23 years. However, the current data set covers only four times the possible period of 23 years, so more data is needed to confirm this 23 years period.

7.3.2.4 OJ287

The periodicity of this object has been extensively studied in the literature. To systematically study the periodicity of BL Lac objects, we also list this object here. The period of 11.6 ± 0.5 yr found in the historical data of OJ287 for more than 100 hundred years has been confirmed (Sillanpää *et al.* 1996; Kidger et al 1992 and references therein). The light curve may be maintained at high level as long as 4 years, or more, with underlying sinusoidal variation of period of 11.6 years.

7.3.2.5 Other BL Lac Objects

No period of less than five years has been found in AGNs (Kidger *et al.* 1992). Smith and Nair (1995) analyze the optical light curves of 36 BL Lac objects, based on the 25-year data base of the Rosemary Hill Observatory. They make a conclusion that about 28 % of the sample show well-defined, easily measured cyclic changes in base-level of the light curves. The averaged cyclic time scale is about 7.2 years or 5.1 years on the rest frames. However, we have to note that the data base is very short and is only one or two times the claimed periods. It will be necessary to collect much more data to confirm their results.

7.3.2.6 The Robustness of the Jurkevich Method

In order to test the robustness of the Jurkevich method and to investigate intermediate time-scale periods, we exclude Miller's data for Mkn421 and use the observational data from 1972 to 1991 . During this period, Mkn421 was more extensively monitored and thus has sufficient data for our analysis to give reliable results. The result of the analysis, with $m = 5$, is shown in Fig. (7.7) (*lower*).

Although the time interval considered in this case is less than that covered in the *upper* panel of Fig. (7.7) (only 19 years), a minimum at $P = 13.7 \pm 2.0$ years is quite significant and broad, and consistent with the results given in Fig. (7.7) (*upper*). This shows that the method can properly treat data with different qualities and different observational time interval and consistently give a result of the period of 15.3 years.

In addition to the minimum at 13.7 years, a second minimum at period $P = 6.0$ years with $V_m^2 = 0.606$ and $f = 0.649$ is found to be significant and broad. Its relative depth, however, is only about 8 times the nearby noise. Although the period of $P = 6.0$

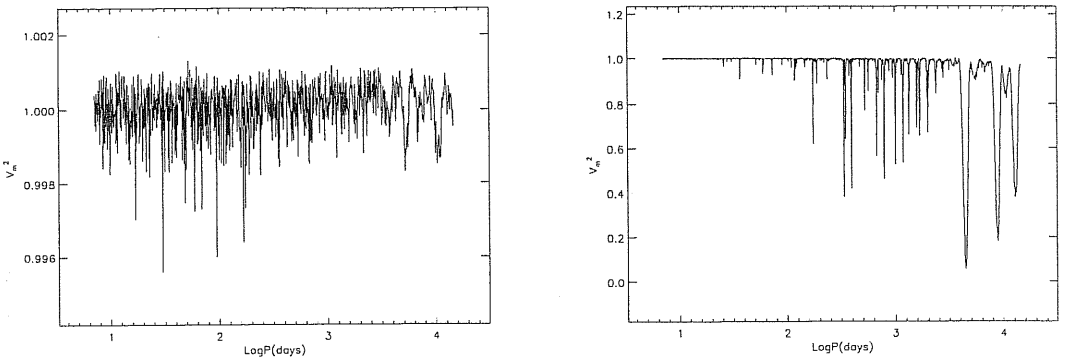

Figure 7.8: (*left*) The plot of normalized V_m^2 vs trial period for a test object with random variations . No significant minima corresponding to trial periods of one year and multiples are found in the plot. (*right*) Same as Fig. 4 for a test object with sinusoidal variation of a period of 12.5 years. In addition to the minima at $P = 12.5$ years and multiple, the minima corresponding to a period of one year and multiples are significant and are artifacts of the method used.

years is about half the period of $P = 13.7$ year, we cannot be sure of the reality of the former, as it does not appear in upper panel of Fig. (7.7).

In addition to the possible periods of 22.8, 13.7 and 6.0 years, the plots also show minima at $P = 1.1$, 2.2 and 3.4 years in both figures in Fig. (7.7) with relatively less significance. A one-year period was also found in the light curves of ON231 (see Sec. (7.3.2.2)) and 3C120 (Jurkevich *et al.* 1971). A one year period is doubtful as the astronomical cycle is one year. In order to check whether the period is a spurious result of the Jurkevich method, we did the following test.

To test the method, we take an object with only random variations with an amplitude of 4 magnitudes. To mimic real observations, we make a further assumption that the object can been observed only from the beginning of January to the end of March every year and that the available data cover a one hundred year range. We also assume that, for moon light reasons for example, it can been observed only for 10, 20 or 30 days a month. Under these assumptions, the number of data points (one point a day) would be 3000, 6000 and 9000. The result of the Jurkevich analysis for the 20-day case is shown in Fig. (7.8) (*left*). The results for the others are similar. No significant minima are found at one year and multiples of this. When we change the assumption from three months to four months and do the test again, the conclusion is unchanged. Now, we assume that the source varies sinusoidally with a period of 12.5 years and we keep all the other assumptions . The result of the analysis for the 20-day case is shown in Fig. (7.8) (*right*). In addition to the minima at 12.5 years and multiples, the minima at one year and multiples become very significant. If we assume we could observe 12 months a year, the minimum corresponding to a period of one year does not exist any more in the V_m^2 plot. We conclude that the Jurkevich method does not give

a spurious period of one year for a randomly variable source, but if there exists a long term period in the light curve of the source, a spurious period of around one year will appear. This can probably be understood as follows: when the trial period is slightly different from half a year, one year and multiples, some of the ten groups (m) contain very few observational points which leads to a very small variance V_i^2 and therefore small V_m^2 (cf. Fig. (7.8) *right*). So the minima at about one year and multiples might be taken as another signal of the existence of a long time-scale period in the light curve.

7.3.3 The Thermal Instability of Slim Accretion Disks

We have collected the data for two BL Lac objects and analyzed their long term variabilities and periodicities. We find that periodic variation with periods of about ten years exist in their light curves. The results are summarized in Table (7.2).

Sillanpää *et al.* (1988) explained the periodic variation in the intensity of the BL Lac object OJ287 with binary black hole models, while Katz (1997) modeled the same object with a precessing disk. For a binary black hole system with about a ten year period, its life is only about $\sim 10^6$ years (Sillanpää *et al.* 1988). Such a system must be very rare and OJ287 must be unusual. However, its other characteristic are not special. From our results and those obtained by other authors (e.g. Hagen-Thorn *et al.* (1997) suggested a 15 yr period in 3C120), the periodicity may also exist in some other objects.

The binary black-hole model also suffers from several difficulties: the observed periodicity is not exact, the period in OJ287 corresponding to the minimum of brightness is quite doubtful, observed burst structures are very broad, the system is short-lived due to gravitational radiation and dynamically unstable due to the interaction between the secondary black hole and the disk. We suggest an alternative explanation for the periodicity in BL Lac objects.

Simulations show that slim disks with high viscosity parameter α can indeed perform limit-cycle type oscillations similar to thin disks in dwarf novae, but the oscillation behavior is different (Honma *et al.* 1991; Szuszkiewicz & Miller 1998). The basic characteristics of these thermal limit cycles strongly depend on $\alpha_{0.1} = \alpha/0.1$, central black hole mass $M_6 = M/10^6 M_\odot$, accretion rate $\dot{M}$ and the generalized stress tensor parameter μ, which is defined by a relation between a generalized stress tensor and pressures

$$\tau_{r\varphi} = \alpha \left[P_g^{\mu} P^{2-\mu} \right]^{1/2} \quad , \tag{7.2}$$

where P and P_g is the total and gas pressures, respectively, and μ is non-negative. However, the time duration of bursts is almost independent of both μ and $\dot{M}$, and may empirically be written as (Honma *et al.* 1991)

$$t_{burst} \simeq 4.5 \alpha_{0.1}^{-0.62} M_6^{1.37} \quad \text{yrs}, \tag{7.3}$$

Table 7.2: The model results. $\alpha = 0.1$, $\mu = 0.5$, $\dot{M} = 0.2\dot{M}_c$.

Object	ON231	Mkn421	OJ287	Group
P_{obs}	13.6	23.1	11.6	7.2
P_{res}	12.3	22.4	8.9	5.1
$M(10^6 M_\odot)$	1.3	2.0	1.0	0.7

when $\mu = 0.5$ and $\dot{M} \simeq 0.2\dot{M}_c$, where the critical accretion rate $\dot{M}_c$ is

$$\dot{M}_c = \dot{M}_E/\epsilon \quad , \tag{7.4}$$

$\dot{M}_E$ being the Eddington accretion rate and ϵ the accretion efficiency (Honma *et al.* 1991). The time interval between subsequent bursts depends strongly on μ, but weakly on $\dot{M}$. Both the origin and the appearance of the presumed viscosity in accretion disks are unknown at present, but a hydro-magnetic origin is one of the options. Having $\mu \simeq 0.5$ may hold if the escape rate of the magnetic field is low (Horiuchi & Kato 1990). The sinusoidal result for these parameters gives a typical period about $2t_{burst}$,

$$t_{cyc} \sim 9.0\alpha_{0.1}^{-0.62} M_6^{1.37} \quad \text{yrs} \quad . \tag{7.5}$$

A slight modification of μ strongly affects the value of the cyclic time-scale t_{cyc}. We fix α, μ and $\dot{M}$ and change the mass M to fit all the sample. The results are given in table 7.2. The mass of black holes in the center of BL Lac objects are estimated to be about $10^6 M_\odot$. The fact that no period of less than five years has been found indicates that the mass of black hole in AGNs is not smaller than $10^6 M_\odot$. This mass is reasonable if the parent galaxies of BL Lac objects are FR I radio galaxies.

7.3.4 Conclusion

We have compiled a catalogue of extragalactic radio jets known to us upto the end of October 1998 and have discussed their statistical properties.

We have collected all of the available observational data for two BL Lac objects and constructed their historical light curves. We analyzed the characters of the light curves with the Jurkevich method. We also discussed the robustness of the method. We found that in these two BL Lac objects the variabilities are complicated with the superposition of two kinds of variation in flux. One is sporadic and the other may be periodic. The period is about ten years. The periodicity in the two objects and also in others can be explained as related to the thermal instability of accretion disks with high α. Such models give a fitted central black hole mass of order $10^6 M_\odot$.

Bibliography

Abramovici, M., *et al.* 1992, Science, 256, 325

Abramowicz, M.A., Calvani, M., Nobili, L. 1983, Nature, 302, 597

Abramowicz M.A., Czerny B., Lasota J.-P., Szuszkiewicz E. 1988, ApJ, 332, 646

Abramowicz, M.A., Karas, V., Lanza, A. 1998, A&A, 331, 1143

Akerlof C.W., *et al.* 1999, Nature, in press

Aloy, M.A., Ibáñez, J.M$^{\underline{a}}$., Martí, J.M$^{\underline{a}}$., Müller, E. 1999, astro-ph/9903352

Amenomori, M., Cao, Z., Dai, B.Z., Ding, L.K., Feng, Y.X., Feng, Z.Y., Hibino, K., *et al.* 1996, A&A, 311, 919

Anderson, S.B., *et al.* 1990, Nature, 346, 42

Anile, A.M. 1989, Relativistic Fluids and Magnetofluids. Combridge University Press

Antonucci R. 1993, ARA&A 31, 473

Atteia, J.-L., Boer, M., Hurley, K., Niel, M., Vedrenne, G., Fenimore, E.E., Klebesadel, R.W., *et al.* 1987, ApJ, 320, L105

Babuska, I., Bieteman, M. 1985, J. Comp. Phys., 63, 33

Bailes, M. 1996, in Compact Stars in Binaries, eds. van Paradijs, J.A., van den Heuvel, E.P.J., and Kuulkers, E., Kluwer, Dordrecht, p.213

Balázs, L.G., Mészáros, A., Horváth, I. 1998, A&A, 339, 1

Balbus, S.A., Hawley, J.F. 1991, ApJ, 376, 214

Band, D.L., Matteson, J., Ford, L., Schaefer, B., Palmer, D., *et al.* 1993, ApJ, 413, 281.

Band, D.L., Ryder, S., Ford, L.A., Matteson, J.L., Palmer, D.M., Teegarden, B.J., Briggs, M.S., Paciesas, W.S., Pendleton, G.N., Precce, R.D. 1996, ApJ, 458, 746

Band, D.L., Ford, L.A., Matteson, J.L., Briggs, M.S., Paciesas, W.S., Pendleton, G.N., Precce, R.D. 1997, ApJ, 485, 747

Band, D., Hartmann, D.H. 1998, ApJ, 493, 555

Barat, C., Chambon, G., Hurley, K., Niel, M., Vedrenne, G., Estulin, I.V., Kuznetsov, A.V., and Zenchenko, V.M. 1981, Ap&SS, 75, 83

Barbieri, C., Cappellaro, E., Romano, G., Turatto, M., Szuszkiewice, E. 1988, A&AS, 76, 477

Barbieri, C., Romano, G. 1981, A&AS, 44, 159

Bardachia, C., *et al.* 1990, Nucl. Inst. A., 289, 518

Battistini, P., Braccesi, A. 1974, A&A, 35, 93

Begelman, M.C., Blandford, R., Rees, M.J. 1984, Rev. Mod. Phys., 56, 255

Belinikov, S.I., Novikov, I.D., Tsvetkov, D.Yu. 1984, Sv. Astron. Let., 10, 177

Belli, B.M. 1995, Ap&SS, 231, 43

Benton, M.J. 1995, Science, 268, 52

Berger, M.J., Oliger, J. 1984, J. Comput. Phys., 53, 484

Bhat, P.N., *et al.* 1992, Nature, 359, 217

Bickert, K.F., Greiner, J. 1993, COMPTON Gamma-Ray Observatory Symp., AIP Conf. Proc. 280, eds. Friedlander M., Gehrels N., & Macomb D., (AIP:New York), p.1059

Binney, J., Tremaine, S. 1987, Galactic Dynamics, Princeton: Princeton University Press

Birand, F. 1971, Nature, 232, 178

Bisnovatyi-Kogan, G.S., Murzina, M.V.A. 1995, PRD, 52, 4380

Blackman, E.G., Yi, I., Field, G.B. 1996, ApJ, 473, L79

Blaes, O. 1987, MNRAS, 227, 975

Blaes, O.M., Webster, R.L. 1992, ApJ, 391, L63

Blandford, R.D., Helfand, D.J. 1999, MNRAS, submitted; astro-ph/9902004

Blandford, R.D., Znajek, R.L. 1977, MNRAS, 179, 433

Blinnikov, S.I., Kozyreva, A.V., Panchenko, I.E. 1999, astro-ph/9902378

Bloom, J.S., *et al.* 1999, GCN208

Bloom, J.S., *et al.* 1999a, GCN240

Bloom, J.S., Djorgovski, S.G., Kulkarni, S.R., Frail, D.A. 1998b, ApJ, 507, L25

Bloom, J.S., Frail, D.A., Kulkarni, S.R., Djorgovski, S.G., Halpern, J.P., Marzke, R.O., Patton, D.R., Oke, J.B., Horne, K.D., Gomer, R., Goodrich, R., Campbell, R., Moriarty-Schieven, G.H., Redman, R.O., Feldman, P.A., Costa, E., Masetti, N. 1998a, astro-ph/9808319

Bloom, J.S., Kulkarni, S.R., Harrison, F., Prince, T., Phinney, E.S., Frail, D.A. 1998, ApJ, 506, L105

Bloom, J.S., Odewahn, S.C., Djorgovski, S.G., Kulkarni, S.R., Harrison, F.A., Koresko, C., Neugebauer, G., Armus, L., Frail, D.A., Gal, R.R., Sari, R., Squires, G., Illingworth, G., Kelson, D., Chaffee, F., Goodrich, R., Feroci, M., Costa, E., Piro, L., Frontera, F., Mao, S., Akerlof, C., McKay, T.A. 1999b, submitted to ApJ Letters; astro-ph/9902182

Boella, G., Chiappetti, L., Conti, G., Cusumano, G., Del Sordo, S., La Rosa, G., Maccarone, M.C., Mineo, T., *et al.* 1997, A&AS, 122, 299

Böttcher, M., Dermer, C.D., Crider, A.W., Liang, E.P. 1999, A&A, in press; astro-ph/9809156

Bradaschia, C., *et al.* 1990, Nucl. Instr. Methods A, 289, 518

Brainerd, J.J. 1992, Nature, 355, 522

Brainerd, J.J. 1994, ApJ, 428, L1

Brainerd, J.J. 1997, ApJ, 487, 96

Brainerd, J.J., Meegan, C.A., Briggs, M.S., Pendelton, G.N., Brock, M.N. 1995, ApJ, 441, L39

Brandt, A. 1977, Math. Comput. 31, 333

Brandt, S., Lund, N., Castro-Tirado, A. 1994, Proc. Gamma-Ray Burst Workshop, Huntsville, AL AIP Conf. Proc. 307, eds. Fishman G.J., Braiderd J.J., Hurley K., New York: Am. Inst. Phys., p.13

Bridle A.H. 1986, Can. J. Phys. 64, 353

Bridle A.H., Perley R.A. 1984, ARA&A 22, 319

Bridle, A.H., Hough, D.H., Lonsdale, C.J., Burns, J.O., Laing, R.A. 1994, AJ 108, 766

Briggs, M.S. 1993, ApJ, 407, 126

Briggs, M.S., Band, D.L., Preece, R.D., Paciesas, W.S., Pendleton, G.N. 1999, astro-ph/9901224

Briggs, M.S., Paciesas, W.S., Pendleton, G.N., Meegan, C.A., Fishman, G.J., Horack, J.M., Brock, M.N., Kouveliotou, C., Hartmann, D.H., Hakkila, J. 1996, ApJ, 459, 40

Burderi, L., King, A.R., Szuszkiewicz, E. 1998, ApJ, 509, 85

Burenin, R.A., Vikhlinin, A.A., terekhov, O.V., sazonov, S.Yu. 1998, Astron. Lett., in press; astro-ph/9804274

Carigan, B.J., Katz, J.L. 1992, ApJ, 399, 100

Carilli C.L., Röttgering H.J.A., van Ojik R., Miley G.K., van Breugel W.J.M. 1997, ApJS 109, 1

Carter, B. 1992, ApJ, 391, L67

Castro-Tirado, A., Brandt, S., Lund, N., Guziy, S., *et al.* 1994, in Proc. Gamma-Ray Burst Workshop, Huntsville, AL AIP Conf. Proc. 307, eds. Fishman G.J., Braiderd J.J., Hurley K., New York: Am. Inst. Phys., p.404

Castro-Tirado, A.J., *et al.* 1998, Science, 387, 1011

Cen, R. 1998, astro-ph/9809022

Chandrasekhar, S. 1939, An Introduction to the theory of Stellar Structure, Chicago: University of Chicago Press

Chandrasekhar, S. 1975, ApJ, 202, 809

Che, H., Yang, Y., Nemiroff, R.J. 1998, submitted to ApJ; astro-ph/9810220

Choptuik, M.W. 1989, in Frontiers in Numerical Relativity, (eds.) C.R. Evans, L.S. Finn, D.W. Hobill, Cambridge University Press, Cambridge, p.206

Choptuik, M.W. 1993, PRL, 44(9), ??

Clark, J.P.A., van den Heuvel, E.P.J., Sutantyo, W. 1979, A&A, 72, 120

Clark, J.P.A., Eardley, D. 1977, ApJ, 215, 311

Clarke, T.E., Blaes, O., Tremaine, S. 1994, AJ, 107, 1873

Cohen, E., Katz, J.I., Piran, T., Sari, R., Preece, R.D., Band, D.L. 1997, ApJ, 488, 330

Cohen, E., Piran, T. 1995, ApJ, 444, L25

Cohen, E., Piran, T. 1997, ApJ, 488, L7

Cohen, E., Piran, T., Sari, R. 1999, astro-ph/9803258

Cohen, E., Narayan, R., Piran, T. 1998, ApJ, 500, 888

Colgate, S.A. 1974, ApJ, 187, 333

Collar, J.I. 1996, PRL, 76, 999

Colella, P., Woodward, P.R. 1984, J. Comp. Phys., 54, 174

Connolly, A.J., Szalay, A.S., Dickinson, M., Subbarao, M.U., Brunner, R.J. 1997, ApJ, 486, L11

Cordes, J.M., Chernoff, D.F. 1997, ApJ, 482, 971

Cordes, J.M., Chernoff, D.F. 1997, ApJ, 505, 315

Costa, E., Frontera, F., Heis, J., Feroci, M., in't Zand, J., Fiore, F., Cinti, M.N., Dal Fiume, D., Nicastro, L., Orlandini, M., Palazzi, Z., Rapisarda, M., Zavattini, G., Jager, R., Parmar, A., Owens, A., Molendi, S., Cusumano, G., Maccarone, M.C., Giarrusso, S., Coletta, A., Antonelli, L.A., Giommi, P., Muller, J.M., Piro, L., and Butler, R.C. 1997, Nature, 387, 783

Courant, R., Friedrich, K.O. 1976, Supersonic Flow and Shock Waves, Springer-Verlag, New York

Craine, E.R., Warner, J.W. 1973, ApJ, 179, L53

Crider, A., Liang, E.P., Preece, R.D. 1997, astro-ph/9711100

Crider, A., Liang, E.P., Preece, R.D., Briggs, M.S., Pendleton, G.N., Paciesas, W.S., Band, D.L., Matteson, J.L. 1998, ApJ, submitted; astro-ph/9810326

Crider, A., Liang, E.P., Smith, I.A., Preece, R.D., Briggs, M.S., Pendleton, G.N., Paciesas, W.S., Band, D.L., Matteson, J.L. 1997, ApJ, 479, L39

Cruz-Gonzalez, I., Huchra, J.P. 1984, AJ, 89, 441

Curran, S.J., Lorimer, D.R. 1995, MNRAS, 276, 347

Daigne, F., Mochkovitch, R. 1997, MNRAS, 285, L15

Dar, A. 1998, ApJ, 500, L93

Dar, A. 1998a, Nature, submitted; astro-ph/9809163

Dar, A. 1998b, PRL, in press; astro-ph/9811196

Dar, A. 1999, submitted to ApJ Let.; astro-ph/9901005

Dar, A. 1999b, astro-ph/9902017

Dar, A., Laor, A., Shaviv, N.J. 1998, PRL, 80, 5813

Davies, M.B., Benz, W., Piran, T., Thielemann, F.K. 1994, ApJ, 431, 742

Dermer, C.D. 1992, PRL, 68, 1799

Dermer, C.D., Chiang, J. 1999, astro-ph/9810222

Dermer, C.D., Mitman, K.E. 1999, ApJ, 513, L5

Dezalay, J.P., Lestrade, J.P., Barat, C. *et al.* 1996, ApJ, 471, L27

Djorgovski, S.G., Kulkarni, S.R., Bloom, J.S., Goodrich, R., Frail, D.A., Piro, L., Palazzi, E. 1998, ApJ, 508, L17

Djorgovski, S.G., Kulkarni, S.R., Bloom, J.S., Frail, D.A., Chaffee, F., Goodrich, R. 1999, GCN Circ. 189

Djorgovski, S.G., *et al.* 1999a, GCN Circ. 250

Dokuchaev, V.I., Eroshenko, Yu.N., Ozernoy, L.M. 1998, ApJ, 502, 192

Drozdova, D.N., Panchenko, I.E. 1997, A&A, 324, L17

Duncan, R.C., Thompson, C. 1992, ApJ, 392, L9

Duncan, R.C., Li, H. 1997, ApJ, 484, 720

Eberl, T., Ruffert, M., Janka, H.-Th. 1998, A&A, in press

Eichler D., Levinson, A. 1999, astro-ph/9903103

Eichler D., Livio, M., Piran, T., Schramm, D. 1989, Nature, 340, 126

Eichler D., Silk, J. 1992, Science, 257, 937

Erwin, D.H. 1996, Sci. Am., 275, No. 1, 56

Evans, W.D., Fenimore, E.E., Klebesadel, R.W., Laros, J.G., Terrell, N.J. 1981, Ap&SS, 75, 35

Fabian, A.C., Podsiadlowski, P. 1993, MNRAS, 263, 49

Fenimore, E.E., *et al.* 1988, ApJ, 335, L71

Fenimore, E.E., *et al.* 1993a, Nature, 366, 40

Fenimore, E.E., Bloom, J.S. 1995, ApJ, 453, 25

Fenimore, E.E., Cooper, C., Ramirez, E., Sumner, M.C., Yoshida, A., Namiki, M. 1999, ApJ, 512, 683

Fenimore, E.E., Epstein, R.L., Ho, C. 1993, A&AS, 97, 59

Fenimore, E.E., *et al.* 1995, ApJ, 448, L101

Fenimore, E.E., in't Zand, J.J.M., Norris, J.P., Bonell, J.T., Nemiroff, R.J. 1996, ApJ, 448, L101

Fenimore, E.E., Klebesadel, R.W., Laros, J.G., Stockdale, R.E., Kane, S.R. 1982, Nature, 297, 665

Fenimore, E.E., Madras, C., Nayakshin, S. 1996a, ApJ, 473, 998

Feroci, M. *et al.* IAU Circ. No.7095

Finn, L.S., Chernoff, D. 1993, PRD, 47, 2198

Fishman, G.J., Meegan, C.A.A. 1995, ARA&A, 33, 415

Fishmann, G.J., Meegan, C.A., Wilson, R.B., Brock, M.N., Horack, J.M., Kouveliotou, C., Howard, S., Paciesas, W., Briggs, M.S., Pendleton, G.N., Koshut, T.M., Mallozzi, R.S., Stollberg, M., and Lestrade, J.P., 1994, ApJS, 92, 229

Fishman, G.J., *et al.* 1989, Proc. Gamma Ray Observatory Science Workshop, ed. W.N. Johnson, NASA, Greenbert, MD, Vol.2, P.39

Ford, L.A., *et al.* 1995, ApJ, 439, 307

Forrest, D.J., Chupp, E.L., Ryan, J.M., Cherry, M.L., Gleske, I.U., Reppin, C., Pinkau, K., Rieger, E., Kanbach, G., Kinzer, R.L., Share, G.H., Johnson, W.N., and Kurfess, J.D. 1980, Solar Phys., 65, 15

Forrest, D., Vestand, W., McConnell, M., Ryan, J., and Owens, A. 1995, Ap&SS, 231, 459

Frail, D.A., Kulkarni, S.R. 1999, GCN211

Frail, D.A., Kulkarni, S.R., Nicastro, L., Feroci, M., Taylor, G.B. 1997, Nature, 389, 261

Frontera, F., Costa, E., Piro, L., Muller, J.M., Amati, L., Feroci, M., Fiore, F., Pizzichini, G., Tavani, M., Castro-Tirado, A., Cusumano, G., Dal Fiume, D., Heise, J., Hurley, K., Nicastro, L., Orlandini, M., Owens, A., Palazzi, E., Parmar, A.N., in't Zand, J., Zavattini, G. 1997, ApJ, 493, L67

Fruchter, A.S. 1998, Submitted to ApJ Letter; astro-ph/9810224

Fruchter, A.S., Pian, E., Thorsett, S.E., Bergerson, L.E., González R.A., Metzger, M., *et al.* 1998, submitted to ApJ; astro-ph/9807295

Fruchter, A.S., Thorsett, S.E., Metzger, M., Sahu, K.C., Petro, L., Livio, M., Ferguson, H., Pian, E., Hogg, D.W., Galama, T., Gull, T.R., Kouvelioutou, C., Macchetto, D., van Paradijs, J., Pedersen, H., Smette, A. 1999, astro-ph/9902236

Fryer, C.L., Kalogera, V. 1997, ApJ, 489, 244

Fryer, C.L., Woosley, S.E. 1998, ApJ, 501, 780

Fryer, C.L., Woosley, S.E. 1998a, ApJ, 502, L9; astro-ph/9804167

Fryer, C.L., Woosley, S.E., Hartmann, D.H. 1998a, ApJ submitted

Fryer, C.L., Woosley, S.E., Herant, M., Davies, M.B. 1998, ApJ, in press; astro-ph/9808094

Fuller, G.M., Shi, X. 1998, ApJ, 502, L5

Gagen-Torn, V.A., Marchenko, S.G., Smekhacheva, R.I. Yakovleva, V.A. 1983, Astrophysics, 19, 111

Gaisser, T.K. 1990, Cosmic Rays and Particle Physics (Cambridge Univ. Press, Cambridge).

Gal, R., *et al.* 1999, GCN207

Galama, T.J., *et al.* 1997, Nature, 387, 497

Galama, T.J., *et al.* 1999a, submitted to A&AS, Proc. of the Conference "Gamma-Ray Bursts in the Afterglow Era", held in Rome, 1998 November 3-6

Galama, T.J., Briggs, M.S., Wijers, R.A.M.J., Vreeswijk, P.M., Rol, E., Band, D., van Paradijs, J., Kouvelioutou, C., Preece, R.D., Bremer, M., Smith, I.A., Tilanus, R.P.J., de Bruyn, A.G., *et al.* 1999, Nature, in press; astro-ph/9903021

Galama, T.J., Groot, P.J., van Paradijs, J., Kouveliotou, C., Strom, R.G., Wijers, R.A.M.J., *et al.* 1998, ApJ, 497, L13

Galama, T.J., Vreeswijk, P.M., van Paradijs, J., Kouveliotou, C., Augusteijn, T., Hainaut, O.R., Patat, F., Böhnhardt, H., Brewer, J., Doublier, V., Gonzalez, J.-F., *et al.* 1998a, Nature, 395, 670

Galama, T.J., Wijers, R.A M.J., Bremer, M., Groot, P.J., Strom, R.G., Kouveliotou, C., van Paradijs, J. 1998b, ApJ, 500, L97

Galama, T.J., Wijers, R.A M.J., Bremer, M., Groot, P.J., Strom, R.G.,), de Bruyn, A.G., Kouveliotou, C., Robinson, C.R., van Paradijs, J. 1998c, ApJ, 500, L101

Garnavich, P., *et al.* 1999, GCN215

Garrington, S.T., Conway, R.G., Leahy, J.P. 1991, MNRAS, 250, 171

Ghisellini, G., Celotti, A. 1998, astro-ph/9812079

Ghisellini, G., G., Haardt, F., Campana, S., Lazzati, D., Covino, S. 1999, ApJ, in press; astro-ph/9808156

Ghosh, P., Abramowicz, M.A. 1997, MNRAS, 292, 887

Guilbert, P.W., Fabian, A.C., Rees, M.J. 1983, MNRAS, 205, 593

Godunov, S.K. 1959, Mat. Sb., 47, 271

Golenetskii, S.V., *et al.* 1983, Nature, 306, 451

Goodman, J. 1986, ApJ, 308, L47

Goodman, J. 1997, New Astronomy, 2, 449

Goodman, J., Dar, A., Nussinov, S. 1987, ApJ, 314, L7

Gorosabel, J., Castro-Tirado, A.J. 1997, ApJ, 483, L83

Gorosabel, J., Castro-Tirado, A.J., Brandt, S., Lund, N. 1998, A&A, in press; astro-ph/9804239

Granot, J., Piran, T., Sari, R. 1998, astro-ph/9808007

Granot, J., Piran, T., Sari, R. 1999, astro-ph/9806192

Graziani, C., Lamb, D.Q., Marion, G.H. 1998, ApJ, in press; astro-ph/9810374

Greiner, J. *et al.* 1994, A&A, 302, 1216

Greisen, K. 1966, PRL, 16, 748

Grindlay, J.E. 1998, ApJ, in press; astro-ph/9808242

Groot, P.J., Galama,T.J., van Paradijs, J., Kouveliotou, C., Wijers, R.A.M.J., Bloom, J., Tanvir, N., Vanderspek, R., Greiner, J., Castro-Tirado, A.J., Gorosabel, J., von Hippel, T., Lehnert, M., Kuijken, K., Hoekstra, H., Metcalfe, N., Howk, C., Conselice, C., Telting, J., Rutten, R.G.M., Rhoads, J., Cole, A., Pisano, D.J., Naber, R., Schwarz, R. 1998, ApJ, 493, L27

Grossman, S.A., Nowak, M.A. 1994, ApJ, 435, 548

Gruz-Gonzalez, Z., Huchra, J.P. 1984, AJ, 89, 441

Haensel, P., Paczyński, B., Amsterdamski, P. 1992, ApJ, 375, 209

Hakkila, J., Meegan, C., Pendleton, G., Fishman, G., Wilson R., *et al.* 1994, ApJ, 422, 659

Halpern, J., *et al.* 1997, IAU Circ. 6788

Halpern, J.P., Kemp, J., Piran, T., Bershday, M.A. 1999, ApJ, sumitted; astro-ph/9903418

Hamade, R.S., Stewart, J.M. 1996, Class. Quantum Grav., 13, 497

Harris, M. 1997, 4th Huntsville Symposium on Gamma-Ray Bursts, AIP Conf. Proc., no. 428, eds. C. Meegan, R. Preece, and T. Koshut, New York: AIP, in press

Hartmann, D.H. 1992, Comm. in Astrophys., 16, 231

Hartmann, D.H., Blumenthal, G.R. 1989, ApJ, 342, 521

Hartmann, D.H., Greiner, J., Briggs, M.S. 1995, A&A, 303, L65

Hartmann, D.H., *et al.* 1995, Ap&SS, 231, 361

Hawking, S.W. 1974, Nature, 248, 30

Hayashida, N., Honda, K., Honda, M., Inoue, N., Kadota, K., Kakimoto, F., Kamata, K., Kawaguchi, S., *et al.* 1996, PRL, 77, 1000

Heise, J., *et al.* 1997, IAU Circ. 6787

Higdon, J.C., Lingenfelter, R.E., 1990, ARA&A, 28, 401

Hillas, A.M. 1984, ARA&A, 22, 42

Hills, D., Bender, P.L., Webbink, R.F. 1991, ApJ, 369, 675

Hjorth, J., Anderson, M.J., Cairos, L.M., Caon, N., Zapatero Osorio, M., Pederson, H., Lindgren, B., Castro-Tirado, A.J., Perez, E. 1999, GCN Circ. 219

Hjorth, J., Anderson, M.J., Pederson, H., Zapatero-Osorio, M.R., Perez, E., Castro-Tirado, A.J. 1999a, GCN Circ. 249

Höflich, P., Wheeler, J.C., Wang, L. 1998, astro-ph/9807345

Hogg, D.W., Fruchter, A.S. 1998, astro-ph/9807262

Holz, D.E., Miller, M.C., Quashnock, J.M. 1999, ApJ, 510, 54

Holz, D.E., Wald, R. 1998, Phys. Rev. D, 58, 063501

Honma, F., Matsumoto, R., Kato, S. 1991, PASJ, 43, 147

Horack, J.M., Emslie, A.G. 1994, ApJ, 428, 620

Horack, J.M., Koshut, T.M., Mallozzi, R.S., Storey, S.D., Emslie, A.G. 1994, ApJ, 429, 319

Horiuchi, T., Kato, S. 1990, PASJ, 42, 661

Horváth, I. 1998, ApJ, 508, 757

Horváth, I., Mészáros, P., Mészáros, A. 1996, ApJ, 470, 56

Hulse, R.A., Taylor, J.H. 1975, ApJ, 368, 504

Hurley, K. 1992, Present at the Conference on Gamma-Ray Bursts in the Afterglow Era, Rome; A&AS, in press

Hurley, K. 1994, ApJS, 90, 857

Hurley, K. 1996, 3th Huntsville Symp. on Gamma-Ray Bursts, AIP Conference Proceedings No384, eds. C. Kouvelioutou, M.F. Briggs, and G.J. Fishman, New York:AIP, p.889

Hurley, K. 1999, 4th Huntsville Symposium on Gamma-Ray Bursts, AIP Conf. Proc., no. 428, eds. C. Meegan, R. Preece, and T. Koshut, New York: AIP, in press

Hurley, K., Dingus, B., Mukherjee, R., Sreekumar, P., Kouvelioutou, C., *et al.* 1994, Nature, 372, 652

Hurley, K., Hartmann, D., Kouvelioutou, C., Fishman, G., Laros, J., Cline, T., Boer, M. 1997, ApJ, 479, L113

Hurley, K., Sommer, M., Atteia, J.-L., Boer, M., Cline, T., Cotin, F., Henoux, J.-C., Kane, S., Lowes, P., Niel, M. 1992, A&AS, 92, 401

Hurley, K., Cline, T., Mazets, E., Barthelmy, S., Butterworth, P., Marshall, F., Palmer, D., Apter, R., Golenetskii, S., *et al.* 1999, Nature, 397, 41

Imamura, J.N., Epstein, R.I. 1987, ApJ, 313, 711

Israel, W. 1966, Il Nuovo Cimento, 44, 1

Iwamoto, K., Mazzali, P.A., Nomoto, K., Umeda, H., Nakamura, T., Patat, F., Danziger, I.J., Young, T.R., Suzuki, T., Shigeyama, T., Augusteijn, T., Doublier, V., Gonzalez, J.-F., *et al.* 1998, Nature, 395, 673

Janka, H.-T., Ruffert, M. 1996, A&A, 307, L33

Jaroszynski, M. 1996, A&A, 305, 839

Jeffrey, A. 1978, Quasilinear hyperbolic systems and waves, Pitman:London.

Jurkevich, I. 1971, Ap&SS, 13, 154

Katz, J.I. 1992, Ap&SS, 197, 163

Katz, J.I. 1994a, ApJ, 422, 248

Katz, J.I. 1994, ApJ, 432, L107

Katz, J.I. 1997, ApJ, 490, 633

Katz, J.I. 1997a, ApJ, 478, 527

Katz, J.I., Canel, L.M. 1996, ApJ, 471, 527

Katz, J.I., Piran, T. 1997, ApJ, 490, 772

Katz, J.I., Piran, T., Sari, R. 1998, PRL, 80, 1580

Kelson, D.D., Illingworth, G.D., Franx, M., Magee, D., van Dokkum, G. 1999, IAU Circ. 7096

Kidger, M.R., Takalo, L., Sillanpää, A. 1992, A&A, 264, 32

King, I.R. 1962, AJ, 67, 471

Kippen, R.M., Briggs, M.S., Kommers, J.M., Kouveliotou, C., Hurley, K., Robinson, C.R., van Paradijs, J., Hartmann, D.H., Galama, T.J., Vreeswijk, P.M. 1998, ApJ, 506, L27

Kippen, R.M., Hurley, K., Pendleton, G.N. 1998a, astro-ph/9810040

Kippen, R.M., *et al.* 1998b, GCN Circ. 143

Kippen, R.M., *et al.* 1999, GCN Circ. 224

Kippenhahn, R., Weigert, A. 1990, Stellar Structure and Evolution, Berlin: Springer-Verlag

Kirshner, R.P., *et al.* 1983, ApJ, 88, 1285

Klebesadel, R.W., Evans, W.D., and Laros, J.G. 1981, Ap&SS, 75, 5

Klebesadel, R.W., Strong, I.B., and Olson, R.A. 1973, ApJ, 182, L85

Kluźniak, W., Lee, W.H. 1998, ApJ, 494, L53

Kluźniak, W., Lee, W.H. 1998, preprint

Knight, F.K., Matteson, J.L. and Peterson, L.E. 1981, Ap&SS, 75, 21

Knuth, D.E. 1973, The Art of Computer Programming, Vol. 1: Fundamental Algorithms, Addison-Wesley Publishing Company, Massachusetts

Kobayashi, S., Piran, T., Sari, R. 1997, ApJ, 490, 92

Kobayashi, S., Piran, T., Sari, R. 1999, astro-ph/9803217

Kochaneck, C., Piran, T. 1993, ApJ, 417, L17

Kollat, T., Piran, T. 1996, ApJ, 467, L41

Kommers, J.M., Lewin, W.H., Kouvelioutou, C., van Paradijs, J., Pendleton, G.N., Meegan, C.A., Fishmann, G.J. 1998, submitted to ApJ; astro-ph/9809300

Kompaneetz, D., Stern, B. 1997, 4th Huntsville Symposium on Gamma-Ray Bursts, AIP Conf. Proc., no. 428, eds. C. Meegan, R. Preece, and T. Koshut, New York: AIP, in press

Koshut, T.M., Kouveliotou, C., Paciesas, W., van Paradijs, J., Pendleton, G., Briggs, MS., Fishman, G.J., Meegan, C.A., 1995, ApJ, 452, 145

Kouveliotou, C., *et al.* 1987, ApJ, 332, L21

Kouveliotou, C., Meegan, C., Fishmann, G., Bhat, N., Briggs, M., *et al.* ., 1993, ApJ, 413, L101

Kouveliotou, C., *et al.* 1996, 3rd Hunstville Workshop on Gamma-Ray Bursts, AIP Conf. Proc. 384, eds. C. Kouveliotou, M.S., Briggs, and G.J. Fishmann, New York:AIP, p42

Kretschkmar B., Mereghetti, S., Robert, A., *et al.* 1999, Present at the Conference on Gamma-Ray Bursts in the Afterglow Era, Rome; A&AS, in press.

Krolik, J.H., Pier, E.A. 1991, ApJ, 373, 277

Kröner, D. 1997, Numerical Schemes for Conservation Laws, John Wiley & Sons Ltd and B.G. Teubner Publishers, Chichester

Krumholz, M., Thorsett, S.E., Harrison, F.A. 1998, ApJ, 506, L81

Kukarkin, B.V., Kholopov, P.N., *et al.* 1969, General Catalog of Variable Stars, Vol.1, Moscow

Kulkarni, S.R., *et al.* 1998, Nature, 393, 35

Kulkarni, S.R., Djorgovski, S.G., Odewahn, S.C., Bloom, J.S., Gal, R.R., Koresko, C.D., Harrison, F.A., Lubin, L.M., Neugebauer, G., Armus, L., Sari, R., Illingworth, G.D., Kelson, D.D., Magee, D.K., van Dokkum, P.G., Frail, D.A., Mulchaey, J.S., Malkan, McLean, Teplitz, H.I., Koerner, D., Kirkpatrick, D., Kobayashi, N., Yadigaroglu, I.A., Halpern, J., Piran, T., Goodrich, R., Chaffee, Feroci, M., Costa, E. 1999, Nature submitted; astro-ph/9902272

Kurochkin, N.E. 1971, Astron. Cirk., No.644

Lai, D., Rasio, F.A., Shapiro, S.L. 1994, ApJ, 420, 811

Lamb, D.Q. 1995, PASP, 107, 1152

Lamb, D.Q. 1997, 4th Huntsville Symposium on Gamma-Ray Bursts, AIP Conf. Proc., no. 428, eds. C. Meegan, R. Preece, and T. Koshut, New York: AIP.

Lamb, D.Q., Graziani, C., Smith, I.A., 1993, ApJ, 413, L11

Lamb, D.Q., Quashnock, J.M. 1993, ApJ, 415, L1

Landau L.D., Lifshitz E.M. 1979, Fluid Mechanics, Pergamon Press, Oxford

Larson, S.B., McLean, I.S. 1997, ApJ, 491, 93

Larson, S.B., McLean, I.S., Becklin, E.E. 1996, ApJ, 460, L95

Lasota, J.P., Pelat, D. 1991, A&A, 249, 574

Leahy, J.P., Black, A.R.S., Dennett-Thorpe, J., Hardcastle, M.J., Komissaror, S., Perley, R.A., Riley, J.M., Scheuer, P.A.G. 1997, MNRAS, 291, 20

LeVeque, R.J. 1992, Numerical Methods for Conservation Laws, Birkhäuser Verlag, Basel

Levinson, A., Eichler, D. 1993, ApJ, 418, 386

Loeb, A., Perna, R. 1997, ApJ,

Li, H., Dermer, C. 1992, Nature, 359, 514

Li, L.-X. 1999, astro-ph/9902352

Li, T.-P. 1997, Ap&SS, in press; astro-ph/9704264

Liang, E.P. 1997, ApJ, 491, L15

Liang, E.P., Crider, A., Böttcher, M., Smith, I.A. 1999, astro-ph/9903438

Liang, E.P., Kargatis, V. 1996, Nature, 381, 49

Liang, E.P., Kusunose, M., Smith, I.A., Crider, A. 1997, ApJ, 479, L35

Lilly, S.J., *et al.* 1996, ApJ, 460, L1

Lin, D.N.C., Shields, G.A. 1986, ApJ, 305, 28

Lingenfelter, R.E., Higdon, J.C. 1992, Nature, 356, 132

Link, B., Epstein, R.I., 1996, astro-ph/9601033

Lipunov, V.M., Postnov, K.A., Prokhorov, M.E. 1997, MNRAS, 288, 245

Lipunov, V.M., Postnov, K.A., Prokhorov, M.E. 1998, Astron. Lett., in press

Liu, F.K. 1996, MNRAS, 281, 1197

Liu, F.K. 1996a, in Proceedings of 21st Century Chinese Astronomy Conference, (eds.) Li. Q.B., Cheng K.S., *et al.* HongKong, in press

Liu, F.K. 1999, Astrophys. Let. & Commun., in press; Proc. of the 3rd INTEGRAL Workshop "The Extreme Universe", Taormina/Sicily, Sep 1998.

Liu, F.K., Liu, B.F., Xie, G.Z. 1997, A&AS, 123, 569

Liu, F.K., Xie, G.Z., Bai, J.M. 1995, A&A, 295, 1

Liu, F.K., Zhang, Y.H. 1999, A&AS, submitted

Livio, M., Ogilvie, I., Pringle, J.E. 1998, ApJ Submitted

Loredo, T.J., Wassermann, I.M. 1995a, ApJS, 96, 59

Loredo, T.J., Wassermann, I.M. 1995b, ApJS, 96, 261

Luchkov, B. 1994, Astron. Lett., 20, 253

Luginbuhl, C.B., *et al.* 1995, Ap&SS, 231, 289

Lyne, A.G., Lorimer, D.R. 1994, Nature, 369, 127

Ma, F., Xie, B. 1996, ApJ, 462, L63

MacFadyen, A., Woosley, S.E. 1998, astro-ph/9810274

Madau, P., *et al.* 1996, MNRAS, 283, 1388

Makino, F., Tanako, Y., Matsuoka, M., *et al.* 1987, ApJ, 313, 662

Mallozzi, R.S., Paciesas, W.S., Pendleton, G.N., Briggs, M.S., Preece, R.D., Meegan, C.A., Fishman, G.J. 1995, ApJ, 454, 597

Mao, S. 1992, ApJ, 389, L41

Mao, S., Narayan, R., Piran, T., 1994, ApJ, 420, 171

Mao, S., Paczyński, B. 1992, ApJ, 389, L13

Maoz, E. 1993, ApJ, 414, 877

Marani, G.F., Nemiroff, R.J., Bonnell, J.T., Norris, J.P. 1996, astro-ph/9607158

Marani, G.F., Nemiroff, R.J., Norris, J.P., Hurley, K., Bonnell, J.T. 1998a, astro-ph/9810391

Marani, G.F., *et al.* 1998, 4th Huntsville Symposium on Gamma-Ray Bursts, AIP Conf. Proc., no. 428, eds. C. Meegan, R. Preece, and T. Koshut, New York: AIP, p.166

Marronetti, P., Mathews, G.J., Wilson, J.R. 1998, PRL submitted; gr-qc/9803093

Martí, J.M$^{\underline{a}}$., Müller, E. 1994, J. Fluid Mech., 258, 317

Martí, J.M$^{\underline{a}}$., Müller, E. 1996, J. Comput. Phys., 123, 1

Masetti, M., *et al.* 1999, GCN220

Masso, J., Seidel, E., Walker, P. 1995, in General Relativity MG7 Proceedings, Adaptive Mesh Refinement in Numerical Relativity, (eds.) Ruffini, R., Keiser, M., World Scientific

Masuda, N., Eriguchi, Y. 1997, ApJ, ???

Mathews, G.J., Wilson, J.R. 1997, ApJ, 482, 929

May, M.M., White, R.H. 1967, Methods in Comutational hysics, Vol. 7: Astrophyscis, (Eds.) B. Alder, S. Fernbach, M. Rotenberg, Academic Press, New York, p.219

Mazets, E.P., Golenetskii, S.V. 1981, Ap&SS, 75, 47

Mazets, E.P., Golenetskii, S.V., Aptekar, R.L., Guryan, Y.A., Ilyinski, V.N., 1980, Pisma Astron. Zh., 6, 706

Mazets, E.P., Golenetskii, S.V., Aptekar, R.L., Guryan, Y.A., Ilyinski, V.N., 1981, Nature, 290, 378

McGimsey, B.Q., Smith, A.G., Scott, R.L., Leacock, R.J., Edward, P.L., Hackney, R.L., Hackney, K.R. 1975, AJ, 80, 895

McNamara, B.J., Harrison, T.E., Ryan, J., Kippen, R.M., McConnell, M., Macri, J., Kouveliotou, C., Fishman, G.J., Meegan, C.A., Green, D.A., *et al.* 1996, ApJS, 103, 173

Meegan, C., *et al.* 1998, BATSE catalogue, http://www.batse.msfc.nasa.gov/data/grb/catalog

Meegan, C., Fishman, G.J., Horack, J.M., Brock, M.N., Cole, S., Paciesas, W.S., Briggs, M.S., Pendleton, G.N., Preece, R., Koshut, T.M., Mallozzi, R.S., Kouveliotou, C., Mccollough, M., 1995, electronic catalog available via telnet gross.gsfc.nasa.gov, username gronews.

Meegan, C., Fishman, G., Wilson, R., Paciesas, W., Pendleton, G., Horack, J.M., Brock, M.N., Kouveliotou, C. 1992, Nature, 355, 143

Meegan, C.A., *et al.* 1996a, IAU Circ. 6518

Meegan, C.A., Pendleton, G.N., Briggs, M.S., Kouveliotou, C., Koshut, T.M., Lestrade, J.P., Paciesas, W.S., McCollough, M.L., Brainerd, J.J., Horack, J.M., Hakkila, J., Henze, W., Preece, R.D., Mallozzi, R.S., Fishman, G.J. 1996, ApJS, 106, 65

Melia, F., Fatuzzo, M. 1992, ApJ, 398, L85

Mészáros, A., Mészáros, P. 1996, ApJ, 466, 29

Mészáros, P., Laguna, P., Rees, M.J. 1993, ApJ, 415, 181

Mészáros, P., Mészáros, A. 1995, ApJ, 449, 9

Mészáros, P., Rees, M.J. 1992, ApJ, 397, 570

Mészáros, P., Rees, M.J. 1992a, MNRAS, 258, 41P

Mészáros, P., Rees, M.J. 1993, ApJ, 405, 278

Mészáros, P., Rees, M.J. 1997a, ApJ, 476, 232

Mészáros, P., Rees, M.J. 1997, ApJ, 482, L29

Mészáros, P., Rees, M.J. 1998, ApJ, 502, L105

Mészáros, P., Rees, M.J. 1998a, ApJL, in press; astro-ph/9806183

Mészáros, P., Rees, M.J. 1999, astro-ph/9902367

Mészáros, P., Rees, M.J., Wijers, R.A.M.J. 1998, New Astronomy, in press; astro-ph/9808106

Mészáros, P., Rees, M.J., Wijers, R.A.M.J. 1998b, ApJ, 499, 301

Metzger, M.R., *et al.* 1997, Nature, 387, 878

Meyer-Hofmeister E., 1993, in Proceedings of 4th MPG-CAS Workshop on High Energy Astrophysics and Cosmology, Börner G., Buchert T. (eds.), MPA/P8, P.126

Milgrom, M., Usov, V. 1995, ApJ, 449, L37

Miller, H.R. 1975, ApJ, 201, L109

Miller, H.R., McGimsey, B.Q., Williamon, R.M. 1977, ApJ, 217, 382

Miller, J.C., Pantano, O. 1989, PRD, 40, 1789

Miller, J.C., Pantano, O. 1990, PRD, 42, 3334

Misner, C.W., Thorne, K.S., Wheeler, J.A. 1973, Gravitation, (Freeman:San Francisco).

Mitrofanov, I.G., *et al.* 1984, Sov. Astron., 28, 547

Mitrofanov, I.G., *et al.* 1995, Astronomy Reports, 39, 305

Mitrofanov, I.G., *et al.* 1996, ApJ, 459, 570

Mitrofanov, I.G., Litvak, M.L., Ushakov, D.A. 1997, ApJ, 490, 509

Mitrofanov, I.G., Pozanenko, A.S., Briggs, M.S., Paciesas, W.S., Preece, R.D., Pendleton, G.N., Meegan, C.A. 1998, ApJ, 504, 925

Mochkovitch, R., Hernanz, M., Isern, J., Martin, X. 1993, Nature, 361, 236

Mochkovitch, R., Maitia, V., Marques, R. 1995, in Proceedings of 29th ESLAB Symp., Towards the Source of Gamma-Ray Bursts, (eds.) Bennett, K., & Winkler, C., P.531

Moles, M., Garcia-Pelayo, J., Masegosa, J., Aparicio, A. 1985, ApJS, 58, 255

Mufson, S.L., Wisniewski, W.Z., Wood K., *et al.* 1980, ApJ, 241, 74

Mukherjee, S., Feigelson, E.D., Babu, G.J., Murtagh, F., Fraley, C., Raftery, A. 1998, ApJ, 508, 314

Murakami, T. 1990, Adv. Space Res., 10, 63

Murakami, T., Fujii, M., Hayashida, K., Itoh, M., Nishimura, J., *et al.* 1988, Nature, 335, 234

Murakami, T., *et al.* 1989, PASJ, 41, 405

Murakami, T., *et al.* 1990, A&A, 227, 451

Murakami, T., Inoue, H., Nishimura, J., van Paradijs, J., Fenimore, E., *et al.* 1991, Nature, 350, 592

Narayan, R., Paczyński, B., Piran, T. 1992, ApJ, 395, L83

Narayan, R., Piran, T., Shemi, A. 1991, ApJ, 379, L17

Narayan, R., Wallington, S. 1992, ApJ, 399, 368

Natarajan, P., *et al.* 1997, New Astronomy, 2, 471

Natarajan, P., Lynden-Bell, D. 1997, MNRAS, ???; astro-ph/9604084

Nemiroff, R.J. 1994, Comm. Astrophys., 17, 189

Nemiroff, R.J., *et al.* , 1993, ApJ, 414, 36

Nemiroff, R.J., Norris, J.P., Bonnell, J.T., Wickramasinghe, W.A.D.T., Scargle, J.D., Kouveliotou, C., *et al.* 1994, ApJ, 435, L133

Nemiroff, R.J., Norris, J.P., Kouveliotou, C., Fishmann, G.J., Meegan, C.A., Paciesas, W.S., 1994, ApJ, 423, 432

Nice, D.J., Sayer, R.W., Taylor, 1996, ApJ, 466, L87

Nishida, S., Lanza, A., Eriguchi, Y., Abramawicz, M.A. 1996, MNRAS, 278, L41

Nobili, L., Turolla, R., Zampieri, L. 1993, ApJ, 406, 686

Norris, J.P., *et al.* 1986, ApJ, 301, 213

Norris, J.P., Bonnell, J.T., Nemiroff, R.J., Scargle, J.D., Kouveliotou, C. *et al.* 1995, ApJ, 439, 542.

Norris, J.P., Bonnell, J.T., Watanabe, K. 1999, ApJ, in press; astro-ph/9807322

Norris, J.P., Nemiroff, R.J., Bonnell, J.T., Scargle, J.D., Kouvelioutou, C., Paciesas, W.S., Meegan, C.A., Fishman, G.J. 1996, ApJ, 459, 393

Norris, J.P., Nemiroff, R.J., Scargle, J.D., Kouveliotou, C., Fishman, G.J., *et al.* 1994, ApJ, 423, 432.

O'Dell, S.L., Puschell, J.J., Stein, W.A., Warner, J.W. 1978, ApJS, 38, 267

Odewahn S.D. *et al.* IAUC 7094

Ofek, E., Leibowitz, E., 1999, GCN210

Owen, F.N., Porcas, R.W., Mufson, S.L., Moffett, T.J. 1978, AJ, 83, 685

Paciesas, W.S., Meegan, C.A., Pendleton, G.N., Briggs, M.S., Kouveliotou, C., Koshut, T.M., Lestrade, J.P., McCollough, M., Brainerd, J.J., Hakkila, J., Henze, W., Preece, R.D., Mallozzi, R.S., and Fishman, G.J., 1998, ApJS, in press; at: *http* : *//www.batse.msfc.nasa.gov/data/grb/4bcatalog/*

Paczyński, B., 1986, ApJ, 308, L43

Paczyński, B. 1987, ApJ, 317, L51

Paczyński, B. 1990, ApJ, 363, 218

Paczyński, B. 1991, Acta Astron., 41, 257

Paczyński, B. 1992, Nature, 355, 521

Paczyński, B. 1995, PASP, 107, 1167

Paczyński, B. 1998, ApJ, 494, L45

Paczyński, B., Rhoads, J. 1993, ApJ, 418, L4

Paczyński, B., Xu, G. 1994, ApJ, 427, 709

Page, D.N., Hawking, S.W. 1976, ApJ, 206, 1

Palmer, D.M., *et al.* 1994, ApJ, 433, L77

Panaitescu, A., Mészáros, P. 1998, ApJ, 493, L31

Panaitescu, A., Mészáros, P., Rees, M. 1998, astro-ph/9801258

Papaloizou, J.C.B., Pringle, J.E. 1987, MNRAS, 225, 267

Papathanassiou, H. 1999, astro-ph/9903199

Parma, P., Cameron, R.A., de Ruiter, H.R. 1991, AJ, 102, 1960

Paul, J., *et al.* 1990 Proc. XXVIII COSPAR Meeting, The Hague. ??

Pedersen, H., *et al.* 1998, ApJ, 496, 311

Pendleton, C.N., *et al.* 1994, ApJ, 424, 540

Pendleton, C.N., Paciesas, W.S., Briggs, M.S., Preece, R.D., Mallozzi, R.S., Meegan, C.A., Horack, J.M., Fishman, G.J., Band, D.L., Matteson, J.L., Skelton, R.T., Hakkila, J., Ford, L.A., Kouvelioutou, C., Koshut, T.M. 1997, ApJ, 489, 175

Perna, R., Loeb, A. 1998, ApJ, in press; astro-ph/9810085

Petrosian, V., Efron, B. 1995, ApJ, 441, L37

Phinney, E.S. 1991, ApJ, 380, L17

Pian, E., Amati, L., Antonelli, L.A., Butler, R.C., Costa, E., Cusumano, G., *et al.* 1999, submitted to A&AS, Proc. of the Conference "Gamma-Ray Bursts in the Afterglow Era", held in Rome, 1998 November 3-6

Pian, E., *et al.* 1998a, ApJ, 492, L103

Pian, E., Frontera, F., Antonelli, L.A., Piro, L. 1998, GCN No. 69

Pilla, R., Loeb, A. 1998, ApJ, 494, L167

Piran, T. 1992, ApJ, 389, L45

Piran, T. 1999, Physics Report, in press; astro-ph/9810256

Piran, T., Shemi, A. 1993, ApJ, 403, L67

Piran, T., Shemi, A., Narayan, R. 1993, MNRAS, 263, 861

Piran, T., Singh, A., 1997, ApJ, 483, 552

Piro, L., *et al.* GCN203

Piro, L., Costa, E., Feroci, M., Frontera, F., Amati, L., Dal Fiume, D., Antonelli, L.A., Heise, J., in't Zand, J., Owens, A., Parmar, A.N., Cusumano, G., Vietri, M., Perola, G.C. 1999, ApJ Letter, submitted; astro-ph/9902013

Podsiadlowski, P., Rees, M.J., Ruderman, M. 1995, MNRAS, 373, 755

Podsiadlowski, P., Rees, M.J., Ruderman, M. 1995, MNRAS, 374, 430

Pollock, J.T., Hall, D.L., Ambruster, C., Usher, P.D. 1974, A&A, 30, 41

Pollock, J.T., Pica, A.J., Smith, A.G., Leacock, R.J., Edwards, P.L., Scott, R.L. 1979, AJ, 84, 1658

Popham, R., Woosley, S.E., Fryer, C. 1999, ApJ submitted; astro-ph/9807028

Portegies Zwart, S.F., Spreeuw, J.N. 1996, A&A, 312, 670

Portegies Zwart, S.F., Lee, C-H., Lee, H.K. 1999, astro-ph/9808191

Potter, D. 1973, Computational Physics, John Wiley & Sons, Chichester

Preece, R.B., Briggs, M.S., Mallozzi, R.S., Pendleton, G.N., Paciesas, W.S., Band, D.L. 1998, AJ, 506, L23

Pugliese, G., Falcke, H., Biermann, P.L. 1999, astro-ph/9903036

Puschell, J.J., Stein, W.A. 1980, ApJ, 237, 331

Pye, J.P., McHardy, I.M. 1983, MNRAS, 205, 875

Qian, Y.-Z., Woosley, S.E. 1996, ApJ, 471, 331

Quashnock, J. 1996, ApJ, 461, L69

Quashnock, J., Lamb, D., 1993, MNRAS, 265, L59

Rachen, J.P., Mészáros, P. 1998, corrected version: astro-ph/9811266; 1997, 4th Huntsville Symposium on Gamma-Ray Bursts, AIP Conf. Proc., no. 428, eds. C. Meegan, R. Preece, and T. Koshut, New York: AIP, p.776

Rasio, F.A., Shapiro, S.L. 1992, ApJ, 401, 226

Rasio, F.A., Shapiro, S.L. 1999, Class. Quant. Grav., in press; gr-qc/9902019

Rees, M.J., Mészáros, P. 1992, MNRAS, 258, 41P

Rees, M.J., Mészáros, P. 1994, ApJ, 430, L93

Reichart, D.E., Mészáros, p. 1997, ApJ, 483, 597

Rhoads, J.E., 1997, ApJ, 487, L1

Rhoads, J.E., 1998, Constraining Gamma-ray Burst Beaming, *preprint*

Rhoads, J.E. 1999, astro-ph/9903400

Richardson, L.F. 1910, Phil. Trans. Roy. Soc., 210, 307

Ricker, G. 1999, Present at the Conference on Gamma-Ray Bursts in the Afterglow Era, Rome; A&AS, in press.

Romano, G. 1972, M.S.A.I., 43, 309

Rood, H., Struble, M. 1996, MNRAS, in press

Ruderman, M. 1975, Ann. N.Y. Acad. Sci., 262, 164

Ruffert, M., Janka, H.-T. 1996, A&A, 307, L33

Ruffert, M., Janka, H.-T. 1998, A&A, 338, 535

Ruffert, M., Janka, H.-T. 1998a, A&A submitted; astro-ph/9809280

Ruffert, M., Janka, H.-T., Schaefer, G. 1996, A&A, 311, 532

Ruffert, M., Janka, H.-T., Takahashi, K., Schaefer, G.A., 1997, A&A, 319, 122

Rutledge, R.E., Hui, L., Lewin, W.H.C. 1995, MNRAS, 276, 753

Rybicki, G.B., Lightman, A.P. 1979, Radiative Processes in Astrophysics

Sagar, R., Pandey, A.K., Mohan, V., Yadav, R.K.S., Nilakshi, Bhattacharya, D., Castro-Tirado, A.J. 1999, astro-ph/9902196

Sahu, K., Livio, M., Petro, P., Bond, H.E., Macchetto, F.D., Galama, T.J., Groot, P.J., van Paradijs, J., Kouveliotou, C., 1997, ApJ, 489, L127

Sahu, K., *et al.* 1997a, Nature, 387, 476

Sari, R. 1997, ApJ, 489, L37

Sari, R. 1998, ApJ, 494, L49

Sari, R., Narayan, R., Piran, T. 1996, ApJ, 473, 204

Sari, R., Piran, T. 1995, ApJ, 455, L143

Sari, R., Piran, T. 1997, ApJ, 485, 270

Sari, R., Piran, T. 1999a, astro-ph/9901338

Sari, R., Piran, T. 1999, astro-ph/9902009

Sari, R., Piran, T. 1998, ApJ, 497, L41

Sari, R., Piran, T., Halpern, J.P. 1999, astro-ph/9903339

Sari, R., Piran, T., Narayan, R. 1998, ApJ, 497, L17

Schaefer, B.E. 1980, PASP, 92, 255

Schaefer, B.E. 1990, ApJ, 364, 590

Schaefer, B.E., Cline, T.L., Hurley, K.C., Laros, J.G. 1997, ApJ, 489, 693

Schaefer, B.E., Cline, T.L., Atteia, J.-L., Barat, C., Desai, U., Estulin, I.V., Evans, W.D., Fenimore, E.E., Hurley, K., *et al.* 1989, ApJ, 340, 455

Schaefer, B.E., Cline, T.L., Desai, U., Geegarden, B.J., Atteia, J.-L., Barat, C., Hurley, K., Niel, M., Evans, W.D., Fenimore, E.E., *et al.* 1987, ApJ, 313, 226

Schaefer, B.E., Palmer, D., Dingus, B.L., Schneid, E.J., Schoenfelder, V., Ryan, J., Winkler, C., Hanlon, L., Kippen, R.M., Connors, A. 1998, ApJ, 492, 696

Schartel, N., Andernach, H., Greiner, J. 1997, A&A, 323, 659

Schmidt, W.K.H. 1978, Nature, 271, 525

Schmidt, M., Higdon, J.C., Jueter, G. 1988, ApJ, 329, L85

Schneid, E., Bertsch, D., Fichtel, C., Hartman, R., Hunter, S., *et al.* 1992, A&A, 255, L13

Seguin, F.H. 1975, ApJ, 197, 745

Shapiro, S.L. 1998, PRD, 57, 908

Shaviv, N., Dar, A. 1995, MNRAS, 277, 287

Shemi, A. 1994, MNRAS, 269, 1112

Shemi, A., Piran, T. 1990, ApJ, 365, L55

Shklovskii, I.S., Mitrofanov, I.G. 1985, MNRAS, 212, 545

Shibata, M., Baumgarte, T.W., Shapiro, S.L. 1998, PRD, 58, 23002; gr-qc/9805026

Sigl, G., Schramm, D.N., Lee, S., Coppi, P., Hill, C.P. 1996, astro-ph/9605158

Sigurdsson, S., Rees, M.J. 1997, MNRAS, 284, 318

Sillanpää, A., *et al.* 1996, A&A, 315, L13

Sillanpää, A., Haarala, S., Korhonen, T. 1988, A&AS, 72, 347

Sillanpää, A., Haarala, S., Valtonen, M., Sundelius, B., Byrd, G.G. 1988a, ApJ, 325, 628

Sillanpää, A., Mikkola, S., Valtaoja, L. 1991, A&AS, 88, 225

Sitko, M.L., Schmidt, G.D., Stein, W.A. 1985, ApJS, 59, 323

Sitko, M.L., Stein, W.A., Zhang, Y.-X., Wisniewski, W.Z. 1983, PASP, 95, 724

Sitko, M.L., Sitko, A.K. 1991, PASP, 103, 160

Skamarock, W.C. 1989, Mon. Weather Rev., 117, 872

Smith, I.A., Lamb, D.Q. 1993, ApJ, 410, L23

Smoller, J., Tepmle, B. 1993, Comments Math. Phys., 156, 67

Soffita, P. *et al.* 1998, IAU Circ. No. 6884

Sokolov, V.V., Kopylov, A.I., Zharikov, S.V., Feroci, M., Nicastro, L., Palazzi, E. 1998, A&A, 334, 117

Sokolov, V.V., *et al.* 1999, GCN209

Spruit, H.C. 1998, A&A, in press; astro-ph/9811007

Stern, B.E. 1996, ApJ, 464, L111

Stern, B.E., Poutanen, J., Svensson, R. 1997, ApJ, 489, L41

Strittmatter, P.A., Serkowski, K., CarsWell, R., Stein, W,A., Merrill, K.M., Burbidge, E.M. 1972, ApJ, 175, L7

Stokes, G.H., *et al.* 1991, ApJ, 294, L21

Strohmayer, T.E., Fenimore, E.E., Miralles, J.A. 1994, ApJ, 432, 665

Synge, J.L. 1960, Relativity: The General Theory, North-Holland, Amsterdam.

Szuszkiewicz E., Miller J.C. 1997, MNRAS, 287, 165

Szuszkiewicz E., Miller J.C. 1998, MNRAS, 298, 888

Taam, R.E., Lin, D.N.C. 1984, ApJ, 287, 761

Takalo, L.O. 1991, A&A, 90, 161

Takalo, L.O., Sillanpää, A., Nilsson, K., *et al.* . 1992, A&AS, 94, 37

Tapia, S., Craine, E.R., Jahnson, K. 1976, ApJ, 203, 291

Tavani, M. 1998, ApJ, 497, L21

Taylor, G., *et al.* 1997, Nature, 389, 263

Taylor, J.H., Dewey, R.J. 1988, ApJ, 332, 770

Taylor, J.H., Weisberg, J.M. 1982, ApJ, 253, 908

Taylor, J.H., Weisberg, J.M. 1989, ApJ, 345, 434

Tegmark, M., Hartmann, D.H., Briggs, M.S., Hakkila, J., Meegan, C.A. 1996, ApJ, 466, 757

Terekhov, O.V., Denisenko, D.V., Lobachev V.A., Sunyaev R.A., Kovtun, A.V. 1994, AstrLett, 20, 265

Thompson, C. 1994, MNRAS, 270, 480

Thompson, C. 1996, 3th Huntsville Symp. on Gamma-Ray Bursts, AIP Conference Proceedings No384, eds. C. Kouvelioutou, M.F. Briggs, and G.J. Fishman, New York:AIP, p.897

Thompson, K.W. 1986, J. Flui Mech., 171, 365

Thorne, K.S. 1981, MNRAS, 194, 439

Thorsett, S.E. 1995, ApJ, 444, L53

Tosti G., *et al.* 1998, A&AS, 130, 109

Totani, T. 1997, ApJ, 486, L71

Totani, T. 1998, ApJ, 509, L81

Totani, T. 1998a, PRL, in press; astro-ph/9810207

Trimble, V., McFadden, L.-A. 1998, PASP, 110, 223

Tully, R.B. 1987, ApJ, 323, 1

Turolla, R., Nobili, 1988, MNRAS, 235, 1273

Tutukov, A.V., Yungelson, L.R. 1993, MNRAS, 260, 675

Tutukov, A.V., Yungelson, L.R. 1994, MNRAS, 268, 871

Ulrich, M.-H., Kinman, T.D., Lynds, C.R., Rieke, G.H. 1975, ApJ, 198, 261

Usov, V. 1992, Nature, 357, 472

Usov, V. 1994, MNRAS, 267, 1035

Usov, V. 1998, PRL, 80, 230

Usov, V., Chibisov, G. 1975, Sov. Astron., 19, 115

Usov, V., Smolsky, M.W. 1996, ApJ, 461, 858

van den Heuvel, E.P., Lorimer, D.R. 1996, MNRAS, 283, L37

van Paradijs, J., *et al.* 1997, Nature, 386, 686

Vermeulen, R.C., Cohen, M.H. 1994, ApJ, 430, 467

Veron, P., Veron, M.P. 1975, A&A, 39, 281

Veron, P., Veron, M.P. 1976, A&AS, 25, 287

Vietri, M. 1995, ApJ, 453, 883

Vietri, M. 1996, ApJ, 471, L95

Vietri, M. 1997, ApJ, 478, L9

Vietri, M., Stella, L. 1998, ApJ, 507, L45

Vikhlinin, A. 1998, ApJ,

Vitello, P., Salvati, M. 1976, Phys. Fluids, 19, 1523

Vrba, F.J., Hartmann, D.H., Jennings, M.C. 1995, ApJ, 446, 115

Wald, R.M. 1974, PRD, 10, 1680

Wang, L., Wheeler, J.C. 1998, ApJ, 504, L87

Wang, V.C., Lingenfelter, R.E. 1995, ApJ, 441, 747

Wasserman, I., Salpeter, E.E. 1994, ApJ, 433, 670

Waxman, E. 1995, ApJ, 452, L1

Waxman, E. 1995a, PRL,75, 386

Waxman, E. 1997, ApJ, 485, L5

Waxman, E. 1997a, ApJ, 489, L33

Waxman, E. 1997b, ApJ, 491, L19

Waxman, E., Bachall, J.N. 1997, PRL, 78, 2292

Waxman, E., Coppi, P. 1996, astro-ph/9603144

Waxman, E., Kulkarni, S.R., Frail, D.A. 1998, ApJ, 497, 288

Waxman, E., Piran, T. 1994, ApJ, 433, L85

Webb, J.R., Smith, A.G., Leacock, R.J., Fitzgibbons, G.L., Gombola, P.P., Shepherd, D.W. 1988, AJ, 95, 374

Weinberg, S. 1972, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity General Ralativity*, Wiley; New York.

Weistrop, D., Shaffer, D.B., Hintzen, P., Romanishin, W. 1985, ApJ, 292, 614

Wen, L., Panaitescu, A., Laguna, P. 1997, ApJ, 486, 919

White, R.S. 1993, Ap&SS, 208, 301

Wijers, R.A.M.J., Gamala, T.J. 1998, astro-ph/9805341

Wijers, R.A.M.J., *et al.* 1998, MNRAS, 294, 13

Wijers, R.A.M.J., Paczyński, B. 1994, ApJ, 437, L107

Wijers, R.A.M.J., Rees, M.J., Mézsáros, P. 1997, MNRAS, 288, L51

Wijers, R.A.M.J., van Paradijs, J. 1996, ApJ, 473, L25

Wild, L.A. 1996, PhD Thesis, University of Wales

Williams, L.L.R., Wijers, R.A.M.J. 1997, astro-ph/9701246

Wilson, J.R., Mathews, G.J., Marronetti, P. 1996, PRD, 54, 1317

Wing, R.F. 1973, AJ, 78, 684

Winkler, K.H. 1977, PhD Thesis, Max-Planck-Institute for Physics and Astrophysics

Wolf, M. 1916, Astron. Nachr., 202, 415

Wolszczan, A. 1991, Nature, 350, 688

Woods, E., Loeb, A. 1994, ApJ, 425, L63

Woods, E., Loeb, A. 1995, ApJ, 383, 292

Woodward, P., Colella, P. 1984, J. Comp. Phys., 54, 115

Woosley, S.E. 1993, ApJ, 405, 273

Woosley, S.E., Baron, E. 1992, ApJ, 391, 228

Woosley, S.E., Eastman, R.G., Schmidt, B.P. 1998, ApJ submitted; astro-ph/9806299

Xie, G.Z., Li, K.H., Bao, M.X., *et al.* 1987, A&AS, 67, 17

Xie, G.Z., LU, R.W., Zhou, Y., *et al.* 1988, A&AS, 72, 163

Xie, G.Z., Li, K.H., Liu, F.K., Lu, R.W., Wu, J.X., Fan, J.H., Zhou, Y.Y., Cheng, F.Z. 1992, ApJS, 80, 683

Xie, G.Z., Li, K.H., Cheng, F.Z., Hao, P.J., Li, Z.L., Lu, R.W., Li, G.H. 1990, A&A, 229, 329

Xie, G.Z., Li, K.H., Cheng, F.Z., Lu, R.W., Liu, F.K., Liu, B.F., Hao, P.J., Liu, Z.H. 1991, A&AS, 87, 461

Yadigaroglu, I.A., Halpern, J.P., Uglesich, R., Kemp, J. 1999, GCN242

Yi, I., Blackman, E.G. 1997, ApJ, 482, 383

Yoshida, A., *et al.* 1989, PASJ, 41, 509

Yoshida, A., *et al.* 1999, ApJ, in press

Zampieri, L., Turolla, R., and Treves, A. 1993, ApJ, 419, 311

Zatsepin, G.T., Kuzmin, V.A. 1966, Sov. Phys. JETP Lett., 4, 78

Zdziarski, A.A., Coppi, P.S., Lamb, D.Q. 1990, ApJ, 357, 149

Zekl, H., Klare, G., Appenzeller, I. 1981, A&A, 103, 342

Zel'dovich, Ya.B., Raizer, Yu.P. 1967, hysics of Shock Waves and High-Temperature Hydroynamic Phenomena, (eds.) W.D. Hayes, R.F. Probstein, Academic Press, New York

Zensus, J.A. 1997, ARA&A, 35, 607

Zharikov, S.V., Sokolov, V.V., Baryshev, Yu.V. 1998, A&A, in press; astro-ph/9804309

Zhu, J., Zhang, H.T. 1999, GCN204

Appendix: Polytropic Gas Spheres: an Approximate Analytic Solution of the Lane-Emden Equation

8.1 Introduction

We present here a calculation which, while not directly connected with the other material in the thesis, may be of interest.

As stellar encounters are not very important for galaxies, clusters of galaxies, or globular clusters (Binney & Tremaine 1987), the fundamental dynamics describing these systems is that of a collisionless system. In the general case the collisionless Boltzmann equation cannot be solved because it involves too many independent variables. However, we can get certain exact solutions of the collisionless Boltzmann equation for a subset of possible stellar-dynamical equilibria. The system with a polytropic state equation, corresponding to isotropic velocity dispersion tensors, is one of them. In stellar structure and evolution theory, the polytropic model also plays an important role (Chandrasekhar 1939). However, the solution of the Lane-Emden equation in the polytropic model for general index n cannot be given analytically but only numerically. Here we present a good analytic approximation of the solution.

For a polytropic system, the relation of pressure P and density ρ is given by

$$P = K\rho^{\gamma} \equiv K\rho^{(n+1)/n} \quad , \tag{A.1}$$

where K is the polytropic constant, γ is the adiabatic index, and n the polytropic index. K is fixed for a degenerate gas sphere (e.g. in a white dwarf or in a neutron star) but free in a non-degenerate system. In galactic dynamics n is larger than 1/2 (Binney & Tremaine 1987) which means that no polytropic stellar system can be homogeneous, and n ranges from 0 to ∞ in the case of the theory of stellar structure and evolution (Kippenhahn & Weigert 1990, Chandrasekhar 1939). With the polytropic

relation (A.1), hydrostatic equilibrium, and Poisson's equation for the gravitational potential, we derive the Lane-Emden equation (see, for instance, Binney & Tremaine 1987, Kippenhahn & Weigert 1990) for spherical symmetry:

$$\frac{d^2\omega}{d\xi^2} + \frac{2}{\xi}\frac{d\omega}{d\xi} = -\omega^n \quad . \tag{A.2}$$

The dimensionless variables ω and ξ are defined by

$$\xi = Ar, \quad A^2 = \frac{4\pi G}{(n+1)K}\rho_c^{(n-1)/n}, \quad \omega = \left(\frac{\rho}{\rho_c}\right)^{1/n} \tag{A.3}$$

where ρ_c is the density at the center of the sphere and G the Newtonian gravitational constant and ω corresponds to the dimensionless gravitational potential. We have excluded the isothermal case $n = \infty$, which we will discuss in detail in a later section.

The Lane-Emden equation must be solved with the original central conditions:

$$\omega(0) = 1, \quad \left(\frac{d\omega}{d\xi}\right)_{\xi=0} = 0, \tag{A.4}$$

which will ensure the regularity of the solution at the center. The solution gives ω as a function of ξ. From ω we can get the density profile ρ. Only for the three values of $n =$ 0, 1, and 5 can the solution be given in analytic form. Apart from these three cases the Lane-Emden equation has to be solved numerically. As the Lane-Emden equation has a regular singularity at $\xi = 0$, we expand $\omega(\xi)$ in a power series as

$$\omega(\xi) = 1 - \frac{1}{6}\xi^2 + \frac{n}{120}\xi^4 + \cdots \quad . \tag{A.5}$$

From (A.2) and (A.4) we have

$$\frac{2}{\xi}\frac{d\omega}{d\xi} = -\frac{2}{3} = 2\frac{d^2\omega}{d\xi^2}$$

at $\xi = 0$. We start with the approximation of taking the second derivative term as

$$\frac{\delta}{2}\frac{2}{\xi}\frac{d\omega}{d\xi}$$

to obtain a good approximate analytic solution of the Lane-Emden equation (Liu 1996). We give the analysis for the isothermal Lane-Emden equation in section (8.2), comparing our results with the numerical solution. The general case is dealt with in section (8.3) followed by discussion and conclusion in section (8.4).

8.2 The Lane-Emden Equation for an Isothermal Sphere

The Lane-Emden equation for an isothermal sphere (see Kippenhahn & Weigert 1990, Binney & Tremaine 1987) can be written as

$$\frac{d^2\omega}{d\xi^2} + \frac{2}{\xi}\frac{d\omega}{d\xi} = e^{-\omega} \quad , \tag{A.6}$$

where ω is related to the mass density at dimensionless radius ξ by

$$\rho = \rho_c e^{-\omega} \quad , \tag{A.7}$$

and ρ_c is the density at the origin. It is easy to find (see also Binney & Tremaine 1987) that

$$\omega = -\ln\left(\frac{2}{\xi^2}\right) \quad , \tag{A.8}$$

is one of the solutions of Eq. (A.6). This solution describes a model known as the singular isothermal sphere. Unfortunately, the singular isothermal sphere has infinite density at $\xi = 0$. To obtain a solution that is well behaved at the origin, equation (A.6) has to be integrated with the central conditions

$$\omega(0) = 0, \quad \left(\frac{d\omega}{d\xi}\right)_{\xi=0} = 0 \quad . \tag{A.9}$$

Its solution cannot be given by analytic expressions but only by numerical computation. The Lane-Emden equation has a regular singularity at $\xi = 0$. In order to understand the behavior of the solutions there, a power series expansion similar to (A.5) can be derived and has to be used. For large radius where the effect of the central conditions is very weak the solution should asymptotically approach the singular isothermal solution. The isothermal sphere consisting of an ideal gas has an infinite radius as well as an infinite mass.

At small radius, a useful approximation to $\rho(\xi)$ is the modified Hubble law (Binney & Tremaine 1987), which was introduced empirically by King (1962),

$$\left(\frac{\rho}{\rho_c}\right) = \frac{1}{\left(1+\frac{\xi^2}{9}\right)^{3/2}} \quad . \tag{A.10}$$

Comparing this relation with a numerical solution of (A.6) one can say that for $\xi \lesssim 5$ the relative error is less than 5 %. Expression (A.10) does not fit the isothermal profile well at $\xi \gtrsim 9$ as it approaches asymptotically to a logarithmic slope -3 and not -2, as is required by the isothermal profile.

Since

$$\left(\frac{d^2\omega}{d\xi^2}\right)_{\xi=0} = \frac{1}{3} = \frac{1}{2}\left(\frac{2}{\xi}\frac{d\omega}{d\xi}\right)_{\xi=0}$$

at $\xi = 0$ from (A.6), it is reasonable to approximate the second derivative term in (A.6) with

$$\frac{1}{2}\left(\frac{2}{\xi}\frac{d\omega}{d\xi}\right)$$

as our first approximation and to get a first order differential equation

$$\frac{3}{2}\frac{2}{\xi}\frac{d\omega}{d\xi} = e^{-\omega} \quad , \tag{A.11}$$

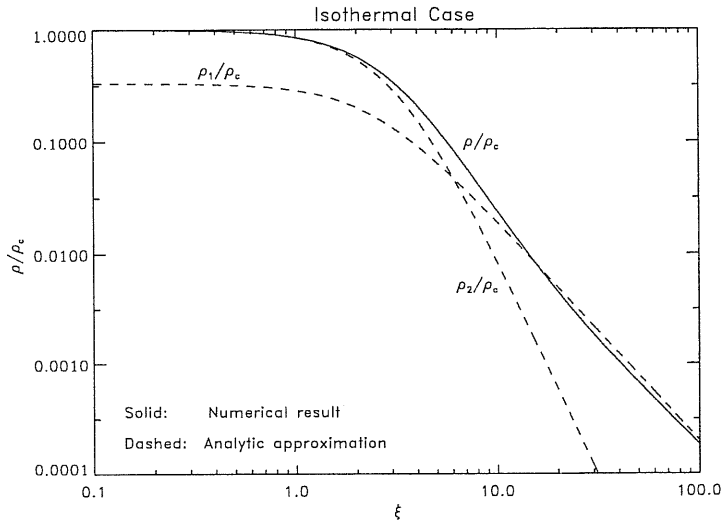

Figure A.1: The first order (dashed line indicated by ρ_1/ρ_c) and the second order (dashed line indicated by ρ_2/ρ_c) approximation of the density for the isothermal case. The first order approximation has been multiplied by 1/3 to show that it has the same tangent as the numerical result (solid line) for large ξ.

which can be integrated immediately, giving

$$\omega_1 = \ln\left(1+\frac{\xi^2}{6}\right) \quad , \tag{A.12}$$

and

$$\frac{\rho_1}{\rho_c} = e^{-\omega_1} = \frac{1}{1+\frac{\xi^2}{6}} \quad , \tag{A.13}$$

where we have used the condition (A.9) to determine the integration constant. The subscript 1 indicates a first approximation.

When ξ is very large, Eq. (A.12) becomes

$$\omega_1 \sim \ln\left(\frac{\xi^2}{6}\right) \quad , \tag{A.14}$$

which has a very similar profile to (A.8) and the same tangent as the numerical solution. Figure (A.1) shows the numerical solution of equation (A.6) and our first approximation (A.13), which shifted by 1/3 approximately reaches the singular solution for large ξ.

Taking the second derivative of ω_1 as our approximation to the second derivative term and replacing the right hand side term in equation (A.6) with (A.13), we obtain the second approximation to the solution of equation (A.6),

$$\omega_2 = 2\ln\left(1+\frac{\xi^2}{6}\right) - \frac{\xi^2}{6}\frac{1}{1+\frac{\xi^2}{6}} + \text{constant} \quad , \tag{A.15}$$

or

$$\frac{\rho_2}{\rho_c} = e^{-\omega_2} = \text{constant} \times \left(1 + \frac{\xi^2}{6}\right)^{-2} \exp\left(\frac{\xi^2/6}{1 + (\xi^2/6)}\right) \quad , \tag{A.16}$$

where subscript 2 indicates the second approximation. The integration constant in (A.15) is determined by the central condition (A.9) and is zero. Equation (A.16) as an approximation to the solution of (A.6) is good for small ξ. For large ξ, the approximation has an asymptotic behavior, $\rho_2/\rho_c \propto \xi^{-4}$, which does not fit the isothermal profile for the same reason as for the modified Hubble law discussed above. We also show the second approximation in Fig. (A.1).

As the first order approximation (A.12) is good for large ξ and the second order approximation (A.15) is good for small ξ, we combine them and obtain a more general approximation

$$\begin{aligned} \omega_0 &= \alpha\omega_2 + (1-\alpha)\omega_1 \\ &= (1+\alpha)\ln\left(1 + \frac{\xi^2}{6}\right) - \frac{\xi^2}{6}\frac{\alpha}{1 + (\xi^2/6)} + \text{constant} \quad , \end{aligned} \tag{A.17}$$

where $0 \leq \alpha \leq 1$, coming from the condition (A.9), which requires $\omega = 0$ at $\xi = 0$. For very large ξ, equation (A.17) asymptotically approaches

$$(1+\alpha)\ln\left(\xi^2/6\right) - \alpha + \text{constant} \quad ,$$

but does not approach equation (A.8). So, we change the constant term in (A.17) to

$$-\alpha \ln\left(1 + \xi^2/A\right)$$

to make equation (A.17) approach (A.8) for very large ξ. From (A.8) and (A.17), A takes the value $3^{1/\alpha}12e$, where e is the natural constant. Defining the general approximation as $\omega = \omega_0 + \Delta\omega$ and substituting it in equation (A.6), we get

$$\Delta\omega \sim \ln\left[1 + \left(2^{-\alpha} - 1\right)\frac{D\xi^2}{1 + D\xi^2}\right] \quad .$$

The final analytic approximation is now given by

$$\begin{aligned} \omega &= \omega_0 + \Delta\omega = (1+\alpha)\ln\left(1 + \frac{\xi^2}{6}\right) - \frac{\xi^2}{6}\frac{\alpha}{1 + (\xi^2/6)} \\ &\quad -\alpha\ln\left(1 + \frac{\xi^2}{3^{1/\alpha}12e}\right) + \ln\left[1 + \left(2^{-\alpha} - 1\right)\frac{D\xi^2}{1 + D\xi^2}\right] \quad , \end{aligned} \tag{A.18}$$

and

$$\frac{\rho}{\rho_c} = e^{-\omega} \quad , \tag{A.19}$$

where D is a constant depending on α. Equation (A.19) reaches the singular solution for $\xi \to \infty$. The discrepancy of (A.19) to the numerical solution is therefore at its

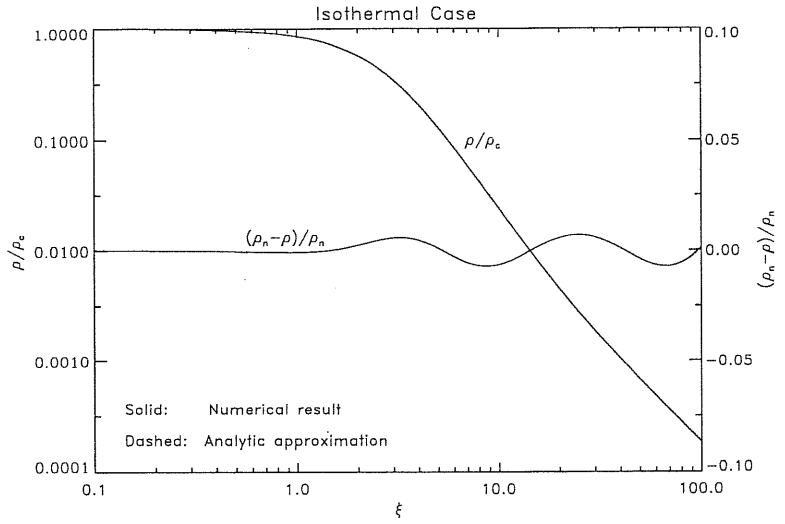

Figure A.2: The combination of the first and second approximations with $\alpha = 0.551$ and $D = 3.84 \times 10^{-4}$ with its relative error to the numerical result for the isothermal case. The relative error is less than 0.72 per cent for the whole solution. The approximation and numerical solution are plotted in the figure with different lines, but they overlap and cannot be distinguished.

largest value when ξ is intermediate. Adjusting α and D, we can modify the goodness of fit for intermediate ξ. One good combination is $\alpha = 0.551$ and $D = 3.84 \times 10^{-4}$, which makes the largest relative error of equation (A.19) to the numerical solution of (A.6) be 0.72 per cent. In Fig. (A.2), we compare the structure of densities obtained both numerically and analytically and show the relative error $(\rho_n - \rho)/\rho_n$, where ρ_n is the numerical computation solution of (6). For $\xi \to \infty$, the relative error approaches zero. As, in the isothermal case ω, has no physical meaning, we give only the result for density ρ. The largest error for ω is also about 0.72 per cent.

8.3 The Lane-Emden Equation for the General Case

We can apply a similar method for the general case. We first discuss the $n \neq 1$ case. The limit case $n = 1$ will be discussed later. For $n \neq 1$, equation (A.2) also has a singular solution as in the isothermal case

$$\omega = \left[\frac{(1-n)^2}{2\,(n-3)}\xi^2\right]^{1/(1-n)} . \tag{A.20}$$

For $1 < n \leq 3$, the singular solution has no physical meaning. When $n > 5$, the solution of equation (A.2) under central condition (A.4) should asymptotically approach (A.20) for $\xi \to \infty$. We approximately solve Eq. (A.2) with the same technique as for the

isothermal sphere. We approximate the second derivative term in Eq. (A.2) with

$$\frac{\delta}{2}\frac{2}{\xi}\frac{d\omega}{d\xi}$$

and get

$$\frac{2+\delta}{2}\frac{2}{\xi}\frac{d\omega}{d\xi} = -\omega^n \quad , \tag{A.21}$$

where δ is a constant; δ can affect the goodness of fit of the approximation strongly for $n < 1$ but not for $n > 1$. We can however get a very good approximation by fixing δ and adjusting other parameters (see below) for any n. Therefore, we will always take $\delta = 1$. Integrating equation (A.21), we get

$$\omega_1 = \left(1 + A_n\xi^2\right)^{1/(1-n)} \quad , \tag{A.22}$$

where A_n is $(n-1)/6$. Here, we have used the central condition (A.4) to determine the integration constant. However, this first analytic approximation does not have the right behavior for large ξ. Equation (A.22) for any $n \geq 1$ cannot reach zero at finite ξ and give a finite radius. Integrating equation (A.2) gives

$$\omega = -\int\frac{\xi}{2}\omega^n\,d\xi - \int\frac{\xi}{2}\frac{d^2\omega}{d\xi^2}\,d\xi \quad . \tag{A.23}$$

Substituting (A.22) in the right-hand side of equation (A.23) and integrating, we get our second approximate analytic solution,

$$\omega_2 = \text{constant} + 2\left(1 + A_n\xi^2\right)^{1/(1-n)} + \frac{\xi^2}{6}\left(1 + A_n\xi^2\right)^{n/(1-n)} \quad , \tag{A.24}$$

where the constant is determined by the central condition (A.4) as -1. This equation, however, reaches zero at finite ξ for any finite n and gives a radius smaller than the one obtained from numerical computation.

Since solution (A.22) gives too large values and (A.24) gives too small values for large ξ, we construct a more general approximation for ω as a linear combination of equations (A.22) and (A.24)

$$\omega_0 = \alpha\omega_2 + (1-\alpha)\omega_1 = -\alpha + (1+\alpha)(1+A_n\xi^2)^{1/(1-n)} + \frac{\alpha\xi^2}{6}(1+A_n\xi^2)^{n/(1-n)} \quad . \tag{A.25}$$

We would like to point out that if we keep δ in (A.21) as a free parameter, then equation (A.25) gives a good approximation to the solution of (A.2) at level < 1 per cent with proper δ and α for $n < 1$. In this case δ varies from 0 to 1. We note that (A.25) asymptotically reaches the constant $-\alpha$ and not zero when $\xi \to \infty$ for $n \geq 5$. As (A.25) should approach the singular solution at very large ξ for $n \geq 5$ (I thank Dr. P. Eggleton for pointing out this) as in the isothermal case, we change the constant $-\alpha$

in equation (A.25) to $-\alpha(1+B_n\xi^2)^{1/(1-n)}$ and let (A.25) for $\xi \to \infty$ be equal to the singular solution (A.20) when $n \gg 5$ and to zero when $n = 5$. We get

$$B_n = \begin{cases} A_n \alpha^{n-1} \left\{ 1 + \alpha \frac{n}{n-1} - \left[\frac{2(n-5)}{9(n+1)} \right]^{1/(n-1)} \right\}^{1-n} & , \quad \text{for } n \geq 5 \quad ; \\ \frac{n(n-1)}{(n+1)^2} \frac{6}{5} \left(\frac{4\alpha}{4+5\alpha} \right)^4 & , \quad \text{for } n < 5 \quad ; \end{cases} \tag{A.26}$$

where we have considered the conditions that B_n should equal zero at $n = 0$ and 1 and continue at $n = 5$ to get the expression for $n < 5$. With B_n given by (A.26), (A.25) gives a good approximation with relative error less than 5 per cent for $n \lesssim 5$ and about 1 per cent for $n \gg 5$. For $n = 1.5$, as an example, the relative error is less than about 1 per cent when $\alpha = 0.44$.

So, if our general solution were

$$\omega = \omega_0 + \Delta\omega \quad , \tag{A.27}$$

$\Delta\omega$ would be very small. From equations (A.2) and (A.25) as well as (A.26), we get at very large ξ for $n = 5$ and at $\xi \to \xi_n$ for $n < 5$,

$$\Delta\omega \sim \frac{1}{\xi} \quad , \tag{A.28}$$

where $\omega_0 \sim 0$. In order to avoid the singularity of $\Delta\omega$ at $\xi = 0$, we set

$$\Delta\omega = \frac{C_n \xi^{2\beta-1}}{(1 + D_n \xi^{\beta})^2} \quad , \tag{A.29}$$

where C_n and D_n are small constants and the index β determined from equation (A.2) is

$$\beta = 6.47 - 7.01\beta_1 + 5.53\beta_1^2 - 25.63\beta_2 + 49.42\beta_2^2 - 26.88\beta_2^3 \quad , \tag{A.30}$$

where $\beta_1 = 1/\left[1 + (n-5)^2\right]$ and $\beta_2 = 1/[1 + (n-3)^2]$. With the same procedure and the same reason as for $n \leq 5$, we have for $n > 5$

$$\Delta\omega \simeq \frac{E_n F_n \xi^{\eta + [2/(1-n)]}}{1 + F_n \xi^{\eta}} + \frac{C_n \xi^{2\beta-1}}{(1 + D_n \xi^{\beta})^2} \quad , \tag{A.31}$$

where the index η is $(n-2)/(n-4)$ and the coefficient E_n determined from equation (A.2) is

$$\begin{aligned} E_n \simeq & \frac{1}{n-4} A_n^{1/(1-n)} \left[\frac{2(n-5)}{9(n+1)} \right]^{(5-n)/(n-1)} \\ & \times \left\{ \left[\frac{n-3}{3(n-1)} \right]^{(n-4)/(n-1)} - \left[\frac{2(n-5)}{9(n+1)} \right]^{(n-4)/(n-1)} \right\} \quad , \end{aligned} \tag{A.32}$$

with which our approximation always asymptotically approaches the singular solution (A.20) at large ξ for $n > 5$. To fit the solution for intermediate ξ, the parameter F_n can be given as

$$F_n \simeq \begin{cases} \frac{(n-5)^6}{(n-3)^5(4n+50)}, & \text{for } n \geq 5 \quad ; \\ 0, & \text{for } n < 5 \quad . \end{cases} \tag{A.33}$$

Finally, we configure a general result for any n

$$\begin{aligned} \omega &= -\alpha(1+B_n\xi^2)^{1/(1-n)} + (1+\alpha)(1+A_n\xi^2)^{1/(1-n)} \\ &\quad + \frac{\alpha}{6}\xi^2(1+A_n\xi^2)^{n/(1-n)} + \frac{E_nF_n\xi^{\eta+[2/(1-n)]}}{1+F_n\xi^\eta} + \frac{C_n\xi^{2\beta-1}}{(1+D_n\xi^\beta)^2} \quad , \end{aligned} \tag{A.34}$$

where η is $(n-2)/(n-4)$, A_n is $(n-1)/6$, B_n is given by (A.26), E_n is given by (A.32) and β is given by (A.30). The free parameters α, F_n, C_n and D_n should be determined by fitting the numerical solution of equation (A.2). Equation (A.34) with proper values of the parameters α, C_n, D_n and F_n can give a very good approximation solution of equation (A.2) under the central condition (A.4).

For $n \gg 5$, equation (A.34) tends to the singular solution as $\xi \to \infty$. In this case, (A.34) always approximates the solution of equation (A.2) with error < 0.1 per cent. Especially when $n > 10$, the last term, and when $n \gtrsim 500$ the last two terms, in equation (A.34) are not important and can be ignored and the approximation is still within the level 0.1 per cent level. As n decreases, (A.34) slowly deviates from the singular solution to finally reach $\omega = 1/(1+\xi^2/3)^{1/2}$ at ξ very large for $n = 5$ where the sum of the first three terms in (A.34) approaches zero in accordance with (A.25) and (A.26). Equation (A.34) reaches zero at infinity for $n \geq 5$ and at finite ξ_n for $n < 5$. All the properties of (A.34) discussed above are also properties of the exact solution of (A.2). [for a detailed discussion of the general properties of the solution of equation (A.2), see Binney & Tremaine (1987), Kippenhahn & Weigert (1990).]

Equation (A.34) is the exact solution for $n = 0$. For general n, we have to adjust α, C_n, D_n and F_n to get a good approximation. Equation (A.34) gives a good approximation to the solution of equation (A.2) for small and large ξ, which is insensitive to the parameters within proper ranges of values. For intermediate ξ, however, special α, C_n, D_n and F_n are needed to reduce the error of the approximation. In Fig. (A.3) we show potential ω, density ρ and relative error $(\omega_n - \omega)/\omega_n$ as functions of radius ξ for $n = 3/2$ and for $n = 5$, respectively. For $n = 3/2$, the relative error is larger near the finite radius ξ_n and oscillates around the mean value zero. Such oscillating behavior also exists for $n = 5$. In fact, the oscillation always exists for any n. We give the values of the parameters used to get Fig. (A.3) and the largest relative fitting errors ϵ_{max} in Table (A.1). In Table (A.1) we also give the values of the parameters for several polytropic indices n and the largest fitting errors ϵ_{max}, as well as the best-fitting value

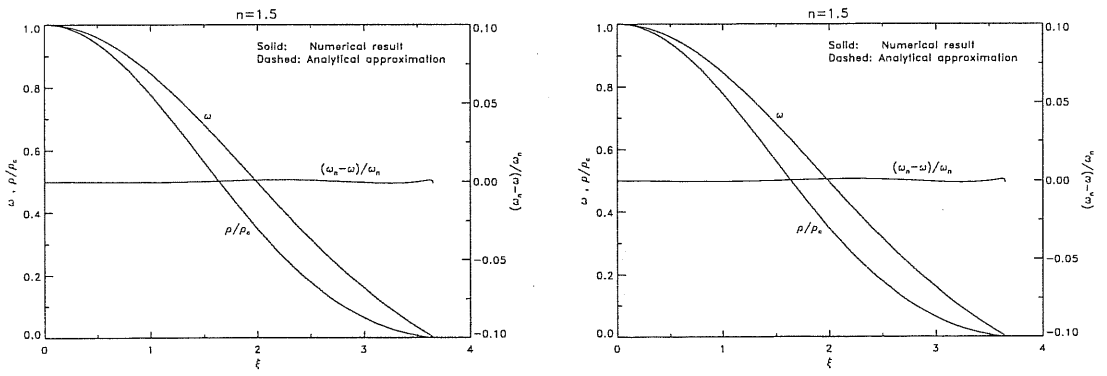

Figure A.3: The analytic approximation for $n = 3/2$ (*left*) and for $n = 5$ (*right*). The figure shows the variation of potential ω and density ρ with dimensionless radius ξ. The relative approximation error of ω to exact solution ω_n is also shown. The largest errors in the whole domain of the solutions are 0.12 per cent for the case $n = 3/2$ and 0.34 per cent for the case $n = 5$, respectively. In the case $n = 5$, when $\xi \to 0$ or $\xi \to \infty$, the difference approaches zero. As for the isothermal case, the approximation and exact solution are plotted in the figure with different lines, but they overlap and cannot be distinguished.

for F_n. With the values of α, C_n and D_n and the value of F_n given by equation (A.33), equation (A.34) can also give an approximation at the level $\epsilon_{max} < 1$ per cent.

For $n < 5$, the relative error at $\xi \to \xi_n$ is larger as the numerical solution ω_n approaches zero at the surface of the sphere. Equation (A.34) with B_n given by (A.26) reaches zero at finite ξ and is the exact solution for $n = 0$. For $n \geq 5$, the error is at its maximum when ξ is in the range of around 1–20, depending on index n. When ξ is very large, the approximation approaches the singular solution. With increasing distance away from the center of the gas sphere, the effect of the central boundary condition (A.4) on the structure of the sphere decreases. Therefore, the solution under the central boundary condition (A.4) approaches the singular solution. The approximation is very good even for $\xi \to \infty$. For $n > 10$, C_n can be set to zero. For $n > 500$, E_n is less than 10^{-3} and $\Delta\omega$ given by equation (A.31) becomes unimportant.

For $n = 1$, equation (A.2) under the central boundary condition (A.4) has a simple analytic solution:

$$\omega = \frac{\sin(\xi)}{\xi} \quad . \tag{A.35}$$

Dealing with equation (A.2) for the case $n = 1$ with the same procedure as for the general case n, we can get a highly accurate approximate solution which can also be obtained from the limit of equation (A.34) for $n \to 1$. In Table (A.1), we give a set of values of C_n and D_n and the largest error of equation (A.34) for $n = 1$. For $n = 1$, the largest relative error of the approximation is only 0.15 per cent.

Table A.1: Parameters for several polytropic indices n: α is the mixing parameter of the first and second approximations; ϵ_{max} (given in per cent) is the maximum relative error of the analytic approximation to the numerical calculation result. For other parameters, see the text.

n	α	C_n	D_n	F_n	$\epsilon_{max}(\%)$
0.5	0.5	1.30659×10^{-4}	2.3×10^{-4}	0	0.37
1	0.455	1.27746×10^{-4}	0	0	0.15
1.5	0.481	4.61841×10^{-4}	1.1×10^{-3}	0	0.12
2	0.512	8.3218×10^{-4}	2.56×10^{-2}	0	0.28
3	0.53	5.56215×10^{-4}	2.745×10^{-2}	0	0.60
5	0.545	1.9415×10^{-3}	3.348×10^{-2}	0	0.34
6	0.56	2.0×10^{-4}	3.0×10^{-2}	5.3×10^{-5}	0.50
7	0.526	1.53×10^{-7}	1.1×10^{-3}	7.40×10^{-4}	0.39
8	0.540	6.1×10^{-8}	1.0×10^{-3}	2.7×10^{-3}	0.35
9	0.546	3×10^{-8}	9×10^{-4}	5.4×10^{-3}	0.26
10	0.538	1.12×10^{-8}	9×10^{-4}	8.91×10^{-3}	0.24
20	0.613	0	$\cdots$	4.2×10^{-2}	0.12
50	0.65	0	$\cdots$	0.101	0.060
500	0.58	0	$\cdots$	0.19	0.021

8.4 Discussion and Conclusion

We have given analytic approximate solutions to the Lane-Emden equations for the isothermal sphere in Sec. (8.2) and for general polytropic case in Sec. (8.3). From the analytic results we can derive some useful relations.

From (A.3) and (A.2) for the general case and from (A.7) and (A.6) for the isothermal case, the mass contained in a sphere of radius r may be given as

$$m = \int_0^r 4\pi r^2 \rho \, dr = 4\pi \rho_c r^3 \left(\pm \frac{1}{\xi} \frac{d\omega}{d\xi} \right) \quad , \tag{A.36}$$

where the $-$ sign is for the general case and the $+$ is for the isothermal case. Dimensional r and non-dimensional ξ are related, for the general case, by equation (A.3) and for the isothermal case by

$$\xi = \left(\frac{4\pi G \rho_c}{K} \right)^{1/2} r = Ar \quad . \tag{A.37}$$

For the general case and $n \neq 1$, it follows from (A.34) that

$$\begin{aligned} m(r) \;=\; & 4\pi \rho_c r^3 \left[\frac{2\alpha B_n}{1-n} (1 + B_n \xi^2)^{n/(1-n)} + \frac{1}{3} (1 + A_n \xi^2)^{n/(1-n)} \right. \\ & \left. + \frac{n\alpha}{18} \xi^2 (1 + A_n \xi^2)^{[n/(1-n)]-1} - E_n F_n \xi^{\eta + [2/(1-n)]-2} \, \frac{\eta + \frac{2}{1-n} + \frac{2}{1-n} F_n \xi^\eta}{(1 + F_n \xi^\eta)^2} \right] \end{aligned}$$

$$-C_n\xi^{2\beta-3}\,\frac{2\beta-1-D_n\xi^{\beta}}{(1+D_n\xi^{\beta})^3}\Bigg] \quad . \tag{A.38}$$

For $n < 5$, the star has a finite radius R or ξ_n, which is calculated from (A.34) by setting ω to zero, and a finite mass M. From (A.38) we can get the expression for the mass M from which the central density ρ_c is determined. However, for $n \geq 5$ the polytropic gas sphere has infinite radius and infinite mass. In this case we cannot estimate the central density directly. From equations (A.1) and (A.3) we have

$$P = K\rho_c^{(n+1)/n}\omega^{n+1} \quad , \tag{A.39}$$

where ω is given by (A.34), K is determined by equation (A.3) and A is given by $A = \xi_n/R$ for $n < 5$. For an ideal gas the temperature is given as

$$T = \frac{\mu m_p}{k}\frac{P}{\rho} = \frac{\mu m_p}{k}K\rho_c^{1/n}\omega \quad , \tag{A.40}$$

where μ is the mean molecular weight, m_p is the mass of atomic Hydrogen, and k is the Boltzmann constant. From the definition of potential energy within radius r of the sphere

$$E_g \equiv -\int_0^m \frac{Gm'}{r'}\,dm' \quad , \tag{A.41}$$

the polytropic relation (A.1), and the hydrostatic equilibrium equation, we get for $n \neq 5$

$$E_g = -\frac{3}{5-n}\frac{Gm^2}{r} - \frac{3}{5-n}\Phi m - \frac{n+1}{5-n}4\pi r^3 P \quad , \tag{A.42}$$

where Φ is gravitational potential, given by

$$\Phi = -(n+1)K\rho^{1/n} \quad . \tag{A.43}$$

This means that for $n < 5$ the polytropic sphere has finite radius and the potential is set to zero at the surface. For $n \geq 5$ the potential becomes zero at infinity.

In polytropic models, $n = 3$ ($\gamma = 4/3$) and $n = 3/2$ ($\gamma = 5/3$) are two important cases. We have given $n = 3/2$ as an example before. Here, as another example, we construct a polytropic model of index 3 for the sun ($M = 1.989 \times 10^{33}$ g and $R = 6.96 \times 10^{10}$ cm) and compare our results with those from numerical computation. For $n = 3$, from Table (A.1) and equations (A.26) and (A.30), we have $\alpha = 0.53$, $\beta = 2.199$, $B_n = 4.648 \times 10^{-3}$, $C_n = 5.56215 \times 10^{-4}$ and $D_n = 2.745 \times 10^{-2}$. Then we get $\xi_3 = 6.897$ from (A.34) and $A = \xi_3/R = 9.895 \times 10^{-11}$. Substituting these in equation (A.38) we get central density $\rho_c = 76.96\ \mathrm{g\,cm^{-3}}$, $K = 3.87 \times 10^{14}$ from equation (A.3) and, consequently, $P_c = 1.26 \times 10^{17}$ dyn/cm^2 from equation (A.39) as $\xi = 0$ at the center. For the ideal gas $\mu = 0.62$ we get from equation (A.40) the central temperature $T_c = 1.2 \times 10^7$ K. From Kippenhahn and Weigert (1990,

p.179), numerical computation for this model gives $\xi_3 = 6.897$, $\rho_c = 76.39$ g cm^{-3}, $P_c = 1.24 \times 10^{17}$ dyn/cm^2 and $T_c = 1.2 \times 10^7$ K. We see that our approximation for $n = 3$ gives a quite good result at an error of level of 0.7 per cent for density, 1.6 per cent for pressure and 0 per cent for temperature.

For the isothermal sphere, the mass from equations (A.18) and (A.36) is

$$m = 8\pi\rho_c r^3 \left[\frac{1+\alpha}{6+\xi^2} - \frac{6\alpha}{(6+\xi^2)^2} - \frac{\alpha}{3^{1/\alpha}12e+\xi^2} + \frac{D(2^{-\alpha}-1)}{(1+2^{-\alpha}D\xi^2)(1+D\xi^2)}\right] \tag{A.44}$$

where $\alpha = 0.551$ and $D = 3.84 \times 10^{-4}$. When $\xi \gg \sqrt{6}$ the mass changes with radius as $m \propto r$. The pressure is given by polytropic relation (A.1) for $n = \infty$. The constant K is related to the central density ρ_c by (A.37). However, as the isothermal sphere does not have a finite radius, the constant A cannot be determined as in the case $n < 5$. From definition (A.41) and the hydrostatic equilibrium equation, we find the potential energy within radius r for the isothermal sphere

$$E_g = 4\pi r^3 K\rho - 3Km \quad , \tag{A.45}$$

where ρ is given by (A.19) and m by (A.44).

We have given a good analytic approximate solution of the Lane-Emden equation with which we have obtained analytic expressions for the mass contained within radius r, the pressure, temperature, and gravitational potential energy within radius r (Liu 1996). It would be interesting to apply the approximation in modeling the structures of stars and galaxies.

After this work had been done, Natarajan & Lynden-Bell (1997) also obtained an analytic approximation to isothermal sphere with a similar technique. However, they were only interested in the central part of the sphere and their approximation is not good at large radius. In this case, they could give an analytic approximation with a simpler form.