

The physics of gamma-ray bursts

Tsvi Piran*

Racah Institute for Physics, The Hebrew University, Jerusalem, 91904, Israel

(Published 28 January 2005)

Gamma-ray bursts (GRB's), short and intense pulses of low-energy γ rays, have fascinated astronomers and astrophysicists since their unexpected discovery in the late sixties. During the last decade, several space missions—BATSE (Burst and Transient Source Experiment) on the Compton Gamma-Ray Observatory, BeppoSAX and now HETE II (High-Energy Transient Explorer)—together with ground-based optical, infrared, and radio observatories have revolutionized our understanding of GRB's, showing that they are cosmological, that they are accompanied by long-lasting afterglows, and that they are associated with core-collapse supernovae. At the same time a theoretical understanding has emerged in the form of the fireball internal-external shocks model. According to this model GRB's are produced when the kinetic energy of an ultrarelativistic flow is dissipated in internal collisions. The afterglow arises when the flow is slowed down by shocks with the surrounding circumburst matter. This model has had numerous successful predictions, like the predictions of the afterglow itself, of jet breaks in the afterglow light curve, and of the optical flash that accompanies the GRB's. This review focuses on the current theoretical understanding of the physical processes believed to take place in GRB's.

CONTENTS

I. Introduction	1144	C. Relativistic beaming and the patchy shell model	1162
II. Observations	1145	V. Physical Processes	1163
A. Prompt emission	1145	A. Relativistic shocks	1163
1. Spectrum	1145	B. Particle acceleration	1164
2. Temporal structure	1147	C. Synchrotron	1164
3. Populations	1148	1. Frequency and power	1165
a. Long and short bursts	1148	2. The optically thin synchrotron spectrum	1165
b. X-ray flashes	1149	a. Fast cooling ($\gamma_{e,c} < \gamma_{e,min}$)	1166
c. Polarization	1149	b. Slow cooling ($\gamma_{e,c} > \gamma_{e,min}$)	1166
d. Prompt optical flashes	1150	3. Synchrotron self-absorption	1167
e. The GRB-afterglow transition—observations	1150	D. Inverse Compton scattering	1168
B. The afterglow	1150	E. Quasithermal Comptonization	1169
1. The x-ray afterglow	1151	F. Polarization from relativistically moving sources	1169
2. Optical and IR afterglow	1151	VI. The GRB and its Prompt Emission	1171
3. Dark GRB's	1152	A. Internal vs external shocks	1172
4. Radio afterglow	1153	1. General considerations	1172
C. Hosts and distribution	1153	2. Caveats and complications	1172
1. Hosts	1153	3. External shocks on a clumpy medium	1173
2. The spatial distribution	1154	4. The shotgun model	1173
3. GRB rates and the isotropic luminosity function	1154	5. Relativistic turbulence	1174
4. Association with supernovae	1156	B. Internal shocks	1174
a. GRB 980425 and SN98bw	1156	1. Hydrodynamics of internal shocks	1174
b. Red bumps	1156	2. The efficiency of internal shocks	1174
c. GRB 030329 and CN 2003dh	1157	3. Light curves from internal shocks	1175
d. X-ray lines	1157	C. External shocks	1175
D. Energetics	1157	1. Hydrodynamics	1175
III. The Global Picture—Generally Accepted Ingredients	1160	2. Synchrotron spectrum from external shocks	1177
IV. Relativistic Effects	1161	a. Forward shock	1178
A. Compactness and relativistic motion	1161	b. Reverse shock	1178
B. Relativistic time effects	1162	D. The transition from internal shocks to external shocks	1178
		E. Prompt polarization	1179
		VII. The Afterglow	1179
		A. Relativistic blast waves and the Blandford-McKee solution	1179
		B. Light curves for the “standard” adiabatic synchrotron model	1180
		C. Light curves for the early radiative phase	1182

*Electronic address: tsvi@phys.huji.ac.il

D. Light curve during the Newtonian transition	1182
E. Generalizations: I. Winds	1183
F. Generalizations: II. Energy injection and refreshed shocks	1183
G. Generalizations: III. Inhomogeneous density profiles	1184
1. The light curve of a Blandford-McKee solution	1184
2. The light curve with a variable density or energy	1185
H. Generalizations: IV. Jets	1186
I. Generalizations: V. Angle-dependent jets and the universal structured jet model	1188
J. Afterglow polarization—A tool that distinguishes between the different jet models	1190
K. Orphan afterglows	1191
1. Optical orphan afterglow	1191
2. Radio orphan afterglow	1193
L. Generalizations: VI. Additional physical processes	1193
1. Preacceleration	1193
2. Neutron decoupling and decay	1195
VIII. Additional Emission from GRB's	1195
A. TeV γ rays	1195
B. Neutrinos	1196
C. Cosmic rays and ultrahigh-energy cosmic rays	1196
D. Gravitational radiation	1197
1. Gravitational radiation from neutron star mergers	1197
2. Gravitational radiation from collapsars	1197
3. Gravitational radiation from a supranova	1198
4. Gravitational radiation from a gamma-ray burst	1198
IX. Models of Inner Engines	1198
A. Black-hole accretion	1199
B. The pulsar model	1199
C. Rotating black holes and the Blandford-Znajek mechanism	1200
D. The collapsar model	1200
E. The supranova model	1200
F. Merging neutron stars	1201
X. Open Questions and Future Prospects	1202
Acknowledgments	1203
References	1203

I. INTRODUCTION

Gamma-ray bursts (GRB's) are short and intense pulses of soft γ rays. The bursts last from a fraction of a second to several hundred seconds. GRB's arrive from cosmological distances from random directions in the sky. The overall observed fluences range from 10^{-4} ergs/cm² to 10^{-7} ergs/cm² (the lower limit depends, of course, on the characteristics of the detector and not on the bursts themselves). This corresponds to isotropic luminosity of 10^{51} – 10^{52} ergs/sec, making GRB's the most luminous objects in the sky. However, we know today that most GRB's are narrowly beamed and the corresponding energies are “only” around 10^{51} ergs (Frail *et al.*, 2001; Panaitescu and Kumar, 2001; Piran *et al.*, 2001), making them comparable to supernovae in total energy release.

GRB's are followed by an afterglow—lower-energy, long-lasting emission in the x-ray, optical, and radio wavelengths. The radio afterglow has been observed in some cases several years after the bursts. Accurate afterglow positions have made possible the identification of host galaxies in almost all cases when afterglow was detected, and this in turn enabled the determination of the corresponding redshifts, which range from 0.16 (or possibly even down to 0.0085) to 4.5. Within the host galaxies there is evidence that (long-duration) GRB's arise within star-forming regions, and there is evidence that they follow the star-formation rate.

While not all observed features are understood there is an overall agreement between the observations and the fireball model. According to the fireball model GRB's are produced when the kinetic energy of an ultrarelativistic flow is dissipated. The GRB itself is produced by internal dissipation within the flow while the afterglow is produced via external shocks with the circumburst medium. I shall focus in this review on this model.

The numerous observations of GRB's and of their afterglows constrain the fireball model that describes the emitting regions. The evidence on the nature of the inner engine that powers the GRB and produces the ultrarelativistic flow is, however, indirect. The energetic requirements and the time scales suggest that GRB's involve the formation of a black hole via a catastrophic stellar collapse event or possibly a neutron star merger. Additional indirect evidence arises from the requirement of the fireball model of long (several dozen seconds) activity of the inner engine. This hints at an inner engine built on an accreting black hole. On the other hand, the evidence of association of GRB's with star-forming regions indicates that GRB progenitors are massive stars. Finally, the appearance of supernova bumps in the afterglow light curve (most notably in GRB 030329)¹ suggests an association with supernovae and stellar collapse.

I review here the theory of GRB, focusing on the fireball internal-external shock model. I begin in Sec. II with a brief discussion of the observations. I turn in Sec. III to some generally accepted properties of GRB models, such as the essential ultrarelativistic nature of this phenomenon. Before turning to a specific discussion of the fireball model I review in Sec. IV several relativistic effects and in Sec. V the physical processes, such as synchrotron emission or particle acceleration in relativistic shocks, that are essential ingredients of this model. In Sec. VI I turn to a discussion of prompt emission and the GRB. In Sec. VII I discuss modeling of the afterglow emission. I consider other related phenomenon—such as TeV γ -ray emission, high-energy neutrinos, ultrahigh-

¹Because typical experiments detect bursts at a rate no more than one every few days, it is convenient to name them according to the date on which they occur, according to the system YYMMDD. The burst seen on 29 March 2003 is therefore called GRB 030329.

energy cosmic rays, and gravitational radiation—in Sec. VIII. Finally, I turn in Sec. IX to an examination of different “inner engines” and their operation. I conclude with a discussion of open questions and observational prospects.

While writing this review I realized how large the scope of this field is and how difficult it is to cover all aspects of this interesting phenomenon. Some important aspects had to be omitted. I also did not attempt to give a complete historical coverage of the field. I am sure that inadvertently I have missed many important references. I refer the reader to several other recent review papers (Fishman and Meegan, 1995; Piran, 1999, 2000; van Paradijs *et al.*, 2000; Mészáros, 2001, 2002; Galama and Sari, 2002; Hurley, Sari, and Djorgovski, 2002) that discuss these and other aspects of GRB theory and observations from different points of view.

II. OBSERVATIONS

I begin with a short review of the basic observed properties of GRB's. A complete review of the observations would require a whole paper by itself. I therefore refer the reader to several review papers for a more detailed treatment of the observations (Fishman and Meegan, 1995; van Paradijs *et al.*, 2000; Galama and Sari, 2002; Hurley, Sari, and Djorgovski, 2002). I divide this section into three parts, beginning with prompt emission—the GRB itself. I continue with properties of the afterglow

and conclude with a discussion of the rates of GRB's, the location of the bursts within their host galaxies, and the properties of the host galaxies.

A. Prompt emission

A GRB consists of both γ rays and any lower-energy emission that occurs simultaneously with them. This includes x-ray emission that generally accompanies the γ -ray emission as a low-energy tail. In some cases, called x-ray flashes (XRF's), the γ -ray signal is so weak that all we have is this x-ray signal. Prompt (operationally defined as the time period when the γ -ray detector detects a signal above background) longer-wavelength emission may also occur at the optical and radio wavelengths, but it is harder to detect. So far optical flashes simultaneous with the γ -ray emission have been observed in only three cases (Akerlof *et al.*, 1999; Fox, Yost, *et al.*, 2003; Li *et al.*, 2003).

1. Spectrum

The spectrum is nonthermal. The energy flux peaks at a few hundred keV and in many bursts there is a long high-energy tail extending up to GeV. The spectrum varies strongly from one burst to another. An excellent phenomenological fit for the spectrum was introduced by Band *et al.* (1993) using two power laws joined smoothly at a break energy $(\tilde{\alpha} - \tilde{\beta})E_0$:

$$N(\nu) = N_0 \times \begin{cases} (h\nu)^{\tilde{\alpha}} \exp\left(-\frac{h\nu}{E_0}\right) & \text{for } h\nu < (\tilde{\alpha} - \tilde{\beta})E_0; \\ [(\tilde{\alpha} - \tilde{\beta})E_0]^{(\tilde{\alpha} - \tilde{\beta})} (h\nu)^{\tilde{\beta}} \exp(\tilde{\beta} - \tilde{\alpha}) & \text{for } h\nu > (\tilde{\alpha} - \tilde{\beta})E_0. \end{cases} \quad (1)$$

I denote the spectral indices here as $\tilde{\alpha}$ and $\tilde{\beta}$ to distinguish them from the afterglow parameters (α and β) discussed later. There is no particular theoretical model that predicts this spectral shape. Still, this function provides an excellent fit to most of the observed spectra. For most observed values of $\tilde{\alpha}$ and $\tilde{\beta}$, $\nu F_\nu \propto \nu^2 N(\nu)$ peaks at $E_p = (\tilde{\alpha} + 2)E_0$. For about 10% of the bursts the upper slope is larger than -2 and there is no peak for νF_ν within the observed spectrum. Another group of bursts, known as NHE bursts (no high energy; Pendleton *et al.*, 1997), does not have a hard component (which is reflected by a very negative value of $\tilde{\beta}$). The “typical” energy of the observed radiation is E_p . E_p defined in this way should not be confused with the commonly used

hardness ratio, which is the ratio of photons observed in two BATSE² channels: channel 3 (100–300 keV) counts divided by channel 2 (50–100 keV) counts. The break frequency and the peak flux frequencies are lower on average for bursts with lower observed flux (Mallozzi *et al.*, 1995, 1998).

Band *et al.* (1993) present a small catalog of the spectra of 52 bright bursts, which they analyze in terms of

²BATSE is the Burst and Transient Source Experiment on the CGRO (Compton Gamma-Ray Observatory), see <http://coss.gsfc.nasa.gov/batse/>. It operated for almost a decade, detecting several thousand bursts, more than any satellite before or after it. The BATSE data were published in several catalogs. See Paciesas *et al.* (1999) for the most recent one.

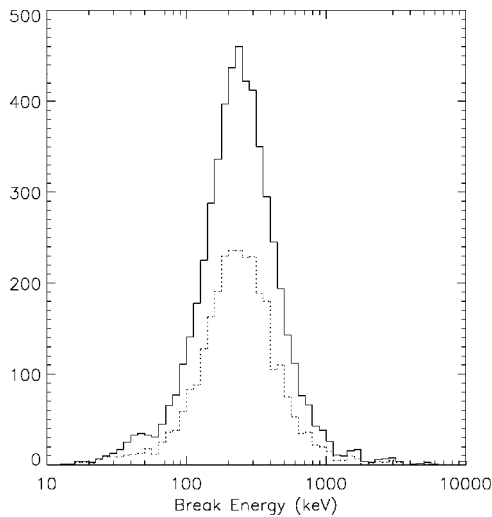


FIG. 1. Distribution of the observed values of the break energy $(\tilde{\alpha}-\tilde{\beta})E_0$ in a sample of bright bursts observed by BATSE: From Preece *et al.*, 2000. Solid line, the whole sample; dashed line, a subset of the data.

the band function. Preece *et al.* (2000) present a larger catalog with 156 bursts selected for either high flux or fluence. They consider several spectral shapes, including the band function.

Figure 1 shows the distribution of observed values of the break energy, $(\tilde{\alpha}-\tilde{\beta})E_0$, in a sample of bright bursts (Preece *et al.*, 2000). Most of the bursts are in the range $100 < (\tilde{\alpha}-\tilde{\beta})E_0 < 400$ keV, with a clear maximum in the distribution around $(\tilde{\alpha}-\tilde{\beta})E_0 \sim 250$ keV. There are not many soft GRB's—that is, GRB's with peak energy in the tens of keV range. However, the discovery (Heise *et al.*, 2001) of XRF's—x-ray flashes with similar temporal structure to GRB's but lower typical energies—shows that the low peak energy cutoff is not real and it reflects the lower sensitivity of BATSE in this range (Kippen *et al.*, 2002).

Similarly, it is debatable whether there is a real paucity of hard GRB's and an upper cutoff to GRB hardness or it just happens that the detection is optimal in this band (a few hundred keV). BATSE triggers, for example, are based mostly on the count rate between 50 and 300 keV. BATSE is therefore less sensitive to harder bursts that emit most of their energy in the MeV range. Using BATSE's observation alone one cannot rule out the possibility that there is a population of harder GRB's that emit equal power in total energy that are not observed because of this selection effect (Piran and Narayan, 1996; Cohen *et al.*, 1997; Higdon and Lingfelter, 1998; Lloyd and Petrosian, 1999). 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 Solar Maximum Mission (SMM) data (Harris and Share, 1998) suggests that there is a deficiency (by at least a factor of 5) of

GRB's with hardness above 3 MeV, relative to GRB's peaking at ~ 0.5 MeV, but these data are consistent with a population of hardness that extends up to 2 MeV.

Overall the narrowness of the hardness distribution is very puzzling. First, as I stressed earlier it is not clear whether it is real or an observational artifact. If it is real, then on the one hand there is no clear explanation of the physical process that controls the narrowness of the distribution (see, however, Guetta *et al.*, 2001a). On the other hand, cosmological redshift effects must broaden this distribution, and it seems likely (though not yet demonstrated) that if the GRB distribution extends to $z=10$ as some suggest (Ciardi and Loeb, 2000; Lamb and Reichart, 2000; Bromm and Loeb, 2002; Lloyd-Ronning *et al.*, 2002), then such a narrow distribution requires a correlation between the intrinsic hardness of the burst and its redshift, namely, that the intrinsic hardness increases with the redshift. There is some evidence for such a correlation between E_p and the observed peak flux (Mallozzi *et al.*, 1995, 1998). More recently Amati *et al.* (2002) reported on a correlation between E_p and the isotropic equivalent energy seen in 12 BeppoSAX³ bursts that they have analyzed. They also report on a correlation between E_p and the redshift, as the bursts with higher isotropic equivalent energy are typically more distant. These three different correlations are consistent with each other if the observed peak flux of bursts is determined by their intrinsic luminosity more than by the distance of the bursts. In such a case (because of the larger volume at larger distances) the more distant bursts observed are on average brighter than nearer ones (see also Sec. II.C).

Even though the burst hardness distribution shows a single population, a plot of the hardness vs temporal duration shows that short bursts (see Fig. 2) are typically harder (Dezalay *et al.*, 1996; Kouveliotou *et al.*, 1996). The correlation is significant. Another interesting subgroup of bursts is the NHE (no high energy) bursts—bursts with no hard component, that is, no emission above 300 keV (Pendleton *et al.*, 1997). This group is characterized by a large negative value of β , the high-energy spectral slope. The NHE bursts have luminosities about an order of magnitude lower than regular bursts, and they exhibit an effectively homogeneous intensity distribution with the average ratio $\langle V/V_{max} \rangle = 0.53 \pm 0.029$. As I discuss later in Sec. II.A.2 most GRB light curves are composed of many individual pulses. It is interesting that in many bursts there are NHE pulses combined with regular pulses.

EGRET (The Energetic Gamma-Ray Experiment Telescope), the high-energy γ -ray detector on the Compton Gamma-Ray Observatory, detected seven

³BeppoSAX is an Italian-Dutch x-ray astronomy spacecraft responsible for the first gamma-ray burst source identifications because it was able to quickly point a comparatively high-resolution x-ray telescope at gamma-ray bursts. See <http://www.asdc.asi.it/beppoSAX/> for information on BeppoSAX and its different instruments.

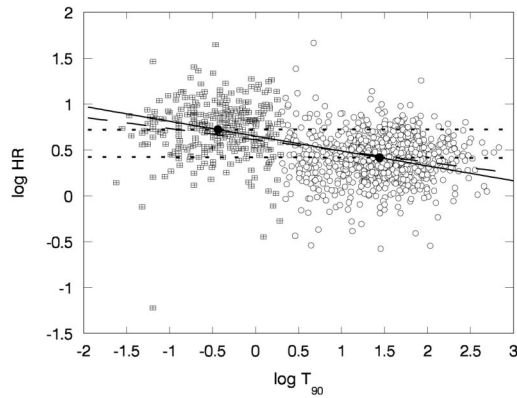


FIG. 2. The hardness-duration correlation for BATSE bursts. HR is the ratio of fluence between BATSE channels 3 and 2. \boxplus , short bursts; \circ , long bursts; solid line, a regression line for the whole sample; dotted lines, the regressions lines for the short and long samples, respectively. From Qin *et al.*, 2000.

GRB's with photon energies ranging from 100 MeV to 18 GeV (Dingus and Catelli, 1998). In some cases this very-high-energy emission was delayed more than an hour after the burst (Hurley, 1994; Sommer *et al.*, 1994). No high-energy cutoff above a few MeV has been observed in any GRB spectrum. Recently, González *et al.* (2003) have combined the BATSE (30 keV–2 MeV) data with the EGRET data for 26 bursts. In one of these bursts, GRB 941017 (according to the common notation GRB's are numbered by the date), they have discovered a high-energy tail that extended up to 200 MeV and looked like a different component. This high-energy component appeared 10–20 sec after the beginning of the burst and displayed a roughly constant flux with a relatively hard spectral slope ($F_\nu \propto \nu^0$) up to 200 sec. At late times (150 sec after the trigger) the very-high-energy (10–200 MeV) tail contained 50 times more energy than the “main” γ -ray energy (30 keV–2 MeV) band. The TeV detector, Milagro, discovered (at a statistical significance of $1.5e-3$ or so, namely, at 3σ) a TeV signal coincident with GRB 970417 (Atkins *et al.*, 2000, 2003). If true, this would correspond to a TeV fluence that exceeds the low-energy γ -ray fluence. However, no further TeV signals were discovered from the other 53 bursts observed by Milagro (Atkins *et al.*, 2000) or from several bursts observed by the more sensitive Milagro (McEney, 2002). One should recall, however, that due to the attenuation of the IR background TeV photons could not be detected from $z > 0.1$. Thus, even if most GRB's emit TeV photons, those photons will not be detected on Earth.

Another puzzle is the low-energy tail. Cohen *et al.* (1997) analyze several strong bursts and find that their low-energy slope is around $1/3$ to $-1/2$. However, Preece *et al.* (1998, 2002) suggest that about $1/5$ of the bursts have a low-energy power spectrum α steeper than $1/3$ (the synchrotron slow-cooling low-energy slope). A larger fraction is steeper than $-1/2$ (the fast-cooling synchrotron low-energy slope). However, this is not seen in

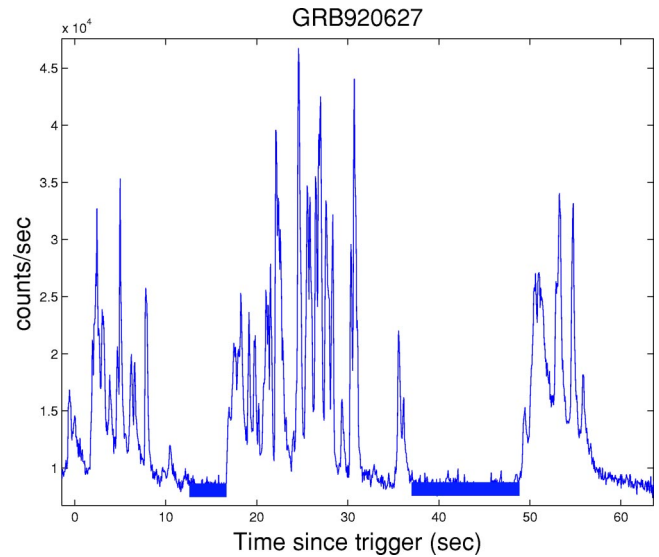


FIG. 3. (Color in online edition) The light curve of GRB 920627. The total duration of the burst is 52 sec, while typical pulses are 0.8 sec wide. Two quiescent periods lasting ~ 10 sec are marked by horizontal solid bold lines.

any of the spectra from HETE⁴ whose low-energy resolution is somewhat better. All HETE bursts have a low-energy spectrum that is within the range $1/3$ to $-1/2$ (Barraud *et al.*, 2003). As both BATSE and HETE use NaI detectors that have a poor low-energy resolution (Cohen *et al.*, 1997), this problem might be resolved only when a better low-energy spectrometer is flown.

2. Temporal structure

The duration of the bursts spans five orders of magnitude, ranging from less than 0.01 sec to more than 100 sec. Common measures for the duration are T_{90} (T_{50}), which corresponds to the time in which 90% (50%) of the counts of the GRB arrive. As I discuss below (see Sec. II.A.3), the bursts are divided into long and short bursts according to their T_{90} . Most GRB's are highly variable, showing 100% variations in flux on a time scale much shorter than the overall duration of the burst. Figure 3 depicts the light curve of a typical variable GRB (GRB 920627). The variability time scale δt is determined by the width of the peaks. δt is much shorter (in some cases by more than a factor of 10^4) than T , the duration of the burst. Variability on a time scale of milliseconds has been observed in some long bursts (McBreen *et al.*, 2001; Nakar and Piran, 2002c). However, only $\sim 80\%$ of the bursts show substantial substructure in their light curves. The rest are rather smooth, typically with a fast-rise exponential decay (FRED) structure.

⁴HETE II is a dedicated GRB satellite that aims at quickly locating bursts with high positional accuracy. It is properly known as HETE II because HETE I was lost upon launch when its rocket failed. See <http://space.mit.edu/HETE/> for a description of HETE II and its instruments.

Fenimore and Ramirez-Ruiz (2000; see also Reichart *et al.*, 2001) discovered a correlation between the variability and the luminosity of the bursts. This correlation (as well as the lag-luminosity relation discussed later) allows us to estimate the luminosity of bursts that do not have a known redshift.

The bursts seem to be composed of individual pulses, with a pulse being the “building block” of the overall light curve. Individual pulses display a hard-to-soft evolution with the peak energy decreasing exponentially with the photon fluence (Ford *et al.*, 1995; Liang and Kargatis, 1996; Norris *et al.*, 1996). The pulses have the following temporal and spectral features.

- (i) The light curve of an individual pulse is a FRED—fast-rise exponential decay—with an average rise-to-decay ratio of 1:3 (Norris *et al.*, 1996).
- (ii) The low-energy emission is delayed compared to the high-energy emission⁵ (Norris *et al.*, 1996). Norris *et al.* (2000) have found that these spectral lags are anticorrelated with the luminosity of the bursts: Luminous bursts have long lags. This lag-luminosity relation provides another way to estimate the luminosity of a burst from its (multispectral) light curve.
- (iii) The pulses’ low-energy light curves are wider compared to the high-energy light curves. The width goes as $\sim E^{-0.4}$ (Fenimore *et al.*, 1995).
- (iv) There is a width-symmetry-intensity correlation. High intensity pulses are (statistically) more symmetric (lower decay-to-rise ratio) and with shorter spectral lags (Norris *et al.*, 1996).
- (v) There is a hardness-intensity correlation. The instantaneous spectral hardness of a pulse is correlated to the instantaneous intensity (the pulse becomes softer during the pulse decay) (Borgonovo and Ryde, 2001).

The pulse widths δt and the pulse separation Δt have rather similar log-normal distributions. However, the pulse separation distribution has an excess of long intervals (Fig. 4; Nakar and Piran, 2002c). These long intervals can be classified as quiescent periods (Ramirez-Ruiz and Merloni, 2001), relatively long periods of several dozen seconds with no activity. When these quiescent periods are excluded, both distributions are log-normal with comparable parameters (Nakar and Piran, 2002c; Quilligan *et al.*, 2002). The average pulse interval, $\bar{\Delta t} = 1.3$ sec, is larger by a factor of 1.3 than the average pulse width $\bar{\delta t} = 1$ sec. One also finds that the pulse widths are correlated with the preceding interval (Nakar and Piran, 2002c). Ramirez-Ruiz and Fenimore (2000)

⁵Low/high energy implies the low vs the high BATSE channels. The four BATSE channels are 20–50 keV, 50–100 keV, 100–300 keV, and >300 keV.

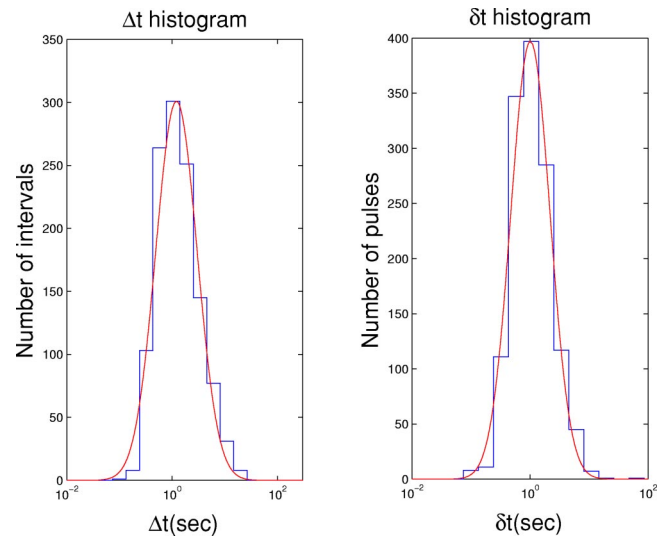


FIG. 4. (Color in online edition) Pulses within γ -ray bursts: right, the pulse width distribution; left, the distribution of intervals between pulses. From Nakar and Piran, 2002c.

found that the pulse width does not vary along the bursts.

One can also analyze the temporal behavior using the traditional Fourier-transform method. The power density spectra of the light curves show a power-law slope of $\sim -5/3$ and a sharp break at 1 Hz (Beloborodov *et al.*, 2000).

The results described so far are for long bursts. The variability of short ($T < 2$ sec) bursts is more difficult to analyze. The duration of these bursts is closer to the limiting resolution of the detectors. Still most ($\sim 66\%$) short bursts are variable, with $\delta t/T < 0.1$ (Nakar and Piran, 2002b). These variable bursts are composed of multiple subpulses.

3. Populations

a. Long and short bursts

The clearest classification of bursts is based on their duration. Kouveliotou *et al.* (1993) showed that GRB can be divided into two distinct groups: long bursts with $T_{90} > 2$ sec and short bursts with $T_{90} < 2$ sec. Note that it was suggested (Horváth, 1998; Mukherjee *et al.*, 1998) that there is a third, intermediate class with $2.5 < T_{90} < 7$ sec. However, it is not clear whether this division into three classes is statistically significant (Hakkila *et al.*, 2000).

An interesting question is whether short bursts could arise from single peaks of long bursts in which the rest of the long burst is hidden by noise. Nakar and Piran (2002b) have shown that in practically all long bursts the second-highest peak is comparable in height to the first one. Thus if the highest peak is above the noise the second one should be, as well. Short bursts are a different entity. This is supported by the observation that short bursts are typically harder (Dezalay *et al.*, 1996; Kouveliotou *et al.*, 1996). The duration-hardness distribution

(see Fig. 2) shows clearly that there are not soft short bursts.

The spatial distribution of the observed short bursts is clearly different from the distribution of the observed long one. A measure of the spatial distribution is the average ratio $\langle V/V_{\max} \rangle \equiv \langle (C/C_{\min})^{-3/2} \rangle$, where C is the count rate and C_{\min} is the minimal rate required for triggering. In a uniform Euclidian sample this ratio equals 0.5 regardless of the luminosity function. One of the first signs that GRB's might have a cosmological origin was the deviation of this value from 0.5 for the BATSE sample (Meegan *et al.*, 1992). The $\langle V/V_{\max} \rangle$ of the BATSE short-burst sample (Mao *et al.*, 1994; Katz and Canel, 1996; Piran, 1996) is significantly higher than $\langle V/V_{\max} \rangle$ of the long-burst sample. Note that more recently Schmidt (2001a) suggested that the two values are similar and the distribution of long and short bursts is similar. However, Guetta and Piran (2004) find $\langle V/V_{\max} \rangle_{\text{long}} = 0.282$ and $\langle V/V_{\max} \rangle_{\text{long}} = 0.390$ (I discuss this point further in Sec. II.C.3). This implies that the population of observed short bursts is nearer on average than the population of the observed long ones. This is not necessarily a statement on the location of short vs long bursts. Instead it simply reflects the fact that it is more difficult to detect a short burst. For a short burst one has to trigger on a shorter (and hence noisier) window, and the detector (specifically BATSE, which triggers on 64 msec for short bursts and on 1 sec for long ones) is less sensitive to short bursts. I discuss later, in Sec. II.C.3, the question of rates of long vs short bursts.

So far afterglow has been detected only from long bursts. It is not clear whether this is an observational artifact or a real feature. However, there was no x-ray afterglow observed for the only well-localized short hard burst, GRB 020531 (Hurley, Cline, *et al.*, 2002a). Chandra observations show an intensity weaker by at least a factor of 100–300 than the intensity of the x-ray afterglow from long bursts at a similar time (Butler *et al.*, 2002). Neither was afterglow observed in other wavelengths (Klotz *et al.*, 2003).

As identification of hosts and redshifts depends on the detection of afterglow, this implies that nothing is known about the distribution, progenitors, environment, etc. of short bursts. These bursts are still waiting for their afterglow revolution.

b. X-ray flashes

X-ray flashes (XRF's) are x-ray bursts with a similar temporal structure to GRB's but lower typical energies. Heise *et al.* (2001) discovered these flashes by comparing the triggering of GRB monitors with sensitivity above 40 keV and (wide-field cameras) on BeppoSAX. In 39 cases the wide-field cameras were triggered while the GRB monitors were not, implying that these flashes do not have any hard component and most of their flux is in the x-ray region of the spectrum. The duration of 17 of these transients (out of 39), denoting XRF's, is comparable to the duration of the x-ray emission accompanying GRB's. The peak fluxes of the XRF's are similar to

the x-ray fluxes observed by wide-field cameras during GRB's ($\sim 10^{-8}$ ergs/sec/cm²) but their peak energy is clearly below 40 keV. These findings confirmed the detection by Strohmayer *et al.* (1998) of seven GRB's with $E_p < 10$ keV and five additional GRB's with $E_p < 50$ keV in the Ginga⁶ data.

Barraud *et al.* (2003) analyze 35 bursts detected on HETE II. They find that XRF's lie on the extension of all the relevant GRB distributions. That is, there is a continuity from GRB's to XRF's. Detailed searches in the BATSE data (Kippen *et al.*, 2002) have revealed that some of these bursts were also detected by BATSE. Using a complete search in 90% of the wide-field camera data available, Heise (2003) finds that the observed frequency of XRF's is approximately half of the GRB frequency: In six years of BeppoSAX observations they have observed 32 XRF's above a threshold peak luminosity of 5×10^{-9} erg/sec/cm² in the 2–25-keV range compared with 54 GRB's (all GRB's above BATSE thresholds are observed if in the field of view).

Soderberg *et al.* (2002) discovered optical afterglow from XRF 020903 and suggested that the burst was at $z=0.25$. They also suggested a hint of an underlying supernova signal (see Sec. II.C.4) peaking between 7 and 24 days after the initial XRF trigger. Afterglow was discovered from XRF 030723 as well (Fox, Kaplan, *et al.*, 2003).

4. Polarization

Recently, Coburn and Boggs (2003) reported on a detection of a very high ($80\% \pm 20\%$) linear polarization during the prompt γ -ray emission of GRB 021206. This burst was extremely powerful. The observed fluence of GRB 021206 was 1.6×10^{-4} ergs/cm² at the energy range of 25–100 keV (Hurley, Cline, *et al.*, 2002b, 2002c). This makes GRB 021206 one of the most powerful bursts (a factor of 2–3 above GRB 990123) and the most powerful one after correcting for the fact that it was observed only in a narrow band (compared to the wide BATSE band of 20–2000 keV). Coburn and Boggs (2003) analyzed the data recorded by the Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI). The polarization is measured in this detector by the angular dependence of the number of simultaneous pairs of events that are most likely caused by a scattering of the detected γ rays within the detector. The data analysis is based on 12 data points, which are collected over 5 sec. Each of these points is a sum of several independent observations taken at different times. Thus the data are a kind of convolution of the polarization over the whole duration of the burst.

Coburn and Boggs (2003) tested two hypotheses. First they tested the null hypothesis of no polarization. This hypothesis was rejected at a confidence level of 5.7σ .

⁶The Ginga x-ray astronomy satellite (Ginga is Japanese for “galaxy”) was a joint Japanese/UK mission in orbit from 1987 to 1991.

Second they estimated the modulation factor assuming a constant polarization during the whole burst. The best fit to the data was achieved with $\Pi = (80 \pm 20)\%$. However, Coburn and Boggs (2003) found that the probability that χ^2 is greater than the value obtained with this fit is 5%. The model of constant polarization is consistent with the analysis and observations only at the 5% level.

Rutledge and Fox (2004) reanalyzed these data and found several inconsistencies within the methodology of Coburn and Boggs (2003). They pointed out that if we rely on the instrumental correction factor of Coburn and Boggs, the upper limit on polarization is greater than 220%, implying that the polarization would not be detectable, even if the GRB were 100% polarized. In their rebuttal Boggs and Coburn (2003) did not provide a clear answer to the criticism of the methodology raised by Rutledge and Fox. Furthermore, the results of Rutledge and Fox (2004) were recently corroborated by Wigger *et al.* (2004), who also reanalyzed the same data and concluded that the polarization level is $40\% \pm 60\%$, i.e., there is no significant evidence for polarization.

5. Prompt optical flashes

The robotic telescope ROTSE (Robotic Optical Transient Search Experiment) detected a ninth-magnitude optical flash that was concurrent with the GRB emission from GRB 990123 (Akerlof *et al.*, 1999). The six snapshots began 40 sec after the trigger and lasted until 3 min after the burst. The second snapshot, which took place 60 sec after the trigger, recorded the ninth-magnitude flash. While the six snapshots do not provide a “light curve,” it is clear that the peak optical flux did not coincide with the peak γ -ray emission, which took place around the first ROTSE snapshot. This suggests that the optical flux is not the “low-energy tail” of the γ -ray emission. Recently, Fox, Yost, *et al.* (2003) reported the detection of a 15.45-magnitude optical signal from GRB 021004, 193 sec after the trigger. This is just 93 sec after the 100-sec-long burst stopped being active. Shortly afterwards Li *et al.* (2003) reported on the detection of a 14.67-magnitude optical signal from GRB 021211, 105 sec after the trigger. Finally, Price *et al.* (2003) detected a 12th-magnitude prompt flash, more than 1.5 h after the trigger. No similar prompt signal was observed from any other burst, in spite of extensive searches that provided upper limits. Kehoe *et al.* (2001) searched five bright bursts and found single-image upper limits ranging from 13th to 14th magnitude around 10 sec after the initial burst detection and from 14 to 15.8 magnitudes 1 h later. These upper limits are consistent with the two recent detections, which are around 15th magnitude. The recent rapid detections suggest that we should expect many more such discoveries in the near future.

6. The GRB-afterglow transition—observations

There is no direct correlation between the γ -ray fluxes and the x-ray (or optical) afterglow fluxes. Extrapolation

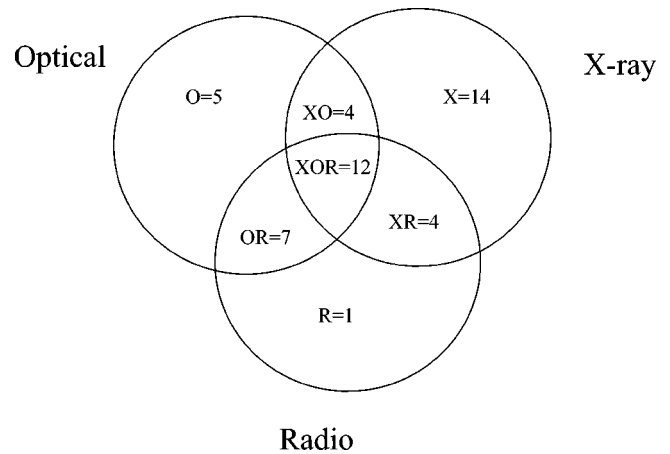


FIG. 5. Venn diagram of the distribution of 47 afterglows observed in different wavelengths between 1997 and 2001. From a talk given by D. Frail at the Sackler GRB Workshop, Harvard, May 2001.

of the x-ray afterglow fluxes backwards generally do not fit the γ -ray fluxes. Instead they fit the late prompt x-ray signal. These results are in nice agreement with the predictions of the internal-external shocks scenario in which the two phenomena are produced by different effects and one should not expect a simple extrapolation to work.

The expected GRB-afterglow transition has been observed in several cases. The first observation took place in 1992 (Burenin *et al.*, 1999), but was not reported until much later. BeppoSAX data show a rather sharp transition in the hardness that takes place several tens of seconds after the beginning of the bursts. This transition is seen clearly in the different energy-band light curves of GRB 990123 and in GRB 980923 (Giblin *et al.*, 1999). Connaughton (2002) averaged the light curves of many GRB's and discovered long and soft tails: the early x-ray afterglow. Additional evidence for a transition from the GRB to the afterglow can be observed in the different spectra within the GRB (Preece *et al.*, 2002).

B. The afterglow

Until 1997 there were no known counterparts to GRB's in other wavelengths. On February 28, 1997, the Italian-Dutch satellite BeppoSAX detected an x-ray afterglow from GRB 970228 (Costa *et al.*, 1997). The exact position given by BeppoSAX led to the discovery of an optical afterglow (van Paradijs *et al.*, 1997). Radio afterglow was detected in GRB 970508 (Frail *et al.*, 1997). More than 40 x-ray afterglows have now been observed (see <http://www.mpe.mpg.de/~jcg/grb.html> for complete up-to-date tables of well localized GRB's with or without afterglow. Another useful web page is <http://grad40.as.utexas.edu/grblog.php>). About half of these GRB's have optical and radio afterglows (see Fig 5). The accurate positions given by the afterglows enabled the identification of the host galaxies of many bursts. In 20 or so cases the redshift has been measured. The ob-

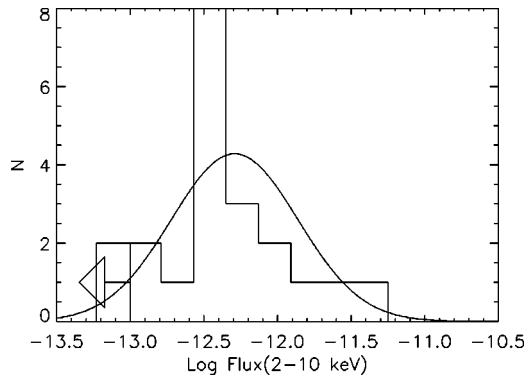


FIG. 6. The distribution of x-ray fluxes (2–10 keV) at $t=11$ h after the GRB in 21 afterglows observed by BeppoSAX. The sample includes all the fast observations performed by BeppoSAX on GRB's from January 1997 to October 1999. No x-ray afterglow was detected in GRB 990217 to the limiting instrumental sensitivity of 10^{-13} erg cm $^{-2}$ sec $^{-1}$, 6 h after the burst. In the case of GRB 970111 a candidate was detected, but evidence of fading behavior is marginal, so both cases are considered as upper limits (indicated by the arrow). From Piran *et al.*, 2001.

served redshifts range from 0.16 for GRB 030329 (or 0.0085 for GRB 980425) to a record of 4.5 for GRB 000131. Even though the afterglow is a single entity I shall follow the astronomical wavelength division and review here the observational properties of x-ray, optical, and radio afterglows.

1. The x-ray afterglow

The x-ray afterglow is the first and strongest, but shortest signal. In fact, it seems to begin while the GRB is going on (see Sec. II.A.6 for a discussion of the GRB-afterglow transition). The light curve observed several hours after the burst can usually be extrapolated to the late parts of the prompt emission.

The x-ray afterglow fluxes from GRB's have a power-law dependence on ν and on the observed time t (Piro, 2001): $f_\nu(t) \propto \nu^{-\beta} t^{-\alpha}$ with $\alpha \sim 1.4$ and $\beta \sim 0.9$. The flux distribution, when normalized to a fixed hour after the burst has a rather narrow distribution. A cancellation of the k corrections and the temporal decay makes this flux, which is proportional to $(1+z)^{\beta-\alpha}$ insensitive to the redshift. Using 21 BeppoSAX bursts (Piro, 2001) Piran *et al.* (2001) find that the 1–10 keV flux, 11 h after the burst is 5×10^{-13} ergs/cm 2 sec. The distribution is log normal, with $\sigma_{f_x} \approx 0.43 \pm 0.1$ (see Fig. 6). De Pasquale *et al.* (2003) find a similar result for a larger sample. However, they find that the x-ray afterglows of GRB's with optical counterparts are on average five times brighter than those of dark GRB's (GRB's with no detected optical afterglow). The overall energy emitted in an x-ray afterglow is generally a few percent of the GRB energy. Berger *et al.* (2003) find that the x-ray luminosity is indeed correlated with the opening angle, and when taking the beaming correction into account they find

that $L_X = f_b L_{X,iso}$ is approximately constant, with a dispersion of only a factor of 2.

X-ray lines have been seen in seven GRB's: GRB 970508 (Piro *et al.*, 1999), GRB 970828 (Yoshida *et al.*, 1999), GRB 990705 (Amati *et al.*, 2000), GRB 991216 (Piro *et al.*, 2000), GRB 001025a (Watson *et al.*, 2002), GRB 000214 (Antonelli *et al.*, 2000), and GRB 011211 (Reeves *et al.*, 2002). The lines were detected using different instruments: BeppoSAX, ASCA (Advanced Satellite for Cosmology and Astrophysics), Chandra, and XMM-Newton (x-ray Multi Mirror-Newton). The lines were detected around 10 h after the burst. The typical luminosity in the lines is around 10^{44} – 10^{45} ergs/sec, corresponding to a total fluence of about 10^{49} ergs. Most of the lines are interpreted as emission lines of Fe $K\alpha$. However, there is also a radiative-recombination-continuum line edge, as well as $K\alpha$ lines of lighter elements like Si, S, Ar, and Ca (all seen in the afterglow of GRB 011211; Reeves *et al.*, 2002). In one case (GRB 990705; Amati *et al.*, 2000) there is a transient absorption feature within the prompt x-ray emission, also corresponding to Fe $K\alpha$. The statistical significance of the detection of these lines is of some concern (2 – 5σ), and even though the later instruments are much more sensitive than the early ones all detections remain at this low significance level. Rutledge and Sako (2003) and Sako *et al.* (2003) expressed concern about the statistical analysis of the data showing these lines and claimed that none of the observed lines is statistically significant. The theoretical implications are far reaching. Not only do the lines require, in most models, a very large amount of iron at rest (the lines are quite narrow), they most likely require (Ghisellini *et al.*, 2002) a huge energy supply ($>10^{52}$ ergs), 20 times larger than the typical estimated γ -ray energy ($\sim 5 \times 10^{50}$ ergs).

2. Optical and IR afterglow

About 50% of well-localized GRB's show optical and IR afterglow. The observed optical afterglow is typically around 19–20 mag one day after the burst (see Fig. 7). The signal decays, initially, as a power law in time $t^{-\alpha}$, with a typical value of $\alpha \approx 1.2$ and large variations around this value. In all cases the observed optical spectrum is also a power law $\nu^{-\beta}$. Generally absorption lines are superimposed on this power law. The absorption lines correspond to absorption on the way from the source to Earth. Typically the highest-redshift lines are associated with the host galaxy, providing a measurement of the redshift of the GRB. In a few cases emission lines, presumably from excited gas along the line of sight, were also observed.

Technical difficulties caused a gap of several hours between the burst and the detection of the optical afterglow, which could be found only after an accurate position was available. The rapid localization provided by HETE II helped to close this gap and an almost complete light curve from 193 sec after the trigger (≈ 93 sec after the end of the burst) is available now for GRB 021004 (Fox, Yost, *et al.*, 2003).

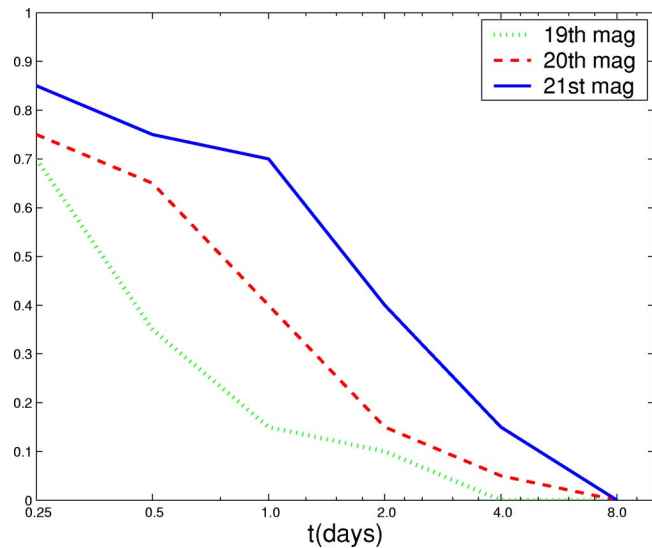


FIG. 7. (Color in online edition) The fraction of bursts with optical afterglow above three limiting magnitudes as a function of time (compared to the total number of bursts with optical afterglow).

Many afterglow light curves show an achromatic break to a steeper decline with $\alpha \approx 2$. The classical example of such a break was seen in GRB 990510 (Harrison *et al.*, 1999; Stanek *et al.*, 1999), shown here in Fig. 8. It is common to fit the break with the phenomenological formula: $F_\nu(t) = f_*(t/t_*)^{-\alpha_1} [1 - \exp[-(t/t_*)^{\alpha_1 - \alpha_2}]] (t/t_*)^{\alpha_1 - \alpha_2}$. This break is commonly interpreted as a jet break that allows us to estimate the opening angle of the jet (Rhoads, 1999; Sari *et al.*, 1999) or the viewing angle within the universal

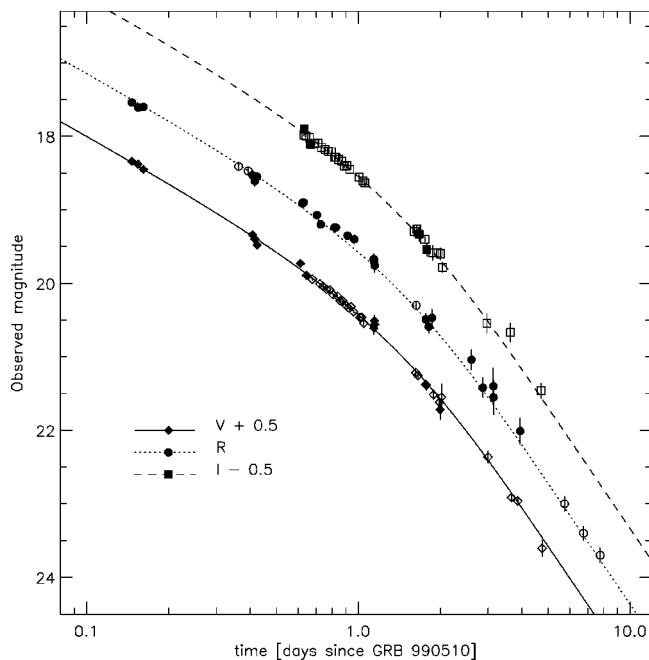


FIG. 8. Optical light curves of GRB 990510. A fit for the observed optical light curves is obtained with $\alpha_1 = 0.82 \pm 0.02$, $\alpha_2 = 2.18 \pm 0.05$ and $t_* = 1.2 \pm 0.08$ days. From Harrison *et al.*, 1999.

standard jet model (Rossi *et al.*, 2002; see Sec. II.D below).

The optical light curve of the first detected afterglow (from GRB 970228) could be seen for more than half a year (Fruchter *et al.*, 1998). In most cases the afterglow fades faster and cannot be followed for more than several weeks. At this stage the afterglow becomes significantly dimmer than its host galaxy and the light curve reaches a plateau corresponding to the emission of the host.

In several cases, e.g., GRB 980326 (Bloom *et al.*, 1999), GRB 970228 (Reichart, 1999), GRB 011121 (Bloom, Kulkarni, Price, *et al.*, 2002; Garnavich *et al.*, 2003), red bumps are seen at late times (several weeks to a month). These bumps are usually interpreted as evidence for an underlying supernova. A most remarkable supernova signature was seen recently in GRB 030329 (Hjorth *et al.*, 2003; Stanek *et al.*, 2003), having the same signature as SN 1981bw, which was associated with GRB 990425 (see Sec. II.C.4).

Finally, I note that varying polarization at optical wavelengths has been observed in GRB afterglows at the level of a few to ten percent (Covino *et al.*, 1999, 2002; Wijers *et al.*, 1999; Rol *et al.*, 2000; Bersier *et al.*, 2003; Greiner *et al.*, 2003). These observations are in agreement with rough predictions (Ghisellini and Lazzati, 1999; Sari, 1999b) of the synchrotron emission model, provided that there is a deviation from spherical symmetry (see Sec. V.F below).

3. Dark GRB's

Only $\sim 50\%$ of well-localized GRB's show optical transients successive to the prompt gamma-ray emission, whereas an x-ray counterpart is present in 90% of cases (see Fig. 5). Several possible explanations have been suggested for this situation. It is possible that late and shallow observations could not detect the optical transients in some cases; several authors argue that dim and/or rapidly decaying transients could bias the determination of the fraction of truly obscure GRB's (Fynbo *et al.*, 2001; Berger *et al.*, 2002). However, recent reanalysis of optical observations (Ghisellini *et al.*, 2001; Reichart and Yost, 2001; Lazzati *et al.*, 2002) has shown that GRB's without optical transient detection [called dark GRB's or FOA's (failed optical afterglows) or GHOST's (gamma-ray burst hiding an optical source transient)] have had on average weaker optical counterparts, at least two magnitudes in the *R* band, than GRB's with optical transients. Therefore they appear to constitute a different class of objects, although there could be a fraction undetected because of bad imaging.

The nature of dark GRB's is not clear. So far three hypotheses have been put forward to explain their behavior. The first is that they are similar to the other bright GRB's, except for the fact that their lines of sight pass through large and dusty molecular clouds, which cause high absorption (Reichart and Price, 2002). The second is that they are more distant than GRB's with optical transients, at $z \geq 5$ (Fruchter *et al.*, 1999; Lamb

and Reichart, 2000), so that the Lyman break is redshifted into the optical band. Nevertheless, the distances of a few dark GRB's have been determined, and they do not imply high redshifts (Antonelli *et al.*, 2000; Djorgovski, Frail, *et al.*, 2001; Piro *et al.*, 2002). A third possibility is that the optical afterglow of dark GRB's is intrinsically much fainter than (2–3 mag below) that of other GRB's.

De Pasquale *et al.* (2003) find that GRB's with optical transients show a remarkably narrow distribution of flux ratios, which corresponds to an average optical-to-x-ray spectral index 0.794 ± 0.054 . They find that, while 75% of dark GRB's have flux ratio upper limits still consistent with those of GRB's with optical transients, the remaining 25% are 4–10 times weaker in the optical than in the x-ray wavelengths. This result suggests that the afterglows of most dark GRB's are intrinsically fainter in all wavelengths relative to the afterglows of GRB's with observed optical transients. As for the remaining 25%, here the spectrum (optical-to-x-ray ratio) must be different from those of other afterglows with a suppression of the optical band.

4. Radio afterglow

Radio afterglow has been detected in $\sim 50\%$ of the well-localized bursts. Most observations are done at about 8 GHz since the detection falls off drastically at higher and lower frequencies. The observed peak fluxes are at the level of 2 mJy. A turnover is seen around 0.2 mJy and the undetected bursts have upper limits of the order of 0.1 mJy. As the localization is based on the x-ray afterglow (and as practically all bursts have an x-ray afterglow) almost all these bursts were detected in the x-ray domain. About 80% of bursts with radio afterglow also have optical afterglow. The rest are optically dark. Similarly $\sim 80\%$ of the optically observed afterglows also have a radio component (see Fig. 5).

Several bursts (GRB's 980329, 990123, 91216, 000926, 001018, 010222, 011030, and 011121) were detected at around 1 day. Recent radio observations begin well before that, but do not get a detection until about 24 h after a burst. The earliest radio detection took place in GRB 011030 at about 0.8 days after the burst (Taylor *et al.*, 2001). In several cases (GRB's 990123, 990506, 991216, 980329, and 020405) the afterglow was detected early enough to indicate emission from the reverse and a transition from the reverse shock to the forward shock.

The radio light curve of GRB 970508 (see Fig. 9) depicts early strong fluctuations (of order unity) in the flux (Frail *et al.*, 1997). Goodman (1997) suggested that these fluctuations arise due to scintillations and that the decrease (with time) in their amplitude comes from a transition from strong to weak scintillations. Frail *et al.* (1997) used this to infer the size of the emitting region of GRB 970508 at ~ 4 weeks after the burst as $\sim 10^{17}$ cm. These observations provided the first direct proof of relativistic expansion in GRB's.

The self-absorbed frequencies fall in the centimeter-to-meter wavelength radio regime and hence the lower

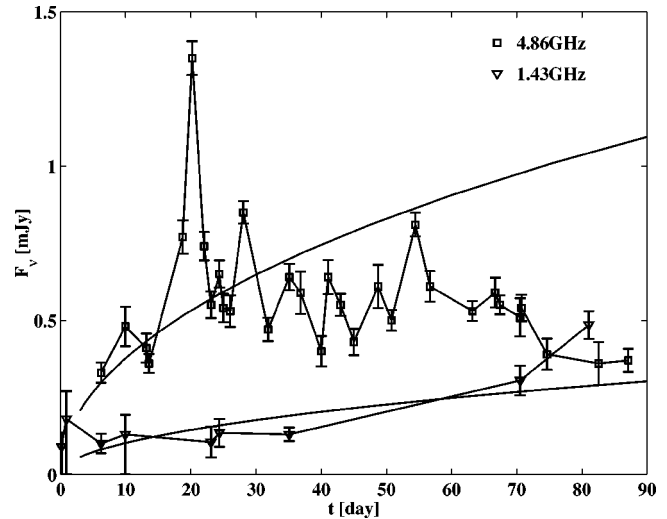


FIG. 9. Light curves of the radio afterglow of GRB 970508 at 4.86 and 1.43 GHz, compared with the predictions of the adiabatic fireball model. From Frail *et al.*, 1997.

radio emission is within the self-absorption part of the spectrum (see Sec. V.C.3 later). In this case the spectrum rises as ν^2 (Katz and Piran, 1997). The fact that the spectral shape is determined by the optical thickness of the system enables us (using similar arguments to those of a simple blackbody emission) to determine the size of the emitting region. In GRB 990508 this has led to an estimate of $\sim 10^{17}$ cm, comparable to the estimate derived from scintillations.

The long-lived nature of the radio afterglow allows for unambiguous calorimetry of the blast wave to be made when its expansion has become subrelativistic and quasi-spherical. The light curves evolve on a longer time scale in the radio. Some GRB afterglows have been detected years after the burst, even after the relativistic to Newtonian transition (see Sec. VII.D). At this stage the expansion is essentially spherical and this enables a direct “calorimetric” estimate of the total energy within the ejecta (Waxman *et al.*, 1998).

C. Hosts and distribution

1. Hosts

By now (early 2004) host galaxies have been observed for all but one or two bursts with optical, radio, or x-ray afterglow localization with arcsec precision (Hurley, Sari, and Djorgovski, 2002). The no-host problem, which was much discussed in the nineties has disappeared. GRB's are located within host galaxies (see Djorgovski, Frail, *et al.*, 2001; Djorgovski, Kulkarni, *et al.*, 2001; and Hurley, Sari, and Djorgovski, 2002 for detailed reviews). While many researchers believe that the GRB host population is likely to be representative of the normal star-forming field galaxy population at comparable redshifts, others argue that GRB host galaxies are significantly bluer than average and their star formation rate is much higher than average.

The host galaxies are faint, with median apparent magnitude $R \approx 25$. Some faint hosts are at $R \approx 29$. Down to $R \approx 25$ the observed distribution is consistent with deep field galaxy counts. Jimenez *et al.* (2001) find that the likelihood of finding a GRB in a galaxy is proportional to the galaxy's luminosity.

The magnitude and redshift distribution of GRB host galaxies are typical for normal, faint field galaxies, as are their morphologies (Odewahn *et al.*, 1998; Djorgovski, Kulkarni, *et al.*, 2001; Holland, 2001; Bloom, Kulkarni, and Djorgovski, 2002; Hurley, Sari, and Djorgovski, 2002). While some researchers argue that the broadband optical colors of GRB hosts are not distinguishable from those of normal field galaxies at comparable magnitudes and redshifts (Sokolov *et al.*, 2001; Bloom, Kulkarni, and Djorgovski, 2002), others (Fruchter *et al.*, 1999) assert that the host galaxies are unusually blue and that they are strongly star forming. Le Floc'h *et al.* (2003) argue that R-K colors of GRB hosts are unusually blue and the hosts may be of low metallicity and luminosity. This suggests (Le Floc'h, 2004) that hosts of GRB's might be different from the sites of the majority of star-forming galaxies, which are luminous, reddened, and dust-enshrouded infrared starbursts (Elbaz and Cesarsky, 2003, and references therein). Le Floc'h (2004) also suggests that this difference might arise due to an observational bias and that GRB's that originate in dust-enshrouded infrared starbursts are dark GRB's whose afterglow is not detectable due to obscuration. Whether this is true or not is very relevant to the interesting question to what extent GRB's follow the star formation rate and to what extent they can be used to determine the star formation rate at high redshifts.

Totani (1997), Wijers *et al.* (1998), and Paczynski (1998) suggested that GRB's follow the star formation rate. As early as 1999 Fruchter *et al.* (1999) noted that all four early GRB's with spectroscopic identification or deep multicolor broadband imaging of the host (GRB's 970228, 970508, 971214, and 980703) lie in rapidly star-forming galaxies. Within the host galaxies the distribution of GRB-host offset follows the light distribution of the hosts (Bloom, Kulkarni, and Djorgovski, 2002). The light is roughly proportional to the density of star formation. Spectroscopic measurements suggest that GRB's are within galaxies with a higher star formation rate. However, this is typical for normal field galaxy populations at comparable redshifts (Hurley, Berger, *et al.*, 2002). There are some intriguing hints. In particular, the flux ratios of [Ne III] 3859 to [OII] 3727 are on average a factor of 4 to 5 higher in GRB hosts than in star-forming galaxies at low redshifts (Djorgovski, Kulkarni, *et al.*, 2001). This may be indirect evidence linking GRB's with massive-star formation. The link between GRB's and massive stars has been strengthened with the centimeter and submillimeter discoveries of GRB host galaxies (Berger *et al.*, 2001; Frail *et al.*, 2002) undergoing prodigious star formation ($\text{SFR} \sim 10^3 M_\odot \text{ yr}^{-1}$), which remains obscured at optical wavelengths.

Evidence for the different characteristics of GRB host galaxies is presented in the work of Fynbo *et al.* (2002,

2003), who find that GRB host galaxies "always" show Lyman-alpha emission in cases where a suitable search has been conducted. This backs up the claim for active star formation and at most moderate metallicity in GRB hosts. It clearly distinguishes GRB hosts from the Lyman-break galaxy population, in which only about 1/4 of galaxies show strong Lyman-alpha emission.

2. The spatial distribution

BATSE's discovery that x-ray bursts are distributed uniformly on the sky (Meegan *et al.*, 1992) was one of the first indications of the cosmological nature of GRB's. The uniform distribution indicated that GRB's are not associated with the Galaxy or with "local" structure in the near universe.

Recently there have been several claims that subgroups of the whole GRB population show a deviation from a uniform distribution. Mészáros *et al.* (2000a, 2000b), for example, find that the angular distribution of the intermediate subgroup of bursts (more specifically of the weak-intermediate subgroup) is not random. Magliocchetti *et al.* (2003) reported that the two-point angular correlation function of 407 short BATSE GRB's reveals a $\sim 2\sigma$ deviation from isotropy on angular scales 2° – 4° . This result is consistent with the possibility that observed short GRB's are nearer and the angular correlation is induced by the large-scale structure correlations on this scale. These claims are important as such an angular correlation could arise only if these bursts were relatively nearby. Alternatively, the deviation from isotropy could indicate repetition of these sources (Magliocchetti *et al.*, 2003). Any such deviation would imply that these subgroups are associated with different objects than the main GRB population.

Cline *et al.* (2003) studied the shortest GRB population, bursts with typical durations of several dozen msec. They found a significant angular asymmetry and suggested that the $\langle V/V_{\text{max}} \rangle$ distribution provides evidence for a homogeneous source distribution. Cline *et al.* (2003) proposed that these features are best interpreted as sources of a galactic origin. However, one has to realize that there are strong selection effects involved in the detection of this particular subgroup.

3. GRB rates and the isotropic luminosity function

There have been many attempts to determine the GRB luminosity function, that is, the distribution function with respect to luminosity and rate from the BATSE peak flux distribution, with different levels of statistical sophistication and different physical assumptions on the evolution of the rate of GRB's with time.

Roughly speaking the situation is as follows: More than 30 redshifts have now been measured. The median redshift is $z \approx 1$, and the range is from 0.16 (or even 0.0085 if the association of GRB 980425 with SN98bw is also considered) to 4.5 (for GRB 000131). Direct estimates from the sample of GRB's with determined redshifts are contaminated by observational biases and are insufficient to determine the rate and luminosity func-

tion. An alternative approach is to estimate these quantities from the BATSE peak flux distribution. However, the observed sample with known redshifts clearly shows that the luminosity function is wide. With a wide luminosity function, the rate of GRB is only weakly constrained by the peak flux distribution. The analysis is further complicated by the fact that the observed peak luminosity, at a given detector with a given observation energy band, depends also on the intrinsic spectrum. Hence different assumptions on the spectrum yield different results. This situation suggests that there is no point in employing sophisticated statistical tools (see, however, Loredó and Wasserman, 1995; Piran, 1999 for a discussion of these methods), and a simple analysis is sufficient to obtain an idea of the relevant parameters.

I shall not attempt to review the various approaches here.⁷ Instead I shall just quote some estimates of the rates and luminosities of GRB's. The simplest approach is to fit $\langle V/V_{max} \rangle$, which is the first moment of the peak flux distribution. Schmidt (1999, 2001a, 2001b) found that using $\langle V/V_{max} \rangle$ of the long-burst distribution and assuming that the bursts follow the SFR2 star formation rate of Porciani and Madau (2001), the present local rate of long observed GRB's would be $\approx 0.15 \text{ Gpc}^{-3} \text{ yr}^{-1}$ (Schmidt, 2001a). Note that this rate from Schmidt (2001a) is smaller by a factor of 10 than the earlier rate of Schmidt (1999)! This estimate corresponds to a typical (isotropic) peak luminosity of $\sim 10^{51} \text{ ergs/sec}$. These are the observed rate and isotropic peak luminosity.

Recently Guetta, Piran, and Waxman (2004) have repeated these calculations. They use both the Rowan-Robinson (1999) star formation rate,

$$R_{GRB}(z) = \rho_0 \begin{cases} 10^{0.75z}, & z < 1 \\ 10^{0.75z_{peak}}, & z > 1, \end{cases} \quad (2)$$

and SFR2 from Porciani and Madau (2001). Their best-fit luminosity function (per logarithmic luminosity interval $d \ln L$) is

$$\Phi_o(L) = c_o \begin{cases} (L/L^*)^\alpha, & L^*/30 < L < L^* \\ (L/L^*)^\beta, & L^* < L < 30L^*, \end{cases} \quad (3)$$

and 0 otherwise with a typical luminosity, $L^* = 1.1 \times 10^{51} \text{ ergs/sec}$, $\alpha = -0.6$ and $\beta = -2$, and c_o is a normalization constant so that the integral over the luminosity function equals unity. The corresponding local GRB rate is $\rho_0 = 0.44 \text{ Gpc}^{-3} \text{ yr}^{-1}$. There is an uncertainty of a factor of ~ 2 in the typical energy L^* and in the local rate. I shall use these numbers as the “canonical” values in the rest of this review.

The observed (BATSE) rate of short GRB's is smaller by a factor of 3 than that of long ones. However, this is not the ratio of the real rates as (i) the BATSE detector is less sensitive to short bursts than to long ones; and (ii)

the true rate depends on the spatial distribution of the short bursts. So far no redshift has been detected for any short bursts and hence this distribution is uncertain. For short bursts we can only resort to estimates based on the peak flux distribution. There are indications that the value $\langle V/V_{max} \rangle$ of short bursts is larger (and close to the Euclidian value of 0.5) than that of long ones, which is around 0.32. This implies that the observed short bursts are nearer to us than the long ones (Mao *et al.*, 1994; Katz and Canel, 1996; Tavani, 1998), possibly with all observed short bursts at $z < 0.5$. However, Schmidt (2001a) finds for short bursts $\langle V/V_{max} \rangle = 0.354$, which is rather close to the value of long bursts. Assuming that short GRB's also follow the star formation rate, he obtains a local rate of $0.075 \text{ Gpc}^{-3} \text{ yr}^{-1}$ —a factor of 2 below the rate of long GRB's. The (isotropic) peak luminosities are comparable. This result differs from a recent calculation of Guetta and Piran (2004), who find for short bursts $\langle V/V_{max} \rangle = 0.390$ and determine from this a local rate of $0.1\text{--}0.8 \text{ Gpc}^{-3} \text{ yr}^{-1}$, which is comparable to the rate of long bursts. This reflects the fact that the observed short GRB's are significantly nearer than the observed long ones.

These rates and luminosities are assuming that the bursts are isotropic. Beaming reduces the actual peak luminosity and increases the implied rate by a factor $f_b^{-1} = 2/\theta^2$. There is now evidence that GRB's are beamed and moreover that the total energy is narrowly distributed (Frail *et al.*, 2001; Panaitescu and Kumar, 2001). There is also good evidence that the corrected peak luminosity is much more narrowly distributed than the isotropic peak luminosity (van Putten and Regimbau, 2003; Guetta, Piran, and Waxman, 2004). The corrected peak luminosity is $L_{peak}(\theta^2/2) \sim \text{const}$. Frail *et al.* (2001) suggest that the true rate is larger by a factor of 500 than the observed isotropic estimated rate. However, Guetta, Piran, and Waxman (2004) have repeated this calculation, performing a careful average over the luminosity function, and find that the true rate is only a factor of $\sim 75 \pm 25$ times the isotropic estimate. Overall, the true rate is $33 \pm 11 h_{65}^3 \text{ Gpc}^{-3} \text{ yr}^{-1}$.

As the number of GRB's with redshifts increases, it may soon be possible to determine the GRB redshift distribution directly from the data. However, it is not clear which observational biases influence this data set, and one needs a homogenous data set in order to perform the calculation. Alternatively one could try to determine luminosity estimators (Fenimore and Ramirez-Ruiz, 2000; Norris *et al.*, 2000; Schaefer *et al.*, 2001; Schafer, 2003) from the subset with known redshifts and use them to obtain a redshift distribution for the whole GRB sample. Lloyd-Ronning *et al.* (2002) find, using the Fenimore and Ramirez-Ruiz (2000) sample, that this method implies that

- (i) the rate of GRB's increases backwards with time even for $z > 10$;
- (ii) the luminosity of GRB's increases with redshift as $(1+z)^{1.4 \pm 0.5}$;

⁷A partial list of calculations includes the work of Piran, 1992, 1999; Cohen and Piran, 1995; Fenimore and Bloom, 1995; Loredó and Wasserman, 1995, 1998; Horack and Hakkila, 1997; Schmidt, 1999, 2001a, 2001b; Sethi and Bhargavi, 2001.

(iii) hardness and luminosity are strongly correlated.

It is not clear how these features, which clearly depend on the inner engine, could depend strongly on the redshift. Note that in view of the luminosity-angle relation (see Sec. II.D below) the luminosity depends mostly on the opening angle. An increase of the luminosity with redshift would imply that GRB's were more narrowly collimated at earlier times.

4. Association with supernovae

The association of GRB's with star-forming regions and the indications that GRB's follow the star formation rate suggest that GRB's are related to stellar death, namely, to supernovae (Paczynski, 1998). Additionally there is some direct evidence of association of GRB's with supernovae.

a. GRB 980425 and SN98bw

The first indication of an association between GRB's and supernovae was found when SN 1998bw was discovered within the error box of GRB 980425 (Galama *et al.*, 1998). This was an unusual type-Ic supernova, much brighter than most. Typical ejection velocities in the supernova were also larger than usual ($\sim 2 \times 10^4$ km/sec), corresponding to a kinetic energy of 2×10^{52} ergs, more than ten times greater than previously known supernova energies (Iwamoto *et al.*, 1998). Additionally radio observations suggested a component expanding subrelativistically with $v \sim 0.3c$ (Kulkarni *et al.*, 1998). Thus SN 1998bw was significantly more powerful than comparable supernovae. This may imply that GRB's are associated with more powerful supernovae. Indeed all other observations of supernova signatures in GRB afterglow light curves use a SN 1998bw template. The accompanying GRB, 980425, was also unusual in that it had a smooth FRED light curve and no high-energy component in its spectrum. Other bursts like this exist but they are rare. The redshift of SN 1998bw was 0.0085, implying an isotropic equivalent energy of $\sim 10^{48}$ ergs, weaker by several orders of magnitude than a typical GRB.

The BeppoSAX wide-field cameras have localized GRB 980425 to within an 8-arcmin-radius accuracy. In this circle, the BeppoSAX NFI (Narrow-Field Instrument) has detected two sources, S1 and S2. The NFI could associate with each of these two sources an error circle of 1.5-arcmin radius. The radio and optical position of SN 1998bw were consistent only with the NFI error circle of S1, and were out of the NFI error circle of S2. Therefore Pian *et al.* (2000) identified S1 with x-ray emission from SN 1998bw, although this was of course no proof of association between the supernova and the GRB. It was difficult, based only on the BeppoSAX NFI data, to characterize the behavior and variability of S2, and it could not be excluded that S2 was the afterglow of GRB 980425. The XMM observations of March 2002 (Pian *et al.*, 2003) seem to have brought us closer to the solution. XMM detects S1 well, and its flux is lower than in 1998: the supernova emission has evidently decreased.

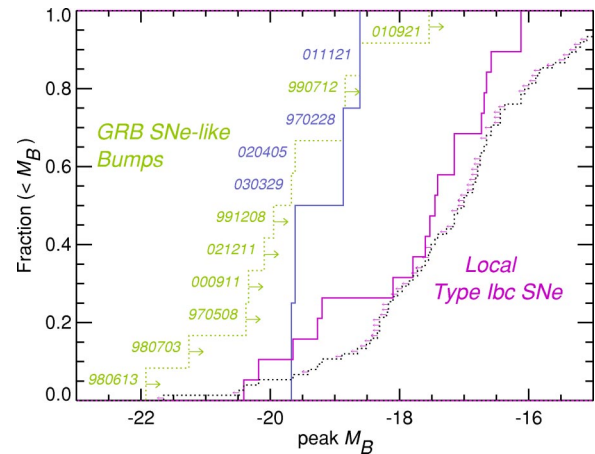


FIG. 10. (Color in online edition) Comparison of the peak magnitudes of GRB supernovae with those of local type-Ib/Ic supernovae. Left, solid cumulative histogram for those GRB's with a believable detection of a supernova bump; dotted histogram, the brightness of all other claimed GRB supernova peaks or reported upper limits. Right, solid histogram for those local Ib/c supernovae for which the supernova was observed before peak; all others were discovered after peak. The faintest GRB-supernova nondetection (010921) probes only the top ~ 40 th percentile of local type-Ib/Ic supernovae. It is clear that the current GRB-supernovae population may have revealed only the tip of the iceberg; plausibly, then, supernovae could accompany all long-duration GRB's. From Bloom, 2003.

XMM, having a better angular resolution than the BeppoSAX NFI's, seems to have resolved S2 into a number of sources. In other words, S2 appears to be not a single source, but a group of small faint sources. Their random variability (typical fluctuations of x-ray sources close to the level of the background) may have caused the flickering detected for S2. This demolishes the case for the afterglow nature of S2 and strengthens the case for an association between GRB 980425 and SN 1998bw.

b. Red bumps

Late red bumps (see Sec. II.B.2) have been discovered in several GRB light curves (Bloom *et al.*, 1999; Reichart, 1999; Bloom, Kulkarni, Price, *et al.*, 2000; Garnavich *et al.*, 2003). These bumps involve both a brightening (or a flattening) of the afterglow as well as a transition to a much redder spectrum and have been generally interpreted as due to an underlying supernova (Bloom *et al.*, 1999). In all cases the bumps have been fit with a template of SN 1998bw, which was associated with GRB 980425. Esin and Blandford (2000) proposed that the bumps are produced by light echoes on surrounding dust (but see Reichart, 2001). Waxman and Draine (2000) proposed an alternative explanation based on dust sublimation.

For most GRB's there is an upper limit to the magnitude of the bump in the light curve. A comparison of these upper limits (see Fig. 10) with the maximal magnitudes of type-Ib/c supernovae shows that the faintest

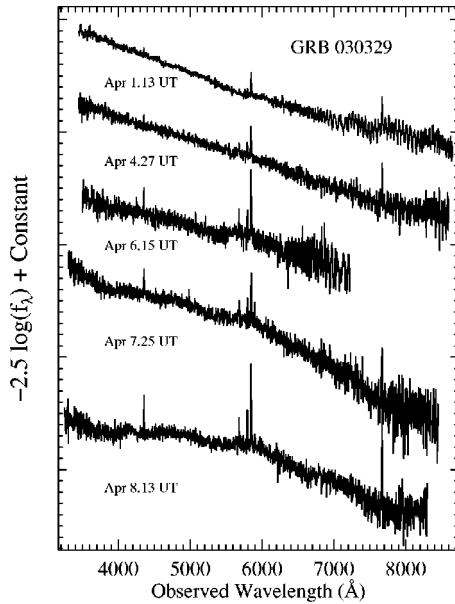


FIG. 11. Evolution of the GRB 030329/SN2004dh spectrum from April 1.13 UT (2.64 days after the burst) to April 8.13 UT (6.94 days after the burst). The early spectra consist of a power-law continuum ($F_\nu \propto \nu^{-0.9}$) with narrow emission lines originating from HII regions within the host galaxy at redshift $z=0.168$. Spectra taken after April 5 show the development of broad peaks characteristic of a supernova. From Stanek *et al.*, 2003.

GRB-SN nondetection (GRB 010921) probes only the top ~ 40 th percentile of local type-Ib/Ic supernovae. It is clear that the current GRB-supernova population may have revealed only the tip of the iceberg; plausibly, then, supernovae could accompany all long-duration GRB's.

c. GRB 030329 and SN 2003dh

Recently the association of a GRB with a supernova was dramatically confirmed (Hjorth *et al.*, 2003; Stanek *et al.*, 2003) in the very bright GRB 030329, which is associated with SN 2003dh (Chornock *et al.*, 2003). A bump begun to be noticed six days after the burst and its SN 1999bw-like spectrum dominated the optical light curve at later times (see Fig. 11). The spectral shapes of SN 2003dh and SN 1998bw are quite similar, although there are also differences. In Sec. II.D below, we estimate a somewhat larger expansion velocity for SN 2003dh. Additionally the x-ray signal was much brighter (but this could be purely afterglow).

For most researchers in the field this discovery provided the final link between supernovae and GRB's (at least, long GRB's). As the supernova signature coincides with the GRB, this observation also provides evidence against the supranova interpretation, in which the GRB arises from the collapse of a neutron star sometime after the supernova in which the neutron star was born—see Sec. IX.E. (The supranova scenario might still apply if there were a variety of supranova types, some with long delay and others with short delay, between the first and the second collapses.) The spectral shapes of SN 2003dh

and SN 1998bw are quite similar, although there are also differences. For example, there is a slightly larger expansion velocity for 2003dh. It is interesting that while not as weak as GRB 990425, the accompanying GRB 990329 is significantly weaker than average. The implied opening angle reveals that the prompt γ -ray energy output, E_γ , and the x-ray luminosity at 10 h, L_X , are a factor of ~ 20 and ~ 30 , respectively, below the average value around which most GRB's are narrowly clustered (see Sec. II.D below).

It is interesting to compare SN 1999bw and SN 2003dh. Matheson *et al.* (2003) find that basically, at all epochs the best fit to the spectra of 2003dh is given by 1998bw at about the same age. The light curve is harder, as the afterglow contribution is significant, but using spectral information they find that 2003dh had basically the same light curve as 1998bw. Mazzali *et al.* (2003) model the spectra and find the same similarity. They find some differences, but some of these might be due to a somewhat different approach to spectral decomposition, which gives somewhat fainter supernovae.

d. X-ray lines

The appearance of iron x-ray lines (see Sec. II.B.1) has been interpreted as additional evidence for a link with supernovae. One has to be careful with this interpretation, as the iron x-ray lines are seen as if emitted by matter at very low velocities and at rather large distances. This is difficult to achieve if the supernova is simultaneous with the GRB, as the supernova bumps imply. These x-ray lines might be consistent with the supranova model (Vietri and Stella, 1998) in which the supernova takes place a month before the GRB. However, in this case there will not be a supernova bump in the light curve. Rees and Mészáros (2000), Mészáros and Rees (2001), and Kumar and Narayan (2003) suggest alternative interpretations which do not require a supranova.

While the association of supernovae with GRB's is generally accepted now, I should point out that as early as 1994 Katz (1994b) predicted that GRB's might be accompanied by events like type-II supernovae. According to Katz's model these events arise due to the deposition of some of the GRB energy into a nearby cloud and not due to core collapse. Depending on the energy deposited, the resulting optical emission would be expected to have a brightness comparable to the usual supernova or even brighter. Such a scenario would explain the red bumps maintained in Sec. II.B.2

D. Energetics

Before redshift measurements were available, GRB energy was estimated from the BATSE catalog by fitting an (isotropic) luminosity function to the flux distribution (see, for example, Cohen and Piran 1995; Loredano and Wasserman, 1998; Schmidt, 1999, 2001a, 2001b; Guetta, Piram, and Waxman, 2004; and many others). This led to

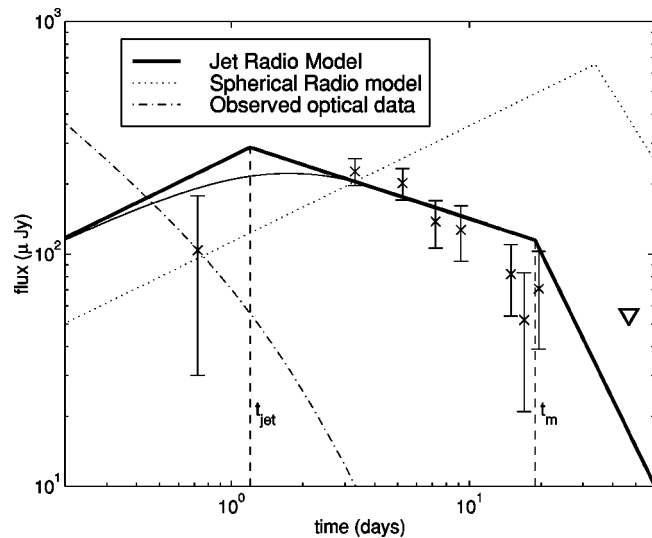


FIG. 12. Observed and predicted light curve at 8.6 GHz light curves of GRB 990510. The different behaviors of the optical and radio light curves after the jet break are clearly seen. From Harrison *et al.*, 1999.

a statistical estimate of the luminosity function of a distribution of bursts.

These estimates were revolutionized with the direct determination of the redshift for individual bursts. Now the energy could be estimated directly for specific bursts. Given an observed γ -ray fluence and the redshift to a burst, one could easily estimate the energy emitted in γ rays, $E_{\gamma,iso}$ assuming that the emission is isotropic (see Bloom *et al.*, 2001 for a detailed study including k correction). The inferred energy $E_{\gamma,iso}$ was the isotropic energy, that is, the energy assuming that the GRB emission was isotropic in all directions. The energy of the first burst with a determined redshift, GRB 970508, was around 10^{51} ergs. However, as afterglow observations proceeded, alarmingly large values (e.g., 3.4×10^{54} ergs for GRB 990123) were measured for $E_{\gamma,iso}$. The variance was around three orders of magnitude.

It turned out (Rhoads, 1999; Sari *et al.*, 1999) that GRB's are beamed, and thus $E_{\gamma,iso}$ is not a good estimate for the total energy emitted in γ rays. Instead, $E_{\gamma} \equiv (\theta^2/2)E_{\gamma,iso}$ was proposed, where the angle θ is the effective angle of γ -ray emission. It can be estimated from t_b , the time of the break in the afterglow light curve (Sari *et al.*, 1999):

$$\theta = 0.16(n/E_{k,iso,52})^{1/8} t_{b,days}^{3/8} = 0.07(n/E_{k,\theta,52})^{1/6} t_{b,days}^{1/2}, \quad (4)$$

where $t_{b,days}$ is the break time in days. $E_{k,iso,52}$ is the “isotropic equivalent” kinetic energy, discussed below, in units of 10^{52} ergs, while $E_{k,\theta,52}$ is the real kinetic energy in the jet, i.e., $E_{k,\theta,52} = (\theta^2/2)E_{k,iso,52}$. One has to be careful which of the two energies one discusses. In the following I shall usually consider, unless specifically noted otherwise, $E_{k,iso,52}$, which is also related to the energy per unit solid angle as $E_{k,iso,52}/4\pi$. The jet break is observed in both the optical and the radio frequencies.

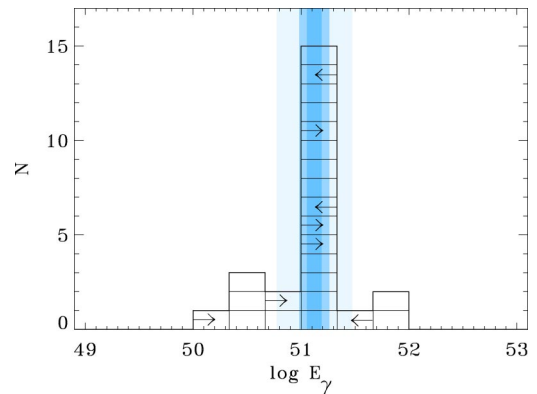


FIG. 13. (Color in online edition) Histogram of GRB energies (E_{γ}) with three equal logarithmic spacings per decade. The histogram shows a narrow distribution of GRB energies about the standard energy $E_{\gamma} = 1.33 \times 10^{51}$ ergs, with an error of $\sigma = 0.07$ dex. The observed burst-to-burst rms spread is 0.35 dex (a factor of 2.23) about this value. Bands of 1, 2, and 5 σ about the standard energy are shown. There are five identifiable outliers, which lie more than 5 σ from the mean. However, there is currently no basis other than discrepant energy to exclude these bursts from the sample. From Bloom *et al.*, 2003.

Note that the observational signature in the radio differs from that at optical and x-ray wavelengths (Harrison *et al.*, 1999; Sari *et al.*, 1999; see Fig. 12) and this provides an additional confirmation for this interpretation.

Frail *et al.* (2001) estimated E_{γ} for 18 bursts, finding typical values around 10^{51} ergs (see also Panaitescu and Kumar, 2001). Bloom *et al.* (2003) found $E_{\gamma} = 1.33 \times 10^{51} h_{65}^{-2}$ ergs and a burst-to-burst variance about this value ~ 0.35 dex, a factor of 2.2. This is three orders of magnitude smaller than the variance in the isotropic equivalent E_{γ} . A compilation of the beamed energies from Bloom *et al.* (2003) is shown in Figs. 13 and 14, nicely demonstrating this phenomenon. The constancy

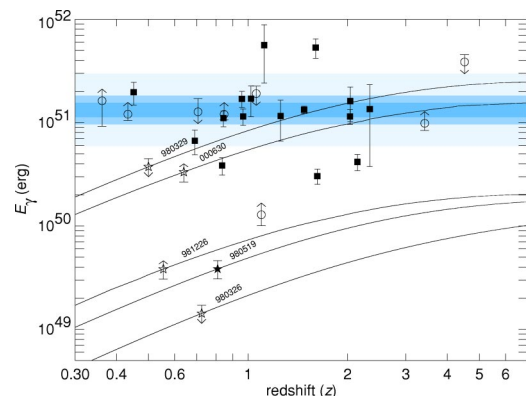


FIG. 14. (Color in online edition) GRB energy release vs redshift. Bands of 1, 2, and 5 σ about the mean energy $E_{\gamma} = 1.33 \times 10^{51}$ ergs are shown. Star symbols: trajectories of five GRB's with no known spectroscopic redshift. While the energies of GRB 980329 and GRB 000630 could be consistent with the standard value at redshifts beyond $z \sim 1.5$, the energies of GRB 980326 and GRB 980519 were consistent at no redshift. From Bloom *et al.*, 2003.

of E_γ is remarkable, as it involves a product of a factor inferred from the GRB observation (the γ -ray flux) with a factor inferred from the afterglow observations (the jet opening angle). However, E_γ might not be a good estimate for E_{tot} , the total energy emitted by the central engine. First, an unknown conversion efficiency of energy to γ rays has to be considered: $E_{\text{tot}} = \varepsilon^{-1} E_\gamma = \varepsilon^{-1} (\theta^2/2) E_{\gamma, \text{iso}}$. Second, the large Lorentz factor during the γ -ray emission phase makes the observed E_γ rather sensitive to angular inhomogeneities of the relativistic ejecta (Kumar and Piran, 2000a). The recent early observations of the afterglow of GRB 021004 indicate that indeed a significant angular variability of this kind exists (Nakar and Piran, 2003a; Nakar, Piran, and Granot, 2003).

The kinetic energy of the flow during the adiabatic afterglow phase, E_k , is yet another energy measure. This energy (per unit solid angle) can be estimated from the afterglow light curve and spectra. Specifically, it is rather closely related to the observed afterglow x-ray flux (Kumar, 2000; Freedman and Waxman, 2001; Piran *et al.*, 2001). As this energy is measured when the Lorentz factor is smaller, it is less sensitive than E_γ to angular variability. The constancy of the x-ray flux (Piran *et al.*, 2001) suggests that this energy is also constant. Estimates of $E_{k, \theta}$ (Panaitescu and Kumar, 2001) show that $\bar{E}_\gamma \approx 3 \bar{E}_{k, \theta}$, that is, the observed “beamed” GRB energy is larger than the estimated “beamed” kinetic energy of the afterglow. Frail *et al.* (2001), however, find that $\bar{E}_\gamma \approx \bar{E}_{k, \theta}$, that is, that the two energies are comparable.

An alternative interpretation to the observed breaks is that we are viewing a universal angle-dependent structured jet from different viewing angles (Lipunov *et al.*, 2001; Rossi *et al.*, 2002; Zhang and Mészáros, 2002). The observed break corresponds in this model to the observing angle θ and not to the opening angle of the jet. This interpretation means that the GRB beams are wide and hence the rate of GRB’s is smaller than the rate implied by the usual beaming factor. On the other hand, it implies that GRB’s are more energetic. Guetta, Piran, and Waxman (2004) estimate this factor (the ratio of the fixed energy of a “structured” jet to the energy of a uniform jet) to be ~ 7 . However, they find that the observing angle distribution is somewhat inconsistent with the simple geometric one that should arise in universal structured jets (see also Perna *et al.*, 2003; Nakar, Granot, and Guetta, 2004). The energy-angle relation discussed earlier requires (see Sec. VII.I below) an angle-dependent jet with $E(\theta) \propto \theta^{-2}$.

Regardless of the nature of the jet (universal structured jet or uniform one with an opening angle that differs from one burst to another), at late times it becomes nonrelativistic and spherical. With no relativistic beaming, every observer detects emission from the whole shell. Radio observations at this stage enable us to obtain a direct calorimetric estimate of the total kinetic energy of the ejecta at late times (Frail, Waxman, and Kulkarni, 2000). Estimates performed in several cases yield comparable values for the total energy.

If GRB’s are beamed, we should expect orphan afterglows (see Sec. VII.K), events in which we would miss the GRB but observe the late afterglow that is not so beamed. A comparison of the rate of orphan afterglows to GRB’s would give us a direct estimate of the beaming of GRB’s (and hence of their energy). Unfortunately there are not even good upper limits on the rate of orphan afterglows. Veerswijk (2003) considers the observations within the Faint Sky Variability Survey (FSVS) carried out on the wide-field camera on the 2.5-m Isaac Newton Telescope on La Palma. This survey mapped 23 square degrees down to a limiting magnitude of about $V=24$. One object was found which faded and was not detectable after a year. However, its colors suggest that it was a supernova and not a GRB. Similarly, Vanden Berk *et al.* (2002) found a single candidate within the Sloan Digital Sky Survey. Here the colors were compatible with an afterglow. However, it was later revealed that this was a variable active galactic nucleus and not an orphan afterglow. As I discuss later these limits are still far from constraining the current beaming estimates (see Sec. VII.K).

One exception is late radio emission for which there are some limits (Perna and Loeb, 1998; Levinson *et al.*, 2002). Levinson *et al.* (2002) show that the number of orphan radio afterglows associated with GRB’s that should be detected by a flux-limited radio survey is smaller for a smaller jet opening angle θ . This might seem at first sight contrary to expectation, as narrower beams imply more GRB’s. But, on the other hand, with narrower beams each GRB has a lower energy and hence its radio afterglow is more difficult to detect. Overall the second factor wins. Using the results of the Faint Images of the Radio Sky at Twenty-centimeters (FIRST) and NRAO VLA Sky Survey (NVSS) surveys, they find nine afterglow candidates. If all candidates are associated with GRB’s, then there is a lower limit on the beaming factor of $f_b^{-1} \equiv (\theta^2/2) > 13$. If none is associated with a GRB they find $f_b^{-1} > 90$. This gives immediately a corresponding upper limit on the average energies of GRB’s. Guetta, Piran, and Waxman (2004) revise these values, using this method, in view of a recent estimate of the correction to the rate of GRB’s to $f_b^{-1} = 40$.

When considering the energy of GRB’s one has to remember the possibility, as some models suggest, that additional energy is emitted that is not involved in the GRB itself or in the afterglow. van Putten and Levinson (2001), for example, suggest that a powerful Newtonian wind collimates the less powerful relativistic one. The “universal standard jet” model also suggests a large amount of energy emitted sideways with a lower energy per solid angle and a lower Lorentz factor. It is interesting to note that the calorimetric estimates mentioned earlier limit the total amount of energy ejected regardless of the nature of the flow. More generally, typically during the afterglow matter moving with a lower Lorentz factor emits lower frequencies. Hence by comparing the relative beaming of afterglow emission in different wavelengths one can estimate the relative beaming

factors $f_b^{-1}(E)$ at different wavelengths and hence at different energies. Nakar and Piran (2003b) use various x-ray searches for orphan x-ray afterglows to limit the (hard) x-ray energy to be at most comparable to the γ -ray energy. This implies that the total energy of matter moving at a Lorentz factor of ~ 40 is at most comparable to the energy of matter moving with a Lorentz factor of a few hundred and producing the GRB itself. At present, limits on optical orphan afterglow are insufficient to set significant limits on matter moving at a slower rate, while as mentioned earlier radio observations already limit the overall energy output.

These observations will not, of course, limit the energy emitted in gravitational radiation, neutrinos, cosmic rays, or very-high-energy photons that may be emitted simultaneously by the source and influence the source's energy budget without influencing the afterglow.

III. THE GLOBAL PICTURE—GENERALLY ACCEPTED INGREDIENTS

There are several generally accepted ingredients in practically all current GRB models.

- **Relativistic motion:** Practically all current GRB models involve relativistic motion with a Lorentz factor $\Gamma > 100$. This is essential to overcome the compactness problem (see Sec. IV.A below). At first this understanding was based only on theoretical arguments. However, now there are direct observational proofs of this concept: It is now generally accepted that both radio scintillation (Goodman, 1997) and lower-frequency self-absorption (Katz and Piran, 1997) provide independent estimates of the size of the afterglow, $\sim 10^{17}$ cm, two weeks after the burst. These observations imply that the afterglow has indeed expanded relativistically. Sari and Piran (1999b) suggested that the optical flash accompanying GRB 990123 provided direct evidence for ultrarelativistic motion, with $\Gamma \sim 100$. Soderberg and Ramirez-Ruiz (2003) found a higher value: 1000 ± 100 . However, these interpretations were model dependent. Recently Taylor *et al.* (2004) detected superluminal expansion in the afterglow of GRB 030329 (see also Oren, Nakar, and Piran, 2004).

Relativistic motion implies that we are observing blueshifted photons that are significantly softer in the moving rest frame. It also implies that when the objects have a size R the observed emission arrives on a typical time scale of $R/c\Gamma^2$ (see Sec. IV.B). Relativistic beaming also implies that we observe only a small fraction ($1/\Gamma$) of the source. As I discussed earlier (see Secs. II.D and IV.C) this has important implications for our ability to estimate the total energy of GRB's.

While all models are based on ultrarelativistic motion, none explains convincingly (this is clearly a subjective statement) how this relativistic motion is attained. There is no agreement even on the nature of the relativistic flow. While in some models the energy

is carried out in the form of kinetic energy of baryonic outflow, in others it is a Poynting-dominated flow or both.

- **Dissipation:** In most models the energy of the relativistic flow is dissipated and this provides the energy needed for the GRB and the subsequent afterglow. The dissipation is in the form of (collisionless) shocks, possibly via plasma instability. There is a general agreement that the afterglow is produced via external shocks with circumburst matter (see Sec. VII). There is convincing evidence (see, for example, Fenimore *et al.*, 1996; Sari and Piran, 1997b; Ramirez-Ruiz and Fenimore, 2000; Piran and Nakar, 2002, and Sec. VI.A below) that in most bursts the dissipation during the GRB phase takes place via internal shocks, that is, shocks within the relativistic flow itself. Some disagree with this statement (e.g., Dermer and Mitman, 1999; Heinz and Begelman, 1999; Ruffini *et al.*, 2001; Dar, 2003).
- **Synchrotron radiation:** Most models (both of the GRB and of the afterglow) are based on synchrotron emission from relativistic electrons accelerated within the shocks. There is a reasonable agreement between the predictions of the synchrotron model and afterglow observations (Granot *et al.*, 1999a; Wijers and Galama, 1999; Panaitescu and Kumar, 2001). These are also supported by measurements of linear polarization in several optical afterglows (see Sec. II.B.2). As for the GRB itself, there are various concerns about the validity of this model. In particular there are some inconsistencies between the observed spectral slopes and those predicted by the synchrotron model (see Preece *et al.*, 2002 and Sec. II.A.1). The main alternative to synchrotron emission is synchrotron self-Compton (Waxman, 1997b; Ghisellini and Celotti, 1999) or inverse Compton scattering of external light (Brainerd, 1994; Shemi, 1994; Shaviv and Dar, 1995; Lazzati *et al.*, 2003). The last model requires, of course, a reasonable source of external light.
- **Jets and collimation:** Monochromatic breaks appear in many afterglow light curves. These breaks are interpreted as “jet breaks” due to the sideways beaming of the relativistic emission (Panaitescu and Mészáros, 1999; Rhoads, 1999; Sari *et al.*, 1999)—when the Lorentz factor drops below $1/\theta_0$ the radiation is beamed outside of the original jet, reducing the observed flux—or due to sideways spreading of a beamed flow (Rhoads, 1999; Sari *et al.*, 1999). An alternative interpretation is that we have different viewing angles of “universal structured jets” (Lipunov *et al.*, 2001; Rossi *et al.*, 2002; Zhang and Mészáros, 2002) whose energy varies with the angle. Both interpretations suggest that GRB's are beamed. However, they give different estimates of the overall rate and the energies of GRB's (see Sec. VII.I below). In either case the energies involved are smaller than the naively interpreted isotropic energy, and the rate is higher than the observed rate.

- A (newborn) compact object: If one accepts the beaming interpretation of the breaks in the optical light curve, the total energy release in GRB's is $\sim 10^{51}$ ergs (Frail *et al.*, 2001; Panaitescu and Kumar, 2001). It is higher if, as some models suggest, the beaming interpretation is wrong or if a significant amount of additional energy (which does not contribute to the GRB or to the afterglow) is emitted from the source. This energy, $\sim 10^{51}$ ergs, is comparable to the energy released in a supernova. It indicates that the process must involve a compact object. No other known source can release so much energy within such a short time scale. The process requires a dissipation of $\sim 0.1M_\odot$ within the central engine over a period of a few seconds. The sudden appearance of so much matter in the vicinity of the compact object suggests a violent process, one that most likely involves the birth of the compact object itself.
- Association with star formation and supernovae: Afterglow observations, which exist for a subset of relatively bright long bursts, show that GRB's arise within galaxies with a high star formation rate (see Djorgovski, Frail, *et al.*, 2001 and Sec. II.C.1). Within galaxies the burst distribution follows the light distribution (Bloom, Kulkarni, and Djorgovski, 2002). This has led to the understanding that (long) GRB's arise from the collapse of massive stars (see Sec. IX.D). This understanding has been confirmed by the appearance of supernova bumps in the afterglow light curve (see Sec. II.C.4 earlier) and in particular by the association of SN 1999bw with GRB 980425 and of SN 2003dh with GRB 030329.
- Summary: Based on these generally accepted ideas, one can sketch the following generic GRB model: GRB's are a rare phenomenon observed within star-forming regions, associated with the death of massive stars and the birth of compact objects. The γ -ray emission arises from internal dissipation within a relativistic flow. This takes place at distances of $\sim 10^{13}$ – 10^{15} cm from the central source that produces the relativistic outflow. Subsequent dissipation of the remaining energy due to interaction with the surrounding circumburst matter produces the afterglow. The nature of the "inner engine" is not yet resolved, but the association with supernovae (like 1998bw and 2003dh) shows that long GRB's involve a collapsing star. Much less is known about the origin of short GRB's.

IV. RELATIVISTIC EFFECTS

A. Compactness and relativistic motion

The first theoretical clues to the necessity of relativistic motion in GRB's arose from the compactness problem (Ruderman, 1975). The conceptual argument is simple. GRB's show a nonthermal spectrum with a significant high-energy tail (see Sec. II.A.1), yet a naive calculation implies that the source is optically thick. The

fluctuations over a time scale δt imply that the source is smaller than $c\delta t$. Given an observed flux F , a duration T , and distance d , we can estimate the energy E at the source. For a typical photon energy \bar{E}_γ this yields a photon density $\approx 4\pi d^2 F / \bar{E}_\gamma c^3 \delta t^2$. Now, two γ rays can annihilate and produce e^+e^- pairs, if the energy in their center-of-mass frame is larger than $2m_e c^2$. The optical depth for pair creation is

$$\tau_{\gamma\gamma} \approx \frac{f_{e^\pm} \sigma_T 4\pi d^2 F}{\bar{E}_\gamma c^2 \delta t}, \quad (5)$$

where f_{e^\pm} is a numerical factor denoting the average probability that one photon will collide with another photon whose energy is sufficient for pair creation. For typical values and cosmological distances, the resulting optical depth is extremely large, $\tau_{\gamma\gamma} \sim 10^{15}$ (Piran, 1997). This is, of course, inconsistent with the nonthermal spectrum.

The compactness problem can be resolved if the emitting matter is moving relativistically towards the observer. I denote the Lorentz factor of the motion by Γ . Two corrections appear in this case. First, the observed photons are blueshifted, and therefore their energy at the source frame is lower by a factor Γ . Second, the implied size of a source moving towards us with a Lorentz factor Γ is $c\delta t\Gamma^2$ (see Sec. IV.B below). The first effect modifies f_{e^\pm} by a factor $\Gamma^{-2\alpha}$, where α is the photon's index of the observed γ rays (namely, the number of observed photons per unit energy is proportional to $E^{-\alpha}$). The second effect modifies the density estimate by a factor Γ^{-4} and influences the optical depth as Γ^{-2} . Taking both effects together, one finds that for $\alpha \sim 2$ one needs $\Gamma \gtrsim 100$ to obtain an optically thin source.

The requirement that the source be optically thin can be used to obtain from specific bursts direct limits on the minimal Lorentz factor within those bursts (Krolik and Pier, 1991; Fenimore *et al.*, 1993; Woods and Loeb, 1995; Baring and Harding, 1997; Piran, 1997, 1999; Lithwick and Sari, 2001). A complete calculation requires a detailed integration over angular integrals and over the energy-dependent pair-production cross section. The minimal Lorentz factor depends also on the maximal photon energy E_{\max} , the upper energy cutoff of the spectrum. Lithwick and Sari (2001) provide a detailed comparison of the different calculations and point out various flaws in some of the previous estimates. They find that

$$\tau_{\gamma\gamma} = \frac{11}{180} \frac{\sigma_T d^2 (m_e c^2)^{-\alpha+1} \mathcal{F}}{c^2 \delta T (\alpha-1)} \times \left(\frac{E_{\max}}{m_e c^2} \right)^{\alpha-1} \Gamma^{-(2\alpha+2)} (1+z)^{-2\alpha-2}, \quad (6)$$

where the high end of the observed photon flux is given by $\mathcal{F}E^{-\alpha}$ (photons per cm^2 per sec per unit photon energy). A lower limit on Γ is obtained by equating Eq. (6) to unity.

B. Relativistic time effects

Consider first a source moving relativistically with a constant velocity along a line towards the observer and two photons emitted at R_1 and R_2 . The first photon (emitted at R_1) will reach the observer at time $(R_2 - R_1)/v - (R_2 - R_1)/c$ before the second photon (emitted at R_2). For $\Gamma \gg 1$ this equals $\approx (R_2 - R_1)/2c\Gamma^2$. This allows us to associate an “observer time” $R/2c\Gamma^2$ with the distance R , and for this reason I have associated a scale $c\delta t\Gamma^{-2}$ with fluctuations on a time scale δt in the optical depth equation earlier (see Sec. IV.A). This last relation should be modified if the source moves at a varying velocity [$v = v(R)$]. Now

$$\delta t_{12} \approx \int_{R_1}^{R_2} \frac{dR}{2c\Gamma^2(R)}, \quad (7)$$

which reduces to

$$T_R \approx R/2c\Gamma^2 \quad (8)$$

for motion with a constant velocity. The difference between a constant-velocity source and a decelerating source introduces a numerical factor of order 8, which is important during the afterglow phase (Sari, 1997).

Consider now a relativistically expanding spherical shell, or at least a shell that is locally spherical (on a scale larger than $1/\Gamma$). Emission from parts of the shell moving at angle θ relative to the line of sight to the observer will arrive later with a time delay $R(1 - \cos \theta)/c$. For small angles this time delay equals $R\theta^2/2c$. As the radiation is beamed with an effective beaming angle $\approx 1/\Gamma$ most of the radiation will arrive within a typical angular time scale

$$T_{ang} \equiv R/2c\Gamma^2. \quad (9)$$

The combination of time delay and blueshift implies that if the emitted spectrum is a power-law spectrum with a spectral index α , then the observed signal from the instantaneous emission of a thin shell will decay at late time as a power law with $t^{-(2-\alpha)}$ (Fenimore *et al.*, 1996; Nakar and Piran, 2003b). The observed pulse from an instantaneous flash from a thin shell is shown in Fig. 15.

As I discuss later (see Sec. VI.A) the similarity between the angular time scale and the radial time scale plays a crucial role in GRB models.

C. Relativistic beaming and the patchy shell model

The radiation from a relativistic source is beamed with a typical beaming angle $1/\Gamma$. This implies that, if the source is expanding radially with an ultrarelativistic speed, a given observer “sees” radiation only from a region that is within Γ^{-1} from its line of sight to the source. If the radius of the emitting region is R , the observer will see radiation from a region of size R/Γ . Since Γ is extremely large during the GRB, we observe emission only from a small fraction of the emitting shell. It is possible, and even likely, that the conditions within the small region that we observe will be different from the average

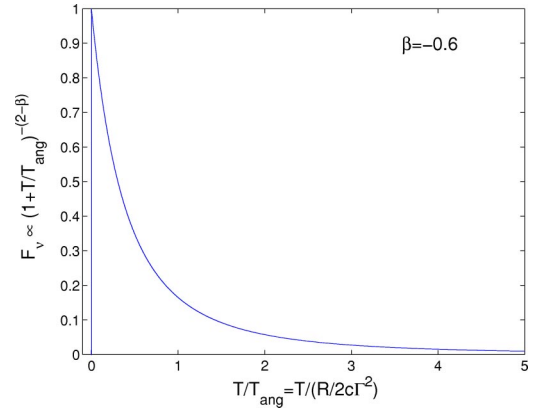


FIG. 15. (Color in online edition) Observed pulse from an instantaneous flash from a spherical relativistic thin shell moving relativistically and emitting with a power law $\nu^{-0.6}$.

ones across the shell. This means that the conditions that we infer will not reflect the true average conditions within this particular GRB.

An interesting point related to the internal-shock model (discussed later) in this context is the following. According to the internal-shock model, individual pulses are obtained by collisions between individual shells. Here the inhomogeneity of individual shells could be wiped out when the contributions of different hot spots from different shells is added. Alternatively, the “inner engine” might produce a consistent angular pattern in which the hot spot was in the same position in all shells, and in this case averaging would not lead to a cancellation of the patchy shell structure.

Within the internal-external model, the GRB is produced by internal shocks in which only the relative motion within the flow is dissipated. The bulk Lorentz factor remains unchanged. During the afterglow the shell is slowed down by external shocks. As the Lorentz factor decreases with time [see Eq. (78)] we observe a larger and larger fraction of the emitting region until $\Gamma \approx \theta^{-1}$, where θ is the angular size of the whole emitting region—the GRB jet (see Sec. VII.H). This has several inevitable implications. If the initial relativistic flow is inhomogeneous on a small angular scale, then different observers looking at the same GRB (from different viewing angles) would see different γ -ray light curves. A strong burst to one observer might look weak to another if it were located at an angle larger than $1/\Gamma$ from the first. The two observers will see similar conditions later on, during the afterglow, as then they will observe the same angular regions. This has the following implications:

- (i) Given that the GRB population originates from some “typical” distribution, we expect that fluctuation between different bursts at early times during the GRB will be larger than fluctuations observed at late times during the afterglow (Kumar and Piran, 2000a). A direct consequence of this behavior is the appearance of a bias in the observations of GRB’s. As we are more likely to

detect stronger events, we will tend to identify bursts in which a “hot spot” was pointing towards us during the GRB phase. If the original GRB shells are inhomogeneous, this would inevitably lead to a bias in the estimates of the GRB emission as compared to the kinetic energy during the afterglow.

- (ii) As the afterglow slows down, we observe a larger and larger region. The angular structure would produce a variability in the light curve with a typical time scale of t , the observed time. These fluctuations will decay later as the Lorentz factor decreases and the observations are averaged over a larger viewing angle. Nakar, Piran, and Granot (2003) have suggested that this is the source of the early fluctuations in the light curve of GRB 021004. Nakar and Oren (2003) modeled this process with a numerical simulation. They found that the fluctuation light curve of GRB 021004 could be nicely fitted by this model and that it also explained the correlated fluctuations in the polarization (see also Granot, 2003).

V. PHYSICAL PROCESSES

The observed prompt emission must be generated by energetic particles that have been accelerated within collisionless shocks. The most likely process is synchrotron emission, even though there is some evidence that a simple synchrotron spectrum does not fit all bursts (Preece *et al.*, 2002; but see Barraud *et al.*, 2003, who find consistency with the synchrotron model). I consider here the different physical ingredients that determine the emission process: particle acceleration, magnetic-field amplification, synchrotron emission, and inverse Compton emission, which could be relevant in some cases.

A. Relativistic shocks

Shocks involve sharp jumps in the physical conditions. Conservation of mass, energy, and momentum determine the Hugoniot shock jump conditions across the relativistic shocks for the case when the upstream matter is cold (see Blandford and McKee, 1976):

$$\begin{aligned} n_2 &= 4\Gamma n_1, \\ e_2 &= 4\Gamma n_1 m_p c^2, \\ \Gamma_{sh}^2 &= 2\Gamma^2, \end{aligned} \quad (10)$$

where $n_{1,2}, e_{1,2}$ are the number density and the energy density (measured in the local rest frame) of the matter upstream (region 1) and downstream (region 2). I have assumed that the energy density in region 1 is very small compared to the rest mass density. Γ is the Lorentz factor of the fluid just behind the shock, and Γ_{sh} is the Lorentz factor of the shock front (both measured in the rest frame of the upstream fluid). Matter is compressed by a factor Γ across a relativistic shock. The pressure, or

the internal energy density behind the shock, is of order $\Gamma^2 n_1 m_p c^2$. Thus in the shock's rest frame the relative “thermal” energy per particle (downstream) is of the same order as the kinetic energy per particle (ahead of the shock) upstream. Put differently, the shock converts the “ordered” kinetic energy to a comparable random kinetic energy. In an ultrarelativistic shock the downstream random velocities are ultrarelativistic.

Similar jump conditions can be derived for the magnetic fields across the shock. A parallel magnetic field (parallel to the shock front) B_{\parallel} is compressed and amplified:

$$B_{\parallel 2} = \Gamma B_{\parallel 1}. \quad (11)$$

A perpendicular magnetic field B_{\perp} remains unchanged.

The energy distribution of the (relativistic) electrons and the magnetic field behind the shock are needed to calculate the synchrotron spectrum. In principle, these parameters should be determined from the microscopic physical processes that take place in the shocks. However, it is difficult to estimate them from first principles. Instead I define two dimensionless parameters, ϵ_B and ϵ_e , that incorporate our ignorance and uncertainties (Paczynski and Rhoads, 1993; Piran, 1994; Sari *et al.*, 1996). It is commonly assumed that these energies are a constant fraction of the internal energy behind the shock (see, however, Daigne and Mochkovitch, 2003). I denote by ϵ_e and ϵ_B the ratio between these energies and the total internal energy:

$$\begin{aligned} e_e &\equiv \epsilon_e e = 4\Gamma_{sh}^2 \epsilon_e n_1 m_p c^2, \\ e_B &= B^2/8\pi \equiv \epsilon_B e = 4\Gamma_{sh}^2 \epsilon_B n_1 m_p c^2. \end{aligned} \quad (12)$$

One usually assumes that these factors ϵ_e, ϵ_B , are constant throughout the burst evolution. One may even expect that they should be constant from one burst to another, as they reflect similar underlying physical processes. However, it seems that a simple model that assumes that these parameters are constant during the prompt burst cannot reproduce the observed spectrum (Daigne and Mochkovitch, 2003). This leads us to explore models in which the equipartition parameters $\epsilon_{e,B}$ depend on the physical conditions within the matter.

In GRB's, as well as in supernova remnants, the shocks are collisionless. The densities are so low that the mean free path of the particles for collisions is larger than the typical size of the system. However, one expects that ordered or random magnetic fields or alternatively plasma waves will play the role of particle collisions in these shocks. One can generally use in these cases the Larmor radius as a typical shock width. A remarkable feature of the above shock jump conditions is that, as they arise from general conservation laws, they are independent of the detailed conditions within the shocks and hence are expected to hold within collisionless shocks as well. See, however, Mitra (1996) for a discussion of the conditions for collisionless shocks in GRB's.

B. Particle acceleration

It is now generally accepted that cosmic rays (more specifically the lower-energy component below 10^{15} eV) are accelerated within (collisionless) shocks in supernova remnants in the Galaxy (see Gaisser, 1991). A beautiful demonstration of this effect arises in the observation of synchrotron emission from supernova remnants, which shows x-ray emission from these accelerated particles within the shocks.

The common model for particle shock acceleration is the diffuse shock acceleration model. According to this model the particles are accelerated when they repeatedly cross a shock. Magnetic-field irregularities keep scattering the particles back so that they keep crossing the same shock. The competition (Fermi, 1949) between the average energy gain E_f/E_i per shock crossing cycle (upstream-downstream and back) and the escape probability per cycle, P_{esc} , leads to a power-law spectrum $N(E)dE \propto E^{-p}dE$ with

$$p = 1 + \ln[1/(1 - P_{esc})]/\ln[\langle E_f/E_i \rangle]. \quad (13)$$

Note that within the particle acceleration literature this index p is usually denoted as s . Our notation follows the common notation within the GRB literature.

Blandford and Eichler (1987) review the theory of diffuse shock acceleration in nonrelativistic shocks. However, in GRB's the shocks are relativistic—mildly relativistic in internal shocks and extremely relativistic in external shocks. Acceleration in ultrarelativistic shocks have been discussed by several groups (Heavens and Drury, 1988; Bednarz and Ostrowski, 1998; Gallant and Achterberg, 1999a; Kirk *et al.*, 2000; Achterberg *et al.*, 2001; Vietri, 2003). In relativistic shocks the considerations are quite different from those in nonrelativistic ones. Using the relativistic shock jump conditions [Eq. (11)] and kinematic considerations one can find (see Vietri, 1995; Gallant and Achterberg, 1999b; Achterberg *et al.*, 2001) that the energy gain in the first shock crossing is of the order Γ_{sh}^2 . However, subsequent shock crossings are not as efficient and the energy gain is of order unity $\langle E_f/E_i \rangle \approx 2$ (Gallant and Achterberg, 1999b; Achterberg *et al.*, 2001).

The deflection process in the upstream region is due to a large-scale smooth background magnetic field perturbed by magnetohydrodynamic (MHD) fluctuations. A tiny change of the particle's momentum in the upstream region is sufficient for the shock to overtake the particle. Within the downstream region the momentum change should have a large angle before the particle overtakes the shock and reaches the upstream region. As the shock moves with a subrelativistic velocity ($\approx c/\sqrt{3}$) relative to this frame, it is easy for a relativistic particle to overtake the shock. A finite fraction of the particles reach the upstream region. Repeated cycles of this type (in each of which the particles gain a factor of ~ 2 in energy) lead to a power-law spectrum with $p \approx 2.2$ – 2.3 (for $\Gamma_{sh} \gg 1$). As in nonrelativistic shock this result is fairly robust, and it does not depend on specific assumptions about the scattering process. It has been

obtained by several groups using different approaches, including both numerical simulations and analytic considerations. The insensitivity of this result arises naturally from the logarithmic dependence in Eq. (13) and from the fact that both the denominator and the numerator are of order unity. This result agrees nicely with that inferred from the GRB spectrum (Sari and Piran, 1997a) or the afterglow spectrum (Panaitescu and Kumar, 2001). Ostrowski and Bednarz (2002) point out, however, that this result requires highly turbulent conditions downstream of the shock. If the turbulence is weaker, the resulting energy spectrum could be much steeper. Additionally as internal shocks are only mildly relativistic the conditions in these shocks might be different.

The maximal energy of the shock-accelerated particles can be obtained by comparing the age of the shock R/c (in the upstream frame) with the duration of an acceleration cycle. For a simple magnetic deflection, this later time is just half of the Larmor time, E/Zq_eB (in the same frame). The combination yields

$$E_{max} \approx Zq_eBR = 10^{20} \text{ eV} B_3 R_{15}, \quad (14)$$

where the values that I have used in the last equality reflect the conditions within the reverse external shocks where ultrahigh-energy cosmic rays can be accelerated (see Sec. VIII.C below). For particle diffusion in a random upstream field (with a diffusion length l) one finds that R in the above equation is replaced by $\sqrt{Rl/3}$.

The acceleration process has to compete with radiation losses of the accelerated particles. Synchrotron losses are inevitable, as they occur within the same magnetic field that is essential for deflecting the particles. Comparing the energy-loss rate with the energy gain, one obtains a maximal energy of

$$E_{max} \approx mc^2 \left(\frac{4\pi q_e \Gamma_{sh}}{\sigma_T B} \right)^{1/2} \approx 5 \times 10^{17} \text{ eV} (m/m_p) \Gamma_{100}^{1/2} B^{-1/2}. \quad (15)$$

The corresponding Lorentz factor is of the order of 10^8 for $\Gamma_{sh} = 100$ and $B = 1$ G. Note that this formula assumes that the acceleration time is the Larmor time and hence that the synchrotron cooling time is equal to the Larmor time. Obviously it should be modified by a numerical factor, which is most likely of order unity.

C. Synchrotron

Synchrotron radiation most likely plays an important role in both the GRB and its afterglow. An important feature of synchrotron emission is its polarization (see Sec. V.F). Observations of polarization in GRB afterglows and in one case in the prompt emission support the idea that synchrotron emission is indeed taking place there (note, however, that inverse Compton scattering also produces polarized emission). I review here the basic features of synchrotron emission, focusing on aspects relevant to GRB's. I refer the reader to Rybicki and Lightman (1979) for a more detailed discussion.

1. Frequency and power

The typical energy of synchrotron photons as well as the synchrotron cooling time depends on the Lorentz factor γ_e of the relativistic electron under consideration and on the strength of the magnetic field. If the emitting material moves with a Lorentz factor Γ , the photons are blueshifted. The characteristic photon energy in the observer frame is given by

$$(h\nu_{syn})_{obs} = \frac{\hbar q_e B}{m_e c} \gamma_e^2 \Gamma, \quad (16)$$

where q_e is the electron's charge.

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

$$P_{syn} = \frac{4}{3} \sigma_T c U_B \gamma_e^2, \quad (17)$$

where $U_B = B^2/8\pi = \epsilon_B e$ is the magnetic energy density and σ_T is the Thompson cross section. The cooling time of the electron in the fluid frame is then $\gamma_e m_e c^2 / P$. The observed cooling time t_{syn} is shorter by a factor of Γ ,

$$t_{syn}(\gamma_e) = \frac{3m_e c}{4\sigma_T U_B \gamma_e \Gamma}. \quad (18)$$

Substituting the value of γ_e from Eq. (16) into the cooling rate [Eq. (18)] one obtains the cooling time scale as a function of the observed photon energy:

$$t_{syn}(\nu) = \frac{3}{\sigma_T} \sqrt{\frac{2\pi c m_e q_e}{B^3 \Gamma}} \nu^{-1/2}. \quad (19)$$

Since γ_e does not appear explicitly in this equation, t_{syn} at a given observed frequency is independent of the electrons' energy distribution within the shock. This is provided, of course, that there are electrons with the required γ_e so that there will be emission in the frequency considered. As long as there is such an electron the cooling time is "universal." This equation shows a characteristic scaling of $t_{syn}(\nu) \propto \nu^{-1/2}$. This is not very different from the observed relation $\delta T \propto \nu^{-0.4}$ (Fenimore *et al.*, 1995). However, it is unlikely that cooling, and not a physical process, determines the temporal profile.

The cooling time calculated above sets a lower limit to the variability time scale of a GRB because the burst cannot possibly contain spikes that are shorter than its cooling time. Observations of GRB's typically show asymmetric spikes in the intensity, where a peak generally has a fast rise and a slower decay. A plausible explanation of this observation is that the shock heating of the electrons happens rapidly (though episodically) and that the rise time of a spike is related to the heating time. The decay time is then set by the cooling, so that the widths of spikes directly measure the cooling time. However, it seems that there are problems with this simple explanation. First, when plugging in reasonable parameters one finds that the decay time as implied by this equation is too short. Second, if the cooling time is long, the shocked region would suffer adiabatic losses

and this would reduce the efficiency of the process. Thus it is unlikely that the pulse shapes can be explained by synchrotron physics alone.

2. The optically thin synchrotron spectrum

The instantaneous synchrotron spectrum of a single relativistic electron with an initial energy $\gamma_e m_e c^2$ is approximately a power law with $F_\nu \propto \nu^{1/3}$ up to $\nu_{syn}(\gamma_e)$ and an exponential decay above it. The peak power occurs at $\nu_{syn}(\gamma_e)$, where it has the approximate value

$$P_{\nu, max} \approx \frac{P(\gamma_e)}{\nu_{syn}(\gamma_e)} = \frac{m_e c^2 \sigma_T \Gamma B}{3q_e}. \quad (20)$$

Note that $P_{\nu, max}$ does not depend on γ_e , whereas the position of the peak does.

If the electron is energetic, it will cool rapidly until it reaches $\gamma_{e,c}$, the Lorentz factor of an electron that cools on a hydrodynamic time scale. For a rapidly cooling electron we have to consider the time-integrated spectrum. For an initial Lorentz factor γ_e , $F_\nu \propto \nu^{-1/2}$ for $\nu_{syn}(\gamma_{e,c}) < \nu < \nu_{syn}(\gamma_e)$.

To calculate the overall spectrum due to the electrons one needs to integrate over the electrons' Lorentz factor distribution. I consider first, following Sari *et al.* (1998), a power-law distribution, a power index p , and a minimal Lorentz factor $\gamma_{e,min}$. This is, of course, the simplest distribution and as discussed in Sec. V.B it is the expected distribution of shock-accelerated particles:

$$N(\gamma_e) \sim \gamma_e^{-p} \text{ for } \gamma_e > \gamma_{e,min}. \quad (21)$$

The condition $p > 2$ is required so that the energy does not diverge at large γ_e (Bhattacharya, 2001). Dai and Cheng (2001) also consider distributions with $2 > p > 1$ with a maximal energy cutoff (see below). The minimum Lorentz factor $\gamma_{e,min}$ of the distribution is related to the electron energy density e_e and the electron number density n_e as

$$\gamma_{e,min} = \frac{p-2}{p-1} \frac{e_e}{n_e m_e c^2} = \frac{p-2}{p-1} \langle \gamma_e \rangle. \quad (22)$$

The minimal Lorentz factor plays an important role as it characterizes the typical electron's Lorentz factor and the corresponding "typical" synchrotron frequency $\nu_m \equiv \nu_{syn}(\gamma_{e,min})$. Interestingly the upper energy cutoff (which essentially exists somewhere) does not play a critical role in the spectrum for $p > 2$. Of course it will lead to a high-frequency cutoff of the spectrum around ν_{syn} that corresponds to this energy. However, quite generally, this happens at the high-energy tail far from where the peak flux or the peak energy are emitted.

A simple modification of the above idea arises if only a fraction ξ_e of the electrons is accelerated to high energies and the rest of the electrons remain cold (Bykov and Meszaros, 1996; Guetta *et al.*, 2001a). If a small fraction of electrons shares the energy e_e then the typical Lorentz factor would be $\xi_e^{-1} \gamma_{e,min}$, where $\gamma_{e,min}$ is calculated from Eq. (22) above. All the corresponding places where $\gamma_{e,min}$ is used should be modified according to this

factor. At the same time fewer electrons will be radiating. This will introduce a factor ξ_e that should multiply the total emitted flux. In the following discussion I will not add this factor into the analysis. Similarly in situations when multiple pairs are formed (Ghisellini and Celotti, 1999) the electron energy is shared by a larger number of electrons. In this case ξ_e is larger than unity and similar modifications of the spectrum apply.

The lowest part of the spectrum (strictly speaking, the lowest part of the optically thin spectrum, as at very low frequencies self-absorption sets in; see Sec. V.C.3 below) is always the sum of the contributions of the tails of all the electrons' emissions: $F_\nu \propto \nu^{1/3}$. This is typical of synchrotron radiation (Meszaros and Rees, 1993; Katz, 1994a; Cohen *et al.*, 1997) and is independent of the exact shape of the electron distribution. Tavani (1996a, 1996b), for example, obtains such a low-energy spectrum both for a Gaussian and for a Gaussian and a high-energy power-law tail. The observation of bursts (about 1/5 of all bursts) with a steeper spectrum at the lower-energy part, i.e., below the "synchrotron line of death" (Preece *et al.*, 1998, 2002) is one of the problems that this model faces. The problem is made even more severe by the need for the GRB to be in the fast-cooling regime in order to radiate efficiently, and there the relevant low-energy spectrum will be $\propto \nu^{-1/2}$ (Cohen *et al.*, 1997; Ghisellini *et al.*, 2000). However, as stressed earlier (see Sec. II.A.1), this problem is not seen in any of the HETE spectra whose low-energy tails are always in the proper synchrotron range with a slope (Barraud *et al.*, 2003). Hence this problem might be an artifact of the low-energy resolution of BATSE in this energy range (Cohen *et al.*, 1997).

On the other hand, the most energetic electrons will always be cooling rapidly (independently of the behavior of the "typical electron"). These electrons emit practically all their energy $m_e c^2 \gamma$ at their synchrotron frequency. The number of electrons with Lorentz factors $\sim \gamma$ is $\propto \gamma^{1-p}$ and their energy $\propto \gamma^2$. As these electrons cool, they deposit most of their energy into a frequency range $\sim \nu_{\text{syn}}(\gamma) \propto \gamma^2$ and therefore $F_\nu \propto \gamma^{-p} \propto \nu^{-p/2}$. Thus the uppermost part of the spectrum will satisfy

$$F_\nu = N[\gamma(\nu)] m_e c^2 \gamma(\nu) d\gamma/d\nu \propto \nu^{-p/2}. \quad (23)$$

In the intermediate-frequency region the spectrum varies between "slow cooling," if the "typical" electrons with $\gamma_{e,\text{min}}$ do not cool on a hydrodynamic time scale, and "fast cooling" if they do. The critical parameter that determines whether the cooling is fast or slow is $\gamma_{e,c}$, the Lorentz factor of an electron that cools on a hydrodynamic time scale. To estimate $\gamma_{e,c}$ compare t_{syn} [Eq. (18)] with t_{hyd} , the hydrodynamic time scale (in the observer's rest frame):

$$\gamma_{e,c} = \frac{3m_e c}{4\sigma_T U_B \Gamma t_{\text{hyd}}}. \quad (24)$$

For fast cooling $\gamma_{e,\text{min}} < \gamma_{e,c}$, while for slow cooling $\gamma_{e,\text{min}} > \gamma_{e,c}$. In the following discussion two important frequencies play a dominant role:

$$\nu_m \equiv \nu_{\text{syn}}(\gamma_{e,\text{min}}); \quad (25)$$

$$\nu_c \equiv \nu_{\text{syn}}(\gamma_{e,c}).$$

These are the synchrotron frequencies of electrons with $\gamma_{e,\text{min}}$ and with $\gamma_{e,c}$.

a. Fast cooling ($\gamma_{e,c} < \gamma_{e,\text{min}}$)

The typical electron is cooling rapidly, hence $\nu_c < \nu_m$. The low-frequency spectrum $F_\nu \propto \nu^{1/3}$ extends up to ν_c . In the intermediate range between ν_c and ν_m , we observe the energy of all the cooling electrons. The energy of an electron $\propto \gamma$, its typical frequency $\propto \gamma^2$, and the flux per unit frequency is $\propto \gamma^{-1} \propto \nu^{-1/2}$. Overall the observed flux F_ν is given by

$$F_\nu \propto \begin{cases} (\nu/\nu_c)^{1/3} F_{\nu,\text{max}}, & \nu < \nu_c, \\ (\nu/\nu_c)^{-1/2} F_{\nu,\text{max}}, & \nu_c < \nu < \nu_m, \\ (\nu_m/\nu_c)^{-1/2} (\nu/\nu_m)^{-p/2} F_{\nu,\text{max}}, & \nu_m < \nu, \end{cases} \quad (26)$$

where $\nu_m \equiv \nu_{\text{syn}}(\gamma_{e,\text{min}})$, $\nu_c \equiv \nu_{\text{syn}}(\gamma_{e,c})$ and $F_{\nu,\text{max}}$ is the observed peak flux. The peak flux is at ν_c $F_{\nu,\text{max}} \equiv N_e P_{\nu,\text{max}} / 4\pi D^2$ (where D is the distance to the source and I ignore cosmological corrections). The power emitted is simply the power given to the electrons, that is, ϵ_e times the power generated by the shock, dE/dt :

$$P_{\text{fast}} = \epsilon_e \frac{dE}{dt}. \quad (27)$$

The peak energy emitted (which corresponds to the peak of νF_ν) is at ν_m . The resulting spectrum is shown in Fig. 16.

b. Slow cooling ($\gamma_{e,c} > \gamma_{e,\text{min}}$)

Now only the high-energy tail of the distribution (those electrons above $\gamma_{e,c}$) cools efficiently. The electrons with $\gamma_e \sim \gamma_{e,\text{min}}$, which form the bulk of the population, do not cool. Now $f_\nu \propto \nu^{1/3}$ up to ν_m , and $F_\nu \propto \nu^{-p/2}$ above ν_c . In the intermediate region between these two frequencies,

$$F_\nu = N[(\gamma(\nu))] P[(\gamma(\nu))] d\gamma/d\nu \propto \nu^{-(p-1)/2}, \quad (28)$$

where $\gamma(\nu)$ is the Lorentz factor for which the synchrotron frequency equals ν , $N[\gamma]$ is the number of electrons with a Lorentz factor γ , and $P[\gamma]$ the power emitted by an electron with γ . Overall one finds

$$F_\nu \propto \begin{cases} (\nu/\nu_m)^{1/3} F_{\nu,\text{max}}, & \nu < \nu_m, \\ (\nu/\nu_m)^{-(p-1)/2} F_{\nu,\text{max}}, & \nu_m < \nu < \nu_c, \\ (\nu_c/\nu_m)^{-(p-1)/2} (\nu/\nu_c)^{-p/2} F_{\nu,\text{max}}, & \nu_c < \nu. \end{cases} \quad (29)$$

The peak flux is at ν_m , while the peak energy emitted is at ν_c . The emitted power is determined by the ability of the electrons to radiate their energy:

$$P_{\text{slow}} = N_e P_{\text{syn}}(\gamma_{e,\text{min}}), \quad (30)$$

where N_e is the number of electrons in the emitting region and where $P_{\text{syn}}(\gamma_{e,\text{min}})$, the synchrotron power of an electron with $\gamma_{e,\text{min}}$, is given by Eq. (17).

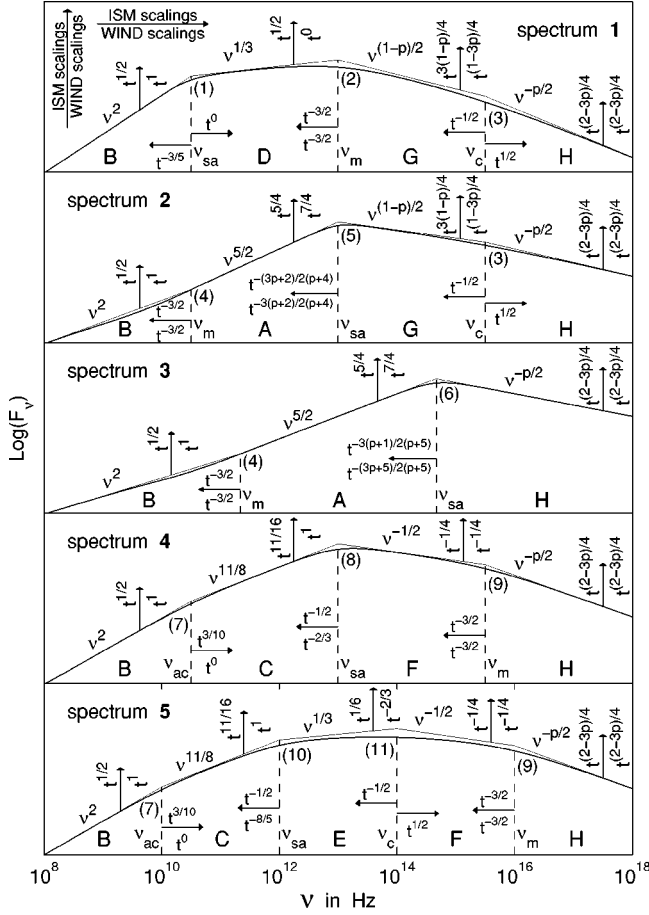


FIG. 16. Different possible broadband synchrotron spectra from a relativistic blast wave, which accelerates the electrons to a power-law distribution of energies. Thin solid line: asymptotic power-law segments and their points of intersection, where the break frequencies ν_b and the corresponding flux densities $F_{\nu_b, \text{ext}}$ are defined. The different PLS's are labeled A through H, while the different break frequencies are labeled 1–11. The temporal scalings of the power-law segments and the break frequencies, for an ISM ($k=0$) or stellar wind ($k=2$) environment, are indicated by the arrows. The different spectra are labeled 1–5, from top to bottom. The relevant spectrum is determined by the ordering of the break frequencies. The top two panels (spectra 1 and 2) correspond to slow cooling ($\nu_m < \nu_c$). Spectrum 1 applies when $\nu_{sa} < \nu_m$, while spectrum 2 applies when $\nu_m < \nu_{sa} < \nu_c$. The two bottom panels (spectra 4 and 5) correspond to fast cooling ($\nu_c < \nu_m$). Spectrum 5 applies when $\nu_{sa} < \nu_c$, and spectrum 4 applies when $\nu_c < \nu_{sa} < \nu_m$. Spectrum 3 (middle panel) applies when $\nu_{sa} > \nu_m, \nu_c$, in which case the relative ordering of ν_c and ν_m is unimportant (i.e., spectrum 3 may apply both to slow cooling or fast cooling). From Granot and Sari, 2002.

Typical spectra corresponding to fast and slow cooling are shown in Fig. 16. The light curve depends on the hydrodynamic evolution, which in turn determines the time dependence of ν_m , ν_c , and $F_{\nu, \text{max}}$. The spectra presented here are composed of broken power laws. Granot and Sari (2002) present a more accurate spectrum in which the asymptotic power-law segments are connected by smooth curves. They fit the transitions by

$[(\nu/\nu_b)^{-n\beta_1} + (\nu/\nu_b)^{-n\beta_2}]^{-1/n}$. The parameter n estimates the smoothness of the transition with $n \approx 1$ for all transitions.

Fast cooling must take place during the GRB itself: the relativistic shocks must emit their energy effectively—otherwise there will be a serious inefficiency problem. Additionally the pulse will not be variable if the cooling time is too long. The electrons must cool rapidly and release all their energy. It is most likely that during the early stages of an external shock (that is, within the afterglow phase, provided that it arises due to external shocks) there will be a transition from fast to slow cooling (Katz and Piran, 1997; Meszaros and Rees, 1997a; Waxman, 1997a, 1997b; Meszaros *et al.*, 1998).

Tavani (1996a, 1996b) discusses the synchrotron spectrum from a Gaussian electron distribution and from a Gaussian electron distribution with a high-energy tail. As mentioned earlier, the Gaussian (thermal) distribution has a typical low-frequency $\nu^{1/3}$ spectrum. However, as expected, there is a sharp exponential cutoff at high frequencies. Without a high-energy tail this spectrum does not fit the observed GRB spectra of most GRB's (see Sec. II.A.1). Note, however, that it may fit a small subgroup with no high energy (Pendleton *et al.*, 1997). With an electron distribution composed of a Gaussian and an added high-energy tail the resulting spectrum has the typical $\nu^{1/3}$ component and an additional high-energy tail that depends on the electron power-law index. Such a spectrum fits several observed GRB spectra (Tavani, 1996a, 1996b).

Another variant is the synchrotron spectrum from a power-law electron distribution with $1 < p < 2$ (Bhattacharya, 2001; Dai and Cheng, 2001). In this case there must be a high-energy cutoff $\gamma_{e, \text{max}}$ and the typical electron energy corresponds to this upper cutoff. A possible cutoff can arise from synchrotron losses at the energy where the acceleration time equals the energy loss time [see, for example, de Jager *et al.* (1996) and the discussion in Sec. V.B]:

$$\gamma_{e, \text{Max}} \approx 4 \times 10^7 B^{-1/2}. \quad (31)$$

The resulting “typical” Lorentz factor $\gamma_{e, \text{min}}$ differs now from the one given by Eq. (22). Bhattacharya (2001) and Dai and Cheng (2001) find that it is replaced by

$$\gamma_{e, \text{min}} = \left[\left(\frac{2-p}{p-1} \right) \left(\frac{m_p}{m_e} \right) \epsilon_e \Gamma \gamma_{e, \text{Max}}^{p-2} \right]^{1/(p-1)}. \quad (32)$$

The resulting spectrum is similar to that obtained for fast or slow cooling with the new critical frequencies ν_m given by using the result of Eq. (32) in Eq. (26).

3. Synchrotron self-absorption

At low frequencies synchrotron self-absorption may take place. It leads to a steep cutoff of the low-energy spectrum, either as the commonly known $\nu^{5/2}$ or as ν^2 . To estimate the self-absorption frequency one needs the optical depth along the line of sight. A simple approximation is $\alpha'_\nu R/\Gamma$, where α'_ν is the absorption coefficient (Rybicki and Lightman, 1979):

$$\alpha'_{\nu'} = \frac{(p+2)}{8\pi m_e \nu'^2} \int_{\gamma_{min}}^{\infty} d\gamma_e P'_{\nu',e}(\gamma_e) \frac{n(\gamma_e)}{\gamma_e}. \quad (33)$$

The self-absorption frequency ν_a satisfies $\alpha'_{\nu_a} R/\Gamma = 1$. It can be estimate only after we have a model for the hydrodynamics and know how R and γ vary with time (Granot *et al.*, 1999c; Wijers and Galama, 1999).

The spectrum below the self-absorption frequency depends on the electron distribution. One obtains the well-known (Rybicki and Lightman, 1979) $\nu^{5/2}$ when the synchrotron frequency of the electron emitting the self-absorbed radiation is inside the self-absorption range. One obtains ν^2 if the radiation within the self-absorption frequency range is due to the low-energy tail of electrons that are radiating effectively at higher energies. For this latter case, which is more appropriate for GRB afterglow (for slow cooling with $\nu_m < \nu_c$; Meszaros and Rees, 1993; Paczynski and Rhoads, 1993; Katz, 1994a; Katz and Piran, 1997):

$$F_\nu \propto \nu^2 [k_B T_e / (\Gamma m_p c^2)] R^2, \quad (34)$$

where R is the radius of the radiating shell and the factor $k_B T_e / (\Gamma m_p c^2)$ describes the degree of electron equipartition in the plasma shock-heated to an internal energy per particle $m_p c^2$ and moving with Lorentz factor γ .

The situation is slightly different for a shock-heated fast-cooling plasma, i.e., if $\nu_c < \nu_m$ (Granot *et al.*, 2000). In this case we expect the electron distribution to be inhomogeneous, as electrons near the shock have not yet cooled but electrons further downstream are cool. This leads to a new spectral range, $\nu_{sa} < \nu < \nu_{sa'}$, with $F_\nu \propto \nu^{11/8}$ (see Fig. 16).

Synchrotron self-absorption is probably irrelevant during the GRB itself. Note, however, that under extreme conditions the self-absorption frequency might be in the low x-ray region of the spectrum and this may explain the steep low-energy spectra seen in some bursts. These extreme conditions are needed in order to make the system optically thick to synchrotron radiation while keeping it optically thin to Thompson scattering and pair creation (Granot *et al.*, 2000). Self-absorption appears regularly during the afterglow and is observed typically in radio emission (Katz, 1994a; Katz and Piran, 1997; Waxman, 1997a; Granot *et al.*, 1999c; Wijers and Galama, 1999). The expected fast-cooling self-absorbed spectrum may arise in the early radio afterglow. So far it has not been observed.

D. Inverse Compton scattering

Inverse Compton scattering may modify our analysis in several ways. It can influence the spectrum even if the system is optically thin (as it must be) to Compton scattering (see, for example, Rybicki and Lightman, 1979). In view of the high energies involved, a photon is inverse Compton scattered only once. After a single scattering the photon's energy is so high that in the electron rest frame it is above the Klein-Nishina energy ($m_e c^2 \sim 0.5$ MeV), and the decrease of the Compton cross sec-

tion in this energy range makes a second scattering unlikely. Note that in some cases (e.g., in forward external shocks) even the first scattering may suffer from this problem. The effect of inverse Compton scattering depends on the Comptonization parameter $Y = \gamma^2 \tau_e$. For fast cooling one can show (Sari *et al.*, 1996) that Y satisfies

$$Y = \epsilon_e / U_B \quad \text{if} \quad U_e \ll U_B,$$

$$Y = \sqrt{U_e / U_B} \quad \text{if} \quad U_e \gg U_B, \quad (35)$$

where U_e and U_B are the energy densities of the electrons and of the magnetic field, respectively. Inverse Compton scattering is unimportant if $Y < 1$, and in this case it can be ignored.

If $Y > 1$, which corresponds to $U_e > U_B$ (or to $\epsilon_e > \epsilon_B$) and to $Y = \sqrt{U_e / U_B}$, then a large fraction of the low-energy synchrotron radiation will be up-scattered by inverse Compton scattering and a large fraction of the energy will be emitted via the inverse Compton processes. Those photons might be too energetic, that is, their energy may be far beyond the observed energy range. In this case inverse Compton scattering will not influence the observed spectra directly. However, it will take a significant fraction of the energy of the cooling electrons and hence it will influence the observations in two ways. First, it will shorten the cooling time (the emitting electrons will be cooled by both synchrotron and the inverse Compton process). Second, assuming that the observed γ -ray photons result from synchrotron emission, inverse Compton scattering will influence the overall energy budget and reduce the efficiency of the production of the observed radiation. I turn now to each of these cases.

An inverse Compton scattering boosts the energy of the photon by a factor γ_e^2 . Typical synchrotron photons that have been scattered once by inverse Compton scattering will be observed at the energy

$$(h\nu_{IC})_{obs} = \frac{\hbar q_e B}{m_e c} \gamma_e^4 \Gamma. \quad (36)$$

The electrons are cooled both by synchrotron radiation and by the inverse Compton process. The latter is more efficient and the cooling is enhanced by the Compton parameter Y . The cooling time scale is

$$t_{IC} = \frac{6\pi c^{3/4} \sqrt{U_B / U_e} \hbar^{1/4} m_e^{3/4} q_e^{1/4}}{B^{7/4} (h\nu)^{1/4} \Gamma^{3/4} \sigma_T}. \quad (37)$$

The conditions needed to produce the observed emission using inverse Compton scattering are probably not fulfilled in either external or internal shocks (see, however, Ghisellini and Celotti, 1999 and the discussion in Sec. V.E below). However even if the scattering does not produce the observed γ -ray photons, it still influences the process if $Y > 1$. First it will add an ultrahigh-energy component to the GRB spectrum. This component will typically be around γ_e^2 times the observed photons, ~ 100 keV that is, in the GeV–TeV range (Vietri, 1997;

Bottcher and Dermer, 1998 and the discussion in Sec. VIII.A). This component may already have been observed in some GRB's during the early afterglow (see Sec. II.A.1). Inverse Compton scattering will also speed up the cooling of the emitting regions and shorten the cooling time t_{syn} estimated earlier [Eq. (19)] by a factor of Y . At the same time this also reduces the efficiency (for producing the observed γ rays) by the same factor.

E. Quasithermal Comptonization

Ghisellini and Celotti (1999) suggested that prompt GRB emission arises in a quasithermal Comptonization process. In their model the optical depth within the emitting region (of internal shocks) is of order unity, leading to copious pair production. The system is optically thick to synchrotron emission. The self-absorbed synchrotron emission is the seed for an inverse Compton emission produced by the pairs. The effective Compton parameter in the new system, \tilde{Y} , is

$$\tilde{Y} \equiv 4\tau \left(\frac{kT'}{m_e c^2} \right) (1 + \tau) \left[1 + 4 \left(\frac{kT'}{m_e c^2} \right) \right], \quad (38)$$

where T' is the effective temperature of the pair and τ is the total cross section for scattering. The pairs act as a thermostat, controlling the effective temperature within the emitting region to 30–300 keV (Svensson, 1982, 1984). The resulting spectrum from this model is a flat spectrum, $F_\nu \propto \nu^0$, between the $h\nu_{\text{sa}}\Gamma$ and $kT'\Gamma$ (Ghisellini and Celotti, 1999). The spectrum will evolve rapidly during the burst while the pairs are being created and the effective temperature decreases.

F. Polarization from relativistically moving sources

Polarization can provide information on both the emission process and the geometry of the emitting regions. Usually the observed polarization is obtained by first integrating the Stokes parameters of the radiation emitted by the individual electrons over the electron distribution. This yields the local polarization. Then we integrate over the emitting region to obtain the global polarization. In GRB's (both in the prompt emission and in the afterglow) the emitting regions move relativistically towards the observer. The implied Lorentz transformations play a very important role in the second integration, as they change the direction of propagation of the photons and hence the direction of the local polarization. The final results are sometimes surprising and counterintuitive. For example, even if the intrinsic (local) emission is 100% polarized in the same direction, the integration over the emitting region would reduce this to 70% polarization. I consider polarization from synchrotron emission here, but the results can be easily applied to inverse Compton scattering, as well. I apply

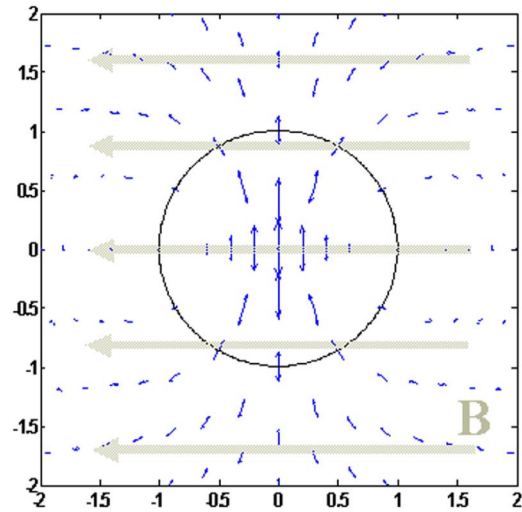


FIG. 17. (Color in online edition) Polarization from a uniform magnetic field (following Granot and Königl, 2003). The circle marks the angle where the matter moves at an angle Γ^{-1} from the observer. Note that the polarization is maximal along the line perpendicular to the uniform field, but it vanishes in the other direction near the Γ^{-1} circle.

the results derived in this section to the possible polarization from the prompt emission and from the afterglow in the corresponding sections VI.E and VII.J.

As an example, let us consider synchrotron emission. Synchrotron emission is polarized with—and the intrinsic local polarization level depends on—the spectral index of the energy distribution of the emitting electrons, p (Rybicki and Lightman, 1979). For typical values ($2 < p < 3$) it can reach 75%. The polarization vector is perpendicular to the magnetic field and, of course, to the direction of the emitted radiation. The formalism can be

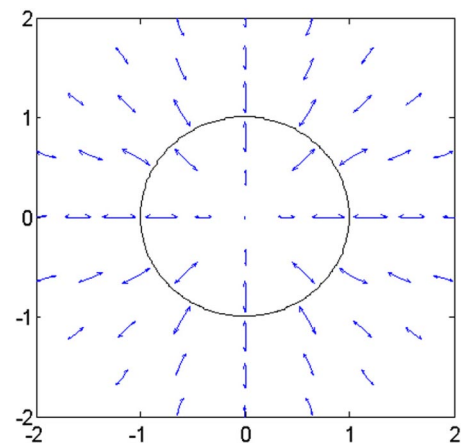


FIG. 18. (Color in online edition) Polarization from a random magnetic field in the plane of the shock. The solid circle marks the angle where the matter moves at an angle Γ^{-1} from the observer. Polarization may arise if we view a part of a jet (dashed line), or if the emission is dominated by hot spot (dash-dotted regions), or if there is an overall gradient of the emissivity as would arise in the standard jet model. From Narkar, Piran, and Waxman, 2003.

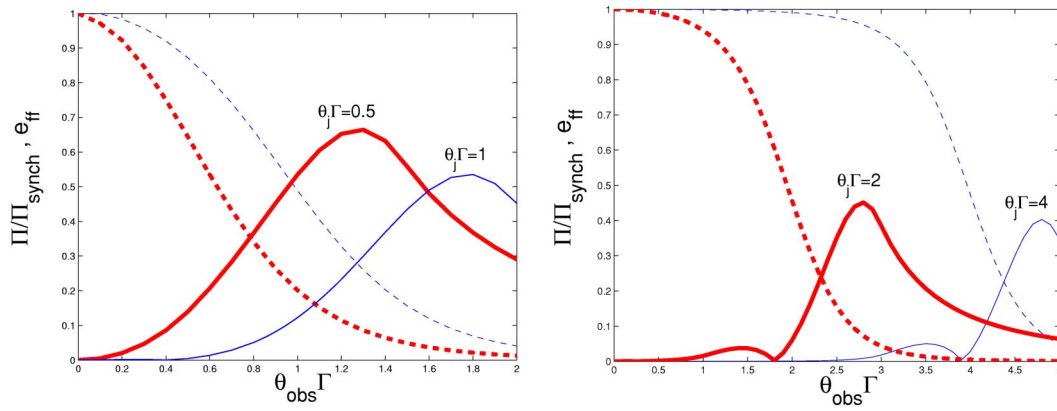


FIG. 19. (Color in online edition) Time integrated polarization (solid lines) and the efficiency (dashed lines) as a function of $\theta_{\text{obs}}\Gamma$; (a) for two different values and (b) for two other different values of θ_j for a random magnetic field.

easily adapted to inverse Compton scattering, for which the intrinsic local polarization is higher and could reach 100% when the photons are scattered at 90° .

Consider first a case in which the magnetic field is uniform locally (over a region of angular size Γ^{-1}). This could happen, for example, if we had an ordered magnetic field along the ϕ direction and the observer was more than Γ^{-1} away from the symmetry axis, as would be the case within internal shocks if the magnetic field were dragged from the source or within several Poynting-flux-dominated models. The locally emitted polarization is uniform, is in the plane of the sky, and is perpendicular to the direction of the magnetic field. In a Newtonian system it would combine so that the observed polarization would equal the emitted one. However, the Lorentz transformations induce their own signature on the observed polarization (Granot, 2003; Granot and Königl, 2003). This is depicted in Fig. 17, which clearly shows that the polarization vector varies along the observed region (whose angular size is $1/\Gamma$). Consequently the observed global polarization will be smaller than the local polarization.

The observed Stokes parameters are weighted averages of the local Stokes parameters at different regions of the shell. The instantaneous polarization is calculated using the instantaneous observed flux $F_\nu(y, T) \propto (1+y)^{-(3+\alpha)}$, with α the relevant spectral index for this segment, as the weights, where $y \equiv (\Gamma\theta)^2$ and T is the observer time. The time-integrated polarization is calculated using the fluences as weights: $\int_0^\infty F_\nu(y, T) dT \propto (1+y)^{-(2+\alpha)}$.

The fluxes depend on how the intensity varies with the magnetic field. For $I_\nu \propto B^0$, which is relevant for fast cooling⁸ (and the prompt GRB), the time-integrated

Stokes parameters (note that $V=0$ as the polarization is linear) and polarization are given by

$$\frac{\begin{Bmatrix} Q \\ U \end{Bmatrix}}{I} = \Pi_{\text{synch}} \frac{\int_0^{2\pi} \int_0^\infty (1+y)^{-(2+\alpha)} \begin{Bmatrix} \cos(2\theta_p) \\ \sin(2\theta_p) \end{Bmatrix} dy d\phi}{\int_0^{2\pi} \int_0^\infty (1+y)^{-(2+\alpha)} dy d\phi}, \quad (39)$$

and the relative polarization is given by

$$\Pi = \frac{\sqrt{U^2 + Q^2}}{I}, \quad (40)$$

where $\theta_p = \phi + \arctan[(1-y)/(1+y)\cot\phi]$ (Granot and Königl, 2003; see also Lyutikov *et al.*, 2003). For $\alpha=1$ Eqs. (39) and (40) yield a polarization level of $\Pi/\Pi_{\text{synch}} \approx 60\%$, i.e., 60% of the maximal synchrotron polarization, or an overall polarization of $\sim 45\%$. Taking the exact values of α and the dependence of I_ν on B for fast cooling and $p=2.5$ results in an overall polarization of $\sim 50\%$ (Granot and Königl, 2003; Nakar, Piran, and Waxman, 2003).

It turns out that one can get a polarized emission even from a random magnetic field (Gruzinov and Waxman, 1999; Medvedev and Loeb, 1999). This happens if the system has nonspherical geometry. Consider a two-dimensional random magnetic field that is in the plane of the shock and assume that the correlation length of this magnetic field is very short compared to all other length scales in the system. The Lorentz transformation induces in this case a radial polarization pattern going out from the center (where the velocity of the matter is towards the observer and the polarization vanishes).

⁸See, however, Granot (2003).

This polarization pattern is shown in Fig. 18. It is clear that a simple integration over this pattern will lead to a vanishing polarization.

However, a net polarization can occur in several cases if the overall symmetry is broken. Polarization will occur if (see Fig. 18)

- We observe a jet at an angle so that only a part of the jet is within an angle of Γ^{-1} .
- The emission is nonuniform and there are stronger patches with angular size smaller than Γ^{-1} from which most of the emission comes.

- We observe a standard jet whose emission is angle dependent, and this dependence is of the order of Γ^{-1} .

Ghisellini and Lazzati (1999), Gruzinov (1999), Sari (1999b), and Waxman (2003) suggested that polarization can arise from a jet even if the magnetic field is random. Nakar, Piran, and Waxman (2003) considered a random magnetic field that remains planar in the plane of the shock (for a three-dimensional random magnetic field the polarization essentially vanishes). For $I_\nu \propto B^0$ the degree of observed polarization of the emission from a small region at angle y is $\Pi(y)/\Pi_{synch} = \min(y, 1/y)$. The overall time-integrated Stokes parameters are

$$\frac{\begin{Bmatrix} Q \\ U \end{Bmatrix}}{I} = \Pi_{synch} \frac{\int_0^{2\pi} \int_0^\infty P'_{\nu',m} (1+y)^{-(2+\alpha)} \min(y, 1/y) \begin{Bmatrix} \cos(2\phi) \\ \sin(2\phi) \end{Bmatrix} dy d\phi}{\int_0^{2\pi} \int_0^\infty P'_{\nu',m} (1+y)^{-(2+\alpha)} dy d\phi}, \quad (41)$$

where $P'_{\nu',m} = P'_{\nu',m}(y, \phi)$ is the emitted power at the synchrotron frequency in the fluid rest frame. For a top-hat jet with sharp edges, $P'_{\nu',m}$ is constant for any y and ϕ within the jet and zero otherwise. For a structured jet $P'_{\nu',m}$ depends on the angle from the jet axis.

The maximal polarization is observed when one sees the edge of the jet. The probability to see the edge of a top-hat jet with sharp edges and an opening angle $\theta_j \Gamma \gg 1$ is negligible. On the other hand, a jet with $\theta_j \Gamma \ll 1$ is not expected. Thus the only physical cases in which we can expect a large polarization are $1 \lesssim \theta_j \Gamma < \text{a few}$.

Figure 19 depicts the time-integrated polarization and the efficiency from sharp-edged jets with different opening angles as a function of the angle between the jet axis and the line of sight, θ_{obs} . The efficiency e_{ff} is defined to be the ratio between the observed fluence at θ_{obs} and the maximal possible observed fluence at $\theta_{obs}=0$. In all these cases the polarization is peaked above 40%, however the efficiency decreases sharply as the polarization increases. Thus the probability to see high polarization grows when θ_j decreases. The probability that θ_{obs} is such that the polarization is larger than 30% ($\cdot \Pi_{synch}$) while $e_{ff} > 0.1$ is 0.68, 0.41, 0.2, and 0.08 for $\theta_j \Gamma = 0.5, 1, 2$, and 4, respectively. In reality this probability will be smaller, as the chance of observing a burst increases with its observed flux.

These later calculations also apply for inverse Compton emission (Dar and De Rujula, 2003; Lazzati *et al.*, 2003). However, in this case the intrinsic local polarization is around 100% and hence one can reach a maximal polarization of $\sim 70\%$.

Polarization could also arise if the magnetic field were uniform over random patches within a region of size Γ^{-1} .

Here it is difficult, of course, to estimate the total polarization without a detailed model of the structure of the jet (Gruzinov and Waxman, 1999).

VI. THE GRB AND ITS PROMPT EMISSION

We turn now to a discussion of the theory of the GRB and its prompt emission. It is generally accepted that both the GRB and the afterglow arise due to dissipation of the kinetic energy of the relativistic flow. The relativistic motion can be dissipated by either external shocks (Mészáros and Rees, 1992; Rees and Mészáros, 1992; Katz, 1994a) or internal shocks (Narayan *et al.*, 1992; Paczynski and Xu, 1994; Rees and Mészáros, 1994). The first involve slowing down by the external medium surrounding the burst. This would be the analog of a supernova remnant in which the ejecta are slowed down by the surrounding interstellar medium. As in SNR's, external shocks can dissipate all the kinetic energy of the relativistic flow. On the other hand, internal shocks are shocks within the flow itself. These take place when faster-moving matter overtakes a slower-moving shell.

Sari and Piran (1997b) have shown that external shocks cannot produce variable bursts (see also Fenimore *et al.*, 1996). By variable I mean here, following Sari and Piran (1997b), that $\delta t \ll T$, where T is the overall duration of the burst (e.g., T_{90}) and δt is the duration of a typical pulse (see Sec. II.A.2). As most GRB's are variable, Sari and Piran (1997b) concluded that most GRB's are produced by internal shocks (Rees and Mészáros, 1994). Internal shocks can dissipate only a fraction of the kinetic energy. Therefore they must be accompanied by external shocks that follow and dissipate the remaining energy. This leads to the internal-external shock sce-

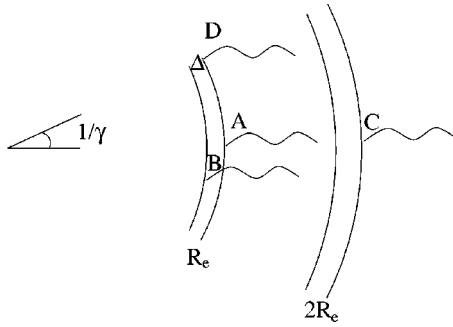


FIG. 20. Different time scales from a relativistic expanding shell in terms of the arrival times (t_i) of various photons: $t_{ang} = t_D - t_A$, $t_R = t_C - t_A$ and $t_\Delta = t_B - t_A$.

nario (Piran and Sari, 1998). GRB's are produced by internal shocks within a relativistic flow. Subsequent external shocks between the flow and the circumburst medium produce a smooth long-lasting emission—the afterglow. Various observations (see Sec. II.A.6) support this picture. I begin with a comparison of internal vs external shocks. I then review the prompt emission from internal shocks, then the prompt emission from external shocks (which includes contributions to the late part of long GRB's and the prompt optical flash). I also discuss the transition from the observations of one shock to the other.

A. Internal vs external shocks

1. General considerations

Consider a quasispherical relativistic emitting shell with a radius R , a width Δ , and a Lorentz factor Γ . This can be a whole spherical shell or a spherical-like section of a jet whose opening angle θ is larger than Γ^{-1} . Because of relativistic beaming, an observer would observe radiation only from a region of angular size $\sim \Gamma^{-1}$. Consider now photons emitted at different points along the shock (see Fig. 20). Photons emitted by matter moving directly towards the observer (point A in Fig. 20) will arrive first. Photons emitted by matter moving at an angle Γ^{-1} (point D in Fig. 20) would arrive after $t_{ang} = R/2c\Gamma^2$. This is also t_R , the time of arrival of photons emitted by matter moving directly towards the observer but emitted at $2R$ (point C in Fig. 20). Thus $t_R \approx t_{ang}$ (Fenimore *et al.*, 1996; Sari and Piran, 1997b). This coincidence is the first part of the argument that rules out external shocks in variable GRB's.

At a given point particles are continuously accelerated and emit radiation as long as the shell with a width Δ is crossing this point. The photons emitted at the front of this shell will reach the observer at a time $t_\Delta = \Delta/c$ before those emitted from the rear (point B in Fig. 20). In fact, photons are emitted slightly longer, as it takes some time for the accelerated electrons to cool. However, for most reasonable parameters the cooling time is much shorter than the other time scales (Sari *et al.*, 1996) and I ignore it hereafter.

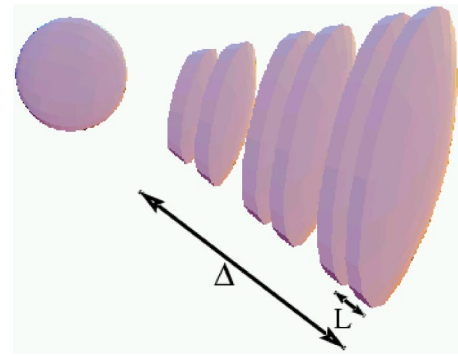


FIG. 21. (Color in online edition) The internal shocks model (from Sari, 1999a). Faster shells collide with slower ones and produce the observed γ rays. The variability time scale is L/c , while the total duration of the burst is Δ/c . From Sari, 1999a.

The emission from different angular points smooths the signal on a time scale t_{ang} . If $t_\Delta \leq t_{ang} \approx t_R$, the resulting burst will be smooth with a width $t_{ang} \approx t_R$. The second part of this argument follows from the hydrodynamics of external shocks. I show later in Sec. VI.C (see also Sari and Piran, 1997b) that for external shocks $\Delta/c \leq R/c\Gamma^2 \approx t_R \approx t_{ang}$ and for a spreading shell $\Delta \approx R/c\Gamma^2$. Therefore external shocks can produce only smooth bursts!

As we find only two time scales and as the emission is smoothed over a time scale t_{ang} , a necessary condition for the production of a variable light curve is that $t_\Delta = \Delta/c > t_{ang}$. In this case t_Δ would be the duration of the burst and t_{ang} the variability time scale. This can be easily satisfied within internal shocks (see Fig. 21). Consider an “inner engine” emitting a relativistic wind active over a time $t_\Delta = \Delta/c$ (where Δ is the overall width of the flow in the observer frame). The source is variable on a scale L/c . Internal shocks will take place at $R_s \approx L\Gamma^2$. At this place the angular time and the radial time satisfy $t_{ang} \approx t_R \approx L/c$. Internal shocks continue as long as the source is active, thus the overall observed duration $T = t_\Delta$ reflects the time that the “inner engine” is active. Note that now $t_{ang} \approx L/c < t_\Delta$ is trivially satisfied. The observed variability time scale, δt , reflects the variability of the source L/c , while the overall duration of the burst reflects the overall duration of the activity of the “inner engine.”

Numerical simulations (Kobayashi *et al.*, 1997) have shown that not only are the time scales preserved but the source's temporal behavior is reproduced on an almost one-to-one basis in the observed light curve. This can be explained (Nakar and Piran, 2002a) by a simple toy model (see Sec. VI.B.3 below).

2. Caveats and complications

Clearly the way to get around the previous argument is to have $t_{ang} < t_R$. In this case one can identify t_R with the duration of the burst and t_{ang} as the variability time scale. The observed variability would require in this case

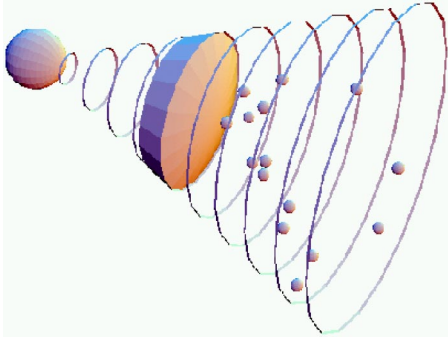


FIG. 22. (Color in online edition) The clumpy ISM model. Note the small covering factor and the resulting “geometrical” inefficiency. From Sari, 1999a.

that $t_{ang}/t_R = \delta t/T$. For this the emitting regions must be smaller than R/Γ .

One can imagine an inhomogeneous external medium which is clumpy on a scale $d \ll R/\Gamma$ (see Fig. 22). Consider a clump located at an angle $\theta \sim \Gamma^{-1}$ to the direction of motion of the matter towards the observer. The resulting angular time, which is the difference in arrival time between the first and the last photons emitted from this clump, would be $\sim d/c\Gamma$. Now $t_{ang} \sim d/c\Gamma < t_R$ and it seems that one can circumvent the argument presented before.

However, Sari and Piran (1997b) have shown that such a configuration would be extremely inefficient. The observations limit the size of the clumps to $d \approx c\Gamma\delta t$ and the location of the shock to $R \approx cTt^2$. The number of clumps within the observed angular cone with an opening angle Γ^{-1} equals the number of pulses which is approximately $T/\delta t$. The *covering factor* of the clumps can be directly estimated in terms of the observed parameters by multiplying the number of clumps ($T/\delta t$) times their area $d^2 = (c\Gamma\delta t)^2$ and dividing by the cross section of the cone $(R/\Gamma)^2$. The resulting covering factor equals $\delta t/T \ll 1$. The efficiency of conversion of kinetic energy to γ rays in this scenario is smaller than this covering factor, which for a typical variable burst could be smaller than 10^{-2} .

I turn now to several attempts to find a way around this result. I will not discuss here the feasibility of the suggested models [that is, the likelihood that the surrounding matter will be clumpy on the needed length scale (Dermer and Mitman, 1999), or that an inner engine could eject “bullets” (Heinz and Begelman, 1999) with an angular width of $\sim 10^{-2}$ degrees or what keeps these bullets so small even when they are shocked and heated]. I examine only the question of whether the observed temporal structure can arise within these models.

3. External shocks on a clumpy medium

Dermer and Mitman (1999) claim that the simple efficiency argument of Sari and Piran (1997b) is flawed. They point out that if the direction of motion of a specific blob is almost exactly towards the observer, the corresponding angular time will be of order d^2/cR and not

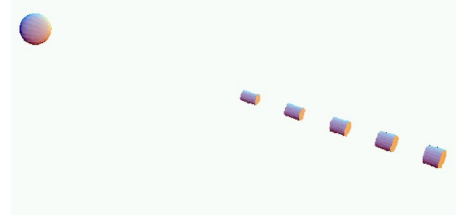


FIG. 23. (Color in online edition) The shotgun model. The inner engine emits narrow “bullets” that collide with the ISM. From Sari, 1999a.

the $d/c\Gamma$ used for a “generic” blob. This is narrower by a factor $d\Gamma/R$ than the angular time across the same blob located at a typical angle of Γ^{-1} . These special blobs would produce strong narrow peaks and would form a small region along a narrow cone with a larger covering factor. Dermer and Mitman (1999) present a numerical simulation of light curves produced by external shocks on a clumpy inhomogeneous medium with $\delta t/T \sim 10^{-2}$ and an efficiency of up to $\sim 10\%$.

A detailed analysis of the light curve, however, poses several problems for this model. While this result is marginal for bursts with $\delta t/T \sim 10^{-2}$, with a modulation of 50% it is insufficient for bursts with $\delta t/T \sim 10^{-3}$ or if the modulation is $\sim 100\%$. Variability on a time scale of milliseconds has been observed (Nakar and Piran, 2002b) in many long GRB’s (namely, $\delta t/T$ can be as small as 10^{-4}). Moreover, in this case one would expect that earlier pulses (that arise from blobs along the direction of motion) would be narrower than later pulses. This is not seen in the observed bursts (Ramirez-Ruiz and Fenimore, 2000).

Finally the arrival time of individual pulses depends on the position of the emitting clumps relative to the observers. Two following pulses would arise from two different clumps that are rather distant from each other. There is no reason why the pulses and intervals should be correlated in any way. Recall (Sec. II.A.2) that the duration of a pulse and the subsequent interval are correlated (Nakar and Piran, 2002c).

4. The shotgun model

Heinz and Begelman (1999) suggested that the “inner engine” operates like a shotgun emitting multiple narrow bullets with an angular size much smaller than Γ^{-1} (see Fig. 23). These bullets do not spread while propagating and they are slowed down rapidly by an external shock with very dense circumburst matter. The pulse width is given by t_{ang} or by the slowing-down time. The duration of the burst is determined by the time that the “inner engine” emits the bullets.

This model can produce the observed variability and, as in the internal shocks model, the observed light curve also represents here the temporal activity of the source. However, in this model the width of the pulses is determined by the angular time or the hydrodynamic time or the cooling time of the shocked material. On the other hand, the intervals between the pulses depend only on

the activity of the inner engine. Again, there is no reason why the two distributions would be similar and why there should be a correlation between them (see Sec. II.A.2 and Nakar and Piran, 2002c).

5. Relativistic turbulence

An interesting alternative to shocks as a way to dissipate kinetic energy is plasma turbulence (Smolsky and Usov, 1996, 2000; Lyutikov and Blandford, 2002, 2003). It has been suggested that in this case the kinetic energy of the shock is dissipated downstream to a combination of macroscopic (relativistic) random motion of plasma and a Lorentz factor Γ_b . Within these blobs the particles also have a (relativistic) random velocity with a Lorentz factor Γ_p , such that $\Gamma_s \approx \Gamma_b \Gamma_p$.

Relativistic turbulence may be the only way to produce variability in a situation when the matter is slowed down by the external medium and not by internal interaction. I stress that in this case the process is not described by regular shocks and hence some of the previous arguments do not hold. Two crucial open questions are (i) whether one can produce the observed correlations between pulses and intervals, and (ii) why there is no spreading of pulses at later times, as would be expected if the emitting region is slowing down and increasing its radius.

B. Internal shocks

1. Hydrodynamics of internal shocks

Internal shocks take place when a faster shell catches up with a slower one, namely, at

$$R_{int} \approx c \delta t \Gamma^2 = 3 \times 10^{14} \text{ cm} \Gamma_{100}^2 \tilde{\delta t}, \quad (42)$$

where Γ_{100} is the typical Lorentz factor in units of 10^2 and $\tilde{\delta t}$ is the time difference between the emission of the two shells. I show later that $\tilde{\delta t}$ as defined here is roughly equal to the observed fluctuations in the light curve of the burst δt . Clearly $R_{int} < R_{ext}$ must hold, otherwise internal shocks will not take place. R_{ext} is defined as the location of efficient extraction of energy by external shocks (see Sec. VI.C). It follows from the discussion in Sec. VI.C that the condition $R_{int} < R_{ext}$ implies

$$\delta \Gamma^2 < \max \left(\frac{l}{\Gamma^{2/3}}, l^{3/4} \Delta^{1/4} \right), \quad (43)$$

where l is defined by Eq. (58) and it is typically of the order of 10^{18} cm, while Δ is the width of the shell and it is of order 10^{12} cm. Both conditions set upper limits on Γ (of the order of a few thousands) for internal shocks. If the initial Lorentz factor is too large, then internal shocks will take place at large radii and external shocks will take place before the internal shocks could take place. It is possible that this fact plays an important role in limiting the relevant Lorentz factors and hence the range of variability of E_p , the peak energy observed in GRB's.

Internal shocks are characterized by a comparable Lorentz factor of order of $1 < \Gamma < 10$, reflecting the relative motion of the shells, and by comparable densities n in both shells. In this case, for an adiabatic index (4/3), the Lorentz factor of the shocked region $\hat{\Gamma}$ satisfies

$$\hat{\Gamma} = \sqrt{(\Gamma^2 + 1)/2}. \quad (44)$$

The shocked density \hat{n} and energy \hat{e} are

$$\hat{n} = (4\hat{\Gamma} + 3)n \approx 4\hat{\Gamma}n; \quad \hat{e} = \hat{\Gamma}\hat{n}m_p c^2. \quad (45)$$

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

2. The efficiency of internal shocks

Consider a collision between two shells with masses m_r and m_s that are moving at different relativistic velocities: $\Gamma_r \gg \Gamma_s \gg 1$. The resulting bulk Lorentz factor Γ_m in an elastic collision is

$$\Gamma_m \approx \sqrt{\frac{m_r \Gamma_r + m_s \Gamma_s}{m_r / \Gamma_r + m_s / \Gamma_s}}. \quad (46)$$

The internal energy \mathcal{E}_{int} in the local frame and E_{int} in the frame of an external observer of the merged shell, $E_{int} = \Gamma_m \mathcal{E}_{int}$, is the difference of the kinetic energies before and after the collision:

$$E_{int} = m_r c^2 (\Gamma_r - \Gamma_m) + m_s c^2 (\Gamma_s - \Gamma_m). \quad (47)$$

The conversion efficiency of kinetic energy into internal energy is (Kobayashi *et al.*, 1997)

$$\epsilon = 1 - \frac{(m_r + m_s)\Gamma_m}{(m_r \Gamma_r + m_s \Gamma_s)}. \quad (48)$$

As can be expected a conversion of a significant fraction of the initial kinetic energy to internal energy requires that the difference in velocities between the shells be significant, $\Gamma_r \gg \Gamma_s$, and that the two masses be comparable, $m_r \approx m_s$ (Kobayashi *et al.*, 1997; Daigne and Mochkovitch, 1998).

Beloborodov (2000) considered internal shocks between shells with a log-normal distribution of $(\Gamma - 1)/(\Gamma_0 - 1)$, where Γ_0 is the average Lorentz factor. The dimensionless parameter A measures the width of the distribution. He showed that the efficiency increased and reached unity when A was of order unity, that is, the distribution of Γ spans a range of an order of magnitude around Γ_0 . Similarly numerical simulations of Guetta *et al.* (2001a) showed that a significant fraction of the wind kinetic energy, on the order of 20%, could be converted to radiation, provided the distribution of Lorentz factors within the wind had a large variance and the minimum Lorentz factor was greater than $\approx 10^{2.5} L_{52}^{2/9}$, where L_{52} is the (isotropic) wind luminosity in units of 10^{52} ergs/sec.

Another problem that involves the efficiency of GRB's is that not all the internal energy generated is emitted. This depends further on ϵ_e , the fraction of energy given to the electron. If this fraction is small and there is no strong coupling between the electrons and

the protons, the thermal energy of the shocked particles (which is stored in this case mostly in the protons) will not be radiated away. Instead it will be converted back to kinetic energy by adiabatic cooling. Kobayashi and Sari (2001) considered a more elaborate model in which colliding shells that do not emit all their internal energy are reflected from each other, causing subsequent collisions and thereby allowing more energy to be emitted. In this case more energy is eventually emitted than would have been emitted if we considered only the first collision. They obtained about 60% overall efficiency even if the fraction of energy that goes to electrons is small, $\epsilon_e = 0.1$. This is provided that the shells' Lorentz factor varies between 10 and 10^4 .

3. Light curves from internal shocks

Both the similarity between the pulse width and the pulse separation distribution and the correlation between intervals and subsequent pulses (Nakar and Piran, 2002c; Quilligan *et al.*, 2002) arise naturally within the internal-shock model (Nakar and Piran, 2002a). In this model both the pulse duration and the separation between the pulses are determined by the same parameter—the interval between the emitted shells. I outline here the main argument (see Nakar and Piran, 2002a, for details). Consider two shells with a separation L . The Lorentz factor of the slower outer shell is $\Gamma_s = \Gamma$ and of the inner faster shell is $\Gamma_f = a\Gamma$ ($a > 2$ for an efficient collision). Both are measured in the observer frame. The shells are ejected at t_1 and $t_2 \approx t_1 + L/c$. The collision takes place at a radius $R_s \approx 2\Gamma^2 L$ (note that R_s does not depend on Γ_2). Omitting the photon flight time and assuming transparent shells, the emitted photons from the collision will reach the observer at time

$$t_o \approx t_1 + R_s/(2c\Gamma^2) \approx t_1 + L/c \approx t_2. \quad (49)$$

The photons from this pulse are observed almost simultaneously with a (hypothetical) photon that was emitted from the “inner engine” together with the second shell (at t_2). This explains why various numerical simulations (Kobayashi *et al.*, 1997; Daigne and Mochkovitch, 1998; Panaitescu *et al.*, 1999) find that for internal shocks the observed light curve replicates the temporal activity of the source.

In order to determine the time between the bursts we should consider multiple collisions. It turns out that there are just three types of collisions that characterize the system and all combinations of multiple collisions can be divided to these three types. Consider four shells emitted at times t_i ($i=1,2,3,4$) with a separation of the order of L between them. In type (i) there are two collisions—between the first and second shells and between the third and the fourth shells. The first collision will be observed at t_2 while the second one will be observed at t_4 . Therefore $\Delta t \approx t_4 - t_2 \approx 2L/c$. A different collision scenario (ii) occurs if the second and first shells collide, and afterward the third shell takes over and collides with them (the fourth shell does not play any role in this case). The first collision will be observed at t_2

while the second one will be observed at t_3 . Therefore $\Delta t \approx t_3 - t_2 \approx L/c$. Numerical simulations (Nakar and Piran, 2002a) show that more than 80% of the efficient collisions follows one of these two scenarios [(i) or (ii)]. Therefore one can estimate

$$\Delta t \approx L/c. \quad (50)$$

Note that this result is independent of the shells' masses.

A third type of multiple collision (iii) arises if the third shell collides first with the second shell and the merged shell then collides with the first one (again the fourth shell does not participate in this scenario). In this case the two pulses merge and will arrive almost simultaneously, at the same time with a (hypothetical) photon that would have been emitted from the inner engine simultaneously with the third (fastest) shell. $t \sim t_3$. Only a 20% fraction exhibits this type of collision.

The pulse width is determined by the angular time (ignoring the cooling time) $\delta t = R_s/(2c\Gamma_s^2)$, where Γ_s is the Lorentz factor of the shocked emitting region. If the shells have an equal mass ($m_1 = m_2$) then $\Gamma_s = \sqrt{a}\Gamma$, while if they have equal energy ($m_1 = am_2$) then $\Gamma_s \approx \Gamma$. Therefore

$$\delta t \approx \begin{cases} R_s/2a\Gamma^2 c \approx L/ac & \text{equal mass,} \\ R_s/2\Gamma^2 c \approx L/c & \text{equal energy.} \end{cases} \quad (51)$$

The ratio of the Lorentz factors a determines the collision's efficiency. For an efficient collision the variations in the shells' Lorentz factor (and therefore a) must be large.

It follows from Eqs. (50) and (51) that for equal-energy shells the Δt - δt similarity and correlation arises naturally from the reflection of the shells' initial separation in both variables. However, for equal-mass shells δt is shorter by a factor of a than Δt . This shortens the pulses relative to the intervals. Additionally, the large variance of a would wipe out the Δt - δt correlation. This suggests that equal-energy shells are more likely to produce the observed light curves.

C. External shocks

1. Hydrodynamics

Consider the situation when a cold relativistic shell (whose internal energy is negligible compared to the rest mass) moves into the cold interstellar medium (ISM). Generally, two shocks form: an outgoing shock that propagates into the ISM or into the external shell, and a reverse shock that propagates into the inner shell, with a contact discontinuity between the shocked material (see Fig. 24).

The dual shock system is divided into four distinct regions (see Fig. 24): the ambient matter at rest (denoted by the subscript 1), the shocked ambient matter which has passed through the forward shock (subscript 2 or f), the shocked shell material which has passed through the reverse shock (subscript 3 or r), and the unshocked material of the shell (subscript 4). The nature of the emitted

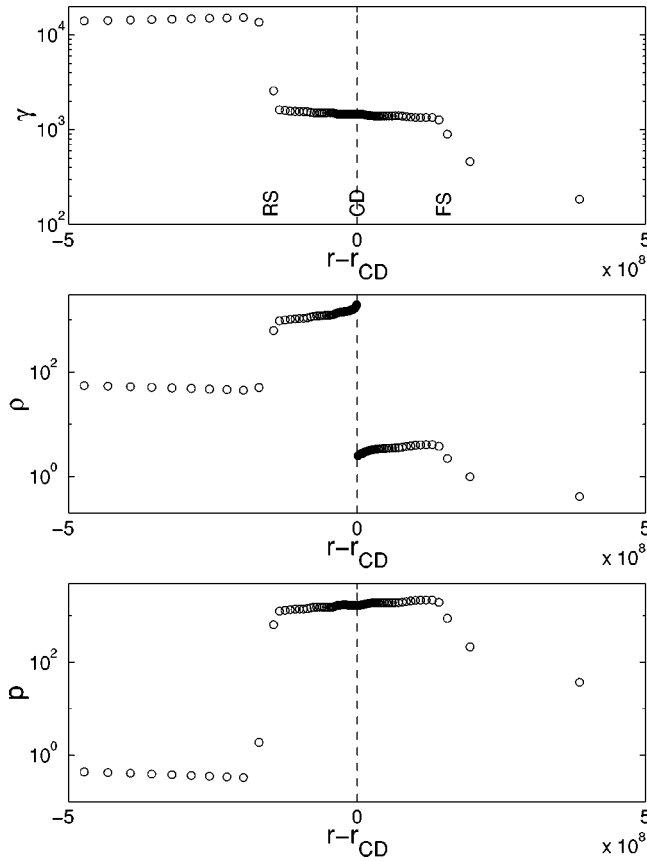


FIG. 24. The Lorentz factor Γ , the density ρ , and the pressure p in the shocks. There are four regions: the ISM (region 1), the shocked ISM (region 2), the shocked shell (region 3), and the unshocked shell (region 4), which are separated by the forward shock (FS), the contact discontinuity (CD), and the reverse shock (RS). From Kobayashi *et al.*, 1999.

radiation and the efficiency of the cooling processes depend on the conditions in the shocked-matter regions 2 and 3. Both regions have the same energy density e . 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 electrons, are different.

Two quantities determine the shocks’ structure: Γ , the Lorentz factor of the motion of the inner expanding matter (denoted 4) relative to the outer matter (the ISM or the outer shell in the case of internal collisions—denoted 1), and the ratio between the particle number densities in these regions, n_4/n_1 . Initially the density contrast between the spherically expanding shell and the ISM is large, specifically $n_4/n_1 > \Gamma^2$. This happens during the early phase of an external shock when the shell is small and dense. This configuration is denoted “Newtonian” because the reverse shock is nonrelativistic at most (or mildly relativistic). In this case all the energy conversion takes place in the forward shock. Only a negligible fraction of the energy is converted to thermal energy in the reverse shock if it is Newtonian (Sari and Piran, 1995). Let Γ_2 be the Lorentz factor of the motion of the shocked fluid relative to the rest frame of the fluid at

region 1, and let $\bar{\Gamma}_3$ be the Lorentz factor of the motion of this fluid relative to the rest frame of the relativistic shell (region 4). Then

$$\Gamma_2 \approx \Gamma; \quad \bar{\Gamma}_3 \approx 1. \quad (52)$$

The particle and energy densities (n, e) in the shocked regions satisfy

$$n_2 \approx 4\Gamma n_1; \quad e \equiv e_2 = 4\Gamma^2 n_1 m_p c^2; \quad n_3 = 7n_4; \quad e_3 = e. \quad (53)$$

Later, the shell expands and the density ratio decreases (as R^{-2} if the width of the shell is constant and as R^{-3} if the shell is spreading) and $n_4/n_1 < \Gamma^2$ (but $n_4/n_1 > 1$). In this case both the forward and the reverse shocks are relativistic. The shock equations between regions 1 and 2 combined with the contact discontinuity between 3 and 2 yield (Blandford and McKee, 1976, 1977; Piran, 1994)

$$\Gamma_2 = (n_4/n_1)^{1/4} \Gamma^{1/2} / \sqrt{2}; \quad n_2 = 4\Gamma_2 n_1; \quad e \equiv e_2 = 4\Gamma_2^2 n_1 m_p c^2. \quad (54)$$

Similar relations hold for the reverse shock:

$$\bar{\Gamma}_3 = (n_4/n_1)^{-1/4} \Gamma^{1/2} / \sqrt{2}; \quad n_3 = 4\bar{\Gamma}_3 n_4. \quad (55)$$

Additionally,

$$e_3 = e; \quad \bar{\Gamma}_3 \equiv (\Gamma/\Gamma_2 + \Gamma_2/\Gamma)/2, \quad (56)$$

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 they are both relativistic.

The interaction between a relativistic flow and an external medium depends on the Sedov length, which is defined generally as

$$E = m_p c^2 \int_0^l 4\pi n(r) r^2 dr. \quad (57)$$

The rest mass energy within the Sedov sphere equals the energy of the explosion. For a homogeneous ISM

$$l \equiv \left(\frac{E}{(4\pi/3) n_{ism} m_p c^2} \right)^{1/3} \approx 10^{18} \text{ cm} E_{52}^{1/3} n_1^{1/3}. \quad (58)$$

Note that in this section E stands for the isotropic equivalent energy. Because of the very large Lorentz factor, angular structure on a scale larger than Γ^{-1} does not influence the evolution of the system, and it behaves as if it is a part of a spherical system. A second length scale that appears in the problem is Δ , the width of the relativistic shell in the observer’s rest frame.

Initially the reverse shock is Newtonian and only a negligible amount of energy is extracted from the shell. At this stage the whole shell acts “together.” Half of the shell’s kinetic energy is converted to thermal energy when the collected external mass is M/Γ , where M is the shell’s mass (Rees and Mészáros, 1992; Katz, 1994a). This takes place at a distance

$$R_\Gamma = \frac{l}{\Gamma^{2/3}} = \left(\frac{E}{n_{ism} m_p c^2 \Gamma^2} \right)^{1/3} = 5.4 \times 10^{16} \text{ cm} E_{52}^{1/3} n_1^{-1/3} \Gamma_{100}^{-2/3}, \quad (59)$$

where E_{52} is the equivalent isotropic energy in 10^{52} ergs, $n_1 = n_{ism}/1$ particle/cm³.

However, the reverse shock might become relativistic before R_Γ . Now energy extraction from the shell is efficient and one passage of the reverse shock through the shell is sufficient for complete conversion of the shell's energy to thermal energy. The energy of the shell will be extracted during a single passage of the reverse shock across the shell. Using the expression for the velocity of the reverse shock into the shell [Eq. (55)] one finds that the reverse shock reaches the inner edge of the shell at R_Δ (Sari and Piran, 1995):

$$R_\Delta = l^{3/4} \Delta^{1/4} \approx 10^{15} \text{ cm} l_{18}^{3/4} \Delta_{12}^{1/4}. \quad (60)$$

The reverse shock becomes relativistic at R_N , where $n_4/n_1 = \Gamma^2$:

$$R_N = l^{3/2} \Delta^{1/2} \Gamma^2. \quad (61)$$

Clearly, if $R_N > R_\Gamma$ then the energy of the shell is dissipated while the shocks are still Newtonian. If $R_N < R_\Gamma$, the reverse shock becomes relativistic. In this case R_Γ loses its meaning as the radius where the energy is dissipated. The energy of the shell is dissipated in this “relativistic” case at r_Δ . The question which of the two conditions is relevant depends on the parameter ξ (Sari and Piran, 1995),

$$\xi \equiv (l/\Delta)^{1/2} \Gamma^{-4/3} = 2(l_{18}/\Delta_{12})^{1/2} \Gamma_{100}^{-4/3}. \quad (62)$$

I have used a canonical value for Δ as 10^{12} cm. It will be shown later that within the internal-external scenario Δ/c corresponds to the duration of the bursts and

10^{12} cm corresponds to a typical burst of 30 sec.

Using ξ one can express the different radii as

$$R_{int}/\xi = R_\Delta/\xi^{3/2} = R_\gamma \xi^2 = R_N/\xi^3. \quad (63)$$

For completeness I have added to this equation R_{int} , where internal shocks take place [see Eq. (42)]. The dimensionless quantity $\xi: \xi \equiv \delta/\Delta$. Thus

$$\begin{cases} R_\Delta < \mathbf{R}_\Gamma < R_N & \xi > 1 \text{ (Newtonian reverse shock)} \\ \mathbf{R}_N < R_\Gamma < R_\Delta & \xi < 1 \text{ (relativistic reverse shock)}. \end{cases} \quad (64)$$

I have marked in boldface the location where the effective energy extraction does take place. With typical values for l , Δ , and Γ ξ is around unity. The radius where energy extraction takes place is marked in boldface.

A physical shell is expected to expand during as it propagates with $\Delta = \Delta_0 + R\Gamma^2$ (Piran *et al.*, 1993). This will lead to a monotonically decreasing ξ . As the value of R_Γ is independent of Δ it does not vary. However, R_Δ and R_N decrease from their initial values. If $\Delta_0 < R_\Gamma \Gamma^2$ (corresponding to $\xi_0 > 1$), then $\xi = 1$ at $R_\Delta = R_\Gamma = R_N$ and all three radii coincide. Given the fact that with typical parameters ξ is of order unity, this seems to be the “typical” case. The reverse shocks become mildly relativistic just when the energy extraction becomes efficient. However, if $\xi_0 \ll 1$ then the shell will not expand enough and there will still be a relativistic reverse shock operating at R_Δ . It is useful to note that in this case the effective energy extraction takes place at R_Δ for all initial values of ξ_0 . In the following I denote by $\tilde{\xi}$ the value of ξ at R_Δ : $\tilde{\xi} \approx \xi_0$ if $\xi_0 < 1$ and otherwise $\tilde{\xi} \approx 1$.

Overall the external shocks take place at

$$R_{ext} = \begin{cases} \max(l/\Gamma^{2/3}, l^{3/4} \Delta^{1/4}) & \text{nonspreading shell,} \\ l/\Gamma^{2/3} \approx l^{3/4} \Delta^{1/4} \approx 5 \times 10^{16} \text{ cm} E_{52}^{1/3} n_1^{1/3} \Gamma_{100}^{-2/3} & \text{spreading shell.} \end{cases} \quad (65)$$

Usually I shall use the second relation (the spreading shell) in the following discussion. Note that in the case of a nonspreading shell one uses the maximum of the two possible radii. For example, in the Newtonian case where the extraction is at $l/\Gamma^{2/3}$ the shocks pass the shell many times and hence $l/\Gamma^{2/3} > l^{3/4} \Delta^{1/4}$.

2. Synchrotron spectrum from external shocks

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 (this would be the case in either a thick shell with $\xi < 1$ or an expanding shell that begins with $\xi_0 > 1$ but reaches ξ

≈ 1 around the time when $R_\Gamma = R_\Delta$ and efficient dissipation takes place). At this radius, the conditions at the forward shock are

$$\Gamma_2 = \Gamma \xi^{3/4}, \quad n_2 = 4\Gamma_2 n_1, \quad e_2 = 4\Gamma_2^2 n_1 m_p c^2, \quad (66)$$

while at the reverse shock, the conditions are

$$\bar{\Gamma}_3 = \xi^{-3/4}, \quad \Gamma_3 = \Gamma \xi^{3/4}, \quad n_3 = 4\xi^{9/4} \Gamma^2 n_1, \quad e_3 = e_2. \quad (67)$$

Substitution of $\Gamma_{sh} = \Gamma_2 = \Gamma \xi^{3/4}$ in Eq. (13) yields, for the equipartition magnetic field,

$$B = \sqrt{32\pi c} \epsilon_B^{1/2} \Gamma \xi^{3/4} m_p^{1/2} n_1^{1/2} = (40 \text{ G}) \epsilon_B^{1/2} \xi^{3/4} \Gamma_{100} n_1^{1/2}. \quad (68)$$

If the magnetic field in region 2 behind the forward shock were obtained purely by shock compression of the ISM field, the field would be very weak, with $\epsilon_B \ll 1$. Such low fields are incompatible with observations of GRB's. I consider therefore the possibility that there may be some kind of turbulent instability that brings the magnetic field to approximate equipartition (Silva *et al.*, 2003; Frederiksen *et al.*, 2004). In the case of the reverse shock, i.e., in region 3, magnetic fields of considerable strength might be present in the preshock shell material if the original exploding fireball were magnetic. The exact nature of magnetic-field evolution during fireball expansion depends on several assumptions. Thompson (1994) found that the magnetic field remained in equipartition if it was originally in equipartition. Mészáros, Laguna, and Rees (1993), on the other hand, estimated that, if the magnetic field were initially in equipartition, then it would be below equipartition by a factor of 10^{-5} by the time the shell expanded to R_Δ . It is uncertain which, if either, of these estimates is right. As in a forward shock, an instability could boost the field back to equipartition. Thus, while both shocks may have $\epsilon_B \ll 1$ with pure flux freezing, both could achieve $\epsilon_B \rightarrow 1$ in the presence of instabilities. In principle, ϵ_B could be different for the two shocks. For simplicity I will consider it to be the same value in the following discussions.

Following the discussion in Sec. V.B, I assume that in both regions 2 and 3 the electrons have a power-law distribution with a minimal Lorentz factor $\gamma_{e,min}$ given by Eq. (22) with the corresponding Lorentz factors for the forward and reverse shocks.

a. Forward shock

The typical energy of synchrotron photons as well as the synchrotron cooling time depend on the Lorentz factor γ_e of the relativistic electrons under consideration and on the strength of the magnetic field. Using Eq. (22) for $\gamma_{e,min}$ and Eq. (16) for the characteristic synchrotron energy for the forward shock, we obtain

$$\begin{aligned} (h\nu_{syn})_{obs|\gamma_{e,min}} &= 160 \text{ keV} \epsilon_B^{1/2} \epsilon_e^2 \Gamma_{2,100}^4 n_1^{1/2} \\ &= 0.5 \text{ keV} (\epsilon_B/0.1)^{1/2} (\epsilon_e/0.1)^2 \tilde{\xi}_0^3 \Gamma_{100}^4 n_1^{1/2} \end{aligned} \quad (69)$$

and

$$\begin{aligned} t_{syn|\gamma_{e,min}} &= 0.085 \text{ sec} \epsilon_B^{-1} \epsilon_e^{-1} \Gamma_{2,100}^{-4} n_1^{-1} \\ &= 0.085 \text{ sec} \epsilon_B^{-1} \epsilon_e^{-1} \tilde{\xi}^{-3} \Gamma_{100}^{-4} n_1^{-1}. \end{aligned} \quad (70)$$

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

The electrons at the forward shock are fast cooling, and the typical cooling frequency is (Sari and Piran, 1999c)

$$\nu_c = 6 \text{ keV} (\epsilon_B/0.1)^{-3/2} (\Gamma_2/100)^{-4} n_1^{-3/2} t_s^{-2}, \quad (71)$$

where t_s is the time in seconds. The photons from the early forward shock are in the low γ -ray to x-ray range, but this depends strongly on the various parameters [note the strong Γ_2^4 dependence in Eq. (69)]. For this set of canonical parameters $\nu_m < \nu_c$. However, the ratio of these two frequencies depends on Γ ! For Γ slightly larger than 100 the inequality will reverse and the system will be in the fast-cooling regime.

b. Reverse shock

The Lorentz factor of the reverse shock, $\bar{\Gamma}_3$, is smaller by a factor of $\xi^{3/2} \Gamma$ than the Lorentz factor of the forward shock, Γ_2 . Similarly the Lorentz factor of a “typical electron” in the reverse shock is lower by a factor $\xi^{3/2} \Gamma$. Therefore the observed energy is lower by a factor $\xi^3 \Gamma^2$. The typical synchrotron frequency of the reverse shock is

$$\nu_{m|reverse \text{ shock}} = 1.3 \times 10^{13} \text{ Hz} (\epsilon_B/0.1)^{1/2} (\epsilon_e/0.1)^2 \Gamma_{100}^2. \quad (72)$$

This is in the IR regions but note again the strong dependence on the Lorentz factor and on ϵ_e , which could easily bring this frequency up to the optical regime. The cooling frequency in the reverse-shock region is the same as the cooling frequency of the forward shock, if both regions have the same ϵ_B (Sari and Piran, 1999c) hence

$$\begin{aligned} \nu_{c|reverse \text{ shock}} &= 8 \times 10^{18} \text{ Hz} (\epsilon_B/0.1)^{-3/2} (\Gamma_2/100)^{-4} n_1^{-3/2} t_s^{-2} \\ &= 8.8 \times 10^{15} \text{ Hz} (\epsilon_B/0.1)^{-3/2} E_{52}^{-1/2} n_1^{-1} t_s^{-1/2}. \end{aligned} \quad (73)$$

In the forward shock ν_m is comparable to or larger than ν_c . In the reverse shock $\nu_m < \nu_c$ and it is usually in the slow-cooling regime. The reverse shock exists for a short time until it reaches the back of the relativistic shell. Then it turns into a rarefaction wave that propagates forwards. After some back-and-forth bouncing of these waves all the matter behind the forward shock organizes itself into the form of the Blandford-McKee self-similar solution discussed latter in Sec. VII.A. The above estimates suggest (Mészáros and Rees, 1997a; Sari and Piran, 1999a, 1999b, 1999c) that during the short phase in which the reverse shock exists it should produce a powerful optical flash. This flash should coincide with the late part of the GRB. Kobayashi (2000) calculated the light curves and typical frequencies of the reverse shock for a variety of conditions.

D. The transition from internal shocks to external shocks

The internal shocks take place at a distance $R_{int} \sim c \delta t^2 \sim (\delta t/0.3 \text{ sec}) \Gamma_2^2 10^{14} \text{ cm}$. These shocks last as long as the inner engine is active. The typical observed time scale for this activity is $\sim 50 \text{ sec}$ for long bursts and $\sim 0.5 \text{ sec}$ for short ones. External shocks begin at $R_{ext} \sim 10^{16} \text{ cm}$. If $R_{ext}/\Gamma^2 \leq T = \Delta/c$, that is, if the burst is long,

the afterglow begins while internal shocks are still going on, and the initial part of the afterglow overlaps the late part of the GRB (Sari, 1997). At an early time the afterglow emission (from the forward shock) peaks in the high x-ray region also contributing to the observed γ -ray flux. One can expect, therefore, a transition within the GRB from a hard (pure GRB) to a softer and smoother (GRB and afterglow) signal. Some observational evidence for this transition was presented in Sec. II.A.6

E. Prompt polarization

In Sec. II.A.4 I discussed the detection of very high linear polarization from GRB 021206 (Coburn and Boggs, 2003). While the data analysis is uncertain, several papers have claimed that this detection has strong implications. First Coburn and Boggs (2003) suggested that this polarization indicates that the emission mechanism is synchrotron radiation. Lyutikov *et al.* (2003) and Granot (2003) suggested further that it implies uniform magnetic fields within the emitting regions, and Lyutikov *et al.* (2003) even conclude that this implies that the relativistic flow is Poynting flux dominated and that the dissipation is in the form of external plasma instability. Waxman (2003) and Nakar, Piran, and Waxman (2003) showed, however, that (i) the random magnetic field in the shock's plane could produce almost as high polarization as a uniform field (provided that the emitting jet is narrow and one is looking along the edge of the jet). (ii) Even if the magnetic field is uniform the flow does not have to be Poynting flux dominated. They also stress that, while in the uniform field case we expect high polarization in almost every burst, in the random field one we can expect high polarization in only a very few bursts. The different time dependences of the polarization (Nakar, Piran, and Waxman, 2003) could also enable us to distinguish between the two possibilities.

Lazzati *et al.* (2003) and Dar and De Rujula (2003) suggested that this polarization implies inverse Compton scattering (which can have in principle higher intrinsic polarization). This shows that even the simplest conclusion (that the polarization confirms synchrotron radiation as the emission mechanism) is uncertain. My overall conclusion is that without further data on other bursts (which is, unfortunately, quite unlikely in the near future) not much can be learned from this tentative detection.

VII. THE AFTERGLOW

It is generally accepted that the afterglow is produced when the relativistic ejecta are slowed down by the surrounding matter (Mészáros and Rees, 1997a). The afterglow begins at R_{ext} where most of the energy of the ejecta is transferred to the shocked external medium. For a long burst this takes place while the burst is still going on (see Sari, 1997, and Sec. VI.D). Initially the process might be radiative, that is, a significant fraction of the kinetic energy is dissipated and the radiation process effects the hydrodynamics of the shock. I discuss

this phase in Sec. VII.C. Later the radiation processes become less efficient and an adiabatic phase begins during which the radiation losses are minor and do not influence the hydrodynamics. The hydrodynamic evolution at this stage is adiabatic. If the ejecta are in the form of a jet with an opening angle θ , then a jet transition will take place when Γ reaches θ^{-1} . A transition into the Newtonian regime takes place when $\Gamma - 1 \approx 0.5$. I begin the discussion of the afterglow with the hydrodynamics of the adiabatic phase and with the resulting synchrotron light curve. I continue with a discussion of the possible early radiative evolution. Then I turn to the jet break and to the Newtonian transition. Various complications and variations on these themes conclude the section.

A. Relativistic blast waves and the Blandford-McKee solution

The theory of relativistic blast waves was worked out in a classic paper by Blandford and McKee in 1976. The Blandford-McKee model is a self-similar spherical solution describing an adiabatic ultrarelativistic blast wave in the limit $\Gamma \gg 1$. This solution is the relativistic analog of the well-known Newtonian Sedov-Taylor solution. Blandford and McKee (1976) also described a generalization for varying ambient mass density, $\rho = \rho_0(R/R_0)^{-k}$, R being the distance from the center. The latter case would be particularly relevant for $k=2$, as expected in the case of wind from a progenitor, prior to the GRB explosion.

The Blandford-McKee solution describes a narrow shell of width $\sim R/\Gamma^2$, in which the shocked material is concentrated. For simplicity I approximate the solution with a thin homogenous shell. Then the adiabatic energy conservation yields

$$E = \frac{\Omega}{3-k} (\rho_0 R_0^k) R^{3-k} \Gamma^2 c^2, \quad (74)$$

where E is the energy of the blast wave and Ω is the solid angle of the afterglow. For a full sphere $\Omega = 4\pi$, but it can be smaller if the expansion is conical with an opening angle θ : $\Omega = 4\pi(1 - \cos \theta) \approx 2\pi\theta^2$ (assuming a double-sided jet). This expression can be simplified using a generalized Sedov scale:

$$l = [(3-k)E/\rho_0 R_0^k c^2]^{1/(3-k)}. \quad (75)$$

If Ω does not change with time, then the blast wave collects ambient rest mass that equals its initial energy at $R=l$. If we take into account sideways expansion (after the jet break) we find that $\Gamma \approx 1$ and the blast wave becomes Newtonian at

$$R = l(\Omega/4\pi)^{1/(3-k)}. \quad (76)$$

Using the approximate (the numerical factor in this equation assumes that the shell is moving at a constant velocity) time-radius relation in Eq. (8) one can invert Eq. (74) [using the definition of l in Eq. (75)] and obtain R and Γ as a function of time:

$$R = \left[\frac{2l^{3-k}}{\Omega} \right]^{1/(4-k)} t^{1/(4-k)}, \quad (77)$$

$$\Gamma = \left[\frac{l^{3-k}}{2^{3-k}\Omega} \right]^{1/2(4-k)} t^{-(3-k)/2(4-k)}.$$

The time in these expressions is the observer time—namely, the time that photons emitted at R arrive at the observer (relative to the time that photons emitted at $R=0$). For spherical (or spherical-like) evolution, Ω in these expressions is a constant. In general it is possible that Ω varies with R or with Γ (as is the case in a sideways expansion of a jet). This will produce, of course, a different dependence of R and Γ on t .

The values of R and Γ from Eq. (78) can now be plugged into the typical frequencies ν_c , ν_m , and ν_{sa} as well as into the different expression for $F_{\nu, \max}$ to obtain the light curve of the afterglow.

Alternatively, one can calculate the light curve using a more detailed integration over the Blandford-McKee density and energy profiles. To perform such an integration recall that the radius of the front of the shock is

$$R = \hat{t}[1 - [2(4-k)\Gamma^2]^{-1}], \quad (78)$$

where $\Gamma(t)$ is the shock's Lorentz factor and \hat{t} is the time since the explosion was in its rest frame. The different hydrodynamic parameters behind the shock can be expressed as functions of a dimensionless parameter χ :

$$\chi \equiv [1 + 2(4-k)\Gamma^2](1 - R/\hat{t}), \quad (79)$$

as

$$n = 2\sqrt{2}n_1\Gamma\chi^{-(10-3k)/(2(4-k))},$$

$$\gamma^2 = \frac{1}{2}\Gamma^2\chi^{-1},$$

$$p = \frac{2}{3}w_1\Gamma^2\chi^{-(17-4k)/(12-3k)}, \quad (80)$$

where n_1 and w_1 are the number density and enthalpy density of the undisturbed circumburst material and n and p are measured in the fluid's rest frame.

The Blandford-McKee solution is self-similar and assumes $\Gamma \gg 1$. Obviously, it breaks down when $R \sim l$. This relativistic-to-Newtonian transition should take place around

$$t_{NR} = lc \approx 1.2 \text{ yr} (E_{\text{iso}, 52}/n_1)^{1/3}, \quad (81)$$

where the scaling is for $k=0$, E_{52} is the isotropic equivalent energy, $E_{\text{iso}} = 4\pi E/\Omega$, in units of 10^{52} ergs, and n_1 is the external density in cm^{-3} . After this transition the solution will turn into the Newtonian Sedov-Taylor solution with

$$R = R_{NR}(t/t_{NR})^{2/5}, \quad (82)$$

$$\nu = \nu_{NR}(t/t_{NR})^{-3/5},$$

$$e = e_{NR}(t/t_{NR})^{-6/5}. \quad (83)$$

The adiabatic approximation is valid for most of the duration of the afterglow. However, during the first hour or so (or even for the first day for $k=2$), the system could be radiative (provided that $\epsilon_e \approx 1$) or partially radiative. During a radiative phase the evolution can be approximated as

$$E = \frac{\Omega}{3-k} AR^{3-k} \Gamma_0 c^2, \quad (84)$$

where Γ_0 is the initial Lorentz factor. Cohen *et al.* (1998) derived an analytic self-similar solution describing this phase.

Cohen and Piran (1999) described a solution for the case when energy is continuously added to the blast wave by the central engine, even during the afterglow phase. A self-similar solution arises if the additional energy deposition behaves like a power law. This would arise naturally in some models, e.g., in a pulsarlike model (Usov, 1994).

B. Light curves for the “standard” adiabatic synchrotron model

In Sec. V.C.2 we considered the instantaneous synchrotron spectrum. The light curve that corresponds to this spectrum depends simply on the variation of the $F_{\nu, \max}$ and the break frequencies as a function of the observer time (Mészáros and Rees, 1997a; Sari *et al.*, 1998). This in turn depends on the variation of the physical quantities along the shock front. For simplicity I approximate here the Blandford-McKee solution as a spherical homogeneous shell in which the physical conditions are determined by the shock jump between the shell and the surrounding matter. As in Sec. V.C.2 the calculation is divided into two cases: fast cooling and slow cooling.

Sari *et al.* (1998) estimated the observed emission as a series of power-law segments in time and frequency:⁹

$$F_\nu \propto t^{-\alpha} \nu^{-\beta}, \quad (85)$$

which are separated by break frequencies, across which the exponents of these power laws change: the cooling frequency ν_c , the typical synchrotron frequency ν_m , and the self-absorption frequency ν_{sa} . To estimate the rates one inserts the expressions for Γ and R as a function of the observer time [Eq. (78)], using for a homogenous external matter $k=0$:

$$R(t) \equiv (17Et/4\pi m_p n c)^{1/4},$$

⁹The following notation appeared in the astro-ph version of Sari *et al.* (1998). Later during the proofs the authors realized that α is often used in astrophysics to denote a spectral index, and in the Astrophys. J. version of Sari *et al.* (1998) the notations were changed to $F_\nu \propto t^{-\beta} \nu^{-\alpha}$. However, in the meantime the astro-ph notation became generally accepted. I use these notations here.

$$\Gamma(t) \equiv (17E/1024\pi n m_p c^5 t^3)^{1/8}, \quad (86)$$

into the expressions of the cooling frequency ν_c , the typical synchrotron frequency ν_m , and the self-absorption frequency ν_{sa} [Eq. (26)] and into the expression for the maximal flux [Eq. (29) for slow cooling and Eq. (26) for fast cooling]. Note that the numerical factors in the above expressions arise from an exact integration over the Blandford-McKee profile. This procedure results in

$$\begin{aligned} \nu_c &= 0.85 \times 10^{14} \text{ Hz} (\epsilon_B/0.1)^{-3/2} E_{52}^{-1/2} n_1^{-1} t_d^{-1/2}, \\ \nu_m &= 1.8 \times 10^{12} \text{ Hz} (\epsilon_B/0.1)^{1/2} (\epsilon_e/0.1)^2 E_{52}^{1/2} t_d^{-3/2}, \\ F_{\nu, \max} &= 0.35 \times 10^5 \mu\text{J} (\epsilon_B/0.1)^{1/2} E_{52} n_1^{1/2} D_{28}^{-2}. \end{aligned} \quad (87)$$

A nice feature of this light curve is that the peak flux is constant and does not vary with time (Meszaros and Rees, 1997a) as it moves to lower and lower frequencies.

At sufficiently early times $\nu_c < \nu_m$, i.e., fast cooling, while at late times $\nu_c > \nu_m$, i.e., slow cooling. The transition between the two occurs when $\nu_c = \nu_m$. This corresponds (for adiabatic evolution) to

$$t_0 = 0.5 \text{ h} (\epsilon_B/0.1)^2 (\epsilon_e/0.1)^2 E_{52} n_1. \quad (88)$$

Additionally one can translate Eq. (87) to the time in which a given break frequency passes a given band. Consider a fixed frequency $\nu = \nu_{15} 10^{15} \text{ Hz}$. There are two critical times, t_c and t_m , when the break frequencies ν_c and ν_m cross the observed frequency ν :

$$\begin{aligned} t_c &= 0.2 \text{ h} (\epsilon_B/0.1)^{-3} E_{52}^{-1} n_1^{-2} \nu_{15}^{-2}, \\ t_m &= 0.2 \text{ h} (\epsilon_B/0.1)^{1/3} (\epsilon_e/0.1)^{4/3} E_{52}^{1/3} \nu_{15}^{-2/3}. \end{aligned} \quad (89)$$

In the Rayleigh-Jeans part of the blackbody radiation $I_\nu = kT(2\nu^2/c^2)$ so that $F_\nu \propto kT\nu^2$. Therefore, in the part of the synchrotron spectrum that is optically thick to synchrotron self-absorption, we have $F_\nu \propto kT_{\text{eff}}\nu^2$. For slow cooling $kT_{\text{eff}} \sim \gamma_m m_e c^2 = \text{const}$ throughout the whole shell of shocked fluid behind the shock, and therefore $F_\nu \propto \nu^2$ below ν_{sa} where the optical depth to synchrotron self-absorption equals one, $\tau_\nu = 1$. For fast cooling, as we go down in frequency, the optical depth to synchrotron self-absorption first equals unity due to absorption over the whole shell of shocked fluid behind the shock, most of which is at the back of the shell and has $kT_{\text{eff}} \sim \gamma_c$. The observer is located in front of the shock, and the radiation that escapes and reaches the observer is from $\tau_\nu \sim 1$. As ν decreases below ν_{sa} the location where $\tau_\nu \sim 1$ moves from the back of the shell toward the front of the shell, where the electrons suffer less cooling so that $kT_{\text{eff}}(\tau_\nu=1) \propto \nu^{-5/8}$. Consequently $F_\nu \propto \nu^{11/8}$. At a certain frequency $\tau_\nu \sim 1$ at the location behind the shock where electrons with γ_m start to cool significantly. Below this frequency (ν_{ac}), even though $\tau_\nu \sim 1$ closer and closer to the shock with decreasing ν , the effective temperature at that location is constant: $kT_{\text{eff}} \sim \gamma_m m_e c^2 = \text{const}$, and therefore $F_\nu \propto \nu^2$ for $\nu < \nu_{ac}$, while $F_\nu \propto \nu^{11/8}$ for $\nu_{ac} < \nu < \nu_{sa}$. Overall the expression for the self-absorption fre-

TABLE I. α and β for fast cooling ($\nu_a < \nu_c < \nu_m$) into a constant-density ISM.

	α	β
$\nu < \nu_a$	1	2
$\nu_a < \nu < \nu_c$	1/6	1/3
$\nu_c < \nu < \nu_m$	-1/4	-1/2
$\nu_m < \nu$	-(3p-2)/4	-p/2 = (2α-1)/3

quency depends on the cooling regime. It divides into two cases, denoted ν_{sa} and ν_{ac} , for fast cooling, and both expressions are different from the slow cooling (Granot *et al.*, 2000). For fast cooling,

$$\begin{aligned} \nu_{ac} &= 1.7 \times 10^9 \text{ Hz} (\epsilon_B/0.1)^{-2/5} (\epsilon_e/0.1)^{-8/5} \\ &\times E_{52}^{-1/10} n_1^{3/10} (t/100 \text{ sec})^{3/10}, \end{aligned} \quad (90)$$

$$\nu_{sa} = 1.8 \times 10^{10} \text{ Hz} (\epsilon_B/0.1)^{6/5} E_{52}^{7/10} n_1^{11/10} (t/100 \text{ sec})^{-1/2}. \quad (91)$$

For slow cooling,

$$\nu_{sa} = 1.24 \times 10^9 \text{ Hz} \frac{(p-1)^{3/5}}{(3p+2)^{3/5}} (1+z)^{-1} \epsilon_e^{-1} \epsilon_B^{1/5} n_0^{3/5} E_{52}^{1/5}. \quad (92)$$

For a given frequency either $t_0 > t_m > t_c$ (which is typical for high frequencies) or $t_0 < t_m < t_c$ (which is typical for low frequencies). The results are summarized in Tables I and II describing α and β for fast and slow cooling. The different light curves are depicted in Fig. 16.

These results are valid only for $p > 2$ (and for γ_{\max} , the maximal electron energy, much higher than γ_{\min}). If $p < 2$ then γ_{\max} plays a critical role. The resulting temporal and spectral indices for slow cooling with $1 < p < 2$ are given by Dai and Cheng (2001) and by Bhattacharya (2001) and summarized in Table III. For completeness I also include in this table the cases of propagation into a wind (see Sec. VII.E) and a jet break (see Sec. VII.H).

The simple solution, based on a homogeneous shell approximation, can be modified by using the full Blandford-McKee solution and integrating over the entire volume of shocked fluid (Granot *et al.*, 1999a). Following Nakar and Piran (2003a) I discuss in Sec. VII.G.1 a simple way to perform this integration. The detailed integration yields a smoother spectrum and light curve

TABLE II. α and β for slow cooling ($\nu_a < \nu_m < \nu_c$) into a constant-density ISM.

	α	β
$\nu < \nu_a$	1/2	2
$\nu_a < \nu < \nu_m$	1/2	1/3
$\nu_m < \nu < \nu_c$	-3(p-1)/4	-(p-1)/2 = 2α/3
$\nu_c < \nu$	-(3p-2)/4	-p/2 = (2α-1)/3

TABLE III. α for slow cooling ($\nu_a < \nu_m < \nu_c$) into a constant-density ISM, wind, and jet for electron distribution with $1 < p < 2$.

	ISM	Wind	Jet
$\nu < \nu_a$	$(17p-26)/16(p-1)$	$(13p-18)/18(p-1)$	$3p-2)/4(p-1)$
$\nu_a < \nu < \nu_m$	$(p+1)/8(p-1)$	$5(2-p)/12(p-1)$	$(8-5p)/6(p-1)$
$\nu_m < \nu < \nu_c$	$-3(p+2)/16$	$-(p+8)/8$	$-(p+6)/4$
$\nu_c < \nu$	$-(3p+10)/16$	$-(p+6)/8$	$-(p+6)/4$

near the break frequencies, but the asymptotic slopes away from the break frequencies, and the transition times remain the same as in the simpler theory. Granot and Sari (2002) described a detailed numerical analysis of the smooth afterglow spectrum including a smooth approximation for the spectrum over the transition regions (see also Gruzinov and Waxman, 1999). They also described additional cases of ordering of typical frequencies that were not considered earlier.

A final note on this “standard” model is that it assumes adiabaticity. However, in reality a fraction of the energy is lost, and over the long run this influences the hydrodynamic behavior. This could easily be corrected by an integration of the energy losses and addition of a variable energy to Eq. (74), followed by the rest of the procedure described above (Panaitescu and Kumar, 2000).

C. Light curves for the early radiative phase

If the electron energy is large (that is, if ϵ_e is not far from unity), then early on during the first few hours of the afterglow there will be a radiative phase in which a significant fraction of the kinetic energy is lost via the radiative processes. One can generalize the Blandford-McKee solution to this radiative stage (see Cohen *et al.*, 1998, and Sec. VII.A). The essence of the radiative phase is that in this case the energy varies as $E \propto \Gamma$, where $\Gamma \equiv (R/L)^{-3}$. Note that L is calculated in terms of M and the initial energy of the explosion, E_0 , via $M = E_0/\Gamma_0 c^2$, where Γ_0 is the initial Lorentz factor of the ejecta:

$$\begin{aligned} R(t) &\cong (4ct/L)^{1/7} L, \\ \Gamma(t) &\cong (4ct/L)^{-3/7}. \end{aligned} \quad (93)$$

The transition from the radiative to the adiabatic phase takes place when the radiation losses become negligible. This happens at time

$$t_{rad} = 0.17 \text{ h} (\epsilon_B/0.1)^{7/5} (\epsilon_e/0.1)^{7/5} E_{52}^{4/5} (\Gamma/100)^{-4/5} n_1^{3/5}. \quad (94)$$

Following Sari *et al.* (1998) one can use the above expressions to show the different typical frequencies and fluxes as

$$\nu_c = 4.1 \times 10^{14} \text{ Hz} (\epsilon_B/0.1)^{-3/2} E_{52}^{-4/7} (\Gamma/100)^{4/7} n_1^{-13/14} t_d^{-2/7},$$

$$\begin{aligned} \nu_m &= 3.8 \times 10^{11} \text{ Hz} (\epsilon_B/0.1)^{1/2} (\epsilon_e/0.1)^2 \\ &\times E_{52}^{4/7} (\Gamma/100)^{-4/7} n_1^{-1/14} t_d^{-12/7}, \end{aligned}$$

$$F_{\nu, max} = 1.4 \times 10^3 \mu\text{J} \epsilon_B^{1/2} E_{52}^{8/7} (\Gamma/100)^{-8/7} n_1^{5/14} D_{28}^{-2} t_d^{-3/7}. \quad (95)$$

As in the adiabatic case this can be translated to the times of passage of the break frequencies at a given observed frequency:

$$t_c = 0.05 \times 10^{-7} \text{ days} (\epsilon_B/0.1)^{-21/4} E_{52}^{-2} \Gamma_2^2 n_1^{-13/4} \nu_{15}^{-7/2},$$

$$t_m = 0.01 \text{ days} (\epsilon_B/0.1)^{7/24} (\epsilon_e/0.1)^{7/6} E_{52}^{1/3} \Gamma_2^{-1/3} \nu_{15}^{-7/12} n_1^{-1/24}. \quad (96)$$

Unlike the adiabatic case, here ν_c must be below ν_m . Otherwise the bulk of the electrons do not cool and the system will not be radiative. Indeed at t_{rad} [given by Eq. (94) above] $\nu_c = \nu_m$.

D. Light curve during the Newtonian transition

At $t \approx t_{NR}$ [see Eq. (81)] the afterglow reaches the Newtonian Sedov-Taylor phase. During this phase the adiabatic hydrodynamic is described by Eq. (83). Frail, Waxman, and Kulkarni (2000) calculated the synchrotron spectrum and light curve of the afterglow in this stage. The energy scaling implies that $B \propto t^{-3/5}$ and $\gamma_{e, min} \propto t^{-6/5}$. Combined together this yields $\nu_m \propto t^{-3}$. Using the standard assumptions of equipartition and of a power-law electron distribution they found

$$\nu_c = 10^{13} \text{ Hz} (\epsilon_B/0.3)^{-3/2} E_{51}^{-2/3} n_1^{-5/6} (t/t_{NR})^{-1/5},$$

$$\nu_m = 1 \text{ GHz} (\epsilon_B/0.3)^{1/2} (\epsilon_e/0.3)^2 n_1^{-1/2},$$

$$\begin{aligned} F_{\nu_m < \nu < \nu_c} &= 1 \text{ mJ} (\epsilon_B/0.3)^{3/4} (\epsilon_e/0.3) n_1^{3/4} E_{51} D_{28}^{-2} \\ &\times \nu_{GHz}^{-(p-1)/2} (t/t_{NR})^{-3(p-1)/2+3/5}. \end{aligned} \quad (97)$$

This late-time light curve provides a simple “calorimetric” estimate of the afterglow energy at this stage (see Sec. II.D). Additionally as the radio flux is rather large and as it varies on a scale of several months it can be used to search for orphan radio afterglows (see Levinson *et al.*, 2002, and Sec. VII.K.2).

TABLE IV. α and β for x-ray and optical frequencies from a blast wave into a wind profile when $\nu_a < \nu_c, \nu_m, \nu$ (Chevalier and Li, 2000). Note that the order of the table is according to the evolution of the light curve at a fixed high observed frequency.

	α	β
$\nu_c < \nu < \nu_m$	$-1/4$	$-1/2$
$\nu_m, \nu_c < \nu$	$-(3p-2)/4$	$-p/2 = (2\alpha-1)/3$
$\nu_m < \nu < \nu_c$	$-(3p-1)/4$	$-(p-1)/2 = (2\alpha+1)/3$

E. Generalizations: I. Winds

The simplest generalization of the previous models is to allow a variable circumburst density with $n(R) \propto R^{-k}$. The hydrodynamic evolution of a relativistic blast wave in such a medium was considered in the original paper of Blandford and McKee (1976). The synchrotron light curve was considered first by Meszaros, Rees, and Wijers (1998) and by Dai and Lu (1999).

Chevalier and Li (1999, 2000) stressed the importance of the $n(R) \propto R^{-2}$ case, which arises whenever there is a stellar wind ejected by the GRB's progenitor prior to the burst. This arises naturally in the collapsar model, which is based on the collapse of a massive star. The calculations follow the outlines in the previous sections, except that the relations determining $R(t)$ and $\Gamma(t)$ for a homogeneous circumburst medium, Eqs. (86), should be replaced by Eq. (78) with $k=2$.

High initial densities in a wind density profile imply a low initial cooling frequency. Unlike the constant-density case the cooling frequency here increases with time (Chevalier and Li, 1999). This leads to a different temporal relation between the different frequencies and cooling regimes. For example, it is possible that the cooling frequency would initially be below the synchrotron self-absorption frequency. Chevalier and Li (2000) considered five different evolutions of light curves for different conditions and observed the frequencies. Tables IV and V show the two most relevant cases: Table IV fits the x-ray and optical afterglows, while Table V is typical for the lower radio frequencies.

Note that for $\nu_m, \nu_c < \nu$ both the spectral slope and the temporal evolution are similar for a wind and for a con-

TABLE V. α and β for radio frequencies from a blast wave into a wind profile (Chevalier and Li, 2000). Note that the order of the table is according to the evolution of the light curve at a fixed low observed frequency.

	α	β
$\nu_c < \nu < \nu_a < \nu_m$	$7/4$	$5/2$
$\nu < \nu_c < \nu_a < \nu_m$	2	2
$\nu < \nu_a < \nu_m < \nu_c$	1	2
$\nu_a < \nu < \nu_m < \nu_c$	0	$1/3$
$\nu_a < \nu_m < \nu < \nu_c$	$-(3p-1)/4$	$-(p-1)/2 = (2\alpha+1)/3$

stant density profile. This of course poses, a problem in the interpretation of afterglow light curves.

F. Generalizations: II. Energy injection and refreshed shocks

The simple adiabatic model assumes that the energy of a GRB is constant. However, the energy could change if additional slower material is ejected behind the initial matter. This would be expected generically in the internal shock model. In this model the burst is produced by a series of collisions between shells moving at different velocities. One naturally expects here slower-moving matter that does not catch up initially with the faster-moving matter. However, as the initially faster-moving matter is slowed down by the circumburst matter, the slower matter eventually catches up and produces refreshed shocks (Rees and Mészáros, 1998; Kumar and Piran, 2000b; Sari and Mészáros 2000).

There are two implications for the refreshed shocks. First the additional energy injection will influence the dynamics of the blast wave (Rees and Mészáros, 1998; Sari and Mészáros, 2000). This effect can be modeled by modifying E in Eq. (74) but the effect of additional mass carrying the slower energy must be included in some cases. This would change the decay slope from the canonical one and produce a slower decay in the light curve. In Sec. VII.G I describe a scheme for calculating the light curve resulting from a variable blast-wave energy. If the additional matter is emitted sporadically, then the shell collision could produce initial temporal variability in the early afterglow signal (Kumar and Piran, 2000b). Fox, Yost, *et al.* (2003), for example, suggest that refreshed shocks are the origin of the variability in the early afterglow of GRB 021004.

A second effect is the production of a reverse shock propagating into the slower material when it catches up with the faster one (Kumar and Piran, 2000b). This is of course in addition to the forward shock that propagates into the outer shell. This reverse shock could be episodic or long lasting, depending on the profile of the additional matter. Kumar and Piran (2000b) consider two shells with energies E_1 and E_2 in the outer and inner shells, respectively. The outer shell is moving with a bulk Lorentz factor $\Gamma_{0c} \sim 5(t/\text{day})^{3/8}$ at the (observed) time t of the collision. As the inner shell catches up with the outer one when both shells have comparable Lorentz factors the reverse shock is always mildly relativistic. The calculation of the shock is slightly different from the calculation of a shell propagating into a cold material (another shell or the ISM) discussed earlier. Here the outer shell has already collided with the ISM. Hence it is hot, with internal energy exceeding the rest-mass energy. The reverse shock produces emission at a characteristic frequency that is typically much lower than the peak of the emission from the outer shell by a factor of $\sim 7\Gamma_{0c}^2(E_2/E_1)^{1.1}$, and the observed flux at this frequency from the reverse shock is larger compared to the flux from the outer shell by a factor of $\sim 8(\Gamma_{0c}E_2/E_1)^{5/3}$. This emission is typically in the radio or the FIR range.

Kumar and Piran (2000b) suggest that due to angular spreading the refreshed shocks produce an enhancement, with a typical time scale $\delta t \sim t$. Granot, Nakar, and Piran (2003) stress that because energy necessarily increases in refreshed shocks, the overall light curve must have a stepwise shape (above the continuous power-law decline) with a break at the corresponding shocks. This behavior was seen in GRB 030329. However, there the transitions are fast with $\delta t < t$. Granot, Nakar, and Piran (2003) point out that if the refreshed shocks take place *after* the jet break (as is likely the case in GRB 030329) then if the later shells remain cold and do not spread sideways we would have $\delta t \sim t_{jet} < t$. This explains nicely the fast transitions seen in this burst.

G. Generalizations: III. Inhomogeneous density profiles

An interesting possibility that arose with the observation of the variable light curve of the afterglow of GRB 021004 is that the ejecta encounter surrounding matter with an irregular density profile (Lazzati, 2002; Heyl and Perna, 2003; Nakar, Piran, and Granot, 2003). To explore this situation one can resort to numerical simulation of the propagation of the blast wave into a selected density profile (Lazzati, 2002). Or one can attempt to model this analytically or almost analytically (Nakar and Piran, 2003a). The key for this analytical model is the approximation of the light curve from an inhomogeneous density profile as a series of emissions from instantaneous Blandford-McKee solutions, each with its own external density.

1. The light curve of a Blandford-McKee solution

The observed flux, at an observer's time t , from an arbitrary spherically symmetric emitting region is given by (Granot *et al.*, 1999a)

$$F_\nu(t) = \frac{1}{2D^2} \int_0^\infty dt' \int_0^\infty r^2 dr \int_{-1}^1 d(\cos \theta) \frac{n'(r) P'_\nu(\nu \Lambda, r)}{\Lambda^2} \times \delta\left(t' - t - \frac{r \cos \theta}{c}\right), \quad (98)$$

where n' is the emitter density and P'_ν is the emitted

spectral power per emitter, both are measured in the fluid frame, θ is the angle relative to the line of sight, and $\Lambda^{-1} = 1/\gamma(1 - \nu \cos \theta/c)$ (ν is the emitting matter bulk velocity) is the blueshift factor.

Nakar and Piran (2003a) show¹⁰ that using the self-similar nature of the Blandford-McKee profile (with an external density $\propto r^{-k}$) one can reduce Eq. (98) to

$$F_\nu(t) = \frac{1}{D^2} \int_0^{R_{max}(t)} A_\nu(R) g_\beta(\tilde{t}, k) dR. \quad (99)$$

The integration over R is over the shock front of the BM solution. The upper limit R_{max} corresponds to the shock position from which photons leaving along the line of sight reach the observer at t . The factor D is the distance to the source (neglecting cosmological factors), and β is the local spectral index.

The factor g_β is a dimensionless factor that describes the observed pulse shape of an instantaneous emission from a Blandford-McKee profile. The instantaneous emission from a thin shell produces a finite pulse (see Sec. IV.B and Fig. 15). This is generalized now to a pulse from an instantaneous emission from a Blandford-McKee profile. Note that even though this profile extends from 0 to R most of the emission comes from a narrow region of width $\sim R/\Gamma^2$ behind the shock front. g_β is obtained by integrating Eq. (98) over $\cos \theta$ and r , i.e., over the volume of the Blandford-McKee profile. It depends only on the radial and angular structure of the shell. The self-similar profile of the shell enables us to express g_β as a general function that depends only on the dimensionless parameter $\tilde{t} \equiv [t - t_{los}(R)]/t_{ang}(R)$, where $t_{los}(R)$ is the time in which a photon emitted at R along the line of sight to the center reaches the observer, and $t_{ang} \equiv R/2c\Gamma^2$. The second function, A_ν , depends only on the conditions of the shock front along the line of sight. It includes only numerical parameters that remain after the integration over the volume of the shell.

When all the significant emission from the shell at radius R is within the same power-law segment, β (i.e., ν is far from the break frequencies), then A_ν and g_β are given by

$$A_\nu(R) = H_\nu \times \begin{cases} R^2 n_{ext,0}^{4/3} E_{52}^{1/3} M_{29}^{-1/3} & \nu < \nu_m \\ R^2 n_{ext,0}^{(5+p)/4} E_{52}^p M_{29}^{-p} & \nu_m < \nu < \nu_c \\ R n_{ext,0}^{(2+p)/4} E_{52}^p M_{29}^{-p} & \nu > \nu_c, \end{cases} \quad \frac{\text{erg}}{\text{sec cm Hz}} \quad (100)$$

where R is the radius of the shock front, $n_{ext}(R)$ is the external density, E is the energy in the blast wave, $M(R)$ is the total collected mass up to radius R , and H_ν is a numerical factor that depends on the observed power-law segment (see Nakar and Piran, 2003a for the numerical values). We obtain

¹⁰See Granot *et al.* (1999a) for an alternative method of integrating Eq. (98).

$$g(\tilde{t}, \beta, k) = \begin{cases} \frac{2}{(4-k)} \int_1^{1+2(4-k)\tilde{t}} \chi^{-\mu(\beta, k)} \left(1 - \frac{1}{2(4-k)} + \frac{2(4-k)\tilde{t}+1}{2(4-k)\chi} \right)^{-(2-\beta)} d\chi & \nu < \nu_c \\ (1+\tilde{t})^{-(2-\beta)} & \nu > \nu_c, \end{cases} \quad (101)$$

where

$$\mu(\beta, k) \equiv 3(71 - 17k)/(72 - 18k) - \beta(37 + k)/(24 - 6k). \quad (102)$$

This set of equations is completed with the relevant relations among different variables of the blast wave, the observer time, and the break frequencies. The equations also describe the light curve within one power-law segment of the light curve. Matching between different power laws can be easily done (Nakar and Piran, 2003a). The overall formalism can be used to calculate the complete light curve of a Blandford-McKee blast wave.

2. The light curve with a variable density or energy

The results of the previous section can be applied to study the effect of variations in the external density or in the energy of the blast wave by approximating the solution as a series of instantaneous Blandford-McKee solutions whose parameters are determined by the instantaneous external density and the energy. Both can vary with time. This would be valid, of course, if the variations were not too rapid. The light curve can be expressed as an integral over the emission from a series of instantaneous Blandford-McKee solutions.

When a blast wave at radius R propagates into the circumburst medium, the emitting matter behind the shock is replenished within $\Delta R \approx R(2^{1/(4-k)} - 1)$. This is the length scale over which an external density variation relaxes to the Blandford-McKee solution. This approximation is valid as long as the density variations are on a larger length scale than ΔR . It fails when there is a sharp density increase over a range of ΔR . However, the contribution to the integral from the region on which the solution breaks is small ($\Delta R/R \ll 1$) and the overall light curve approximation is acceptable. Additionally the density variation must be mild enough so that it does not give rise to a strong reverse shock that destroys the Blandford-McKee profile.

A sharp density decrease is more complicated. Here the length scale over which the emitting matter behind the shock is replenished could be of the order of R . As an example, we consider a sharp drop at some radius R_d and a constant density for $R > R_d$. In this case the external density is negligible at first, and the hot shell cools by adiabatic expansion. Later the forward shock becomes dominant again. Kumar and Panaitescu (2000) show that, immediately after the drop, the light curve is dominated by the emission during the adiabatic cooling. Later the observed flux is dominated by emission from $R \approx R_d$, and at the end the new forward shock becomes

dominant. Our approximation includes the emission before the density drop and the new forward shock after the drop, but it ignores the emission during the adiabatic cooling phase.

As an example for this method, Fig. 25 depicts the $\nu_m < \nu < \nu_c$ light curve for a Gaussian ($\Delta R/R = 0.1$) overdense region in the ISM. Such a density profile may occur in a clumpy environment. The emission from a clump is similar to that from a spherically overdense region as long as the clump's angular size is much larger than $1/\Gamma$. Even a mild, short-length-scale, overdense region (with a maximal overdensity of 2) influences the light curve for a long duration (mainly due to the angular spreading). This duration depends strongly on the magnitude of the overdensity.

The calculations presented so far do not account, however, for the reverse shock resulting from density enhancement and its effect on the blast wave. Thus the above models are limited to slowly varying and low contrast density profiles. Now, the observed flux depends on the external density n roughly as $n^{1/2}$. Thus a large contrast is needed to produce a significant rebrightening. Such a large contrast will, however, produce a strong reverse shock, which will sharply decrease the Lorentz factor of the emitting matter behind the shock, Γ_{sh} , caus-

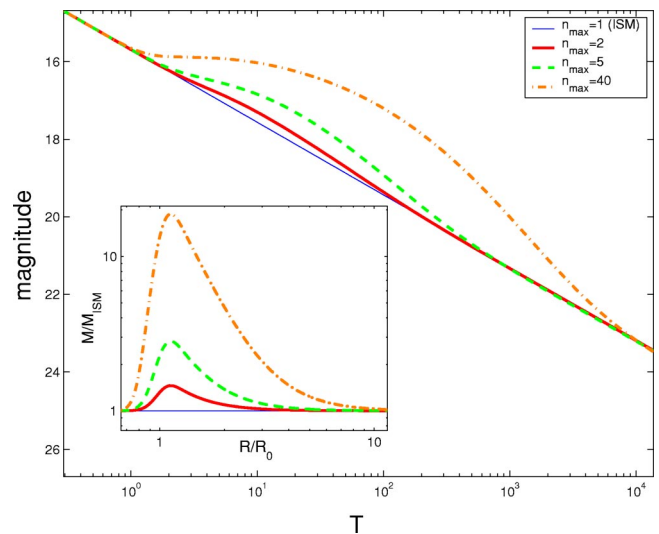


FIG. 25. (Color in online edition) Light curves resulting from a Gaussian ($\Delta R/R = 0.1$) overdense region in the ISM. The different thick lines are for maximal overdensities of 40 (dashed-dot), 5 (dashed), and 2 (solid). The thin line is the light curve for a constant ISM density. The inset depicts the ratio of the mass, $M(R)$ over the mass of an ISM (without the overdense region), $M_{ISM}(R)$. From Nakar and Piran, 2003a.

ing a sharp drop in the emission below ν_c and a long delay in the arrival time of the emitted photons (the observer time is $\propto \Gamma_{sh}^{-2}$). Both factors combine to suppress the flux and to set a strong limit on the steepness of the rebrightening events caused by density variations.

The method can be applied also to variations in the blast wave's energy. Spherically symmetric energy variations are most likely to occur due to refreshed shocks, when new inner shells arrive from the source and refresh the blast wave (Rees and Mészáros, 1998; Kumar and Piran, 2000b; Sari and Mészáros, 2000). Once more, this approximation misses the effect of the reverse shock that arise in this case (Kumar and Piran, 2000b). However, it enables a simple calculation of the observed light curve for a given energy profile.

H. Generalizations: IV. Jets

The afterglow theory becomes much more complicated if the relativistic ejecta are not spherical. The so-called “jets” correspond to relativistic matter ejected into a cone of opening angle θ . I stress that unlike other astrophysical jets this cone is non-steady-state and generally its width (in the direction parallel to the motion) is orders of magnitude smaller than the radius where the jet is. A “flying pancake” is a better description for these jets.

The simplest implication of a jet geometry, which exists regardless of the hydrodynamic evolution, is that once $\Gamma \sim \theta^{-1}$ relativistic beaming of light will become less effective. The radiation was initially beamed locally into a cone with an opening angle Γ^{-1} and remained inside the cone of the original jet. Now with $\Gamma^{-1} > \theta$ the emission is radiated outside of the initial jet. This has two effects: (i) An “on axis” observer, one who sees the original jet, will detect a jet break due to the faster spreading of the emitted radiation. (ii) An “off axis” observer, who could not detect the original emission will now be able to see an “orphan afterglow,” an afterglow without a preceding GRB (see Sec. VII.K). The time of this transition is always given by Eq. (104) below, with $C_2=1$.

Additionally the hydrodynamic evolution of the source changes when $\Gamma \sim \theta^{-1}$. Initially, as long as $\Gamma \gg \theta^{-1}$ (Piran, 1994) the motion would be almost conical. There is not enough time, in the blast wave's rest frame, for the matter to be affected by the nonspherical geometry, and the blast wave will behave as if it were a part of a sphere. When $\Gamma = C_2 \theta^{-1}$, that is, when

$$t_{\text{jet}} = \frac{1}{C_1} \left(\frac{l}{c} \right) \left(\frac{\theta}{C_2} \right)^{2(4-k)/(3-k)} \quad (103)$$

sideways propagation begins.¹¹ The constant C_1 expresses the uncertainty of the relation between the Lor-

TABLE VI. α and β for slow cooling ($\nu_a < \nu_m < \nu_c$) after a jet break.

	α	β
$\nu < \nu_a$	0	2
$\nu_a < \nu < \nu_m$	-1/3	1/3
$\nu_m < \nu < \nu_c$	-p	-(p-1)/2 = ($\alpha+1$)/2
$\nu_c < \nu$	-p	-p/2 = $\alpha/2$

entz factor and the observing time and it depends on the history of the evolution of the fireball. The constant C_2 reflects the uncertainty in the value of Γ , when the jet break begins vs the value of the opening angle of the jet θ . For the important case of constant external density $k=0$ this transition takes place at

$$t_{\text{jet}} = \frac{1 \text{ day}}{C_1 C_2^{8/3}} \left(\frac{E_{\text{iso},52}}{n_1} \right)^{1/3} \left(\frac{\theta}{0.1} \right)^{8/3}. \quad (104)$$

The sideways expansion continues with $\theta \sim \Gamma^{-1}$. Plugging this relation in to Eq. (78) and letting Ω vary as Γ^{-2} one finds that

$$R \approx \text{const}; \quad (105)$$

$$\Gamma \approx (R/2t)^{1/2}.$$

A more detailed analysis (Rhoads, 1997, 1999; Kumar and Panaitescu, 2000; Piran, 2000) reveals that according to the simple one-dimensional analytic models, Γ decreases exponentially with R over a very short length scale.¹²

Table VI describes the parameters α and β for a post-jet-break evolution (Sari *et al.*, 1999). The jet break usually takes place rather late, after the radiative transition. Therefore I include in this table only the slow-cooling parameters.

An important feature of the post-jet-break evolution is that ν_c , the cooling frequency, becomes constant in time. This means that the high-frequency (optical and x-ray) optical spectrum does not vary after the jet break takes place. On the other hand, the radio spectrum varies (see Fig. 12), giving an additional structure that confirms the interpretation of the break as arising due to the sideways expansion of a jet (see Harrison *et al.*, 1999).

Panaitescu and Kumar (2000) find that the jet-break transition in a wind profile will be very long (up to four decades in time) and thus it will be hard to observe a jet break in such a case. On the other hand, it is interesting to note that for typical values of α seen after a jet break ($\alpha \approx -2$) the high-frequency spectral index, $\beta = \alpha/2 \approx -1$, is similar to the one inferred from a spherically symmet-

¹¹The exact values of the uncertain constants C_2 and C_1 are extremely important as they determine the jet opening angle (and hence the total energy of the GRB) from the observed breaks, interpreted as t_{jet} , in the afterglow light curves.

¹²Note that the exponential behavior is obtained after converting Eq. (74) to a differential equation and integrating over it. Different approximations used in deriving the differential equation lead to slightly different exponential behavior; see Piran, 2000.

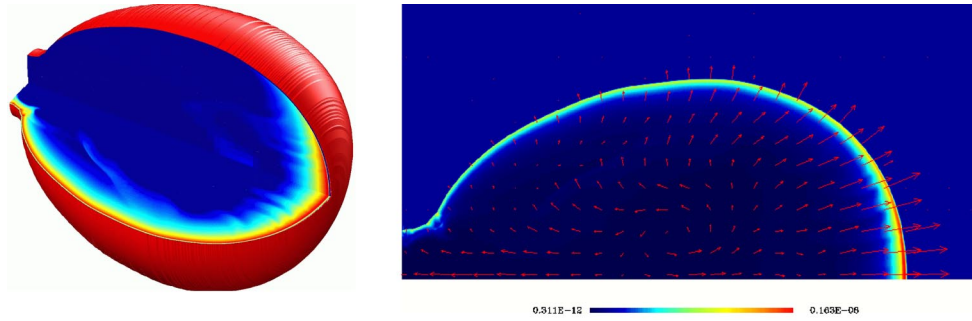


FIG. 26. (Color in online edition) A relativistic jet at the last time step of the simulation: (left) 3D view of the jet. The outer surface represents the shock front, while the two inner faces show the proper number density (lower face) and proper emissivity (upper face) in a logarithmic color scale. (Right) 2D “slice” along the jet axis, showing the velocity field on top of a linear color map of the laboratory frame density. From Granot *et al.*, 2001.

ric wind $\beta = (2\alpha + 1)/3 \approx -1$ (Halpern *et al.*, 1999). Note, however, that the wind interpretation requires a high (≈ 3) p value (which may or may not be reasonable). Still from the optical observations alone it is difficult to distinguish between these two interpretations. Here the radio observations play a crucial role as the radio behavior is very different (Frail, Kulkarni, *et al.*, 2000).

The sideways expansion causes a change in the hydrodynamic behavior and hence a break in the light curve. The beaming outside of the original jet opening angle also causes a break. If the sideways expansion is at the speed of light, then both transitions would take place at the same time (Sari *et al.*, 1999). If the sideways expansion is at the sound speed, then the beaming transition would take place first and only later would the hydrodynamic transition occur (Panaitescu and Mészáros, 1999). This would cause a slower and wider transition with two distinct breaks, first a steep break when the edge of the jet becomes visible and later a shallower break when sideways expansion becomes important.

Analytic or semianalytic calculations of synchrotron radiation from jetted afterglows (Panaitescu and Mészáros, 1999; Rhoads, 1999; Sari *et al.*, 1999; Kumar and Panaitescu, 2000; Moderski *et al.*, 2000) have led to different estimates of the jet-break time t_{jet} and of the duration of the transition. Rhoads (1999) calculated the light curves assuming emission from one representative point, and obtained a smooth jet break, extending ~ 3 – 4 decades in time, after which $F_{\nu > \nu_m} \propto t^{-p}$. Sari *et al.* (1999) assume that the sideways expansion is at the speed of light, and not at the speed of sound ($c/\sqrt{3}$) as others assume, and find a smaller value for t_{jet} . Panaitescu and Mészáros (1999) included the effects of geometrical curvature and the finite width of the emitting shell, along with electron cooling, and obtained a relatively sharp break, extending ~ 1 – 2 decades in time, in the optical light curve. Moderski *et al.* (2000) used a slightly different dynamical model and a different formalism for the evolution of the electron distribution and found that the change in the temporal index α ($F_{\nu} \propto t^{-\alpha}$) across the break was smaller than in analytic estimates ($\alpha = 2$ after the break for $\nu > \nu_m$, $p = 2.4$), while the break extended over two decades in time.

The different analytic or semianalytic models have different predictions for the sharpness of the jet break, the change in the temporal decay index α across the break and its asymptotic value after the break, or even the very existence of a jet break (Huang *et al.*, 2000). All these models rely on some common basic assumptions, which have a significant effect on the dynamics of the jet: (i) the shocked matter is homogeneous; (ii) the shock front is spherical (within a finite opening angle) even at $t > t_{\text{jet}}$; (iii) the velocity vector is almost radial even after the jet break.

However, recent 2D hydrodynamic simulations (Granot *et al.*, 2001) show that these assumptions are not a good approximation of a realistic jet. Using a very different approach, Cannizzo *et al.* (2004) find in another set of numerical simulations a similar result—the jet does not spread sideways as much. Figure 26 shows the jet at the last time step of the simulation of Granot *et al.* (2001). The matter at the sides of the jet is propagating sideways (rather than in the radial direction) and is slower and much less luminous than that at the front of the jet. The shock front is egg shaped, and quite far from being spherical. Figure 27 shows the radius R , Lorentz factor Γ , and opening angle θ of the jet, as a function of the laboratory frame time. The rate of increase of θ with $R \approx ct_{\text{lab}}$, is much lower than the exponential behavior predicted by simple models (Rhoads, 1997, 1999; Kumar and Panaitescu, 2000; Piran, 2000). The value of θ averaged over the emissivity is practically constant, and most of the radiation is emitted within the initial opening angle of the jet. The radius R weighed over the emissivity is very close to the maximal value of R within the jet, indicating that most of the emission originates at the front of the jet,¹³ where the radius is largest, while R averaged over the density is significantly lower, indicating that a large fraction of the shocked matter resides at the sides of the jet, where the radius is smaller. The Lorentz factor Γ averaged over the emissivity is close to its

¹³This may imply that the rate of orphan afterglows should be smaller than estimated assuming significant sideways expansion.

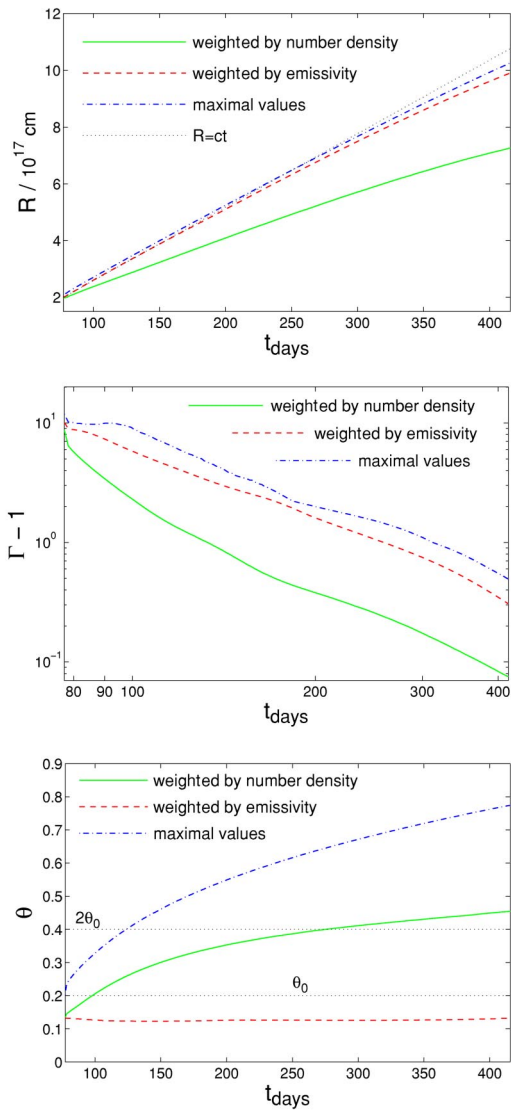


FIG. 27. (Color in online edition) Hydrodynamic simulation of a jet: (a) radius R ; (b) Lorentz factor $\Gamma - 1$; (c) opening angle θ of the jet as a function of the laboratory frame time in days. From Granot *et al.*, 2001.

maximal value (again since most of the emission occurs near the jet axis where Γ is the largest) while Γ averaged over the density is significantly lower, since the matter at the sides of the jet has a much lower Γ than at the front of the jet. The large differences between the assumptions of simple dynamical models of a jet and the results of 2D simulations suggest that great care should be taken when using these models for predicting the light curves of jetted afterglows. As the light curves depend strongly on the hydrodynamics of the jet, it is very important to use a realistic hydrodynamic model when calculating the light curves.

Granot *et al.* (2001) used 2D numerical simulations of a jet running into a constant-density medium to calculate the resulting light curves, taking into account the emission from the volume of the shocked fluid with the

appropriate time delay in the arrival of photons to different observers. They obtained an achromatic jet break for $\nu > \nu_m(t_{\text{jet}})$ (which typically includes the optical and near IR), while at lower frequencies (which typically include the radio) there is a more moderate and gradual increase in the temporal index α at t_{jet} and a much more prominent steepening in the light curve at a later time, when ν_m sweeps past the observed frequency. The jet break appears sharper and occurs at a slightly earlier time for an observer along the jet axis, compared to an observer off the jet axis (but within the initial opening angle of the jet). The value of α after the jet break for $\nu > \nu_m$ is found to be slightly larger than p ($\alpha = 2.85$ for $p = 2.5$). A significant fraction of the jet break occurs due to the relativistic beaming effect (which does not depend on the hydrodynamics). In spite of the different hydrodynamic behavior the numerical simulations show a jet break at roughly the same time as the analytic estimates.

I. Generalizations: V. Angle-dependent jets and the universal structured jet model

In a realistic jet one can expect either a random or a regular angle-dependent structure. Here there are two dominant effects. As the ejecta slow down, their Lorentz factor decreases and an observer will detect radiation from an angular region of size Γ^{-1} (see Sec. IV.C). At the same time mixing within the ejecta will lead to an intrinsic averaging of the angular structure. Thus both effects lead to an averaging over the angular structure at later times.

Several authors (Lipunov *et al.*, 2001; Rossi *et al.*, 2002; Zhang and Mészáros, 2002) have suggested independently a different interpretation of the observed achromatic breaks in the afterglow light curves. This interpretation is based on a jet with a regular angular structure. According to this model all GRB's are produced by jets with a fixed angular structure and the break corresponds to the viewing angle. Lipunov *et al.* (2001) considered a "universal" jet with a three-step profile: a spherical one, a 20° one, and a 3° one. Rossi *et al.* (2002) and Zhang and Mészáros (2002) considered a special profile in which the energy per solid angle $\varepsilon(\theta)$ and the Lorentz factor $\Gamma(t=0, \theta)$ are

$$\varepsilon = \begin{cases} \varepsilon_c, & 0 \leq \theta \leq \theta_c \\ \varepsilon_c \left(\frac{\theta}{\theta_c} \right)^{-a}, & \theta_c \leq \theta \leq \theta_j \end{cases} \quad (106)$$

and

$$\Gamma = \begin{cases} \Gamma_c, & 0 \leq \theta \leq \theta_c \\ \Gamma_c \left(\frac{\theta}{\theta_c} \right)^{-b}, & b > 0 \quad \theta_c \leq \theta \leq \theta_j, \end{cases} \quad (107)$$

where θ_j is a maximal angle and the core angle θ_c is introduced to avoid a divergence at $\theta=0$. The parameters a and b here define the energy and Lorentz-factor angular dependence. This core angle can be taken to be smaller than any other angle of interest. The power-law

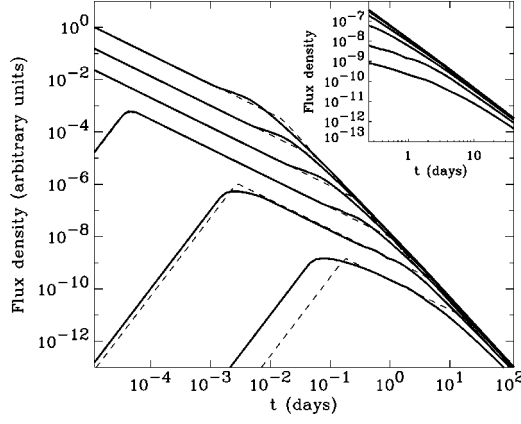


FIG. 28. Light curves of an inhomogeneous jet observed from different angles. From the top, $\theta_o = 0.5^\circ, 1^\circ, 2^\circ, 4^\circ, 8^\circ, 16^\circ$. The break time is related only to the observer angle: $t_b \propto \theta_o^2$. The dashed line is the on-axis light curve of a homogeneous jet with an opening angle $2\theta_o$ and an energy per unit solid angle $\varepsilon(\theta_o)$. The blowup is the time range between 4 h and 1 month, when most of the optical observations are performed. Comparing the solid and dashed lines for a fixed θ_o , it is apparent that one can hardly distinguish the two models by fitting the afterglow data. From Rossi *et al.*, 2002.

index of Γ , b , is not important for the dynamics of the fireball, and the computation of the light curve as long as $\Gamma(t=0, \theta) \equiv \Gamma_0(\theta) > \theta^{-1}$ and $\Gamma_0(\theta) \gg 1$.

To fit the constant-energy result [Frail *et al.* (2001) and Panaitescu and Kumar (2001)]. Rossi *et al.* (2002) considered a specific angular structure with $a=2$. Rossi *et al.* (2002) approximated the evolution assuming that at every angle the matter behaves as if it is a part of a regular Blandford-McKee profile [with the local ε and $\Gamma(t, \theta)$] until $\Gamma(t, \theta) = \theta^{-1}$. Then the matter begins to expand sideways. The resulting light curve is calculated by averaging the detected light resulting from the different angles. They find that an observer at an angle θ_o will detect a break in the light curve that appears around the time that $\Gamma(t, \theta_o) = \theta_o^{-1}$ (see Fig. 28). A simple explanation of break is the following: As the evolution proceeds and the Lorentz factor decreases, an observer will detect emission from larger and larger angular regions. Initially the additional higher energy at small angles $\theta < \theta_o$ compensates over the lower energies at larger angles $\theta > \theta_o$. Hence the observer detects a roughly constant energy per solid angle, and the resulting light curve is comparable to the regular pre-jet-break light curve. This goes on until $\Gamma^{-1}(0) = \theta_o$. After this stage a further increase in the viewing angle Γ^{-1} will result in a decrease in the energy per unit solid angle within the viewing cone and will lead to a break in the light curve.

This interpretation of the breaks in the light curves of standard structured jets in terms of the viewing angles implies a different understanding of total energy within GRB jets and of the rate of GRB's. The total energy in this model is also a constant, but now it is larger as it is the integral of Eq. (106) over all viewing angles. The distribution of GRB luminosities, which is interpreted in

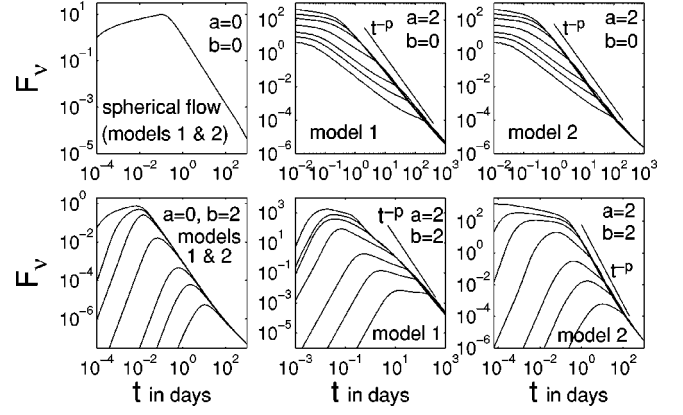


FIG. 29. Light curves for structured jets (initially $\varepsilon \propto \theta^{-a}$ and $\Gamma \propto \theta^{-b}$), for models 1 and 2, in the optical ($\nu = 5 \times 10^{14}$ Hz), for a jet core angle $\theta_c = 0.02$, viewing angles $\theta_{\text{obs}} = 0.01, 0.03, 0.05, 0.1, 0.2, 0.3, 0.5$, $p = 2.5$, $\varepsilon_c = \varepsilon_B = 0.1$, $n = 1 \text{ cm}^{-3}$, $\Gamma_0 = 10^3$, and ϵ_0 was chosen so that the total energy of the jet would be 10^{52} ergs (GK). A power law of t^{-p} is added in some of the panels for comparison. From Granot, Kumar, and Piran, 2003.

the uniform jet picture as a distribution of jet opening angles, is interpreted here as a distribution of viewing angles. As such this distribution is fixed by geometrical reasoning with $P(\theta_o)d\theta_o \propto \sin \theta_o d\theta_o$ (up to the maximal observing angle θ_j). This leads to an implied isotropic energy distribution of

$$P[\ln(E_{\text{iso}})] \propto E_{\text{iso}}^{-1}. \quad (108)$$

Guetta, Piran, and Waxman (2004) and Nakar, Granot, and Guetta (2004) find that these two distributions are somewhat inconsistent with current observations. However, the present data, which suffer from numerous observational biases, are insufficient to reach a definite conclusion.

In order to better estimate the effect of the hydrodynamics on the light curves of a structured jet Granot and Kumar (2003) and Granot, Kumar, and Piran (2003) considered two simple models for the hydrodynamics. In the first (model 1) there is no mixing among matter moving at different angles, i.e., $\varepsilon(\theta, t) = \varepsilon(\theta, t_0)$. In the second (model 2) ε is a function of time and it is averaged over the region to which a sound wave can propagate (this simulates the maximal lateral energy transfer that is consistent with causality). They considered various energy and Lorentz factors profiles and calculated the resulting light curves (see Fig. 29).

Granot and Kumar (2002) found that the light curves of models 1 and 2 were rather similar in spite of the different physical assumptions. This suggests that the widening of the viewing angle has a more dominant effect than the physical averaging. For models with a constant energy and a variable Lorentz factor $[(a, b) = (0, 2)]$ the light curve initially rises and there is no jet break, which is quite different from observations for most afterglows. For $(a, b) = (2, 2), (2, 0)$ they found a jet break at t_j when $\Gamma(\theta_o) \sim \theta_o^{-1}$. For $(a, b) = (2, 2)$ the value, α_1 , of the temporal decay slope at $t < t_j$ increased with θ_o ,

while $\alpha_2 = \alpha(t > t_j)$ decreased with θ_o . This effect was more prominent in model 1, and appeared to a lesser extent in model 2. This suggests that $\delta\alpha = \alpha_1 - \alpha_2$ should increase with t_j , which is not supported by observations. For $(a, b) = (2, 0)$, there is a flattening of the light curve just before the jet break (also noticed by Rossi *et al.*, 2002), for $\theta_o > 3\theta_c$. Again, this effect is larger in model 1 than in model 2, and again this flattening is not seen in the observed data.

Clearly a full solution for an angle-dependent jet requires full numerical simulations. Kumar and Granot (2003) present a simple 1D model for the hydrodynamics that is obtained by assuming axial symmetry and integrating over the radial profile of the flow, thus considerably reducing the computation time. The light curves that they find resemble those of models 1 and 2 above, indicating that these crude approximations are useful. Furthermore they find relatively little change in $\varepsilon(\theta)$ within the first few days, suggesting that model 1 is an especially useful approximation for the jet dynamics at early times, while model 2 provides a better approximation at late times.

J. Afterglow polarization—A tool that distinguishes between the different jet models

Synchrotron emission from a jet (in which the spherical symmetry is broken) would naturally produce polarized emission (Ghisellini and Lazzati, 1999; Gruzinov, 1999; Sari, 1999b). Moreover, the level and direction of the polarization are expected to vary with time and to give observational clues as to the geometrical structure of the emitting jet and our observing angle with respect to it.

The key feature in the determination of polarization during the afterglow is the varying Lorentz factor and (after a jet break) varying jet width. This changes the overall geometry (see Fig. 18) and hence the observer sees different geometries (Sari, 1999b; Hurley, Sari, and Djorgovski, 2002). Initially, the relativistic beaming angle $1/\Gamma$ is narrower than the physical size of the jet θ_o , and the observer sees a full ring; therefore the radial polarization averages out (the first frame, with $\Gamma\theta_o = 4$ of the left plot in Fig. 30). As the flow decelerates, the relativistic beaming angle $1/\Gamma$ becomes comparable to θ_o and only a fraction of the ring is visible; net polarization is then observed. Assuming, for simplicity, that the magnetic field is along the shock, then the synchrotron polarization will be radially outwards. Due to the radial direction of the polarization from each fluid element, the total polarization is maximal when a quarter ($\Gamma\theta_o = 2$ in Fig. 30) or when three quarters ($\Gamma\theta_o = 1$ in Fig. 30) of the ring are missing (or radiate less efficiently) and vanishes for a full and a half ring. The polarization, when more than half of the ring is missing, is perpendicular to the polarization direction when less than half of it is missing.

At late stages the jet expands sideways, and since the offset of the observer from the physical center of the jet is constant, spherical symmetry is regained. The vanish-

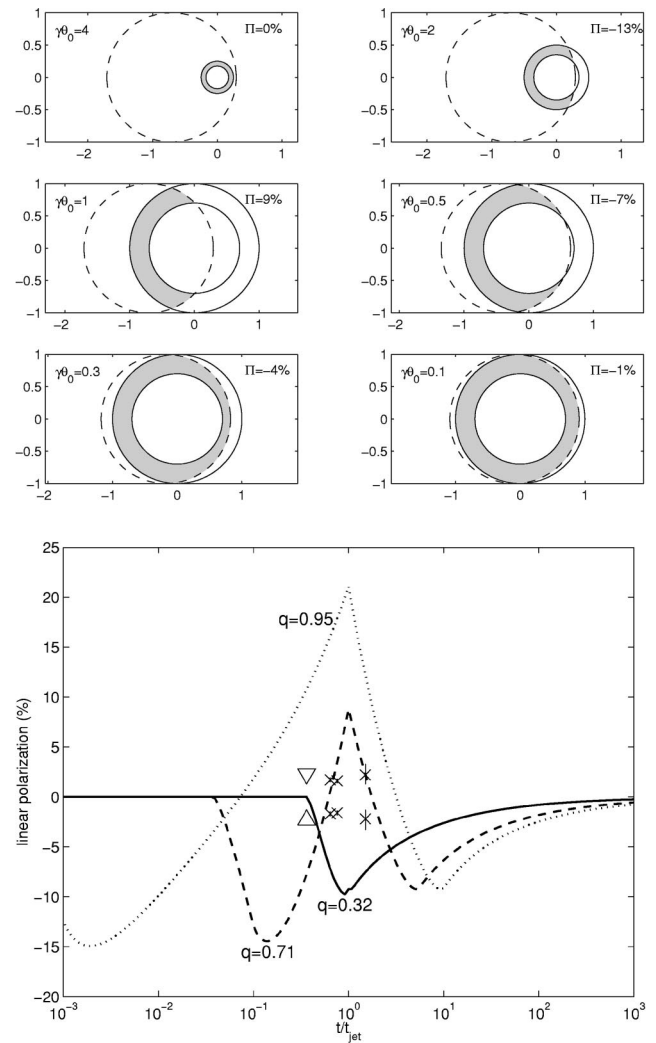


FIG. 30. Polarization geometry from observer's viewpoint: (a) Shape of the emitting region: dashed line, physical extent of the jet; solid lines, viewable region $1/\gamma$. The observed radiation arises from the gray-shaded region. In each frame, the percentage of polarization is given at the top right and the initial size of the jet relative to $1/\gamma$ is given on the left. The frames are scaled so that the size of the jet is unity. (b) Observed and theoretical polarization light curves for three possible offsets of the observer relative to the jet axis. Observational data for GRB 990510 is marked by crosses (\times), assuming $t_{\text{jet}} = 1.2$ days. The upper limit for GRB 990123 is given by a triangle, assuming $t_{\text{jet}} = 2.1$ days. From Sari, 1999b.

ing and reoccurrence of significant parts of the ring results in a unique prediction: there should be three peaks of polarization, with the polarization position angle during the central peak rotated by 90° with respect to the other two peaks. When the observer is very close to the center, more than half of the ring is always observed, and therefore only a single direction of polarization is expected. A few possible polarization light curves are presented in Fig. 30.

The predicted polarization from a structured jet is drastically different from that from a uniform jet, providing an excellent comparison between the two models

(Rossi *et al.*, 2004). Within the structured jet model the polarization arises due to the gradient in the emissivity. This gradient has a clear orientation. The emissivity is maximal at the center of the jet and decreases monotonically outwards. The polarization will be maximal when the variation in the emissivity within the emitting beam is maximal. This happens around the jet break when $\theta_{obs} \sim \Gamma^{-1}$ and the observed beam just reaches the center. The polarization expected in this case is around 20% (Rossi *et al.*, 2004) and it is slightly larger than the maximal polarization from a uniform jet. As the direction of the gradient is always the same (relative to a given observer) there should be no jumps in the direction of polarization.

According to the patchy shell model (Kumar and Piran, 2002a) the jet can include variable emitting hot spots. This could lead to a fluctuation in the light curve (as hot spots enter the observed beam) and also to corresponding fluctuations in the polarization (Granot, 2003; Nakar and Oren, 2003). There is a clear prediction (Nakar and Oren, 2003; Nakar, Piran, and Granot, 2003) that if the fluctuations are angular and have a typical angular scale θ_f then the first bump in the light curve should appear when $\Gamma^{-1} \sim \theta_f$ (the whole hot spot will be within the observed beam). The later bumps in the light curve should decrease in amplitude (due to statistical fluctuations). Nakar and Oren (2004) show analytically and numerically that the jumps in the polarization direction should be random, sharp, and accompanied by jumps in the amount of polarization.

K. Orphan afterglows

Orphan afterglows were predicted as a natural product of GRB jets. The realization that GRB's are collimated with rather narrow opening angles, while the following afterglow could be observed over a wider angular range, led immediately to the search for orphan afterglows, that is, afterglows not associated with observed prompt GRB emission. While the GRB and the early afterglow are collimated to within the original opening angle θ_j , the afterglow can be observed after the jet break, from a viewing angle of Γ^{-1} . The Lorentz factor Γ is a rapidly decreasing function of time. This means that an observer at $\theta_{obs} > \theta_j$ could not see the burst but could detect an afterglow once $\Gamma^{-1} = \theta_{obs}$. As the typical emission frequency and the flux decrease with time, while the jet opening angle θ increases, this implies that observers at larger viewing angles would detect weaker and softer afterglows. X-ray orphan afterglows can be observed several hours or at most a few days after a x-ray burst (depending of course on the sensitivity of the detector). Optical afterglows (brighter than 25th mag) can be detected in the *R* band for a week from small ($\sim 10^\circ$) angles away from the GRB jet axis. On the other hand, at very late times, after the Newtonian break, radio afterglows could be detected by observers at all viewing angles.

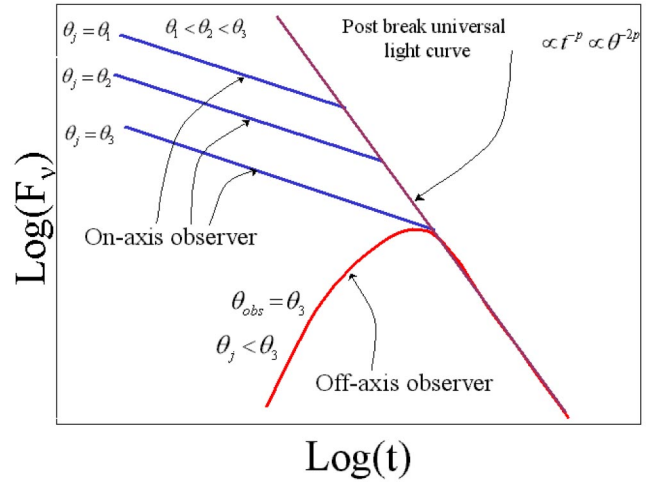


FIG. 31. (Color in online edition) Schematic afterglow light curve. While the bursts differ before the jet break (due to different opening angles), the light curves coincide after the break when the energy per unit solid angle is a constant.

The search for orphan afterglows is an observational challenge. One has to search for a 10^{-12} ergs/sec/cm² signal in the x-ray region of the spectrum, a 23rd or higher magnitude signal in the optical, or a mJy signal in the radio (at GHz) transients. Unlike afterglow searches that are triggered by a well-located GRB, there is no information on where to search for an orphan afterglow and confusion with other transients is rather easy. So far there have been no detections of any orphan afterglows at any wavelength.

Rhoads (1997) was the first to suggest that observations of orphan afterglows would enable us to estimate the opening angles and the true rate of GRB's. Dalal *et al.* (2002) pointed out that as the post-jet-break afterglow light curves decay quickly, most orphan afterglows will be dim and hence undetectable. They commented that if the maximal observing angle, θ_{max} , of an orphan afterglow were a constant factor times θ_j , the ratio of observed orphan afterglows R_{orph}^{obs} to that of GRB's R_{GRB}^{obs} would not tell us much about the opening angles of GRB's or their true rate, $R_{GRB}^{true} \equiv f_b R_{GRB}^{obs}$. However, as we shall see below, this assumption is inconsistent with the constant energy of GRB's, which suggests that all GRB's would be detected up to a fixed angle independent of their jet opening angle.

1. Optical orphan afterglow

An optical orphan afterglow is emitted at a stage when the outflow is still relativistic. The observation that GRB's have a roughly constant total energy (Frail *et al.*, 2001; Panaitescu and Kumar, 2001; Piran *et al.*, 2001) and that the observed variability in the apparent luminosity arises mostly from variation in the jet opening angles leads to a remarkable result: The post-jet-break afterglow light curve is universal (Granot *et al.*, 2002). Figure 31 depicts this universal light curve. This implies that for a given redshift z and a given limiting magnitude m

there will be a fixed $\theta_{\max}(z, m)$ (independent of θ_j , for $\theta_j < \theta_{\max}$) from within which orphan afterglow can be detected.

This universal post-jet-break light curve can be estimated from the observations (Totani and Panaitescu, 2002) or alternatively from first principles (Nakar *et al.*, 2002). An observer at $\theta_{\text{obs}} > \theta_j$ will (practically) observe the afterglow emission only at t_θ when $\Gamma = \theta_{\text{obs}}^{-1}$. Using Eq. (104) and the fact that $\Gamma \propto t^{-1/2}$ after the jet break [Eq. (106)] one can estimate the time t_θ when an emission from a jet would be detected at θ_{obs} :

$$t_\theta = A(\theta_{\text{obs}}/\theta_j)^2 t_{\text{jet}}, \quad (109)$$

where A is a factor of order unity, and t_{jet} is the time of the jet break [given by Eq. (104)]. The flux at this time is estimated by substitution of this value into the post-jet-break light curve (see Nakar *et al.*, 2002 for details):

$$F(\theta_{\text{obs}}) = F_0 f(z) \theta_{\text{obs}}^{-2p}, \quad (110)$$

where F_0 is a constant, $f(z) = (1+z)^{1+\beta} D_{L28}^{-2}$ includes all the cosmological effects, and D_{L28} is the luminosity distance in units of 10^{28} cm. One notices here a very strong dependence on θ_{obs} . The peak flux drops quickly when the observer moves away from the axis. Note also that this maximal flux is independent of the opening angle of the jet, θ_j . The observations of afterglows with a clear jet break—GRB 990510 (Harrison *et al.*, 1999; Stanek *et al.*, 1999) and GRB 000926 (Harrison *et al.*, 2001)—can be used to calibrate F_0 .

Now, using Eq. (110), one can estimate $\theta_{\max}(z, m)$ and more generally the time $t_{\text{obs}}(z, \theta, m)$ that a burst at a redshift z could be seen from an angle θ above a limiting magnitude m :

$$t_{\text{obs}}(z, \theta, m) \approx \frac{A t_{\text{jet}}}{\theta_j^2} (\theta_{\max}^2 - \theta_{\text{obs}}^2). \quad (111)$$

One can then proceed and integrate over the cosmological distribution of bursts (assuming that this follows the star formation rate) and obtain an estimate of the number of orphan afterglows that would appear in a single snapshot of a given survey with a limiting sensitivity F_{lim} :

$$N_{\text{orph}} = \int_0^\infty \frac{n(z)}{(1+z)} \frac{dV(z)}{dz} dz \times \int_{\theta_j}^{\theta_{\max}(z, m)} t_{\text{obs}}(z, \theta, m) \theta d\theta \propto (F_0/F_{\text{lim}})^{2/p}, \quad (112)$$

where $n(z)$ is the rate of GRB's per unit volume and unit proper time and $dV(z)$ is the differential volume element at redshift z . Note that modifications of this simple model may be worked out with more refined models of the jet propagation (Granot *et al.*, 2002; Nakar *et al.*, 2002).

The results of the intergration of Eq. (112) are depicted in Fig. 32. Clearly the rate of a single detection

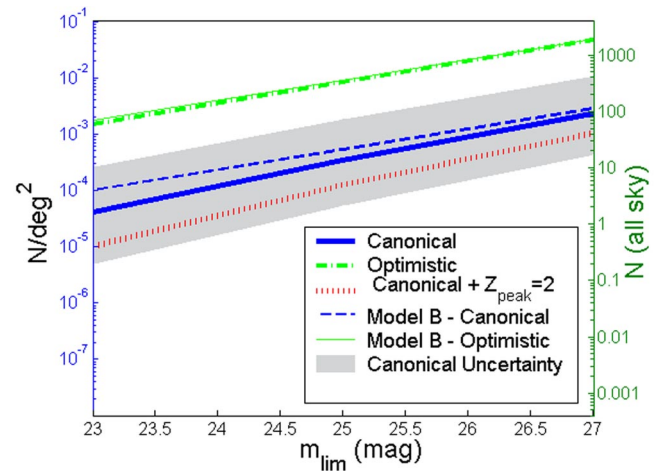


FIG. 32. (Color in online edition) Orphan afterglows: Left vertical scale, number of observed orphan afterglows per square degree; right vertical scale, in the entire sky, in a single exposure, as a function of the limiting magnitude for detection. Thick lines are for model A with three different sets of parameters: solid line, our “canonical” normalization $F_0 = 0.003 \mu\text{Jy}$, $z_{\text{peak}} = 1$, $\theta_j = 0.1$. The gray area around this line corresponds to an uncertainty by a factor of 5 in this normalization. Dashed-dotted line, our most optimistic model with a relatively small $\theta_j = 0.05$ and a large $F_0 = 0.015 \mu\text{Jy}$. Dotted line, the same as our “canonical” model, except for $z_{\text{peak}} = 2$. Thin lines are for model B: solid line, our “optimistic” parameters; dashed line is for our “canonical” parameters. Both models are similar for the “optimistic” parameters, but model B predicts slightly more orphan afterglows than model A for the “canonical” parameters. From Nakar *et al.*, 2002.

with a given limiting magnitude increases with a larger magnitude. However, one should ask what would be the optimal strategy for a given observational facility: short and shallow exposures that cover a larger solid angle or long and deep ones over a smaller area. The exposure time that is required in order to reach a given limiting flux, F_{lim} , is proportional to F_{lim}^{-2} . Dividing the number density of observed orphan afterglows (shown in Fig. 32) by this time factor results in the rate per square degree per hour of observational facility. This rate increases for a shallow survey that covers a large area. This result can be understood as follows. Multiplying Eq. (112) by F_{lim}^2 shows that the rate per square degree per hour of observational facility $\propto F_{\text{lim}}^{2-2/p}$. For $p > 1$ the exponent is positive and a shallow survey is preferred. The limiting magnitude should not however, be, lower than ~ 23 rd, as in this case more transients from on-axis GRB's will be discovered than orphan afterglows.

Using these estimates Nakar *et al.* (2002) find that with their most optimistic parameters 15 orphan afterglows should be recorded in the Sloan Digital Sky Survey (SDSS) (which covers 10^4 square degrees at 23rd magnitude) and 35 transients should be recorded in a dedicated 2m-class telescope operating fulltime for a year in an orphan afterglow search. Totani and Panaitescu (2002) find a somewhat higher rate (a factor ~ 10 above the optimistic rate). About 15% of the transients could

be discovered with a second exposure of the same area provided that it follows after 3, 4, and 8 days for $m_{lim} = 23, 25$, and 27 . This estimate does not tackle the challenging problem of identifying the afterglows within the collected data. Rhoads (2001) suggested identifying afterglow candidates by comparing the multicolor SDSS data to an afterglow template. One orphan afterglow candidate was indeed identified using this technique (Vanden Berk *et al.*, 2002). However, it turned out that it was a variable active galactic nucleus (Gal-Yam *et al.*, 2003). This event will give some idea of the remarkable observational challenge involved in this project.

2. Radio orphan afterglow

After the Newtonian transition the afterglow is expanding spherically. The velocities are at most mildly relativistic, so there are no relativistic beaming effects and the afterglow will be observed from all viewing angles. This implies that observations of the rate of orphan GRB afterglows at this stage could give a direct measure of the beaming factor of GRB's. Upper limits on the rate of orphan afterglows would provide a limit on the beaming of GRB's (Perna and Loeb, 1998). However, as I discuss shortly, somewhat surprisingly, upper limits on the rate of orphan radio afterglow (no detection of orphan radio afterglow) provide a lower (and not upper) limit on GRB beaming (Levinson *et al.*, 2002).

Frail, Waxman, and Kulkarni (2000) estimate the radio emission at this stage using the Sedov-Taylor solution for the hydrodynamics (see Sec. VII.D). They find that the radio emission at GHz will be around 1 mJy at the time of the Newtonian transition (typically three months after the burst) and it will decrease as $t^{-3(p-1)/2+3/5}$ [see Eq. (97)]. Using this limit one can estimate the rate of observed orphan radio afterglows within a given limiting flux. The beaming factor f_b^{-1} arises in two places in this calculation. First, the overall rate of GRB's, $R_{GRB}^{true} \equiv f_b R_{GRB}^{obs}$, increases with f_b . Second, the total energy is proportional to f_b^{-1} ; hence the flux will decrease when f_b increases. The first factor implies that the rate of orphan radio afterglows will increase as f_b . To estimate the effect of the second factor, Levinson *et al.* (2002) used the fact that (for a fixed observed energy) the time that a radio afterglow is above a given flux is proportional to $E^{10/9}$ in units of the NR transition time, which itself is proportional to $E^{1/3}$. Overall this is proportional to $E^{13/9}$ and hence to $f_b^{-13/9}$. To obtain the overall effect of f_b Levinson *et al.* (2002) integrated over the redshift distribution and obtained the total number of orphan radio afterglows as a function of f_b . For a simple limit of a shallow survey (which is applicable to current surveys) typical distances are rather "small," i.e., less than 1 Gpc, and cosmological corrections can be neglected. In this case it is straightforward to carry out the integration analytically and obtain the number of radio orphan afterglows in the sky at any given moment (Levinson *et al.*, 2002):

$$N_R \approx 10^4 f_b^{5/6} (R/0.5) \left(\frac{f_{vmin}}{5 \text{ mJy}} \right)^{-3/2} \left(\frac{\epsilon_e}{0.3} \right)^{3/2} \left(\frac{\epsilon_B}{0.03} \right)^{9/8} n_{-1}^{19/24} \times E_{iso,54}^{11/6} \nu_9^{-3/4} (t_i/3t_{NR})^{-7/20}, \quad (113)$$

where R is the observed rate of GRB's per Gpc³ per year, and t_i is the time in which the radio afterglow becomes isotropic.

Levinson *et al.* (2002) searched the FIRST and NVSS surveys for pointlike radio transients with flux densities greater than 6 mJy. They found nine orphan candidates. However, they argued that the possibility that most of these candidates were radio-loud AGN's could be ruled out without further observations. This analysis sets an upper limit for the all sky number of radio orphans, which corresponds to a lower limit $f_b^{-1} > 10$ on the beaming factor. Rejection of all candidates found in this search would imply $f_b^{-1} > 40$ (Guetta, Piran, and Waxman, 2004).

L. Generalizations: VI. Additional physical processes

With the development of the theory of GRB afterglow it was realized that several additional physical ingredients might influence the observed afterglow light emission. In this section I shall review two such processes: (i) preacceleration of the surrounding matter by the prompt γ -ray emission and (ii) decay of neutrons within the outflow.

1. Preacceleration

The surrounding regular ISM or even stellar wind is optically thin to initial γ -ray pulse. Still the interaction of the pulse and the surrounding matter may not be trivial. Thompson and Madau (2000) pointed out that a small fraction of γ -ray radiation will be Compton scattered on the surrounding electrons. The backscattered photons could now interact with the outward-going γ -ray flux and produce pairs. The pairs will increase the rate of backscattering, and this could lead to an instability. When sufficient numbers of pairs are produced the surrounding matter will feel a significant drag by the γ -ray flux and it will be accelerated outwards (Madau and Thompson, 2000). This preacceleration of the ambient medium could have several implications for the early afterglow (Mészáros *et al.*, 2001; Beloborodov, 2002b).

The key issue is that, while the optical depth of the surrounding medium (as "seen" in the γ -ray photons) is very small, the mean free path of an ambient electron within the γ -ray photons is large (at small enough radius) and each electron scatters many photons. While the medium absorbs only a small fraction of the prompt γ -ray energy, the effect of this energy can be significant. Beloborodov (2002a) characterizes the interaction of the

γ -ray radiation front with the surrounding medium by a dimensionless parameter:¹⁴

$$\eta = \frac{\sigma_T E_{iso}}{4\pi R^2 m_e c^2} = 6.5 E_{52} R_{16}^{-2}, \quad (114)$$

the energy that a single electron scatters relative to its rest mass. Beloborodov (2002b) calculates the Lorentz factor of the ambient medium and the number of pairs per initial electron as functions of η , where $\eta_{load} = 20-30$, depending on the spectrum of the gamma rays, $\eta_{acc} = 5\eta_{load} = 100-150$, and $f_{acc} = [\exp(\eta_{acc}/\eta_{load}) + \exp(-\eta_{acc}/\eta_{load})]/2 = 74$.

If $\eta < \eta_{load} \approx 20-30$, depending on the spectrum of the gamma rays, the medium remains static and e^\pm free. When the front has $\eta > \eta_{load}$, a runaway e^\pm loading occurs. The number of loaded pairs depends exponentially on η as long as $\eta < \eta_{acc} = 5\eta_{load} = 100-150$. The medium is accelerated if $\eta > \eta_{acc}$. η_{acc} is around 100 because the electrons are coupled to the ambient ions, and on the other hand the loaded e^\pm increase the number of scatters per ion. At $\eta = \eta_{gap} \approx 3 \times 10^3$, the matter is accelerated to a Lorentz factor $\Gamma_{ambient}$ that exceeds the Lorentz factor of the ejecta. It implies that the radiation front pushes the medium away from the ejecta and opens a gap.

As the GRB radiation front expands, the energy flux and hence η decreases $\propto R^{-2}$. η passes through η_{gap} , η_{acc} , and η_{load} at R_{gap} , R_{acc} , and R_{load} , respectively. These three characteristic radii define four stages:

- (i) $R < R_{gap} \approx R_{acc}/3$: The ejecta move in a cavity produced by the radiation front with $\Gamma_{ambient} > \Gamma_{ejecta}$.
- (ii) $R_{gap} < R < R_{acc} \approx 3 \times 10^{15} \text{ cm } E_{52}^{1/2}$: The ejecta sweep the e^\pm -rich medium, which has been preaccelerated to $1 \leq \Gamma_{ambient} < \Gamma_{ejecta}$.
- (iii) $R_{acc} < R < R_{load} \approx 2.3 R_{acc}$. The ejecta sweep the “static” medium ($\Gamma_{ambient} \approx 1$), which is still dominated by loaded e^\pm .
- (iv) $R > R_{load}$. The ejecta sweep the static pair-free medium.

This influence of the γ rays on the surrounding matter may modify the standard picture of interaction of external shocks with the surrounding medium (see Sec. VI.C.1. This depends mostly on the relation between R_{ext} and $R_{gap} \approx 10^{15} E_{52}^{1/2} \text{ cm}$. If $R_{ext} > R_{gap}$ this effect will not be important. However, if $R_{ext} < R_{gap}$ then effective deceleration will begin only at R_{gap} . At $R < R_{gap}$ the ejecta move freely in a cavity cleared by the radiation front, and only at $R = R_{gap}$ does the blast wave gently begin to sweep the preaccelerated medium with a small relative Lorentz factor. With increasing $R > R_{gap}$, $\Gamma_{ambient}$ falls off quickly, and it approaches $\Gamma_{ambient} = 1$ at $R = R_{acc} \approx 3 R_{gap}$ as $\Gamma_{ambient} = (R/R_{acc})^{-6}$. Thus, after a delay, the ejecta suddenly “learn” that there is a substantial

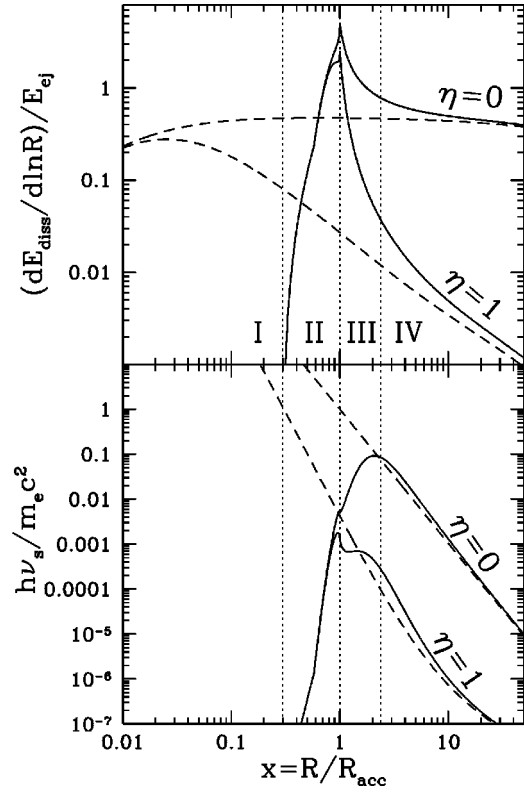


FIG. 33. Afterglow from a GRB ejecta decelerating in a wind of a Wolf-Rayet progenitor, where $\dot{M} = 2 \times 10^{-5} M_\odot \text{ yr}^{-1}$ and $w = 10^3 \text{ km s}^{-1}$. The burst has an isotropic energy $E = 10^{53} \text{ ergs}$ and a thin ejecta shell with kinetic energy $E = 10^{53} \text{ ergs}$ and a Lorentz factor $\Gamma = 200$: dashed curves, prediction of the standard model that neglects the impact of the radiation front; solid curves, actual behavior. Two extreme cases are displayed in the figure: $\eta = 0$ (adiabatic blast wave) and $\eta = 1$ (radiative blast wave). Four zones are marked: I— $R < R_{gap}$ (the gap is opened); II— $R_{gap} < R < R_{acc}$ (the gap is closed and the ejecta sweeps the relativistically preaccelerated e^\pm -loaded ambient medium); III— $R_{acc} < R < R_{load}$ (e^\pm -loaded ambient medium with $\Gamma_{ambient} \approx 1$); and IV— $R > R_{load}$ (pair-free ambient medium with $\Gamma_{ambient} \approx 1$). The radii are measured in units of $R_{acc} \approx 10^{16} \text{ cm}$. Top panel, dissipation rate. Bottom panel, synchrotron peak frequency (assuming $\epsilon_B = 0.1$) in units of $m_e c^2/h$. From Beloborodov, 2002a.

amount of ambient material on its way. This resembles a collision with a wall and results in a sharp pulse (see Fig. 33).

While R_{gap} does not depend on the external density, R_{ext} does [see Eq. (65)]. The condition $R_{ext} < R_{gap}$ implies

$$E_{52}^{1/6} n_1^{1/3} \Gamma_{100}^{2/3} > 0.02. \quad (115)$$

Thus it requires a dense external medium and large initial Lorentz factor. Otherwise R_{gap} is too large and the deceleration takes place after the gap is closed. Hence the conditions for preacceleration will generally occur if the burst takes place in a dense circumburst region, such as a Wolf-Rayet progenitor (Beloborodov, 2002b). Kumar and Panatescu (2004) elaborate on this model and find that the observational limits by the Livermore Optical Transient Imaging System (LOTIS) and ROTSE

¹⁴Note that Beloborodov (2002a) uses the notation χ for this parameter.

on prompt emission from various bursts limit the ambient ISM density (within 10^{16} cm to less than 10^3 rm cm $^{-3}$). Similarly they find that in case of a wind the progenitors' mass loss to the wind's velocity ratio is below $10^{-6}M_{\odot}/yr/(10^3$ km/sec).

2. Neutron decoupling and decay

Derishev *et al.* (1999, 2001) pointed out that neutrons that are included initially in the fireball will change its dynamics and modify the standard afterglow evolution. While the protons slow down due to the interaction with the surrounding matter, the neutrons will coast freely after they decouple with Γ_n , which equals the Lorentz factor while decoupling took place.

At

$$R_{decay} \approx 0.3 \times 10^{16} \text{ cm}(\Gamma_n/100) \quad (116)$$

the neutrons decay. A new baryonic shell forms ahead of the original fireball shell, with energy comparable to the initial energy of the protons' shell (this depends, of course, on the initial ratio of neutrons to protons). At this stage the neutron front, which is not slowed down like the rest of the fireball, is at a distance

$$\Delta R = R(1/2\Gamma^2 - 1/2\Gamma_n^2), \quad (117)$$

from the fireball front, where Γ is the current Lorentz factor of the fireball.

Once more the situation depends on whether R_{decay} is smaller or larger than R_{ext} , the original deceleration radius. If $R_{decay} < R_{ext}$,

$$E_5 2^{1/3} n_1^{-1/3} \Gamma_{100}^{-2/3} (\Gamma_n/300)^{-1} < 0.06, \quad (118)$$

the decaying neutron products will mix with the original protons and will not influence the evolution significantly (apart from adding their energy to the adiabatic fireball energy). Otherwise, the situation depends on Γ_n , the Lorentz factor at decoupling.

Pruet and Dalal (2002) consider a situation in which the neutrons decouple with a low Γ_n . In this case one will get a delayed shock scenario in which the neutron decay products will eventually catch up with the decelerating protons (when their Lorentz factor is of order Γ_n). Along the same line of thought Dalal *et al.* (2002) suggest that a large neutron component, which may exist within the initial fireball material, may help to eliminate the baryon load problem (Shemi and Piran, 1990).

Beloboradov (2003) considers a situation in which $\Gamma_n \approx \Gamma_0$, the initial Lorentz factor of the protons. In this case the decaying neutrons' products will be ahead of the shell of the protons. The decaying products will interact with the surrounding matter and will begin to slow down. There will be a triple interaction between the two shells and the surrounding ambient medium (resembling to some extent the preacceleration scenario described earlier). This will take place at radii of a few times R_{decay} and at an observed time of a few $\times R_{decay}/2c\Gamma^2 \approx$ a few seconds/ $(\Gamma_n/300)$, i.e., extremely early. This will produce brightening when the fronts pass R_{decay} .

Neutrons could also influence the behavior of the relativistic flow during the prompt (internal shocks) phase. Specifically, inelastic collisions between differentially streaming protons and neutrons could produce pions and eventually ν_{μ} of 10 GeV as well as ν_e of 5 GeV (Bahcall and Mészáros, 2000; Mészáros and Rees, 2000). These neutrinos fluxes could produce ~ 7 events/year in km 3 neutrino detectors. GeV photons would also be produced but it is unlikely that they could be detected.

VIII. ADDITIONAL EMISSION FROM GRB'S

A. TeV γ rays

Hurley (1994) reported on detection of 18-GeV photons from GRB 940217. Milagrito, a TeV detector, discovered a possible TeV signal coincident with GRB 970417 (Atkins *et al.*, 2000). González *et al.* (2003) discovered a high-energy tail that extended up to 200 MeV from GRB 941017.

A natural source for high-energy γ rays is the synchrotron self Compton (SSC) component produced by inverse Compton scattering from the burst itself or from the afterglow¹⁵ (Mészáros and Rees, 1994; Meszaros *et al.*, 1994). The photon energy from this process should be γ_e^2 higher than the synchrotron photons. Typical random Lorentz factors of electrons, γ_e , within internal shocks are of order a thousand (in the fluid's rest frame). This implies that if the observed γ -ray emission were produced by synchrotron radiation in internal shocks, then the inverse Compton emission would produce a second peak around a few hundred GeV. This would be the analog of the high-energy component observed in blazars. Note that emission above ~ 10 – 100 GeV might be self-absorbed by pair production within the source (Papathanassiou and Meszaros, 1996; Pilla and Loeb, 1998; Guetta and Granot, 2003b).

The SSC component would be even higher from the early afterglow. The synchrotron emission from the forward shock is expected to be around 10 keV (if the observed early afterglow is indeed produced by external shocks). With a Lorentz factor of a typical electron around 10^5 the expected SSC component should be around 100 TeV. Finally the reverse-shock emission is expected to produce 100-eV photons (Sari and Piran, 1996c). With a typical electron Lorentz factor of a few thousand, this should correspond to SSC photons with typical energy of 100 MeV. Depending on the relevant Y parameter the fluxes of these high-energy components should be comparable to, or even larger than, the prompt GRB γ -ray fluxes. This emission should be simultaneous with the GRB emission. It is also possible

¹⁵In the synchrotron self Compton process, photons are created by synchrotron radiation from relativistic electrons moving in a magnetic field. These photons are then scattered and pushed to higher energies by the same electrons responsible for their creation.

that the forward-shock electrons could cause inverse Compton scattering of reverse-shock photons. It is likely that this is the cause of the high-energy emission seen in GRB 941017 (Pe'er and Waxman, 2004; Piran *et al.*, 2004).

Other mechanisms can produce high-energy emission as well. Vietri (1997) suggested that as GRB's can accelerate protons up to 10^{20} eV (see Sec. VIII.C below), these protons could emit 0.01 of the GRB energy as high-energy γ rays with energies up to 300 GeV. Botcher and Dermer (1998) considered the synchrotron spectrum resulting from high-energy protons and leptons produced in a cascade initiated by photo-pion production. They predicted a significant flux of 10-MeV–100-GeV photons.

While the high-energy photon flux could be significant, these photons might not be detectable on Earth. A high-energy photon flux above 1 TeV would be attenuated significantly due to pair production of such high-energy photons with the intergalactic near-infrared flux (Gould and Schröder, 1967). Dai and Lu (2002) suggested that secondary emission produced via these interactions (upscattering of the cosmic-microwave background by the produced pairs) would still point towards the initial direction and hence might be detectable as a delayed GeV emission. However, even a tiny intergalactic magnetic field ($>10^{-22}$ G) would be sufficient to deflect the electrons and dilute these signals (Guetta and Granot, 2003b).

B. Neutrinos

Neutrinos can be produced in several regions within GRB sources. First, some models, like the collapsar model or the neutron star merger model, predict ample ($\sim 10^{53}$ ergs) production of low-energy (MeV) neutrinos. However, no existing or planned detector could see these from cosmological distances. Furthermore, this signal would be swamped by the much more frequent supernova neutrino signals, which would typically appear closer.

However, GRB's could be detectable sources of high-energy neutrinos, with energies ranging from 10^{14} to 10^{17} eV. These neutrinos would be produced by internal or external shocks of the GRB process itself and hence would be independent of the nature of the progenitor.

To understand the process of neutrino emission, recall that neutrinos are “best” produced in nature following pion production in proton-photon or proton-proton collisions. The proton-photon process requires that the photon energy be around the Δ resonance in the proton's energy frame: that is, at ~ 200 MeV. The resulting pion decays, emitting neutrinos with a typical energy of ~ 50 MeV in the proton's rest frame. If the proton is moving relativistically, with a Lorentz factor γ_p within the laboratory frame, the required photon energy in the laboratory frame is smaller by a factor of γ_p , and the resulting neutrino energy is larger by a factor of γ_p . De-

pending on the surrounding environment, very energetic pions may lose some of their energy before decaying, producing a “cooling break” in the neutrino spectrum. In this case the resulting neutrino energy would be lower than this naive upper limit.

Protons within GRB's are accelerated up to 10^{20} eV (Vietri, 1995; Waxman, 1995). The relevant Lorentz factors of these protons range from Γ up to 10^{11} (at the very-high-energy tail of the proton distribution). Thus we expect neutrinos up to 10^{19} eV provided that there is a sufficient flux of photons at the relevant energies, so that pions can be produced and there are no energy losses to the pions.

Paczynski and Xu (1994) and Waxman and Bahcall (1997) calculated the flux of very-high-energy neutrinos from internal shocks. They found that a significant flux of $\sim 10^{14}$ -eV neutrinos could be produced by interaction of accelerated internal-shock protons with GRB photons. Guetta *et al.* (2001b) estimated that on average each GRB produces a flux of $\sim 10^{-9}$ GeV/cm²/sec/sr, corresponding to 0.01 events in a km³ detector. Calculations of specific fluxes from individual bursts (which factor in the observed γ -ray spectrum) were performed by Guetta, Hooper, *et al.* (2004). Waxman and Bahcall (2000) suggested that protons accelerated by the reverse shock (which arises at the beginning of the afterglow) would interact with the optical-uv flux of the afterglow and produce 10^{18} -eV neutrinos.

Within the collapsar model Mészáros and Waxman (2001) and Razzaque *et al.* (2003b) suggested that, as the jet punches through the stellar shell, it could produce a flux of TeV neutrinos. Within the supranova model the internal-shock protons (Guetta and Granot, 2003a) or external-shock protons (Dermer and Atoyan, 2003) could also interact with external, pulsar-wind-bubble photons producing 10^{16} -eV neutrinos with a comparable detection rate to that obtained from interaction of the internal-shock protons with γ -ray photons. If the external magnetic field were sufficiently large (as in the pulsar wind bubble), external shocks could also accelerate protons to high energy (Vietri *et al.*, 2003). In this case the protons could interact with afterglow photons and produce neutrinos up to 10^{17} eV (Li *et al.*, 2002).

C. Cosmic rays and ultrahigh-energy cosmic rays

As early as in 1990 Shemi and Piran (1990) had noted that a fireball may produce cosmic rays. However, the flux of “low” energy (up to 10^{14} eV) that they considered was smaller by several orders of magnitude than the observed flux of cosmic rays accelerated in SNR's. Hence this component is not important.

Waxman (1995) and Vietri (1995) independently noted that protons could be accelerated up to 10^{20} eV within the relativistic shocks that take place in GRB's. Internal shocks or the reverse shock in GRB's are among the few locations in the universe where the shock acceleration condition [Eq. (14)] needed to accelerate protons up to 10^{20} eV, the Hillas criterion, can be satisfied. Moreover,

to within an order of magnitude the flux of γ rays reaching Earth from GRB's is comparable to the observed flux of ultrahigh-energy cosmic rays (UHECR's) (Waxman, 1995). Thus, if GRB's produce comparable energies in γ rays and in UHECR's, they could be the source of the highest-energy cosmic rays.

Greisen (1966) and Zatsepin and Kuzmin (1966) pointed out that the highest-energy cosmic rays (above $10^{19.5}$ eV) are attenuated as they propagate via the cosmic microwave background. This happens because at these high energies the protons can interact with the photons of the microwave background and produce pions. The typical mean free path of an ultrahigh-energy proton in the cosmic microwave background decreases rapidly with energy, and for a 10^{20} -eV proton it is only several tens of Mpc. Thus the observed UHECR's at energies above the Greisen-Zatsepin-Kuzmin energy ($\sim 10^{19.5}$ eV) must arrive from relatively nearby—on a cosmological scale—sources. However, there are no known steady-state sources within this distance (see, however, Farrar and Piran, 2000). GRB's as transient phenomena could be a “hidden” source of UHECR's. There would not be a direct association between GRB's and arrival of UHECR's as the latter are deflected by the intergalactic magnetic field. This leads to an angular deflection as well as a long-time delay. If GRB's are sources of UHECR's, then we would expect a break in the UHECR spectrum at the Greisen-Zatsepin-Kuzmin energy—below that we would detect UHECR's from the whole universe. Above that energy we would detect only “local” UHECR's from within the nearest several dozen Mpc. Bahcall and Waxman (2003) have suggested that recent observations imply that such a break has been seen. However, the observational situation is not as yet clear, and a final resolution will most likely require the Auger UHECR detector.

D. Gravitational radiation

Like GRB's, typical sources of gravitational radiation involve the formation of compact objects. Hence it is reasonable to expect that gravitational waves would accompany GRB's. This association is indirect: the gravitational waves are not directly related to the GRB's. Additionally, GRB's have their own, albeit weak, gravitational radiation pulse that arises during the acceleration of the jets to relativistic velocities. Unfortunately this signal is weak, and moreover it is perpendicular to the GRB signal.

To estimate the rates of observed gravitational radiation events associated with GRB's we use the rate of long GRB's. The nearest (long) GRB detected within a year would be at 1 Gpc. As GRB's are beamed, the number of gravitational events (which are not strongly beamed) is much larger than the number of observed GRB's, so the nearest (long) event would be much nearer, at $135\theta_{0.1}^2$ Mpc. Only if we were lucky would this nearest burst be directed towards us. Still, a GRB that is beamed away from us is expected to produce an “orphan” afterglow.

The rate of gravitational radiation events from short bursts is less certain because the rate per unit volume of short bursts is not well determined. Schmidt (2001b) estimated that the rate of short GRB's is smaller by a factor of 2 than the rate of long ones. In this case the distances mentioned above should be revised up by a factor of 1.25. However, if the rate of short GRB's were comparable to the rate of long ones, then the corresponding distances should be similar.

1. Gravitational radiation from neutron star mergers

Binary neutron star mergers are the “canonical” sources of gravitational radiation emission. LIGO and VIRGO both aim at detecting these sources. Specifically the goal of these detectors is to detect the characteristic “chirping” signals arising from the in-spiraling phase of these events. The possibility of detection of such signals has been extensively discussed (see, for example, Culter *et al.*, 1993). Such events could be detected up to a distance of several tens of Mpc with LIGO I and up to ~ 100 Mpc with LIGO II.

Comparing this with GRB rates we find that if, as some expect, neutron star mergers are associated with short GRB's and if the rate of short GRB's is indeed large, then we would have one event per year within the sensitivity of LIGO II and marginally detectable by LIGO I. However, this burst would be pointing away from us.

The detection of the chirping merger signal is based on fitting the gravitational radiation signal to precalculated templates. Kochanek and Piran (1993) suggested that the detection of a merger gravitational radiation signal would require a lower signal-to-noise ratio if this signal coincided with a GRB. This would increase somewhat the effective sensitivity of LIGO and VIRGO to such events. Finn *et al.* (1999) suggested using the association of GRB's with sources of gravitational waves in a statistical manner and proposed searching for enhanced gravitational radiation activity towards the direction of a GRB during the short periods when GRB's are detected. Given the huge distances of observed GRB's it is not clear if any of these techniques would be useful.

2. Gravitational radiation from collapsars

The collapsar model (Woosley, 1993; Paczynski, 1998; MacFadyen and Woosley, 1999) is based on the collapse of the core of a massive star to a black hole surrounded by a thick massive accretion disk. As far as gravitational radiation is concerned, this system is very similar to a regular supernova. Rotating gravitational collapse has been analyzed by Stark and Piran (1985). They find that the gravitational radiation emitted in a rotating collapse to a black hole is dominated by the black hole's lowest normal modes, with a typical frequency of $20c^3/GM$. The total energy emitted is

$$\Delta E_{GW} = \epsilon M c^2 = \min(1.4 \times 10^{-3} a^4, \epsilon_{\max}) M c^2, \quad (119)$$

where a is the dimensionless specific angular momentum and ϵ_{\max} is a maximal efficiency of the order a few

$\times 10^{-4}$. The expected amplitude of the gravitational radiation signal, h , would be of the order of $\sqrt{\epsilon GM/c^2 d}$, where d is the distance to the source. Even LIGO II will not be sensitive enough to detect such a signal from a distance of 1 Gpc or even from 100 Mpc.

3. Gravitational radiation from a supranova

According to the supranova model a GRB arises after a neutron star collapses to a black hole. This collapse takes place several weeks or months after the supernova that formed the neutron star (see Sec. IX.E). The expected gravitational waves signal from a supranova (Vietri and Stella, 1998) includes two components. First the signal from the initial supernova is similar to the gravitational waves from the collapsar model. However, here the first collapse (the supernova) takes place several weeks or months before the GRB. Thus there will not be any correlation between the gravitational waves emitted by the first collapse and the GRB. A second component may arise from the second collapse from the supramassive neutron star to a black hole. This signal should coincide with the GRB.

4. Gravitational radiation from a gamma-ray burst

The most efficient generation of gravitational radiation could take place during the acceleration phase of a GRB, in which the mass is accelerated to a Lorentz factor Γ . To estimate this emission I follow Weinberg's (1973) analysis of gravitational radiation emitted from a relativistic collision between two particles. Consider the following simple toy model: two particles at rest with a mass M that are accelerated instantly at $t=0$ to a Lorentz factor Γ and energy E . Conservation of energy requires that some (actually most) of the rest mass be converted to kinetic energy during the acceleration, and the rest mass of the accelerated particle is $m=E/\Gamma=M/\Gamma$. The energy emitted per unit frequency per unit solid angle in the direction at an angle α relative to $\vec{\beta}$ is

$$\frac{dE}{d\Omega d\omega} = \frac{GM^2\beta^2}{c\pi^2} \left[\frac{\Gamma^2(\beta^2 - \cos^2 \alpha)}{(1 - \beta^2 \cos^2 \alpha)^2} + \frac{\cos^2 \alpha}{\Gamma^2(1 - \beta^2 \cos^2 \alpha)^2} \right]. \quad (120)$$

The result is independent of the frequency, implying that the integral overall frequency will diverge. This non-physical divergence arises from the nonphysical assumption that the acceleration is instantaneous. In reality this acceleration takes place over a time δt , which is of order 0.01 sec. This would produce a cutoff $\omega_{\max} \sim 2\pi/\delta t$ above which Eq. (120) is not valid. The angular distribution found in Eq. (120) is disappointing. The electromagnetic emission from the ultrarelativistic source is beamed forwards into a small angle $1/\Gamma$, enhancing the emission in the forward direction by a large factor (Γ^2). The gravitational radiation from this relativistic ejecta is spread rather uniformly in almost all 4π steradians. Instead of beaming there is "antibeaming" with no radia-

tion at all emitted within the forward angle $1/\Gamma$ along the direction of the relativistic motion.

Integration of the energy flux over different directions yields

$$\frac{dE}{d\omega} = \frac{GM^2}{c\pi^2} \left[2\Gamma^2 + 1 + \frac{(1 - 4\Gamma^2)}{\Gamma^2\beta} \arctan(\beta) \right]. \quad (121)$$

As expected, the total energy emitted is proportional to $m^2\Gamma^2$. Further integration over frequencies up to the cutoff $2\pi/\delta t$ yields

$$E \approx \frac{2GM^2\Gamma^2}{c\pi\delta t}. \quad (122)$$

In reality the situation is much more complicated than the one presented here. First, the angular width of the emitted subunits within the flow is larger than $1/\Gamma$. The superposition of emission from different directions washes out the no-emission effect in the forward direction. Additionally according to the internal shocks model the acceleration of different subunits of the flow are accelerated independently. Emission from different internal shocks should be combined to get the actual emission. Both effects reduce the effective emission of gravitational radiation and make the above estimate an upper limit to the actual emission.

The gravitational signal is spread in all directions (apart from a narrow beam along the direction of the relativistic motion of the GRB). It ranges in frequency from 0 to $f_{\max} \approx 100$ Hz. The amplitude of the gravitational radiation signal at the maximal frequency $f_{\max} \approx 100$ Hz would be $h \approx (GM\Gamma^2/c^2 d)$. For typical values of $E = M\Gamma = 10^{51}$ ergs, $\delta t = 0.01$ sec, and a distance of 500 Mpc, $h \approx 0.5 \times 10^{-25}$, far below the sensitivity of planned gravitational radiation detectors. Even if a burst were ten times nearer, this "direct" gravitational radiation signal would still be undetectable.

Some specific models for the inner engines of GRB's predict additional amounts of energy. For example, van Putten (2001; van Putten and Levinson, 2001) suggests a model of a black hole/accretion torus in which a large fraction of the emitted energy of the black-hole/accretion-torus system escapes as gravitational radiation. The radiation arises due to instabilities within the torus that break down the axial symmetry. van Putten and Levinson estimate that as much as 10^{53} ergs would be emitted as gravitational radiation, which would have a characteristic signature corresponding to the normal mode of the black-hole/accretion-torus system with typical frequencies around a few hundred Hz, conveniently within the frequency range of LIGO/VIRGO. If correct, then GRB's are the most powerful burst sources of gravitational waves in the universe (van Putten, 2001).

IX. MODELS OF INNER ENGINES

The fireball model tells us how GRB's operate. However, it does not answer the most interesting astrophysical question: what produces them? Which astrophysical

process generates the energetic ultrarelativistic flows needed for the fireball model? Several observational clues help us answer these questions.

- **Energy:** The total energy involved is large, $\sim 10^{51}$ ergs, a significant fraction of the binding energy of a stellar compact object. The “inner engine” must be able to generate this energy and accelerate $\sim 10^{-5}M_{\odot}$ (or the equivalent in terms of Poynting flux) to relativistic velocities.
- **Collimation:** Most GRB’s are collimated with typical opening angles $1^{\circ} < \theta < 20^{\circ}$. The “inner engine” must be able to collimate the relativistic flow.
- **Long and short bursts:** The bursts are divided into two groups according to their overall duration. Long bursts with $T > 2$ sec and short ones with $T < 2$ sec. As the duration is determined by the inner engine, this may imply that there are two different inner engines.
- **Rates:** GRB’s take place once per 3×10^5 yr per galaxy. GRB’s are very rare, at about 1/3000 the rate of supernovae.
- **Time scales:** The variability time scale δt is at times as short as 1 msec. The overall duration (of long GRB’s) T is of the order of 50 sec. According to the internal shocks model these time scales are determined by the activity of the “inner engine.” $\delta t \sim 1$ msec suggests a compact object. $T \sim 50$ sec is much longer than the dynamical time scale, suggesting a prolonged activity.¹⁶ This requires two (or possibly three (Ramirez-Ruiz and Merloni, 2001; Nakar and Piran, 2002c) different time scales operating within the “inner engine.” This rules out any “explosive” model that releases the energy in a single explosion.

These clues, most specifically the last one, suggest that GRB’s arise due to accretion of a massive ($\sim 0.1M_{\odot}$) disk onto a compact object, most likely a newborn black hole. A compact object is required because of the short time scales. Accretion is needed to produce the two different time scales, and in particular the prolonged activity. A massive ($\sim 0.1M_{\odot}$) disk is required because of the energetics. Such a massive disk can form only simultaneously with the formation of the compact object. This leads to the conclusion that GRB’s accompany the formation of black holes. This model is supported by the observations of relativistic (but not as relativistic as in GRB’s) jets in active galactic nuclei, which are powered by accretion onto black holes.

An important alternative to accretion is Usov’s model (Usov, 1992, 1994), in which the relativistic flow is mostly Poynting flux driven by the magnetic and rotational energies of a newborn rapidly rotating neutron star.

A. Black-hole accretion

Several scenarios could lead to a black-hole/massive-accretion-disk system. This could include mergers such as neutron-star/neutron-star binaries (Eichler *et al.*, 1989; Narayan *et al.*, 1992), neutron-star/black-hole binaries (Paczynski, 1991), white-dwarf/black-hole binaries (Fryer *et al.*, 1999), black-hole/He-star binaries (Fryer and Woosley, 1998), and models based on “failed supernovae” or “collapsars” (Woosley, 1993; Paczynski, 1998; MacFadyen and Woosley, 1999). Narayan *et al.* (2001) have recently shown that accretion theory suggests that from all the above scenarios only collapsars could produce long bursts and only neutron-star/neutron star (or neutron-star/black-hole) mergers could produce short bursts. The basic idea is that the duration of the accretion depends on the size of the disk. So short bursts must be produced by small disks, and those are naturally produced in mergers. On the other hand, long bursts require large disks, which are inefficient. One can overcome this difficulty with a small disk that is fed continuously. In this case the efficiency can be large and the duration long. This happens naturally within the collapsar model.

B. The pulsar model

Several “inner engine” models involve pulsarlike activity of the inner engine which is directly connected to a Poynting-flux-decimated relativistic flow (in a contrast to a baryonic-flux-dominated flow). Energy considerations require an extremely large magnetic field of the order of 10^{15} G within such sources.

Usov (1992) suggested that GRB’s arise during the formation of rapidly rotating highly magnetized neutron stars. Such objects could form by the gravitational collapse of accreting white dwarfs with anomalously high magnetic fields in binaries, as in magnetic cataclysmic binaries. The rapidly rotating and strongly magnetized neutron stars would lose their rotational kinetic energy on a time scale of seconds or less in a pulsarlike mechanism. The energy available in this case would be the rotational and magnetic energies of the neutron star, which are of the order of a few $\times 10^{51}$ ergs for a neutron star rotating near breakup. The rotation of the magnetic field would create a strong electric field and an electron-positron plasma that is initially optically thick and in thermodynamic quasiequilibrium. Additionally a very strong magnetic field would form. The pulsar would produce a relativistic Poynting-flux-dominated flow.

Usov (1994) and Thompson (1994) discussed a scheme in which the energy is dissipated from the magnetic field to the plasma and then via plasma instability to the observed γ rays outside the γ -ray photosphere, which is at around 10^{13} cm. At this distance the MHD approximation of the pulsar wind breaks down, and intense electromagnetic waves are generated. The particles are accelerated by these electromagnetic waves to Lorentz factors of 10^6 and produce the nonthermal spectrum. Smolsky and Usov (1996, 2000) and Drenkhahn and

¹⁶The ratio $\delta t/T \ll 1$ for short bursts as well (Nakar and Piran, 2002b).

Spruit (2002; Spruit *et al.*, 2001) discuss various aspects of the conversion of the Poynting flux energy to γ rays, but these issues are more related to the nature of the emitting regions and only indirectly to the nature of the inner engine.

Usov's model is based on a rotating highly magnetized neutron star, and from this point of view it indeed resembles to a large extent a regular pulsar. Other authors consider pulsarlike activities in other contexts. Katz (1997), for example, considers a black-hole/thick-disk model in which the electromagnetic process turns rotational energy into particle energy in a pulsarlike mechanism. Mészáros and Rees (1997b) discuss a related idea on the formation of a Poynting-flux-dominated flow within a black-hole accretion-disk model.

C. Rotating black holes and the Blandford-Znajek mechanism

It is possible and even likely that the process of energy extraction involves the Blandford-Znajek mechanism (Blandford and Znajek, 1977) in which the black-hole/torus system is engulfed in a magnetic field and the rotational energy of the black hole is extracted via this magnetic field. The exploration of the Blandford-Znajek mechanism involves relativistic MHD considerations which are beyond the scope of this review. I refer the reader to several recent extensive reviews on this subject (e.g., Lee *et al.*, 2000).

D. The collapsar model

The evidence for the association of (long) GRB's with supernovae (see Bloom, Kulkarni, and Djorgovski, 2002 and Sec. II.C.4) provides strong support for the collapsar model. Woosley (1993) proposed that GRB arise from the collapse of a single Wolf-Rayet star endowed with fast rotation a ("failed" type-Ib supernova). Paczynski (1998) pointed out that there is tentative evidence that the GRB's 970228, 970508, and 970828 were close to star-forming regions and this suggests that GRB's are linked to cataclysmic deaths of massive stars. MacFadyen and Woosley (1999) began a series of calculations (Aloy *et al.*, 2000; MacFadyen *et al.*, 2001; Zhang *et al.*, 2003) of a relativistic jet propagation through the stellar envelope of the collapsing star, which is the most important ingredient unique to this model (other features like the accretion process onto the black hole, the corresponding particle acceleration, and to some extent the collimation process are common to other models). The collimation of a jet by the stellar mantle was shown analytically to occur by Mészáros and Rees (2001). Zhang *et al.* (2003) numerically confirmed and extended the basic features of this collimation process.

According to the collapsar model the iron core of a rapidly rotating massive star, of mass $M > 30M_{\odot}$, collapses to a black hole, either directly or during the accretion phase that follows the core collapse. An accretion disk forms around this black hole and a funnel forms along the rotation axis, where the stellar material

has relatively little rotational support. The mass of the accretion disk is around $0.1M_{\odot}$. Accretion of this disk onto the black hole takes place within several dozen seconds and powers the GRB. Energy can be extracted via neutrino annihilation (MacFadyen and Woosley, 1999) or via the Blandford-Znajek mechanism. The energy deposited in the surrounding matter will preferably leak out along the rotation axis, producing jets with opening angles of $< 10^{\circ}$. If the jets are powerful enough they would penetrate the stellar envelope and produce the GRB.

Zhang *et al.* (2003) find that relativistic jets are collimated by their passage through the stellar mantle. Starting with an initial half-angle of up to 20° , the jet emerges with half-angles that, though variable with time, are around 5° . The jet becomes very hot in this phase and has only a moderate Lorentz factor, modulated by mixing, and a very large internal energy (more than 80% of the total energy). As the jet escapes, conversion of the remaining internal energy into kinetic energy gives terminal Lorentz factors along the axis of ~ 150 (depending, of course, on the initial conditions considered). Because of the large ratio of internal to kinetic energy in both the jet and its cocoon, the opening angle of the final jet is significantly greater than at breakout. A small amount of material emerges at large angles, but with a Lorentz factor still sufficiently large to make a weak GRB. When the jet breaks out from the star it may produce a thermal precursor (seen in several GRB's; Paczynski, 1998; Ramirez-Ruiz *et al.*, 2002; Razzaque *et al.*, 2003a). Instabilities in the accretion process, or in the passage of the jet through the stellar envelope (Aloy *et al.*, 2002; Zhang *et al.*, 2003) can produce the required variability in the Lorentz factor that is needed to produce internal shocks.

The processes of core collapse, accretion along the polar column (which is essential in order to create the funnel), and jet propagation through the stellar envelope together take ~ 10 sec (MacFadyen and Woosley, 1999). The duration of the accretion onto the black hole is expected to take several dozen seconds. These arguments imply that collapsars are expected to produce long GRB's (see, however, Zhang *et al.*, 2003 for a suggestion that the breakout of a relativistic jet and its collision with the stellar wind will produce a brief transient with properties similar to the class of "short hard" GRB's).

E. The supranova model

Vietri and Stella (1998) suggested that GRB's take place when a supermassive (or "supramassive" as they call it) neutron star (a neutron star that is above the maximal cold nonrotating neutron star mass) collapses to a black hole. The collapse might take place because the neutron star loses angular momentum via a pulsar wind and loses the extra support of the centrifugal force. Alternatively it might simply cool and become unstable if rotation alone is not enough to support it. The neutron star could also become overmassive and collapse if it slowly accretes matter from a surrounding accretion disk

(Vietri and Stella, 1999). In this latter case the time delay from the supernova could be very large and the supernova remnant will not play any role in the GRB or its afterglow.

The supranova model is a two-step event. First, there is a supernova, which may be more energetic than an average one, in which the supermassive neutron star forms. Then a few weeks or months later this neutron star collapses, producing the GRB. While both the supranova and the collapsar (or hypernova) events are associated with supernovae or a supernovalike event, the details of the models are very different. First, while in the collapsar model one expects a supernova bump on the afterglow light curve, such a bump is not expected in the supranova model unless the time delay is a few days. On the other hand, while it is not clear in the collapsar model how the Fe needed for the Fe x-ray lines reaches the implied large distances from the center, it is obvious in the supernova model, as the supernova shell was ejected to space several month before the GRB. As mentioned earlier (see Sec. II.C.4) the association of GRB 030329 with SN 2003dh (Hjorth *et al.*, 2003; Stanek *et al.*, 2003) is incompatible with the supranova model. Proponents of this model argue, however, that there might be a distribution of delay times between the first and second collapses.

The models are also very different in their physical content. First in the supranova model the GRB jet does not have to punch a hole through the stellar envelope. Instead the ejecta propagate in almost free space, polluted possibly by a pulsar wind (Königl and Granot, 2002; Guetta and Granot, 2003a). In both models, as in many other models, the GRB is powered by accretion of a massive accretion disk surrounding the newborn black hole. This accretion disk forms from the debris of the collapsing neutron star at the same time that the black hole is formed. Again, the time scale of the burst is determined by the accretion time of this disk. Narayan *et al.* (2001; see also Sec. IX.A) point out, however, that long-lived (50-sec) accretion disks must be large and hence extremely inefficient. This may pose a problem for this model.

Königl and Granot (2002), Guetta and Granot (2003a) and Inoue *et al.* (2003) considered the effects of a strong pulsar wind (that may exist after the supernova and before the second collapse) on this scenario. The pulsar wind can have several effects. First it would produce a denser, highly magnetized medium into which the GRB jet propagates. The strong magnetic field would be amplified by the afterglow shock. This resolves the problem of the source of the strong magnetic field needed for the synchrotron afterglow model. It could also explain the high-energy emission detected by EGRET in GRB 940217 (Hurley, 1994 and Sec. II.A.1) by inverse Compton scattering on the pulsar-wind-bubble photons. On the other hand, the density of this wind matter ($\sim 10^3 \text{ cm}^{-3}$) might be too high for the spherical model. Note, however, that this depends on the time delay as t^{-3} . The pulsar wind will not be spherical and one would expect it to form an elongated supernova shell cavity

within which it is bounded. If, as expected, the pulsar jet coincides with the GRB jet, then the relativistic ejecta will move along the elongated direction of this shell.

F. Merging neutron stars

Neutron-star binary mergers (Eichler *et al.*, 1989; Narayan *et al.*, 1992) or neutron-star black-hole binary mergers (Paczynski, 1991), hereafter called mergers, also produce a black-hole/accretion-disk system and are candidates for the inner engines of GRB's, specifically of short GRB's. These mergers take place because of the decay of the binary orbits due to gravitational radiation emission, as was beautifully demonstrated in the famous binary pulsar PSR 1913+16 (Taylor and Weisberg, 1982).

These mergers take place at a rate of $\approx 10^{-6}$ events per year per galaxy (Narayan *et al.*, 1991; Phinney 1991; van den Heuvel and Lorimer, 1996). This is the merger rate of binaries of the type of PSR 1913+16, whose lifetime is of the order of several 10^8 yr. Various population synthesis calculations suggest that there is another population of short-lived binaries (Tutukov and Yungelson, 1993, 1994; Belczynski *et al.*, 2002). These binaries form with very close orbits and hence with short lifetimes of the order of 10^5 yr. Even though the overall rate of such mergers could be comparable to those of the PSR 1913+16 type, one cannot expect to catch a binary in our galaxy in such a stage. Similarly unlike the long-lived mergers that may be kicked from their host galaxy within their long lifetime (Narayan *et al.*, 1992; Bulik *et al.*, 1999) this short-lived population remains within the galaxy when they merge (Belczynski *et al.*, 2002).

Earlier simulations of mergers focused on the gravitational radiation from this system. Davies *et al.* (1994) began a series of numerical simulations of neutron star mergers that focused on GRB-related aspects (Rosswog *et al.*, 1999, 2000; Ayal *et al.*, 2001; Rosswog and Davies, 2002). Using a smoothed particle hydrodynamics scheme they followed neutron star mergers under different assumptions (Newtonian with *ad hoc* addition of gravitational radiation backreaction or post-Newtonian), with different equations of state (adiabatic or realistic) and with different initial spin axis and mass ratios and different estimates of the effects of neutrino cooling. A parallel set of simulations was carried out by Janka and Ruffert (1996; Ruffert *et al.*, 1995; Ruffert and Janka, 1998, 1999, 2001), who used particle-in-cell methods. Both kinds of simulations yielded comparable results. The merger results in a black-hole/accretion-disk system. The mass of the accretion disk is of order $0.1 M_{\odot}$ and it depends of course somewhat on the orientation of the spins and the relative masses of the two neutron stars.

A merger releases $\sim 5 \times 10^{53}$ ergs but most of this energy is in the form of low-energy neutrinos and gravitational waves. Still there is enough energy available to power a GRB, though is not clear how the GRB is produced. A central question is, of course, how does a merger generate the relativistic wind required to power a GRB. Eichler *et al.* (1989) suggested that about one-

thousandth of these neutrinos annihilate and produce pairs that in turn produce gamma rays via $\nu\bar{\nu} \rightarrow e^+e^- \rightarrow \gamma\gamma$. This idea was criticized on several grounds by different authors, the main problem being that it does not produce enough energy. For example, Jaroszynski (1996) pointed out that a large fraction of the neutrinos would be swallowed by the black hole that forms. An alternative source of energy within the merger model is the accretion power of a disk that forms around the black hole. This brings us back to the canonical black-hole/accretion-disk scenario.

Some of these questions have been addressed in the very-high-resolution simulations of Rosswog and his collaborators (Rosswog and Davies, 2002; Rosswog and Ramirez-Ruiz, 2003; Rosswog, Ramirez-Ruiz, and Davies, 2003). These simulations employ smoothed particle hydrodynamics with 10^6 particles and a realistic equation of state (Rosswog and Davies, 2002). Rosswog *et al.* find that sufficient energy to explain the observed isotropic luminosities of short GRB's can be obtained by neutrino annihilation if the outflow is beamed to narrow angles with a solid angle of less than 1%. They argue that the energetic neutrino-driven wind that accompanies the merger could collimate the neutrino annihilation driven jet.

Alternatively, if the central object does not collapse immediately to a black hole, it might produce a large scale dynamo leading to magnetic fields of 10^{17} G, which in turn could power a GRB. Interestingly, the spin down time would be around 0.2 sec, just the typical time scale for a short GRB.

X. OPEN QUESTIONS AND FUTURE PROSPECTS

I believe that overall we have a basic understanding of the GRB phenomenon. As usual some aspects are understood better than others.

There is a very good understanding of the afterglow. Here there are numerous observations over a wide range of wavelengths with which the theory can be confronted. The overall picture, of a decelerating relativistic flow and of synchrotron emission fits the data to a large extent (see Wijers and Galama, 1999; Panaitescu and Kumar, 2001) and there are many other fits of the observations to the model. We have already learned that the “cow is not spherical,” i.e., that the relativistic flow is collimated. New observations, like those of GRB 021004 and GRB 030329, pose at times new puzzles and suggest that the basic simple picture has to be refined. It seems, however, that the answers are within the scope of the current model, with possibilities such as refreshed shocks, patchy shells, and variable external densities. All these phenomena are fairly reasonable in a realistic environment. Within the afterglow, I believe that the x-ray lines pose the greatest current puzzle, in terms of their energy requirements and other implications for the source (see Lazzati, 2002). Another interesting open question is what distinguishes GHOST's and optical transient GRB's—an environmental mechanism (extinc-

tion or “improper conditions within the circumburst matter”) or an intrinsic mechanism?

The main observational challenges concerning the afterglow are the determination of whether short GRB's have afterglow. A wealth of information on long GRB's is available from the information on hosts, environments, and redshifts determined from the afterglow observations. All these are missing for short GRB's. If short GRB's do not have afterglows, then an immediate theoretical question is why? Is it possible that they are produced in a very different environment from that of long ones (such as outside their original galaxies) in a region with no circumburst matter suitable for producing the afterglow? At the moment the observational situation is not clear. The forthcoming Swift satellite may resolve this mystery.

Another important observational question involves the search for orphan afterglows (at either radio or optical wavelengths). Their detection will establish the collimated jets picture. But even firm upper limits will set independent limits on the rates of GRB's. However, as mentioned in Sec. VII.K this is a very challenging observational task, with important implications for the nature of the jets—are GRB jets standard with a fixed angular structure (Lipunov *et al.*, 2001; Rossi *et al.* 2002; Zhang and Mészáros, 2002)? This question is related both to the overall energetics and to the rate of GRB's.

Another interesting challenge will be the resolution of the afterglow image (see Granot *et al.*, 1999b). This may be possible in radio for a nearby burst, and the afterglow of GRB 030329 provides an excellent candidate for that. Polarization measures could pave the way for an understanding of the collimation geometry and for a confirmation of the synchrotron emission process.

As we move backwards in time towards the burst, we encounter the very early afterglow and the optical flash that coincides with the burst itself. Here great progress was made with recent observations triggered by HETE II (e.g., the almost complete light curve of GRB 021004; Fox, Yost, *et al.*, 2003). SWIFT may contribute a lot to this issue. These observations could shed a light on issues like the role of preacceleration and neutrons that are unclear as yet. Here, I stress the importance of early and continuous radio observations, which could determine whether there are refreshed shocks during the early afterglow, which have a clear radio signature (Kumar and Piran, 2000b).

The understanding of the γ -ray-emitting regions is less clear. Here, within the internal shocks model there is a reasonable understanding of the temporal structure (in terms of the activity of the inner engine). However, it is not clear how the observed spectrum is produced and it seems that the simple synchrotron spectrum has to be modified (see Lloyd and Petrosian, 2000 and Medvedev, 2000 for ideas on such modifications). Another possibly related puzzle is the origin of the narrow E_p distributions (see, however, Daigne and Mochkovitch, 1998, 2003; Guetta *et al.*, 2001a). Another set of open questions concerns the origin of the intrinsic correlation between luminosity and the collimation angle (Frail *et al.*,

2001; Panaitescu and Kumar, 2001) discovered by Feniore and Ramirez-Ruiz (2000) or the lag-luminosity relation discovered by Norris *et al.* (2000). Similarly puzzling are the implied correlations between redshift and intrinsic luminosity (Lloyd-Ronning *et al.*, 2002) and between redshift and intrinsic hardness (Schmidt, 2001a; note that this latter correlation is essential in view of the narrow E_p distribution of GRB's). Here pairs (Ghisellini and Celotti, 1999) and inverse Compton scattering can play an important role. Theoretical basic-physics question that arise here (as well as in the theory of the afterglow) deal with the processes of collisionless shocks (see Medvedev, 2001; Nikto and Medvedev, 2001), particle acceleration (see Sec. V.B), and the generation of strong magnetic fields (see Medvedev and Loeb, 1999). Issues like relativistic turbulence and relativistic plasma instabilities might play an important role here (see Lyutikov and Blandford, 2002).

From an observational point of view, it will be a challenge to beat the statistical power of the BATSE data in terms of number of bursts. Somewhat surprisingly many questions are still open: the luminosity function of GRB's, the rate of GRB's as a function of redshift, and to what extent GRB's follow the star formation rate. Detectors with better spectral resolutions could shed some additional light on the spectrum. Another hope for new data, or at least for upper limits, arises from observational windows in higher γ -ray bands. On the low energy side it seems that there is a continuum between x-ray flashes and GRB's (Kippen *et al.*, 2002; Barraud *et al.*, 2003). This result still has to be fully understood in the context of the narrow E_p distribution.

Looking far into the future one can hope to observe neutrinos or gravitational radiation correlated to GRB's. Ultrahigh-energy neutrinos (fluxes of MeV neutrinos would be too weak to be detected from cosmological distances) could confirm that protons are accelerated to ultrahigh energies within GRB's. This in turn would prove (or disprove) the possible role of GRB's as sources of UHECR's. Gravitational radiation could give a direct clue as to the activity of the inner engine (see Sec. VIII.D.4) and identify, for example, merger events.

There is much observational evidence associating long GRB's with core-collapse supernovae. This gives a clear clue about the inner engine of long GRB's. There is no direct or indirect evidence on the progenitors of short GRB's. Even with this clue the situation is far from clear when we turn to the inner engine. Here most models assume some variant of a black-hole/torus system with various energy extraction mechanisms ranging from neutrino annihilation (which is less likely) to variants on the theme of electromagnetic extraction (magnetic turbulence within the accretion disk; the Blandford-Znajek mechanism, which involves a disk/black-hole/magnetic-field interaction; pulsarlike activity). Here there are open questions all around: What is the content of the ultrarelativistic flow—baryonic or Poynting flux? How is the flow accelerated and collimated? What determines the variability of the flow (required for internal shocks) and the different time scales? This part of the model

seems to be in a rather poor shape—but this is understandable as we do not have any direct observations of this inner engine. One hopeful sign is that there seems to be an emerging similarity between GRB's, galactic microquasars, and AGN's. All these systems accelerate collimated flows to relativistic velocities, and they all seem to involve accretion onto black holes. This similarity could lead to a common resolution of how inner engines operate in all those systems.

ACKNOWLEDGMENTS

I would like to thank J. Granot, D. Guetta, P. Kumar, E. Nakar, and R. Sari for many helpful discussions, and J. Bloom, J. Hjorth, P. Mészáros, E. Pian, K. Stanek, and P. Vreeswijk for comments on the manuscript. This research was supported by a grant from the U.S.-Israel Binational Science Foundation.

REFERENCES

- Achterberg, A., Y. A. Gallant, J. G. Kirk, and A. W. Guthmann, 2001, *Mon. Not. R. Astron. Soc.* **328**, 393.
- Akerlof, C., R. Balsano, S. Barthelemy, J. Bloch, P. Butterworth, D. Casperson, T. Cline, S. Fletcher, F. Frontera, G. Gisler, J. Heise, J. Hills, *et al.*, 1999, *Nature (London)* **398**, 400.
- Aloy, M.-A., J.-M. Ibáñez, J.-A. Miralles, and V. Urpin, 2002, *Astron. Astrophys.* **396**, 693.
- Aloy, M. A., E. Müller, J. M. Ibáñez, J. M. Martí, and A. MacFadyen, 2000, *Astrophys. J. Lett.* **531**, L119.
- Amati, L., F. Frontera, M. Tavani, J. J. M. in't Zand, A. Antonelli, E. Costa, M. Feroci, C. Guidorzi, J. Heise, N. Masetti, E. Montanari, L. Nicastro, *et al.*, 2002, *Astron. Astrophys.* **390**, 81.
- Amati, L., F. Frontera, M. Vietri, J. J. M. in't Zand, P. Soffitta, E. Costa, S. Del Sordo, E. Pian, L. Piro, L. A. Antonelli, D. D. Fiume, M. Feroci, *et al.*, 2000, *Science* **290**, 953.
- Antonelli, L. A., L. Piro, M. Vietri, E. Costa, P. Soffitta, M. Feroci, L. Amati, F. Frontera, E. Pian, J. J. M. i. Zand, J. Heise, E. Kuulkers, *et al.*, 2000, *Astrophys. J. Lett.* **545**, L39.
- Atkins, R., W. Benbow, D. Berley, M. L. Chen, D. G. Coyne, B. L. Dingus, D. E. Dorfan, R. W. Ellsworth, D. Evans, A. Falcone, L. Fleyscher, R. Fleyscher, *et al.*, 2000, *Astrophys. J. Lett.* **533**, L119.
- Atkins, R., W. Benbow, D. Berley, M. L. Chen, D. G. Coyne, B. L. Dingus, D. E. Dorfan, R. W. Ellsworth, D. Evans, A. Falcone, L. Fleyscher, R. Fleyscher, *et al.*, 2003, *Astrophys. J.* **583**, 824.
- Ayal, S., T. Piran, R. Oechslin, M. B. Davies, and S. Rosswog, 2001, *Astrophys. J.* **550**, 846.
- Bahcall, J. N., and P. Mészáros, 2000, *Phys. Rev. Lett.* **85**, 1362.
- Bahcall, J. N., and E. Waxman, 2003, *Phys. Lett. B* **556**, 1.
- Band, D., J. Matteson, L. Ford, B. Schaefer, D. Palmer, B. Teegarden, T. Cline, M. Briggs, W. Paciesas, G. Pendleton, G. Fishman, C. Kouveliotou, *et al.*, 1993, *Astrophys. J.* **413**, 281.
- Baring, M. G., and A. K. Harding, 1997, *Astrophys. J.* **491**, 663.
- Barraud, C., J.-F. Olive, J. P. Lestrade, J.-L. Atteia, K. Hurley, G. Ricker, D. Q. Lamb, N. Kawai, M. Boer, J.-P. Dezalay, G. Pizzichini, R. Vanderspek, *et al.*, 2003, *Astron. Astrophys.* **400**, 1021.
- Bednarz, J., and M. Ostrowski, 1998, *Phys. Rev. Lett.* **80**, 3911.

- Belczynski, K., T. Bulik, and V. Kalogera, 2002, *Astrophys. J. Lett.* **571**, L147.
- Beloborodov, A. M., 2000, *Astrophys. J. Lett.* **539**, L25.
- Beloborodov, A. M., 2002a, *astro-ph/0206423*.
- Beloborodov, A. M., 2002b, *Astrophys. J.* **565**, 808.
- Beloborodov, A. M., 2003, *Astrophys. J. Lett.* **585**, L19.
- Beloborodov, A. M., B. E. Stern, and R. Svensson, 2000, *Astrophys. J.* **535**, 158.
- Berger, E., S. R. Kulkarni, J. S. Bloom, P. A. Price, D. W. Fox, D. A. Frail, T. S. Axelrod, R. A. Chevalier, E. Colbert, E. Costa, S. G. Djorgovski, F. Frontera, *et al.*, 2002, *Astrophys. J.* **581**, 981.
- Berger, E., S. R. Kulkarni, and D. A. Frail, 2001, *Astrophys. J.* **560**, 652.
- Berger, E., S. R. Kulkarni, and D. A. Frail, 2003, *Astrophys. J.* **590**, 379.
- Bersier, D., B. McLeod, P. M. Garnavich, M. J. Holman, T. Grav, J. Quinn, J. Kaluzny, P. M. Challis, R. G. Bower, D. J. Wilman, J. S. Heyl, S. T. Holland, *et al.*, 2003, *Astrophys. J. Lett.* **583**, L63.
- Bhattacharya, D., 2001, *Bull. Astron. Soc. India* **29**, 107.
- Blandford, R., and D. Eichler, 1987, *Phys. Rep.* **154**, 1.
- Blandford, R. D., and C. F. McKee, 1976, *Phys. Fluids* **19**, 1130.
- Blandford, R. D., and C. F. McKee, 1977, *Mon. Not. R. Astron. Soc.* **180**, 343.
- Blandford, R. D., and R. L. Znajek, 1977, *Mon. Not. R. Astron. Soc.* **179**, 433.
- Bloom, J., 2003, in *Supernovae (10 years of SN1993J)* IAU Colloquium No. 192: Valencia, Spain, 2003, edited by J. M. Marcaide and K. W. Weiler (IAU, in press).
- Bloom, J. S., D. A. Frail, and S. R. Kulkarni, 2003, *Astrophys. J.* **594**, 674.
- Bloom, J. S., D. A. Frail, and R. Sari, 2001, *Astron. J.* **121**, 2879.
- Bloom, J. S., S. R. Kulkarni, and S. G. Djorgovski, 2002, *Astron. J.* **123**, 1111.
- Bloom, J. S., S. R. Kulkarni, S. G. Djorgovski, A. C. Eichlerberger, P. Cote, J. P. Blakeslee, S. C. Odewahn, F. A. Harrison, D. A. Frail, A. V. Filippenko, D. C. Leonard, A. G. Riess, *et al.*, 1999, *Nature (London)* **401**, 453.
- Bloom, J. S., S. R. Kulkarni, P. A. Price, D. Reichart, T. J. Galama, B. P. Schmidt, D. A. Frail, E. Berger, P. J. McCarthy, R. A. Chevalier, J. C. Wheeler, J. P. Halpern, *et al.*, 2002, *Astrophys. J. Lett.* **572**, L45.
- Boggs, S. E., and W. Coburn, 2003, *astro-ph/0310515*.
- Borgonovo, L., and F. Ryde, 2001, *Astrophys. J.* **548**, 770.
- Bottcher, M., and C. D. Dermer, 1998, *Astrophys. J. Lett.* **499**, L131.
- Brainerd, J. J., 1994, in *Gamma-Ray Bursts*, AIP Conference Proceedings No. 307, edited by G. J. Fishman, J. J. Brainerd, and K. Hurley (AIP, New York), p. 346.
- Bromm, V., and A. Loeb, 2002, *Astrophys. J.* **575**, 111.
- Bulik, T., K. Belczyński, and W. Zbijewski, 1999, *Mon. Not. R. Astron. Soc.* **309**, 629.
- Burenin, R. A., A. A. Vikhlinin, M. R. Gilfanov, O. V. Terekhov, A. Y. Tkachenko, S. Y. Sazonov, E. M. Churazov, R. A. Sunyaev, P. Goldoni, A. Claret, A. Goldwurm, J. Paul, *et al.*, 1999, *Astron. Astrophys.* **344**, L53.
- Butler, N., A. Dullighan, P. Ford, G. Monnelly, G. Ricker, R. Vanderspek, K. Hurley, and D. Lamb, 2002, *GRB Circular Network* 1415.
- Bykov, A. M., and P. Meszaros, 1996, *Astrophys. J. Lett.* **461**, L37.
- Cannizzo, J. K., N. Gehrels, and E. T. Vishniac, 2004, *Astrophys. J.* **601**, 380.
- Chevalier, R. A., and Z. Li, 1999, *Astrophys. J. Lett.* **520**, L29.
- Chevalier, R. A., and Z. Li, 2000, *Astrophys. J.* **536**, 195.
- Chornock, R., R. J. Foley, A. V. Filippenko, M. Papenkova, and W. D., 2003, *GRB Circular Network* 2131.
- Ciardi, B., and A. Loeb, 2000, *Astrophys. J.* **540**, 687.
- Cline, D. B., C. Matthey, and S. Otwinowski, 2003, *Astropart. Phys.* **18**, 531.
- Coburn, W., and S. E. Boggs, 2003, *Nature (London)* **423**, 415.
- Cohen, E., J. I. Katz, T. Piran, R. Sari, R. D. Preece, and D. L. Band, 1997, *Astrophys. J.* **488**, 330.
- Cohen, E., and T. Piran, 1995, *Astrophys. J. Lett.* **444**, L25.
- Cohen, E., and T. Piran, 1999, *Astrophys. J.* **518**, 346.
- Cohen, E., T. Piran, and R. Sari, 1998, *Astrophys. J.* **509**, 717.
- Connaughton, V., 2002, *Astrophys. J.* **567**, 1028.
- Costa, E., F. Frontera, J. Heise, M. Feroci, J. in't Zand, F. Fiore, M. N. Cinti, D. dal Fiume, L. Nicastro, M. Orlandini, E. Palazzi, M. Rapisarda, *et al.*, 1997, *Nature (London)* **387**, 783.
- Covino, S., G. Ghisellini, P. Saracco, G. Tagliaferri, F. Zerbi, S. di Serego, A. Cimatti, F. Fiore, G. L. Israel, L. Stella, N. Kawai, D. Lazzati, *et al.*, 2002, *GRB Circular Network* **1214**, 1.
- Covino, S., D. Lazzati, G. Ghisellini, P. Saracco, S. Campana, G. Chincarini, S. di Serego, A. Cimatti, L. Vanzi, L. Pasquini, F. Haardt, G. L. Israel, *et al.*, 1999, *Astron. Astrophys.* **348**, L1.
- Cutler, C., T. A. Apostolatos, L. Bildsten, L. S. Finn, E. E. Flanagan, D. Kennefick, D. M. Markovic, A. Ori, E. Poisson, and G. J. Sussman, 1993, *Phys. Rev. Lett.* **70**, 2984.
- Dai, Z. G., and K. S. Cheng, 2001, *Astrophys. J. Lett.* **558**, L109.
- Dai, Z. G., and T. Lu, 1999, *Astrophys. J. Lett.* **519**, L155.
- Dai, Z. G., and T. Lu, 2002, *Astrophys. J.* **580**, 1013.
- Daigne, F., and R. Mochkovitch, 1998, *Mon. Not. R. Astron. Soc.* **296**, 275.
- Daigne, F., and R. Mochkovitch, 2003 *Mon. Not. R. Astron. Soc.* **342**, 587.
- Dalal, N., K. Griest, and J. Pruet, 2002, *Astrophys. J.* **564**, 209.
- Dar, A., 2003, *astro-ph/0301389*.
- Dar, A., and A. De Rujula, 2003, *astro-ph/0308248*.
- Davies, M. B., W. Benz, T. Piran, and F. K. Thielemann, 1994, *Astrophys. J.* **431**, 742.
- de Jager, O. C., A. K. Harding, P. F. Michelson, H. I. Nel, P. L. Nolan, P. Sreekumar, and D. J. Thompson, 1996, *Astrophys. J.* **457**, 253.
- De Pasquale, M., L. Piro, R. Perna, E. Costa, M. Feroci, G. Gandolfi, J. J. M. in't Zand, L. Nicastro, L. Frontera, L. A. Antonelli, F. Fiore, and S. G. Stratta, 2003, *Astrophys. J.* **592**, 1018.
- Derishev, E. V., V. V. Kocharovsky, and V. V. Kocharovsky, 1999, *Astrophys. J.* **521**, 640.
- Derishev, E. V., V. V. Kocharovsky, and V. V. Kocharovsky, 2001, *Astron. Astrophys.* **372**, 1071.
- Dermer, C. D., and A. Atoyan, 2003, *Phys. Rev. Lett.* **91**, 071102.
- Dermer, C. D., and K. E. Mitman, 1999, *Astrophys. J. Lett.* **513**, L5.
- Dezalay, J. P., J. P. Lestrade, C. Barat, R. Talon, R. Sunyaev, O. Terekhov, and A. Kuznetsov, 1996, *Astrophys. J. Lett.* **471**, L27.
- Dingus, B. L., and J. R. Catelli, 1998, in *Abstracts of the 19th*

- Texas Symposium on Relativistic Astrophysics and Cosmology*, held in Paris, France, Dec. 1998, edited by J. Paul, T. Montmerle, and E. Aubourg (CEA, Saclay), p. 63.
- Djorgovski, S. G., D. A. Frail, S. R. Kulkarni, J. S. Bloom, S. C. Odewahn, and A. Diercks, 2001, *Astrophys. J.* **562**, 654.
- Djorgovski, S. G., S. R. Kulkarni, J. S. Bloom, D. A. Frail, F. A. Harrison, T. J. Galama, D. Reichart, S. M. Castro, D. Fox, R. Sari, E. Berger, P. Price, *et al.*, 2001, in *Gamma-ray Bursts in the Afterglow Era*, edited by E. Costa, F. Frontera, and J. Hjorth (Springer-Verlag, Berlin), p. 218.
- Drenkhahn, G., and H. C. Spruit, 2002, *Astron. Astrophys.* **391**, 1141.
- Eichler, D., M. Livio, T. Piran, and D. N. Schramm, 1989, *Nature (London)* **340**, 126.
- Elbaz, D., and C. J. Cesarsky, 2003, *Science* **300**, 270.
- Esin, A. A., and R. Blandford, 2000, *Astrophys. J. Lett.* **534**, L151.
- Farrar, G. R., and T. Piran, 2000, *Phys. Rev. Lett.* **84**, 3527.
- Fenimore, E. E., and J. S. Bloom, 1995, *Astrophys. J.* **453**, 25.
- Fenimore, E. E., R. I. Epstein, and C. Ho, 1993, *Astron. Astrophys., Suppl. Ser.* **97**, 59.
- Fenimore, E. E., J. J. M. in't Zand, J. P. Norris, J. T. Bonnell, and R. J. Nemiroff, 1995, *Astrophys. J. Lett.* **448**, L101.
- Fenimore, E. E., C. D. Madras, and S. Nayakshin, 1996, *Astrophys. J.* **473**, 998.
- Fenimore, E. E., and E. Ramirez-Ruiz, 2000, *astro-ph/0004176*.
- Fermi, E., 1949, *Phys. Rev.* **75**, 1169.
- Finn, L. S., S. D. Mohanty, and J. D. Romano, 1999, *Phys. Rev. D* **60**, 121101.
- Fishman, G. J., and C. A. Meegan, 1995, *Annu. Rev. Astron. Astrophys.* **33**, 415.
- Ford, L. A., D. L. Band, J. L. Matteson, M. S. Briggs, G. N. Pendleton, R. D. Preece, W. S. Paciesas, B. J. Teegarden, D. M. Palmer, B. E. Schaefer, T. L. Cline, G. J. Fishman, *et al.*, 1995, *Astrophys. J.* **439**, 307.
- Fox, D. B., D. L. Kaplan, B. Cenko, S. R. Kulkarni, and A. Nechita, 2003, *GRB Circular Network* 2323.
- Fox, D. W., S. Yost, S. R. Kulkarni, K. Torii, T. Kato, H. Yamaoka, M. Sako, F. A. Harrison, R. Sari, P. A. Price, E. Berger, A. M. Soderberg, *et al.*, 2003, *Nature (London)* **422**, 284.
- Frail, D. A., F. Bertoldi, G. H. Moriarty-Schieven, E. Berger, P. A. Price, J. S. Bloom, R. Sari, S. R. Kulkarni, C. L. Gerardy, D. E. Reichart, S. G. Djorgovski, T. J. Galama, *et al.*, 2002, *Astrophys. J.* **565**, 829.
- Frail, D. A., S. R. Kulkarni, S. R. Nicastro, M. Feroci, and G. B. Taylor, 1997, *Nature (London)* **389**, 261.
- Frail, D. A., S. R. Kulkarni, R. Sari, S. G. Djorgovski, J. S. Bloom, T. J. Galama, D. E. Reichart, E. Berger, F. A. Harrison, P. A. Price, S. A. Yost, A. Diercks, *et al.*, 2001, *Astrophys. J. Lett.* **562**, L55.
- Frail, D. A., S. R. Kulkarni, R. Sari, G. B. Taylor, D. S. Shepherd, J. S. Bloom, C. H. Young, L. Nicastro, and N. Masetti, 2000, *Astrophys. J.* **534**, 559.
- Frail, D. A., E. Waxman, and S. R. Kulkarni, 2000 *Astrophys. J.* **537**, 191.
- Frederiksen, J. T., C. B. Heddal, T. Haugboelle, and A. Nordlund, 2004, *Astrophys. J.* **608**, L13.
- Freedman, D. L., and E. Waxman, 2001, *Astrophys. J.* **547**, 922.
- Fruchter, A. S., E. Pian, S. E. Thorsett, L. E. Bergeron, R. A. González, M. Metzger, P. Goudfrooij, K. C. Sahu, H. Ferguson, M. Livio, M. Mutchler, L. Petro, *et al.*, 1999, *Astrophys. J.* **516**, 683.
- Fruchter, A. S., E. Pian, S. E. Thorsett, R. Gonzalez, K. C. Sahu, M. Mutchler, F. Frontera, T. J. Galama, P. J. Groot, R. Hook, C. Kouveliotou, M. Livio, *et al.*, 1998, in *Gamma-Ray Bursts*, AIP Conference Proceedings No. 428, edited by C. A. Meegan, T. M. Koshut, and R. D. Preece (AIP, Woodbury, NY), p. 509.
- Fryer, C. L., and S. E. Woosley, 1998, *Astrophys. J. Lett.* **502**, L9.
- Fryer, C. L., S. E. Woosley, M. Herant, and M. B. Davies, 1999, *Astrophys. J.* **520**, 650.
- Fynbo, J. P. U., P. Jakobsson, P. Möller, J. Hjorth, B. Thomsen, M. I. Andersen, A. S. Fruchter, J. Gorosabel, S. T. Holland, C. Ledoux, H. Pedersen, J. Rhoads, *et al.*, 2003, *Astron. Astrophys.* **406**, L63.
- Fynbo, J. P. U., B. L. Jensen, J. Gorosabel, J. Hjorth, H. Pedersen, P. Möller, T. Abbott, A. J. Castro-Tirado, D. Delgado, J. Greiner, A. Henden, A. Magazzù, *et al.*, 2001, *Astron. Astrophys.* **369**, 373.
- Fynbo, J. P. U., P. Möller, B. Thomsen, J. Hjorth, J. Gorosabel, M. I. Andersen, M. P. Egholm, S. Holland, B. L. Jensen, H. Pedersen, and M. Weidinger, 2002, *Astron. Astrophys.* **388**, 425.
- Gaisser, T. K., 1991, *Cosmic Rays and Particle Physics* (Cambridge University Press, Cambridge, England).
- Gal-Yam, A., E. O. Ofek, A. V. Filippenko, R. Chornock, and W. Li, 2002, *Publ. Astron. Soc. Pac.* **114**, 587.
- Galama, T., and R. Sari, 2002, in *Relativistic Flows in Astrophysics*, edited by A. W. Guthmann, M. Georganopoulos, A. Marcowith, and K. Manolakou (Springer, Berlin), p. 123.
- Galama, T. J., P. M. Vreeswijk, J. van Paradijs, C. Kouveliotou, T. Augusteijn, H. Bohnhardt, J. P. Brewer, V. Doublier, J.-F. Gonzalez, B. Leibundgut, C. Lidman, O. R. Hainaut, *et al.*, 1998, *Nature (London)* **395**, 670.
- Gallant, Y. A., and A. Achterberg, 1999a, *Mon. Not. R. Astron. Soc.* **305**, L6.
- Gallant, Y. A., and A. Achterberg, 1999b, *Mon. Not. R. Astron. Soc.* **305**, L6.
- Garnavich, P. M., K. Z. Stanek, L. Wyrzykowski, L. Infante, E. Bendek, D. Bersier, S. T. Holland, S. Jha, T. Matheson, R. P. Kirshner, K. Krisciunas, M. M. Phillips, *et al.*, 2003, *Astrophys. J.* **582**, 924.
- Ghisellini, G., and A. Celotti, 1999, *Astrophys. J. Lett.* **511**, L93.
- Ghisellini, G., A. Celotti, and D. Lazzati, 2000, *Mon. Not. R. Astron. Soc.* **313**, L1.
- Ghisellini, G., and D. Lazzati, 1999, *Mon. Not. R. Astron. Soc.* **309**, L7.
- Ghisellini, G., D. Lazzati, and S. Covino, 2001, in *Gamma-ray Bursts in the Afterglow Era*, edited by E. Costa, F. Frontera, and J. Hjorth (Springer, Berlin), p. 288.
- Ghisellini, G., D. Lazzati, E. Rossi, and M. J. Rees, 2002, *Astron. Astrophys.* **389**, L33.
- Giblin, T. W., J. van Paradijs, C. Kouveliotou, V. Connaughton, R. A. M. J. Wijers, M. S. Briggs, R. D. Preece, and G. J. Fishman, 1999, *Astrophys. J. Lett.* **524**, L47.
- González, M. M., B. L. Dingus, Y. Kaneko, R. D. Preece, C. D. Dermer, and M. S. Briggs, 2003, *Nature (London)* **424**, 749.
- Goodman, J., 1997, *New Astron.* **2**, 449.
- Gould, R. J., and G. P. Schröder, 1967, *Phys. Rev.* **155**, 1408.
- Granot, J., 2003, *Astrophys. J. Lett.* **596**, L17.
- Granot, J., and A. Königl, 2003, *Astrophys. J. Lett.* **594**, L83.
- Granot, J., and P. Kumar, 2003, *Astrophys. J.* **591**, 1086.
- Granot, J., P. Kumar, and T. Piran, 2003, *astro-ph/0301627*.

- Granot, J., M. Miller, T. Piran, W. M. Suen, and P. A. Hughes, 2001, in *Gamma-ray Bursts in the Afterglow Era*, edited by E. Costa, F. Frontera, and J. Hjorth (Springer-Verlag, Berlin), p. 312.
- Granot, J., E. Nakar, and T. Piran, 2003, *Nature (London)* **426**, 138.
- Granot, J., A. Panaitescu, P. Kumar, and S. E. Woosley, 2002, *Astrophys. J. Lett.* **570**, L61.
- Granot, J., T. Piran, and R. Sari, 1999a, *Astrophys. J.* **513**, 679.
- Granot, J., T. Piran, and R. Sari, 1999b, *Astron. Astrophys., Suppl. Ser.* **138**, 541.
- Granot, J., T. Piran, and R. Sari, 1999c, *Astrophys. J.* **527**, 236.
- Granot, J., T. Piran, and R. Sari, 2000, *Astrophys. J. Lett.* **534**, L163.
- Granot, J., and R. Sari, 2002, *Astrophys. J.* **568**, 820.
- Greiner, J., S. Klose, K. Reinsch, H. Martin Schmid, R. Sari, D. H. Hartmann, C. Kouveliotou, A. Rau, E. Palazzi, C. Straubmeier, B. Stecklum, S. Zharikov, *et al.*, 2003, *Nature (London)* **426**, 157.
- Greisen, K., 1966, *Phys. Rev. Lett.* **16**, 748.
- Gruzinov, A., 1999, *Astrophys. J. Lett.* **525**, L29.
- Gruzinov, A., and E. Waxman, 1999, *Astrophys. J.* **511**, 852.
- Guetta, D., and J. Granot, 2003a, *Phys. Rev. Lett.* **90**, 201103.
- Guetta, D., and J. Granot, 2003b, *Astrophys. J.* **585**, 885.
- Guetta, D., D. Hooper, J. Alvarez-Muniz, F. Halzen, and E. Reuveni, 2004, *Astropart. Phys.* **20**, 429.
- Guetta, D., and T. Piran, 2004, *astro-ph/047063*.
- Guetta, D., T. Piran, and E. Waxman, 2004, *astro-ph/0311488*.
- Guetta, D., M. Spada, and E. Waxman, 2001a, *Astrophys. J.* **557**, 399.
- Guetta, D., M. Spada, and E. Waxman, 2001b, *Astrophys. J.* **559**, 101.
- Hakkila, J., D. J. Haglin, G. N. Pendleton, R. S. Mallozzi, C. A. Meegan, and R. J. Roiger, 2000, *Astrophys. J.* **538**, 165.
- Halpern, J. P., J. Kemp, T. Piran, and M. A. Bershad, 1999, *Astrophys. J. Lett.* **517**, L105.
- Harris, M. J., and G. H. Share, 1998, in *Gamma-Ray Bursts*, AIP Conference Proceedings No. 428, edited by C. A. Meegan, T. M. Koshut, and R. D. Preece (AIP, Woodbury, NY), p. 314.
- Harrison, F. A., J. S. Bloom, D. A. Frail, R. Sari, S. R. Kulkarni, S. G. Djorgovski, T. Axelrod, J. Mould, B. P. Schmidt, M. H. Wieringa, R. M. Wark, R. Subrahmanyam, *et al.*, 1999, *Astrophys. J. Lett.* **523**, L121.
- Harrison, F. A., S. A. Yost, R. Sari, E. Berger, T. J. Galama, J. Holtzman, T. Axelrod, J. S. Bloom, R. Chevalier, E. Costa, A. Diercks, S. G. Djorgovski, *et al.*, 2001, *Astrophys. J.* **559**, 123.
- Heavens, A. F., and L. O. Drury, 1988, *Mon. Not. R. Astron. Soc.* **235**, 997.
- Heinz, S., and M. C. Begelman, 1999, *Astrophys. J. Lett.* **527**, L35.
- Heise, J., 2003, private communication.
- Heise, J., J. in't Zand, R. M. Kippen, and P. M. Woods, 2001, in *Gamma-ray Bursts in the Afterglow Era*, edited by E. Costa, F. Frontera, and J. Hjorth (Springer, Berlin), p. 16.
- Heyl, J. S., and R. Perna, 2003, *Astrophys. J. Lett.* **586**, L13.
- Higdon, J. C., and R. E. Lingenfelter, 1998, in *Gamma-Ray Bursts*, AIP Conference Proceedings No. 428, edited by C. A. Meegan, T. M. Koshut, and R. D. Preece (AIP, Woodbury, NY), p. 40.
- Hjorth, J., J. Sollerman, P. Møller, J. P. U. Fynbo, S. E. Woosley, C. Kouveliotou, N. R. Tanvir, J. Greiner, M. I. Andersen, A. J. Castro-Tirado, J. M. Castro Cerón, A. S. Fruchter, *et al.*, 2003, *Nature (London)* **423**, 847.
- Holland, S., 2001, in *Relativistic Astrophysics: 20th Texas Symposium*, AIP Conference Proceedings No. 586, edited by J. C. Wheeler and H. Martel (AIP, Melville, NY), p. 593.
- Horack, J. M., and J. Hakkila, 1997, *Astrophys. J.* **479**, 371.
- Horváth, I., 1998, *Astrophys. J.* **508**, 757.
- Huang, Y. F., L. J. Gou, Z. G. Dai, and T. Lu, 2000, *Astrophys. J.* **543**, 90.
- Hurley, K., 1994, *Nature (London)* **372**, 652.
- Hurley, K., E. Berger, A. Castro-Tirado, J. M. Castro Cerón, T. Cline, M. Feroci, D. A. Frail, F. Frontera, N. Masetti, C. Guidorzi, E. Montanari, D. H. Hartmann, *et al.*, 2002, *Astrophys. J.* **567**, 447.
- Hurley, K., T. Cline, I. Mitrofanov, D. Anfinov, A. Kozyrev, M. Litvak, A. Sanin, W. Boynton, C. Fellows, K. Harshman, C. Shinohara, R. Starr, *et al.*, 2002a, *GRB Circular Network* **1402**, 1.
- Hurley, K., T. Cline, I. Mitrofanov, D. Anfinov, A. Kozyrev, M. Litvak, A. Sanin, W. Boynton, C. Fellows, K. Harshman, C. Shinohara, and R. Starr, 2002b, *GRB Circular Network* **1727**, 1.
- Hurley, K., T. Cline, I. Mitrofanov, D. Anfinov, A. Kozyrev, M. Litvak, A. Sanin, W. Boynton, C. Fellows, K. Harshman, C. Shinohara, R. Starr, *et al.*, 2002c, *GRB Circular Network* **1728**, 1.
- Hurley, K., R. Sari, and S. G. Djorgovski, 2002, *astro-ph/0211620*.
- Inoue, S., D. Guetta, and F. Pacini, 2003, *Astrophys. J.* **583**, 379.
- Iwamoto, K., P. A. Mazzali, K. Nomoto, H. Umeda, T. Nakamura, F. Patat, I. J. Danziger, T. R. Young, T. Suzuki, T. Shigeyama, T. Augusteijn, V. Doublier, *et al.*, 1998, *Nature (London)* **395**, 672.
- Janka, H.-T., and M. Ruffert, 1996, *Astron. Astrophys.* **307**, L33.
- Jaroszynski, M., 1996, *Astron. Astrophys.* **305**, 839.
- Jimenez, R., D. Band, and T. Piran, 2001, *Astrophys. J.* **561**, 171.
- Katz, J. I., 1994a, *Astrophys. J.* **422**, 248.
- Katz, J. I., 1994b, *Astrophys. J. Lett.* **432**, L27.
- Katz, J. I., 1997, *Astrophys. J.* **490**, 633.
- Katz, J. I., and L. M. Canel, 1996, *Astrophys. J.* **471**, 915.
- Katz, J. I., and T. Piran, 1997, *Astrophys. J.* **490**, 772.
- Kehoe, R., C. Akerlof, R. Balsano, S. Barthelmy, J. Bloch, P. Butterworth, D. Casperson, T. Cline, S. Fletcher, G. Gisler, K. Hurley, M. Kippen, *et al.*, 2001, *Astrophys. J. Lett.* **554**, L159.
- Kippen, R. M., P. M. Woods, J. Heise, J. in't Zand, R. D. Preece, and M. Briggs, 2002, in *American Physical Society, April Meeting*, April 2002, Albuquerque, New Mexico APR02, abstract X6.003, p. 6003.
- Kirk, J. G., A. W. Guthmann, Y. A. Gallant, and A. Achterberg, 2000, *Astrophys. J.* **542**, 235.
- Klotz, A., M. Boër, and J. L. Atteia, 2003, *Astron. Astrophys.* **404**, 815.
- Kobayashi, S., 2000, *Astrophys. J.* **545**, 807.
- Kobayashi, S., T. Piran, and R. Sari, 1997, *Astrophys. J.* **490**, 92.
- Kobayashi, S., T. Piran, and R. Sari, 1999, *Astrophys. J.* **513**, 669.
- Kobayashi, S., and R. Sari, 2001, *Astrophys. J.* **551**, 934.
- Kochanek, C. S., and T. Piran, 1993, *Astrophys. J. Lett.* **417**, L17.

- Königl, A., and J. Granot, 2002, *Astrophys. J.* **574**, 134.
- Kouveliotou, C., T. Koshut, M. S. Briggs, G. N. Pendleton, C. A. Meegan, G. J. Fishman, and J. P. Lestrade, 1996, in *Gamma-ray Bursts*, AIP Conference Proceedings No. 384, edited by C. Kouveliotou, M. F. Briggs, and G. J. Fishman (AIP, Woodbury, NY), p. 42.
- Kouveliotou, C., C. A. Meegan, G. J. Fishman, N. P. Bhat, M. S. Briggs, T. M. Koshut, W. S. Paciesas, and G. N. Pendleton, 1993, *Astrophys. J. Lett.* **413**, L101.
- Krolik, J. H., and E. A. Pier, 1991, *Astrophys. J.* **373**, 277.
- Kulkarni, S. R., D. A. Frail, M. H. Wieringa, R. D. Ekers, E. M. Sadler, R. M. Wark, J. L. Higdon, E. S. Phinney, and J. S. Bloom, 1998, *Nature (London)* **395**, 663.
- Kumar, P., 2000, *Astrophys. J. Lett.* **538**, L125.
- Kumar, P., and J. Granot, 2003, *Astrophys. J.* **591**, 1075.
- Kumar, P., and R. Narayan, 2003, *Astrophys. J.* **584**, 895.
- Kumar, P., and A. Panaitescu, 2000, *Astrophys. J. Lett.* **541**, L9.
- Kumar, P., and A. Panaitescu, 2004, *Mon. Not. R. Astron. Soc.* **354**, 252.
- Kumar, P., and T. Piran, 2000a, *Astrophys. J.* **535**, 152.
- Kumar, P., and T. Piran, 2000b, *Astrophys. J.* **532**, 286.
- Lamb, D. Q., and D. E. Reichart, 2000, *Astrophys. J.* **536**, 1.
- Lazzati, D., S. Covino, and G. Ghisellini, 2002, *Mon. Not. R. Astron. Soc.* **330**, 583.
- Lazzati, D., E. Rossi, S. Covino, G. Ghisellini, and D. Malesani, 2002, *Astron. Astrophys.* **399**, 913.
- Lazzati, D., E. Rossi, G. Ghisellini, and M. J. Rees, 2003, *Mon. Not. R. Astron. Soc.* **347**, L1.
- Le Flo'ch, E., 2004, *New Astron. Rev.* **48**, 601.
- Le Flo'ch, E., P.-A. Duc, I. F. Mirabel, D. B. Sanders, G. Bosch, R. J. Diaz, C. J. Donzelli, I. Rodrigues, T. J.-L. Courvoisier, J. Greiner, S. Mereghetti, J. Melnick, *et al.*, 2003, *Astron. Astrophys.* **400**, 499.
- Lee, H. K., R. A. M. J. Wijers, and G. E. Brown, 2000, *Phys. Rep.* **325**, 83.
- Levinson, A., E. O. Ofek, E. Waxman, and A. Gal-Yam, 2002, *Astrophys. J.* **576**, 923.
- Li, W., A. V. Filippenko, R. Chornock, and S. Jha, 2003, *Astrophys. J. Lett.* **586**, L9.
- Li, Z., Z. G. Dai, and T. Lu, 2002, *Astron. Astrophys.* **396**, 303.
- Liang, E., and V. Kargatis, 1996, *Nature (London)* **381**, 49.
- Lipunov, V. M., K. A. Postnov, and M. E. Prokhorov, 2001, *Astron. Rep.* **45**, 236.
- Lithwick, Y., and R. Sari, 2001, *Astrophys. J.* **555**, 540.
- Lloyd, N. M., and V. Petrosian, 1999, *Astrophys. J.* **511**, 550.
- Lloyd, N. M., and V. Petrosian, 2000, *Astrophys. J.* **543**, 722.
- Lloyd-Ronning, N. M., C. L. Fryer, and E. Ramirez-Ruiz, 2002, *Astrophys. J.* **574**, 554.
- Loredo, T. J., and I. M. Wasserman, 1995, *Astrophys. J., Suppl. Ser.* **96**, 261.
- Loredo, T. J., and I. M. Wasserman, 1998, *Astrophys. J.* **502**, 75.
- Lytikov, M., and R. Blandford, 2002, *astro-ph/0210671*.
- Lytikov, M., and R. Blandford, 2003, *astro-ph/0312347*.
- Lytikov, M., V. I. Pariev, and R. D. Blandford, 2003, *Astrophys. J.* **597**, 998.
- MacFadyen, A. I., and S. E. Woosley, 1999, *Astrophys. J.* **524**, 262.
- MacFadyen, A. I., S. E. Woosley, and A. Heger, 2001, *Astrophys. J.* **550**, 410.
- Madau, P., and C. Thompson, 2000, *Astrophys. J.* **534**, 239.
- Magliocchetti, M., G. Ghirlanda, and A. Celotti, 2003, *Mon. Not. R. Astron. Soc.* **343**, 255.
- Mallozzi, R. S., W. S. Paciesas, G. N. Pendleton, M. S. Briggs, R. D. Preece, C. A. Meegan, and G. J. Fishman, 1995, *Astrophys. J.* **454**, 597.
- Mallozzi, R. S., G. N. Pendleton, W. S. Paciesas, R. D. Preece, and M. S. Briggs, 1998, in *Gamma-Ray Bursts*, AIP Conference Proceedings No. 428, edited by C. A. Meegan, T. M. Koshut, and R. D. Preece (AIP, Woodbury, NY), p. 279.
- Mao, S., R. Narayan, and T. Piran, 1994, *Astrophys. J.* **420**, 171.
- Matheson, T., P. M. Garnavich, K. Z. Stanek, D. Bersier, S. T. Holland, K. Krisciunas, N. Caldwell, P. Berlind, J. S. Bloom, M. Bolte, A. Z. Bonanos, M. J. I. Brown, *et al.*, 2003, *Astrophys. J.* **599**, 394.
- Mazzali, P. A., J. Deng, N. Tominaga, K. Maeda, K. Nomoto, T. Matheson, K. S. Kawabata, K. Z. Stanek, and P. M. Garnavich, 2003, *Astrophys. J. Lett.* **599**, L95.
- McBreen, S., F. Quilligan, B. McBreen, L. Hanlon, and D. Watson, 2001, *Astron. Astrophys.* **380**, L31.
- McEnery, J., 2002, in American Physical Society, April Meeting, April, Albuquerque, New Mexico, abstract 4A3.007, p. A3007.
- Medvedev, M. V., 2000, *Astrophys. J.* **540**, 704.
- Medvedev, M. V., 2001, in *Gamma 2001: Gamma-Ray Astrophysics*, AIP Conference Proceedings No. 587, edited by S. Ritz, N. Gehrels, and C. R. Shrader (AIP, Melville, NY), p. 205.
- Medvedev, M. V., and A. Loeb, 1999, *Astrophys. J.* **526**, 697.
- Meegan, C. A., G. J. Fishman, R. B. Wilson, J. M. Horack, M. N. Brock, W. S. Paciesas, G. N. Pendleton, and C. Kouveliotou, 1992, *Nature (London)* **355**, 143.
- Mészáros, A., Z. Bagoly, I. Horváth, L. G. Balázs, and R. Vavrek, 2000a, *Astrophys. J.* **539**, 98.
- Mészáros, A., Z. Bagoly, and R. Vavrek, 2000b, *Astron. Astrophys.* **354**, 1.
- Mészáros, P., 2001, *Science* **291**, 79.
- Mészáros, P., 2002, *Annu. Rev. Astron. Astrophys.* **40**, 137.
- Mészáros, P., P. Laguna, and M. J. Rees, 1993, *Astrophys. J.* **415**, 181.
- Mészáros, P., E. Ramirez-Ruiz, and M. J. Rees, 2001, *Astrophys. J.* **554**, 660.
- Mészáros, P., and M. J. Rees, 1992, *Mon. Not. R. Astron. Soc.* **257**, 29P.
- Mészáros, P., and M. J. Rees, 1993, *Astrophys. J. Lett.* **418**, L59.
- Mészáros, P., and M. J. Rees, 1994, *Mon. Not. R. Astron. Soc.* **269**, L41.
- Mészáros, P., and M. J. Rees, 1997a, *Astrophys. J.* **476**, 232.
- Mészáros, P., and M. J. Rees, 1997b, *Astrophys. J. Lett.* **482**, L29.
- Mészáros, P., and M. J. Rees, 2000, *Astrophys. J. Lett.* **541**, L5.
- Mészáros, P., and M. J. Rees, 2001, *Astrophys. J. Lett.* **556**, L37.
- Mészáros, P., M. J. Rees, and H. Papathanassiou, 1994, *Astrophys. J.* **432**, 181.
- Mészáros, P., M. J. Rees, and R. A. M. J. Wijers, 1998, *Astrophys. J.* **499**, 301.
- Mészáros, P., and E. Waxman, 2001, *Phys. Rev. Lett.* **87**, 171102.
- Mitra, A., 1996, *Astron. Astrophys.* **313**, L9.
- Moderski, R., M. Sikora, and T. Bulik, 2000, *Astrophys. J.* **529**, 151.
- Mukherjee, S., E. D. Feigelson, G. Jogesh Babu, F. Murtagh, C. Fraley, and A. Raftery, 1998, *Astrophys. J.* **508**, 314.
- Nakar, E., J. Granot, and D. Guetta, 2004, *Astrophys. J.* **606**,

- L37.
- Nakar, E., and Y. Oren, 2004, *Mon. Not. R. Astron. Soc.* **353**, 647.
- Nakar, E., and T. Piran, 2002a, *Astrophys. J. Lett.* **572**, L139.
- Nakar, E., and T. Piran, 2002b, *Mon. Not. R. Astron. Soc.* **330**, 920.
- Nakar, E., and T. Piran, 2002c, *Mon. Not. R. Astron. Soc.* **331**, 40.
- Nakar, E., and T. Piran, 2003a, *Astrophys. J.* **598**, 400.
- Nakar, E., and T. Piran, 2003b, *New Astron.* **8**, 141.
- Nakar, E., T. Piran, and J. Granot, 2002, *Astrophys. J.* **579**, 699.
- Nakar, E., T. Piran, and J. Granot, 2003, *New Astron.* **8**, 495.
- Nakar, E., T. Piran, and E. Waxman, 2003, *J. Cosmol. Astropart. Phys.* **10**, 5.
- Narayan, R., B. Paczynski, and T. Piran, 1992, *Astrophys. J. Lett.* **395**, L83.
- Narayan, R., T. Piran, and P. Kumar, 2001, *Astrophys. J.* **557**, 949.
- Narayan, R., T. Piran, and A. Shemi, 1991, *Astrophys. J. Lett.* **379**, L17.
- Nikto, S., and M. V. Medvedev, 2001, *American Astronomical Society Meeting*, 33198, 824.
- Norris, J. P., G. F. Marani, and J. T. Bonnell, 2000, *Astrophys. J.* **534**, 248.
- Norris, J. P., R. J. Nemiroff, J. T. Bonnell, J. D. Scargle, C. Kouveliotou, W. S. Paciesas, C. A. Meegan, and G. J. Fishman, 1996, *Astrophys. J.* **459**, 393.
- Odehahn, S. C., S. G. Djorgovski, S. R. Kulkarni, M. Dickinson, D. A. Frail, A. N. Ramaprakash, J. S. Bloom, K. L. Adelberger, J. Halpern, D. J. Helfand, J. Bahcall, R. Go-drich, *et al.*, 1998, *Astrophys. J. Lett.* **509**, L5.
- Oren, Y., E. Nakar, and T. Piran, 2004, *Mon. Not. R. Astron. Soc.* **353**, L35.
- Ostrowski, M., and J. Bednarz, 2002, *Astron. Astrophys.* **394**, 1141.
- Paciesas, W. S., C. A. Meegan, G. N. Pendleton, M. S. Briggs, C. Kouveliotou, T. M. Koshut, J. P. Lestrade, M. L. McCollough, J. J. Brainerd, J. Hakkila, W. Henze, R. D. Preece, *et al.*, 1999, *Astrophys. J., Suppl. Ser.* **122**, 465.
- Paczynski, B., 1991, *Acta Astron.* **41**, 257.
- Paczynski, B., 1998, *Astrophys. J. Lett.* **494**, L45.
- Paczynski, B., and J. E. Rhoads, 1993, *Astrophys. J. Lett.* **418**, L5.
- Paczynski, B., and G. Xu, 1994, *Astrophys. J.* **427**, 708.
- Panaiteescu, A., and P. Kumar, 2000, *Astrophys. J.* **543**, 66.
- Panaiteescu, A., and P. Kumar, 2001, *Astrophys. J. Lett.* **560**, L49.
- Panaiteescu, A., and P. Mészáros, 1999, *Astrophys. J.* **526**, 707.
- Panaiteescu, A., M. Spada, and P. Mészáros, 1999, *Astrophys. J. Lett.* **522**, L105.
- Papathanassiou, H., and P. Meszaros, 1996, *Astrophys. J. Lett.* **471**, L91.
- Pe'er, A., and E. Waxman, 2004, *Astrophys. J. Lett.* **603**, L1.
- Pendleton, G. N., W. S. Paciesas, M. S. Briggs, R. D. Preece, R. S. Mallozzi, C. A. Meegan, J. M. Horack, G. J. Fishman, D. L. Band, J. L. Matteson, R. T. Skelton, J. Hakkila, *et al.*, 1997, *Astrophys. J.* **489**, 175.
- Perna, R., and K. Belczynski, 2002, *Astrophys. J.* **570**, 252.
- Perna, R., and A. Loeb, 1998, *Astrophys. J. Lett.* **509**, L85.
- Perna, R., R. Sari, and D. Frail, 2003, *Astrophys. J.* **594**, 379.
- Phinney, E. S., 1991, *Astrophys. J. Lett.* **380**, L17.
- Pian, E., L. Amati, L. A. Antonelli, R. C. Butler, E. Costa, G. Cusumano, J. Danziger, M. Feroci, F. Fiore, F. Frontera, P. Giommi, N. Masetti, *et al.*, 2000, *Astrophys. J.* **536**, 778.
- Pian, E., P. Giommi, L. Amati, E. Costa, J. Danziger, M. Feroci, M. T. Fiocchi, F. Frontera, C. Kouveliotou, N. Masetti, L. Nicastro, E. Palazzi, *et al.*, 2003, *astro-ph/0304521*.
- Pilla, R. P., and A. Loeb, 1998, *Astrophys. J. Lett.* **494**, L167.
- Piran, T., 1992, *Astrophys. J. Lett.* **389**, L45.
- Piran, T., 1994, in *Gamma-Ray Bursts*, AIP Conference Proceedings No. 307, edited by G. F. Fishman, J. J. Brainerd, and K. Hurley (AIP, New York), p. 495.
- Piran, T., 1997, in *Unsolved Problems in Astrophysics*, edited by J. Bahcall and J. Ostriker (Princeton University Press, Princeton, NJ), p. 273.
- Piran, T., 1996, in *Compact Stars in Binaries*, IAU Symposium No. 165, edited by E. P. Van den Heuvel, J. Van Paradijs, and E. Kuulkers (Kluwer Academic, Dordrecht), p. 489.
- Piran, T., 1999, *Phys. Rep.* **314**, 575.
- Piran, T., 2000, *Phys. Rep.* **333**, 529.
- Piran, T., P. Kumar, A. Panaiteescu, and L. Piro, 2001, *Astrophys. J. Lett.* **560**, L167.
- Piran, T. and E. Nakar, 2002, in *Proceeding of the conference on Compact Stars in the QCD Phase Diagram*, edited by R. Ouyed and F. Sannino, eConf C010815, 182.
- Piran, T., E. Nakar, and J. Granot, 2004, in *Gamma-Ray Bursts: 30 Years of Discovery*, AIP Conference Proceedings No. 727, edited by E. E. Fenimore and M. Galassi (AIP, Melville, NY), pp. 181–186.
- Piran, T., and R. Narayan, 1996, in *Gamma-Ray Bursts, 3rd Huntsville Symposium*, AIP Conference Proceedings No. 384, edited by C. Kouveliotou, M. F. Briggs, and G. J. Fishman (AIP, Woodbury, NY), p. 233.
- Piran, T., and R. Sari, 1998, in *Gamma-Ray Bursts*, AIP Conference Proceedings No. 428, edited by C. A. Meegan, T. M. Koshut, and R. D. Preece (AIP, Woodbury, NY), p. 662.
- Piran, T., A. Shemi, and R. Narayan, 1993, *Mon. Not. R. Astron. Soc.* **263**, 861.
- Piro, L., 2001, in *Gamma-ray Bursts in the Afterglow Era*, edited by E. Costa, F. Frontera, and J. Hjorth (Springer, Berlin), p. 97.
- Piro, L., E. Costa, M. Feroci, G. Stratta, F. Frontera, L. Amati, D. dal Fiume, L. A. Antonelli, J. Heise, J. in't Zand, A. Owens, A. N. Parmar, *et al.*, 1999, *Astron. Astrophys., Suppl. Ser.* **138**, 431.
- Piro, L., D. A. Frail, J. Gorosabel, G. Garmire, P. Soffitta, L. Amati, M. I. Andersen, L. A. Antonelli, E. Berger, F. Frontera, J. Fynbo, G. Gandolfi, *et al.*, 2002, *Astrophys. J.* **577**, 680.
- Piro, L., G. Garmire, M. Garcia, G. Stratta, E. Costa, M. Feroci, P. Mészáros, M. Vietri, H. Bradt, D. Frail, F. Frontera, J. Halpern, *et al.*, 2000, *Science* **290**, 955.
- Porciani, C., and P. Madau, 2001, *Astrophys. J.* **548**, 522.
- Preece, R. D., M. S. Briggs, T. W. Gibling, R. S. Mallozzi, G. N. Pendleton, W. S. Paciesas, and D. L. Band, 2002, *Astrophys. J.* **581**, 1248.
- Preece, R. D., M. S. Briggs, R. S. Mallozzi, G. N. Pendleton, W. S. Paciesas, and D. L. Band, 1998, *Astrophys. J. Lett.* **506**, L23.
- Preece, R. D., M. S. Briggs, R. S. Mallozzi, G. N. Pendleton, W. S. Paciesas, and D. L. Band, 2000, *Astrophys. J., Suppl. Ser.* **126**, 19.
- Price, P. A., D. W. Fox, S. R. Kulkarni, B. A. Peterson, B. P. Schmidt, A. M. Soderberg, S. A. Yost, E. Berger, S. G. Djorgovski, D. A. Frail, F. A. Harrison, R. Sari, *et al.*, 2003, *Nature (London)* **423**, 844.

- Pruet, J., and N. Dalal, 2002, *Astrophys. J.* **573**, 770.
- Qin, Y., G. Xie, S. Xue, E. Liang, X. Zheng, and D. Mei, 2000, *Publ. Astron. Soc. Jpn.* **52**, 759.
- Quilligan, F., B. McBreen, L. Hanlon, S. McBreen, K. J. Hurley, and D. Watson, 2002, *Astron. Astrophys.* **385**, 377.
- Ramirez-Ruiz, E., and E. E. Fenimore, 2000, *Astrophys. J.* **539**, 712.
- Ramirez-Ruiz, E., A. I. MacFadyen, and D. Lazzati, 2002, *Mon. Not. R. Astron. Soc.* **331**, 197.
- Ramirez-Ruiz, E., and A. Merloni, 2001, *Mon. Not. R. Astron. Soc.* **320**, L25.
- Razzaque, S., P. Meszaros, and E. Waxman, 2003a, *Phys. Rev. Lett.* **90**, 241103.
- Razzaque, S., P. Mészáros, and E. Waxman, 2003b, *Phys. Rev. D* **68**, 083001.
- Rees, M. J., and P. Mészáros, 1992, *Mon. Not. R. Astron. Soc.* **258**, 41P.
- Rees, M. J., and P. Mészáros, 1994, *Astrophys. J. Lett.* **430**, L93.
- Rees, M. J., and P. Mészáros, 1998, *Astrophys. J. Lett.* **496**, L1.
- Rees, M. J., and P. Mészáros, 2000, *Astrophys. J. Lett.* **545**, L73.
- Reeves, J. N., D. Watson, J. P. Osborne, K. A. Pounds, P. T. O'Brien, A. D. T. Short, M. J. L. Turner, M. G. Watson, K. O. Mason, M. Ehle, and N. Schartel, 2002, *Nature (London)* **416**, 512.
- Reichart, D., and S. Yost, 2001, *astro-ph/0107545*.
- Reichart, D. E., 1999, *Astrophys. J. Lett.* **521**, L111.
- Reichart, D. E., 2001, *Astrophys. J.* **554**, 643.
- Reichart, D. E., D. Q. Lamb, E. E. Fenimore, E. Ramirez-Ruiz, T. L. Cline, and K. Hurley, 2001, *Astrophys. J.* **552**, 57.
- Reichart, D. E., and P. A. Price, 2002, *Astrophys. J.* **565**, 174.
- Rhoads, J. E., 1997, *Astrophys. J. Lett.* **487**, L1.
- Rhoads, J. E., 1999, *Astrophys. J.* **525**, 737.
- Rhoads, J. E., 2001, *Astrophys. J.* **557**, 943.
- Rol, E., R. A. M. J. Wijers, P. M. Vreeswijk, L. Kaper, T. J. Galama, J. van Paradijs, C. Kouveliotou, N. Masetti, E. Pian, E. Palazzi, F. Frontera, and E. P. J. van den Heuvel, 2000, *Astrophys. J.* **544**, 707.
- Rossi, E., D. Lazzati, and M. J. Rees, 2002, *Mon. Not. R. Astron. Soc.* **332**, 945.
- Rossi, E. M., D. Lazzati, J. D. Salmonson, and G. Ghisellini, 2004, *astro-ph/0401124*.
- Rosswog, S., and M. B. Davies, 2002, *Mon. Not. R. Astron. Soc.* **334**, 481.
- Rosswog, S., M. B. Davies, F. Thielemann, and T. Piran, 2000, *Astron. Astrophys.* **360**, 171.
- Rosswog, S., M. Liebendörfer, F. Thielemann, M. B. Davies, W. Benz, and T. Piran, 1999, *Astron. Astrophys.* **341**, 499.
- Rosswog, S., and E. Ramirez-Ruiz, 2003, *Mon. Not. R. Astron. Soc.* **343**, L36.
- Rosswog, S., E. Ramirez-Ruiz, and M. B. Davies, 2003, *Mon. Not. R. Astron. Soc.* **345**, 1077.
- Rowan-Robinson, M., 1999, *Astrophys. Space Sci.* **266**, 291.
- Ruderman, M., 1975, in *7th Texas Symposium on Relativistic Astrophysics*, Dallas, Tx., 1974, *Ann. N.Y. Acad. Sci.* **262**, 164.
- Ruffert, M., and H.-T. Janka, 1998, *Astron. Astrophys.* **338**, 535.
- Ruffert, M., and H.-T. Janka, 1999, *Astron. Astrophys.* **344**, 573.
- Ruffert, M., and H.-T. Janka, 2001, *Astron. Astrophys.* **380**, 544.
- Ruffert, M., H.-T. Janka, and G. Schafer, 1995, *Astrophys. Space Sci.* **231**, 423.
- Ruffini, R., C. L. Bianco, F. Frascchetti, S. Xue, and P. Chardonnet, 2001, *Astrophys. J. Lett.* **555**, L113.
- Rutledge, R. E., and D. B. Fox, 2004, *Mon. Not. R. Astron. Soc.* **350**, 1272.
- Rutledge, R. E., and M. Sako, 2003, *Mon. Not. R. Astron. Soc.* **339**, 600.
- Rybicki, G. B., and A. P. Lightman, 1979, *Radiative Processes in Astrophysics* (Wiley-Interscience, New York).
- Sako, M., F. A. Harrison, and R. E. Rutledge, 2003, *AAS/High Energy Astrophysics Division #7*, #16.10.
- Sari, R., 1997, *Astrophys. J. Lett.* **489**, L37.
- Sari, R., 1999a, Ph.D. thesis (Hebrew University).
- Sari, R., 1999b, *Astrophys. J. Lett.* **524**, L43.
- Sari, R., and P. Mészáros, 2000, *Astrophys. J. Lett.* **535**, L33.
- Sari, R., R. Narayan, and T. Piran, 1996, *Astrophys. J.* **473**, 204.
- Sari, R., and T. Piran, 1995, *Astrophys. J. Lett.* **455**, L143.
- Sari, R., and T. Piran, 1997a, *Mon. Not. R. Astron. Soc.* **287**, 110.
- Sari, R., and T. Piran, 1997b, *Astrophys. J.* **485**, 270.
- Sari, R., and T. Piran, 1999a, *Astron. Astrophys., Suppl. Ser.* **138**, 537.
- Sari, R., and T. Piran, 1999b, *Astrophys. J. Lett.* **517**, L109.
- Sari, R., and T. Piran, 1999c, *Astrophys. J.* **520**, 641.
- Sari, R., T. Piran, and J. P. Halpern, 1999, *Astrophys. J. Lett.* **519**, L17.
- Sari, R., T. Piran, and R. Narayan, 1998, *Astrophys. J. Lett.* **497**, L17.
- Schaefer, B. E., 2003, *Astrophys. J. Lett.* **583**, L67.
- Schaefer, B. E., M. Deng, and D. L. Band, 2001, *Astrophys. J. Lett.* **563**, L123.
- Schmidt, M., 1999, *Astrophys. J. Lett.* **523**, L117.
- Schmidt, M., 2001a, *Astrophys. J. Lett.* **559**, L79.
- Schmidt, M., 2001b, *Astrophys. J.* **552**, 36.
- Sethi, S., and S. G. Bhargavi, 2001, *Astron. Astrophys.* **376**, 10.
- Shaviv, N. J., and A. Dar, 1995, *Astrophys. J.* **447**, 863.
- Shemi, A., 1994, *Mon. Not. R. Astron. Soc.* **269**, 1112.
- Shemi, A., and T. Piran, 1990, *Astrophys. J. Lett.* **365**, L55.
- Silva, L. O., R. A. Fonseca, J. W. Tonge, J. M. Dawson, W. B. Mori, and M. V. Medvedev, 2003, *Astrophys. J. Lett.* **596**, L121.
- Smolsky, M. V., and V. V. Usov, 1996, *Astrophys. J.* **461**, 858.
- Smolsky, M. V., and V. V. Usov, 2000, *Astrophys. J.* **531**, 764.
- Soderberg, A. M., P. A. Price, D. W. Fox, S. R. Kulkarni, S. G. Djorgovski, E. Berger, F. Harrison, S. Yost, M. Hamuy, S. Shectman, N. Mirabal, and J. Halpern, 2002, *GRB Circular Network* 1554.
- Soderberg, A. M., and E. Ramirez-Ruiz, 2003, in *Gamma-Ray Burst and Afterglow Astronomy 2001: A Workshop Celebrating the First Year of the HETE Mission*, AIP Conference Proceedings No. 662, edited by G. R. Ricker and R. K. Van der speck (AIP, Melville, NY), p. 172.
- Sokolov, V. V., T. A. Fatkhullin, A. J. Castro-Tirado, A. S. Fruchter, V. N. Komarova, E. R. Kasimova, S. N. Dodonov, V. L. Afanasiev, and A. V. Moiseev, 2001, *Astron. Astrophys.* **372**, 438.
- Sommer, M., D. L. Bertsch, B. L. Dingus, C. E. Fichtel, G. J. Fishman, A. K. Harding, R. C. Hartman, S. D. Hunter, K. Hurley, G. Kanbach, D. A. Kniffen, C. Kouveliotou, *et al.*, 1994, *Astrophys. J. Lett.* **422**, L63.
- Spruit, H. C., F. Daigne, and G. Drenkhahn, 2001, *Astron. Astrophys.* **369**, 694.
- Stanek, K. Z., P. M. Garnavich, J. Kaluzny, W. Pych, and I.

- Thompson, 1999, *Astrophys. J. Lett.* **522**, L39.
- Stanek, K. Z., T. Matheson, P. M. Garnavich, P. Martini, P. Berlind, N. Caldwell, P. Challis, W. R. Brown, R. Schild, K. Krisciunas, M. L. Calkins, J. C. Lee, *et al.*, 2003, *Astrophys. J. Lett.* **591**, L17.
- Stark, R. F., and T. Piran, 1985, *Phys. Rev. Lett.* **55**, 891.
- Strohmayer, T. E., E. E. Fenimore, T. Murakami, and A. Yoshida, 1998, *Astrophys. J.* **500**, 873.
- Svensson, R., 1982, *Astrophys. J.* **258**, 321.
- Svensson, R., 1984, *Mon. Not. R. Astron. Soc.* **209**, 175.
- Tavani, M., 1996a, *Astrophys. J.* **466**, 768.
- Tavani, M., 1996b, *Phys. Rev. Lett.* **76**, 3478.
- Tavani, M., 1998, *Astrophys. J. Lett.* **497**, L21.
- Taylor, G. B., D. A. Frail, E. Berger, and S. R. Kulkarni, 2004, *Astrophys. J.* **609**, L1.
- Taylor, G. B., D. A. Frail, and D. Fox, 2001, *GRB Circular Network* **1122**, 1.
- Taylor, J. H., and J. M. Weisberg, 1982, *Astrophys. J.* **253**, 908.
- Thompson, C., 1994, *Mon. Not. R. Astron. Soc.* **270**, 480.
- Thompson, C., and P. Madau, 2000, *Astrophys. J.* **538**, 105.
- Totani, T., 1997, *Astrophys. J. Lett.* **486**, L71.
- Totani, T., and A. Panaitescu, 2002, *Astrophys. J.* **576**, 120.
- Tutukov, A. V., and L. R. Yungelson, 1993, *Mon. Not. R. Astron. Soc.* **260**, 675.
- Tutukov, A. V., and L. R. Yungelson, 1994, *Mon. Not. R. Astron. Soc.* **268**, 871.
- Usov, V. V., 1992, *Nature (London)* **357**, 472.
- Usov, V. V., 1994, *Mon. Not. R. Astron. Soc.* **267**, 1035.
- Vanden Berk, D. E., B. C. Lee, B. C. Wilhite, J. F. Beacom, D. Q. Lamb, J. Annis, K. Abazajian, T. A. McKay, R. G. Kron, S. Kent, K. Hurley, R. Kehoe, *et al.*, 2002, *Astrophys. J.* **576**, 673.
- van den Heuvel, E. P. J., and D. R. Lorimer, 1996, *Mon. Not. R. Astron. Soc.* **283**, L37.
- van Paradijs, J., P. J. Groot, T. Galama, C. Kouveliotou, R. G. Strom, J. Telting, R. G. M. Rutten, G. J. Fishman, C. A. Meegan, M. Pettini, N. Tanvir, J. Bloom, *et al.*, 1997, *Nature (London)* **386**, 686.
- van Paradijs, J., C. Kouveliotou, and R. A. M. J. Wijers, 2000, *Annu. Rev. Astron. Astrophys.* **38**, 379.
- van Putten, M. H. P. M., 2001, *Phys. Rep.* **345**, 1.
- van Putten, M. H. P. M., and A. Levinson, 2001, *Astrophys. J. Lett.* **555**, L41.
- van Putten, M. H. P. M., and T. Regimbau, 2003, *Astrophys. J. Lett.* **593**, L15.
- Veerswijk, P. M., 2003, Ph.D. thesis (University of Amsterdam).
- Vietri, M., 1995, *Astrophys. J.* **453**, 883.
- Vietri, M., 1997, *Phys. Rev. Lett.* **78**, 4328.
- Vietri, M., 2003, *Astrophys. J.* **591**, 954.
- Vietri, M., D. De Marco, and D. Guetta, 2003, *Astrophys. J.* **592**, 378.
- Vietri, M., and L. Stella, 1998, *Astrophys. J. Lett.* **507**, L45.
- Vietri, M., and L. Stella, 1999, *Astrophys. J. Lett.* **527**, L43.
- Watson, D., J. N. Reeves, J. Osborne, P. T. O'Brien, K. A. Pounds, J. A. Tedds, M. Santos-Lleó, and M. Ehle, 2002, *Astron. Astrophys.* **393**, L1.
- Waxman, E., 1995, *Phys. Rev. Lett.* **75**, 386.
- Waxman, E., 1997a, *Astrophys. J. Lett.* **489**, L33.
- Waxman, E., 1997b, *Astrophys. J. Lett.* **485**, L5.
- Waxman, E., 2003, *Nature (London)* **423**, 388.
- Waxman, E., and J. Bahcall, 1997, *Phys. Rev. Lett.* **78**, 2292.
- Waxman, E., and J. N. Bahcall, 2000, *Astrophys. J.* **541**, 707.
- Waxman, E., and B. T. Draine, 2000, *Astrophys. J.* **537**, 796.
- Waxman, E., S. R. Kulkarni, and D. A. Frail, 1998, *Astrophys. J.* **497**, 288.
- Weinberg, S., 1973, *Gravitation and Cosmology* (Wiley, New York).
- Wigger, C., W. Hajdas, K. Arzner, M. Güdel, and A. Zehnder, 2004, *astro-ph/0405525*.
- Wijers, R. A. M. J., J. S. Bloom, J. S. Bagla, and P. Natarajan, 1998, *Mon. Not. R. Astron. Soc.* **294**, L13.
- Wijers, R. A. M. J., and T. J. Galama, 1999, *Astrophys. J.* **523**, 177.
- Wijers, R. A. M. J., P. M. Vreeswijk, T. J. Galama, E. Rol, J. van Paradijs, C. Kouveliotou, T. Giblin, N. Masetti, E. Palazzi, E. Pian, F. Frontera, L. Nicastro, *et al.*, 1999, *Astrophys. J. Lett.* **523**, L33.
- Woods, E., and A. Loeb, 1995, *Astrophys. J.* **453**, 583.
- Woosley, S. E., 1993, *Astrophys. J.* **405**, 273.
- Yoshida, A., M. Namiki, C. Otani, N. Kawai, T. Murakami, Y. Ueda, R. Shibata, and S. Uno, 1999, *Astron. Astrophys., Suppl. Ser.* **138**, 433.
- Zatsepin, G. T., and V. A. Kuzmin, 1966, *JETP Lett.* **4**, 78.
- Zhang, B., and P. Mészáros, 2002, *Astrophys. J.* **571**, 876.
- Zhang, W., S. E. Woosley, and A. I. MacFadyen, 2003, *Astrophys. J.* **586**, 356.