Physics – Uspekhi **42** (5) 469 – 480 (1999)

PHYSICS OF OUR DAYS

PACS numbers: 95.85.Pw, 98.70.Rz

Cosmic gamma-ray bursts

K A Postnov

Contents

1. Introduction **469**
2. Observational properties of cosmic GRBs **470**
2.1 BATSE results; 2.2 BeppoSAX results; 2.3 ROTSE and GRB990123
3. Cosmic GRBs — relativistic fireballs **473**
3.1 Compactness problem and the need for relativistic motion; 3.2 Braking of the relativistic fireball in the ambient medium; 3.3 Afterglows of GRBs; 3.4 GRB990123 and the early optical afterglow of GRBs
4. Possible astrophysical sources of cosmological GRBs **477**
5. Conclusion **479**
References **479**

Abstract. **The results of the observation of cosmic gamma-ray bursts are discussed and available theoretical models are presented. Emphasis is placed on a cosmological model in which a gamma burst results from a powerful ($\sim 10^{51} - 10^{53}$ erg) and very short ($\sim 10 - 100$ s) energy release which occurs in a compact ($\sim 10^6 - 10^7$ cm) region and gives rise to a photon–lepton fireball expanding at an ultrarelativistic velocity (Lorentz-factor $\Gamma \gtrsim 100$). The interaction of the relativistic shock wave with its environment produces the observed X-ray and optical afterglows of a burst. Possible physical models of such an energy release event are discussed and some related problems considered.**

1. Introduction

It is not exaggerated to say that the problem of cosmic gamma-ray bursts (GRB) known for three decades has became one of the 'hottest' topics in astrophysics.We recall that GRBs are bursts of hard X-ray and gamma-radiation with photon energies $E > 30 - 50$ keV lasting from a few seconds to hundred seconds, with complex time profiles variable on timescales as short as milliseconds. The bursts are non-periodic and come from different directions with equal probability. The existing detectors register GRBs with fluences from $F_{\min} \simeq 10^{-7}$ erg cm^{-2} to $F_{\max} \sim 5 \times 10^{-4}$ erg cm^{-2} in the energy band $30 - 500$ keV at a rate of roughly one per day.

Unlike the relatively recent great astronomical discoveries such as quasars (1963) or pulsars (1967), the main physical parameters of which were rather quickly understood, the nature of cosmic GRBs discovered at the end of 1960s is still far from being firmly established. For example, in 1996 (see detailed review [1] and references therein) even such a macroscopic characteristic of the phenomenon as the energy released during the burst differed by 20 orders of magnitude (!) in different models. This fact is primarily due to the specifics of astronomical observations in the hard X-ray and gamma-ray bands, more precisely, due to the impossibility of accurately localizing the position of a source in the sky. In the best case, using a triangulation method (i.e. by the difference in the time of arrival of the signal) with several detectors onboard separate spacecraft, it is possible to find the error box of a GRB in the sky to an accuracy of a fraction of degree. From the astronomical point of view, this is a very poor localization: the mean number of stars only from our Galaxy projected onto a square degree is $10^{11}/40000 \sim 2 \times 10^6$ (this number varies by some orders of magnitude depending on the galactic coordinates). Of the same order is the number of galaxies observed up to limiting redshifts. So without any additional information it is impossible to determine at which distance the source of a GRB is located. This explains the aforementioned uncertainty in the GRB energetics: the energy released in the burst is

$$\Delta E_\gamma = F \times 4\pi r^2 \approx 10^{39} (\text{erg s}^{-3}) \left(\frac{F}{10^{-7} \text{ erg s}^{-2}} \right) \left(\frac{r}{10 \text{ kpc}} \right)^2$$

for sources inside the galactic disk, i.e. located within 100 pc of the Solar system[1], and for sources at Gigaparsec distances the energy is greater by $(10^9/100)^2 = 10^{14}$ times, reaching more than 10^{51} erg for cosmological GRBs. Such huge energies are reminiscent of catastrophic events like supernova explosions in which an energy $\sim 0.15 M_\odot c^2 = 3 \times 10^{53}$ erg is released (mainly in the form of neutrinos), or of the coalescence of binary neutron stars or black holes due to gravitational wave induced orbit decay[2].

[1] 1 parsec (pc) $\approx$206265 a.e. $\simeq 3 \times 10^{18}$ cm.

[2] In these estimates we assumed isotropic emission. Clearly, if the energy is beamed inside a narrow cone with an angle θ, the energy release derived from the observed flux will be reduced by a factor $\theta^2/4$. This coefficient for the GRB energy should be taken into account in all formulae below.

K A Postnov M V Lomonosov Moscow State University,
Sternberg Astronomical Institute,
Vorob'evy Gory, 119899 Moscow, Russian Federation
Tel. (7-095) 939 50 06
E-mail: pk@sai.msu.su

Received 17 February 1999
Uspekhi Fizicheskikh Nauk **169** (5) 545 – 558 (1999)
Translated by K A Postnov; edited by S D Danilov

By the beginning of 1999, it was firmly established that at least some fraction of gamma-ray bursts (all?) lie at cosmological distances. This became possible owing to the long-awaited identification of some GRBs with sources at other wavelengths (X-ray, optical, IR, and radio), the so-called *afterglows* of GRBs, which differ from all other known astronomical sources by their time behavior.The identification is being done according to the scheme: detection of a GRB (localization to several square degrees) — X-ray afterglow (localization to several arcminutes) — optical and radio afterglows (localization to several arcseconds) — search for a faint host galaxy at a well-determined position after the afterglow has faded or even simultaneously with it.

The electromagnetic energy release of $10^{51} - 10^{54}$ erg over the characteristic time of order 10 seconds is non-trivial and requires special explanation. Moreover, it cannot pass without trace if such an explosion occurs even in a very rarefied interstellar medium with a density of $0.1 - 10$ atoms per cubic cm.

The aim of the present paper is twofold: first, we wish to present the reader with some of the most recent observational results of the last two years; and second, we would like to consider cosmological GRBs in more detail and to discuss the physical processes associated with an energy release of $10^{51} - 10^{54}$ erg in the ambient interstellar medium. We also consider possible models for GRBs discussed in the literature. The observational data we will use are complete as of the end of 1998, and the relativistic fireball model is considered in more detail in the review by T Piran [2].

The above said may raise the question: if the paper mostly considers cosmological GRBs, why should the paper not be entitled 'Cosmological GRBs'? The answer is that despite enormous recent progress, firm proof of the cosmological nature of *all* GRBs is absent. For example, the error box of GRB 980425 contains a peculiar supernova 1998bw of rare Ic type in a nearby galaxy at 40 Mpc [3]. Although the error box of a fading X-ray afterglow detected by the BeppoSAX satellite was $3'$ away from the supernova, the probability of a random association is 10^{-4}, so the relation of this GRB to the supernova is not totally excluded. If so, the gamma-ray energy release is 10^{47} ergs in this case, which is much smaller than the 'standard' cosmological value. There are also statistical indications that a fraction of the GRBs may be homogeneously distributed in the Euclidean space [4] (clearly, the source distribution at high redshifts differs from that in the flat Euclidean space). These facts may indicate that GRBs can be generated by physically different mechanisms.

Progress in this field of astrophysics is so rapid that any publication on the subject becomes at least partially obsolete by the moment of printing. So we tried to focus on the most reliable, in our opinion, physical processes occuring in GRBs. This firstly relates to the most elaborated qualitative models of a relativistically expanding electron – positron plasma (the so-called fireball model) interacting with the ambient medium [5 – 10]. This model gives a sufficiently adequate description of many observed properties of both GRBs themselves and their afterglows in the soft spectral ranges [11 – 16]. At present, however, there is no definite answer on how such fireballs are created. We also specially avoid considering a small subclass of cosmic GRBs, the so-called soft gamma-ray repeaters, which represent a class of neutron stars with superstrong ($\sim 4 \times 10^{14}$ G) magnetic fields ('magnetars') like SGR $1900 + 14$ (see e.g. Refs [17, 18] and references therein). In addition, some bursts of cosmic gamma-rays originate in 'near space', during solar flares or even in the terrestrial atmosphere. These interesting phenomena deserve a separate discussion.

In Section 2 we briefly enumerate what is known and list the most recent observational data on GRBs and their afterglows. In Section 3 we consider the model of a relativistic fireball. Section 4 is devoted to modern models of cosmic GRBs discussed in the literature.

2. Observational properties of cosmic GRBs

The history of GRB studies can be sharply subdivided into three periods. The first period includes studies before 1991 when the American Compton Gamma-Ray Observatory (CGRO) was launched. The all-sky gamma-ray monitor BATSE (Burst And Transient Search Experiment) is mounted onboard CGRO. It can detect gamma-ray photons with energies from 30 to 500 keV. BATSE remains the most sensitive gamma-ray detector in this energy range capable of detecting fluences down to $F_{\min} = 10^{-7}$ erg cm^{-2} (about one 100 keV quantum). The mean cosmic GRB detection rate by BATSE is about 0.8 events per day.

2.1 BATSE results

The main result obtained with BATSE is the firm proof that GRBs are *isotropically* distributed over the sky. No significant deviations from the isotropic distribution of GRBs have been found [19].The second important fact is the apparent absence of *spatial homogeneity* of GRBs which follows from the dependence $\log(N > S) - \log S$, the integral distribution of the number of sources with a fluence (or peak flux) exceeding a given threshold. The sense of this distribution is simple. If the sources homogeneously fill Euclidean space then $N(> S) \propto S^{-3/2}$, since the number of objects inside a sphere with radius r is proportional to r^3 and the fluence observed from the source $\propto r^{-2}$. Clearly, until the detector sensitivity is smaller than required to register sources lying at the outer boundary of their spatial distribution, $\log(N > S) - \log S$ does not depend on their luminosity function[3]. According to BATSE data [20, 21] $\log(N > S) - \log S$ curve for GRBs from BATSE catalogues deviates significantly from the law $-3/2$ at low count rates. It was noted [22] that the distribution of brighter GRBs registered in Pioneer – Venus – Orbiter (PVO) experiment smoothly transits into the BATSE distribution.

BATSE results have most naturally been explained by the cosmological model for GRB origin [23], although there was a possibility of explaining their high isotropy and apparent inhomogeneity in spatial distribution by an expanded galactic halo model [24, 25]. However, after optical afterglows of GRBs had been discovered and high-redshift host galaxies of some GRBs had been identified, the galactic halo hypothesis for the whole class of GRBs was rejected, and we shall not discuss it further.

2.2 BeppoSAX results

The third period in the newest history of GRB studies was started in April 1996 when the Italian – Dutch satellite

[3] To study statistical properties of GRBS an equivalent differential test $\langle (C_{\min}/C_{\max})^{3/2} \rangle$ is often used, where $C_{\min}$ means the threshold count rate of the detector and $C_{\max}$ is the maximum count rate obtained in a given GRB. For homogeneous Euclidean source distributions $\langle (C_{\min}/C_{\max})^{3/2} \rangle = 0.5$.

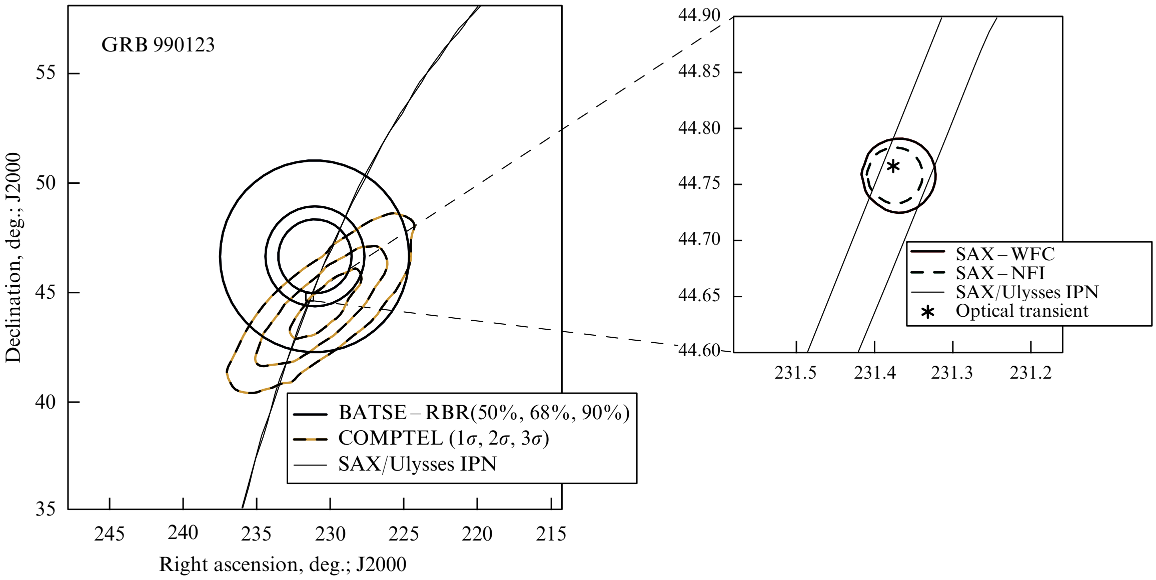

Figure 1. Detection of GRB990123 by CGRO (BATSE, COMPTEL), BeppoSAX (WFC and NFI), and Ulysses and the location of the associated optical transient.

BeppoSAX was launched. The gamma-ray burst detector onboard BeppoSAX has a few times lower sensitivity than BATSE all-sky monitor, but this satellite has one very important advantage. In addition to the gamma-ray detector GRBM (40 – 700 keV), it possesses three small X-ray telescopes: two wide-field cameras WFC (2 – 30 keV), and narrow-field spectrometers LECS (0.1 – 10 keV) and MECS (2-10 keV), commonly named NFI (Narrow Field Instruments). If the GRBM detects a GRB, WFC records are analyzed *post-facto* in order to find the simultaneous X-ray burst. If such a burst is found (the localization accuracy of the X-ray source amounts to several arc minutes), the NFI spectrometers observe this object and can detect X-ray fluxes as small as 10^{12} erg cm^{-2} s^{-1}. The entire process of pointing the X-ray telescopes takes about 8 hours. The position of the X-ray afterglow then is known to an accuracy of one arc minute. The process of registration of a GRB and its afterglow by different instruments is illustrated in Fig. 1 using the recent prominent gamma-ray burst of 23 January 1999 (GRB990123) [4] as an example.

The mean detection rate of GRBs by BeppoSAX satellite is about one burst per several months, and by the middle of 1999 about 17 GRBs had been detected, of them 14 were found to be associated with X-ray afterglows: GRB970111, GRB970228, GRB970402, GRB970508, GRB971214, GRB971227, GRB980329, GRB980425, GRB980515, GRB980519, GRB980613, GRB980703, GRB981226, GRB990123, and GRB990510 (all but the three last GRBs are referenced in Ref. [26]; the most recent information on GRB observations can be found on the GCN page http://gcn.gsfc.nasa.gov).

All observed X-ray afterglows are generally characterized by a *power-law* dependence of the flux on time $\sim t^{-1.3}$. Their spectra are fitted by a power-law with an exponential low absorption cut-off [5] with photon power-law index $\alpha_\Gamma \sim 2$.

[4] It is commonly accepted to denote GRBs by the date of their discovery.
[5] $dN/dE = \exp[-\sigma(E)N_H]N(E)^{-\alpha_\Gamma}$, where $\sigma(E)$ is the absorption cross-section, and N_H is the number of neutral hydrogen atoms along the line of sight.

Since the X-ray error box is small (of the order of one arc minute), deep optical observations by ground-based telescopes start immediately after the X-ray observations in order to identify the X-ray afterglow with some unusual optical object. As a rule, up to several tens of faint optical objects (mostly distant galaxies) fall into such an X-ray error box. The appearance of a new point-like optical source gives all grounds to believe that it is an optical afterglow associated with the GRB. It is in this way that optical afterglows were discovered for 10 out of 15 GRBs listed above: GRB970228, GRB970508, GRB971214, GRB980326, GRB980329, GRB980519, GRB980613, GRB980703, GRB990123, and GRB990510. In some cases (GRB970508, GRB980329, GRB980425, GRB980519, GRB980703, GRB981220 and GRB990123) variable radio afterglows were detected as well. The spectra of several GRB host galaxies obtained on the W Keck 10-m telescope enabled the redshift to be measured: $z = 0.692$ (GRB970228), $z = 0.835$ (GRB970508), $z = 3.42$ (GRB971214), $z = 1.0964$ (GRB980613) and $z = 0.966$ (GRB980703). The redshifts were directly measured in the optical afterglow spectra of GRB970508 ($z = 0.836$), GRB990123 ($z = 1.60$) and GRB990510 ($z = 1.619$). The distinctive feature of optical afterglows is also the power-law decay of the flux with time $\alpha t^{-1.1-1.4}$ and the non-thermal spectrum.

2.3 ROTSE and GRB990123

A genuine sensation at the end of January 1999 was the discovery of the extremely powerful GRB990123 by all the gamma-ray detectors onboard CGRO (BATSE, COMPTEL, OSSE) and by the GRBM detector on BeppoSAX (see Fig. 1).

The burst had a double-peak time profile with a complex fine structure and lasted about 100 s. During the burst, its gamma-ray spectrum became softer. The burst fluence in the energy range 20 keV – 300 MeV was $\sim 3 \times 10^{-4}$ erg cm^{-2}, i.e. it belongs to the top 0.3% of bursts registered by BATSE. It was simultaneously observed with WFC onboard BeppoSAX to have a maximum X-ray (1.5 – 26 keV) flux ~ 3.4 Crab with a total fluence $\sim 7 \times 10^{-6}$ erg cm^{-2}, a few percent of the gamma-ray fluence.

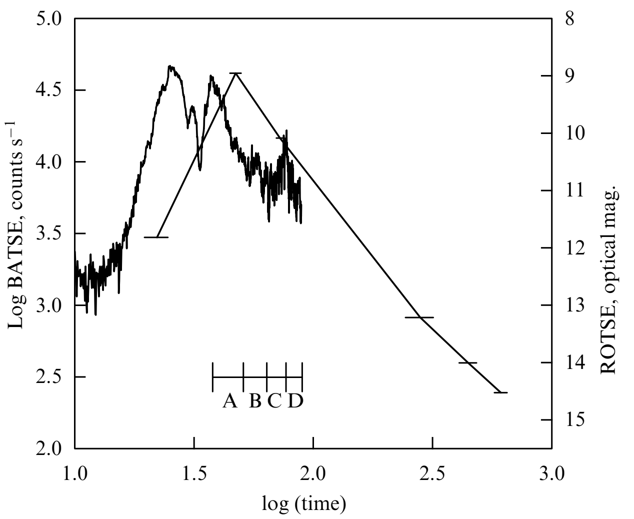

Figure 2. BATSE (50 – 300 keV) and ROTSE (in visual stellar magnitudes) light curves as a function of the logarithm of time (in seconds) after BATSE trigger #7343. The horizontal lines in the ROTSE points show the exposure time. Figure from Ref. [68].

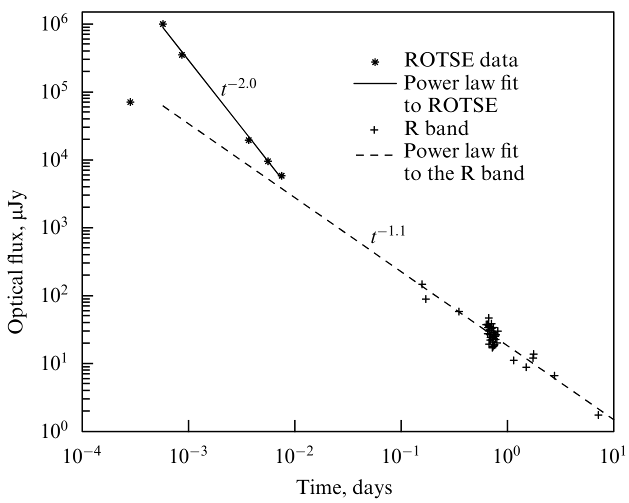

Figure 3. Optical R light curve of GRB990123. Figure from Ref. [67].

However, the most prominent event associated with this GRB was the *simultaneous* registration during the GRB itself of a very bright (up to ~ 9 stellar magnitude at maximum) optical afterglow. For the first time in the history of GRB studies we were able to observe the optical emission from the GRB itself (Fig. 2). This became possible because of using the alert system GCN (GRB Coordinate Network) to which many observatories are connected. This time, the monitor ROTSE (Robotic Optical Transient Search Experiment) in Los-Alamos National Laboratory was lucky in detecting the early optical afterglow [27].

ROTSE consists of four optical cameras 11.1 cm in diameter. The field of view is several degrees, the detection is made by CCD matrices. ROTSE pointed to the error box of GRB990123 in 22.18 seconds after the GRB trigger and took 5-s exposures every 25 seconds and then longer 75-s exposures for 10 minutes after the GRB beginning. Strikingly, the optical transient reached maximum $8^m.95$ (~ 1 Jy[6]) about 50 s after the beginning of the GRB and then started fading as $F_{\mathrm{op1}} \sim t^{-2.0}$. The optical fluence during this phase was $\sim 2.5 \times 10^{-7}$ erg cm^{-2}, i.e. about 7.7×10^{-4} of the gamma-ray fluence. The ROTSE optical light curve superposed on the BATSE gamma-ray light curve is shown in Fig. 2.

Observations of the localization region of GRB990123 by large optical telescopes which started 3 h 46 min after the burst revealed at this position a point-like source of 18th R-magnitude that faded by a shallower law $F_{\mathrm{op2}} \sim t^{-1.13}$ (Fig. 3). The optical spectrum of this source revealed the presence of absorption lines redshifted to $z \approx 1.6$. Two weeks after the burst, when the optical afterglow faded below $V \simeq 25^m$ (i.e. by almost 4×10^6 times!), observations from the Hubble Space Telescope revealed the presence of a faint host galaxy [28]. The galaxy has an irregular shape; the optical transient is $0''.6$ off-side the galaxy center. It is probably a pair of interacting galaxies.

The X-ray afterglow of this burst was very intense as well. 6 h after the burst the X-ray flux was about 1.1×10^{-11} erg cm^{-2} s^{-1} (about 0.8 μJy) and later on faded by a power law with index -1.1. The IR afterglow of the source was observed to display nearly the same behavior. The radio afterglow was registered by the VLA about a day after the burst at a level of 260 ± 32 μJy at the frequency of 8.46 GHz, and later on was not observed.

With a redshift $z = 1.6$ the isotropic energy release in hard X-rays and gamma-rays (20 – 700 keV) from this burst is $\sim 3 \times 10^{54}$ erg (assuming a flat Universe with the Hubble constant $H_0 = 60$ km s^{-1} Mpc^{-1}, the matter density $\Omega_0 = 0.3$, and the cosmological constant $\Omega_\Lambda = 0.7$), which already exceeds the solar rest-mass energy! The maximum gamma-ray luminosity in the burst reached 6×10^{53} erg s^{-1}. Even in the optical range a flux of 1 Jy from such distances indicates an isotropic luminosity of $\sim 10^{51}$ erg s^{-1}, much higher than the maximum known optical luminosities in supernova explosions ($\sim 10^{45}$ erg s^{-1}).

A hypothesis was put forward that this burst may have been gravitationally lensed by an intervening galaxy lying along the line of sight [29], however a detailed analysis [30] shows that the upper limit to possible amplification is $\mu < 60$. No weakly lensed burst has been discovered from this position in subsequent months, so the possible amplification was actually lower than 10.

Therefore, the present observational situation with GRBs can be summarized as follows.

(1) Cosmic GRBs are short (mostly 1 – 100 s) bursts of gamma-rays with energy 30 keV – 100 MeV with complex time profiles and a characteristic variability timescale of ~ 1 ms. The weakest GRBs registered by BATSE have a fluence $S \sim 10^{-7}$ erg cm^{-2}. The spectrum of GRBs is non-thermal, time-variable, with a maximum energy release in the 0.1 – 1 MeV range. Photons with energies as high as 1 GeV and more are usually observed from the brightest GRBs.

(2) GRBs are detected at a rate of ~ 1 per day at the sensitivity level of $\sim 10^{-7}$ erg cm^{-2}, their angular distribution over the celestial sphere is highly isotropic, while the distribution in space significantly deviates from a homogeneous population in Euclidean space (at small fluxes the deficit of weak bursts increases in comparison with a homogeneous Euclidean distribution)[7].

[6] The flux 1 Jy $= 10^{-23}$ erg cm^{-2} s^{-1}.

[7] As shown in Ref. [4], the distributions $\log N - \log S$ for different groups of GRBs separated by the duration and spectral properties have different slopes. Most GRBs (short and long/soft) follow the law $-3/2$ very well, while long hard GRBs most strongly deviate from this law.

(3) In almost all cases when GRBs were registered by the BeppoSAX satellite, X-ray afterglows were detected which faded by a power law $\sim t^{-1.3\pm0.1}$. In about half the cases optical afterglows were detected to fade by a constant power-law for several up to 200 days after the burst. Sometimes, variable radio emission was detected. In GRB990123 a bright early optical afterglow was discovered simultaneously with the GRB itself.

(4) For some GRBs host galaxies have been found. In five cases a measurement of the host galaxy redshift was available giving $0.7 < z < 3.14$. The redshift of the powerful GRB990123 was measured by absorption lines in the optical afterglow spectrum to be $z = 1.6$, so that the total isotropic energy release in gamma-rays for this burst is estimated as $\Delta E_\gamma \sim 3 \times 10^{54}$ erg, which is of order of the Sun's rest mass energy $M_\odot c^2 = 2 \times 10^{54}$ erg.

(5) In one case (GRB980425) the burst coincided in time and location with a peculiar supernova SN 1998 bw at a distance 40 Mpc. This may be not a coincidence.

(6) No gravitational lensing of GRBs has been discovered as yet.

3. Cosmological GRBs — relativistic fireballs

3.1 Compactness problem and the need for relativistic motion

To explain the enormous energy release (10^{51} erg for cosmological gamma-ray bursts) beyond the electron–positron pair creation threshold $E > 500$ keV over a short time $\Delta T \sim 1-10$ s, a hypothesis has been considered since the end of the 70s that gamma-ray bursts may result from expanding lepton–photon plasma fireballs [31–33]. However, for non-relativistic expanding fireballs (both at cosmological distances and in the Galactic halo) an insurmountable compactness problem arises which relates to the explanation of the observed nonthermal spectra of GRBs [34, 35].

In principle, thermal electrons with temperature T_e in an optically thick (relative to scattering) medium can produce power-law spectra up to energies $\sim 3kT_e$ by multiple scattering (see, e.g., Ref. [36]). However the observed power-law spectra from bright GRBs extends up to energies $\gtrsim 1$ GeV, so the electron temperature should be of the same order. But at such high temperatures non-thermal particle distributions are more likely. So hard GRB spectra are commonly considered non-thermal and thus are formed in an optically thin plasma. Here, however, we should note that power-law spectra can be obtained for optically thick relativistically expanding shells as well (see Refs [37, 38]).

Let us consider an energy release ΔE in a small volume. The characteristic time of GRB flux variability $\delta T \sim 10$ ms can be taken as an upper limit on the emitting region size $R_e < c\delta T \approx 3 \times 10^8$ cm (for non-relativistic case). Photons with energy E_γ will interact with lower-energy photons E_t and produce e^+e^- pairs provided that

$$E_\gamma > \frac{2(m_e c^2)^2}{E_t(1-\cos\chi)} \simeq \frac{4(m_e c^2)^2}{E_t \chi^2}\,, \tag{1}$$

where χ is the angle between the wave vectors of the photons. Let the fraction of such photons be f_p. Then the mean optical depth for pair photoproduction $\gamma\gamma \to e^+e^{-1}$ is

$$\tau_{\gamma\gamma} = f_p \frac{\Delta E \sigma_T}{4\pi R_e^2 m_e c^2} \approx 5 \times 10^{14} \left(\frac{\Delta E}{10^{51}\ \mathrm{erg}}\right)\left(\frac{\delta T}{10\ \mathrm{ms}}\right)^{-2}. \tag{2}$$

Here σ_T is the Thompson cross-section. A similar estimate can be obtained from the optical depth of pair photoproduction by very energetic photons ($E_\gamma > 1$ GeV), which are observed in GRB spectra, colliding with softer photons [39, 40].

The compactness problem is related to the very small size of the emitting region and in fact is due to our estimating this size by a non-relativistic formula. If the fireball expands with ultrarelativistic velocities, the compactness problem disappears. This is because in this case the estimate of the characteristic size of the emitting region using the observed variability time changes. Indeed, consider a shell of thickness R moving with a velocity v. Then in the laboratory frame the difference in the time of arrival of two photons emitted along the line of sight simultaneously from inner and outer parts of the expanding shell, δT, relates to the shell size by a simple (non-relativistic!) kinematic expression

$$\delta T = \frac{R}{v} - \frac{R}{c} = \frac{R}{c}\left(\frac{c}{v} - 1\right) = \frac{R}{c}\left(\frac{\Gamma}{\sqrt{\Gamma^2 - 1}} - 1\right) \simeq \frac{R}{2\Gamma^2 c}\,. \tag{3}$$

Here we introduced Lorentz-factor $\Gamma = [1 - (v/c)^2]^{-1/2}$and the last part of the equality holds for ultrarelativistic motion $\Gamma \gg 1$. Note that if the shell is optically thick and the observed variability is due to angular inhomogeneity of the emitting region, the relation between the characteristic variability time and the shell size for large Lorentz-factors is the same [36]. This is simply proved by the well-known fact that the radiation from a relativistically moving source is contained within a narrow angle $\theta \sim 1/\Gamma$.

For a relativistically moving medium the observed photon energy in the laboratory frame increases by Γ times relative to the photon energy in the comoving frame $h\nu_{\mathrm{obs}} \approx \Gamma h\nu_{\mathrm{em}}$. The fraction of photons with energy sufficiently high for pair creation decreases correspondingly. For power-law GRB spectra $\propto \nu^{-\alpha}$ the fraction of photons above the pair creation threshold in the proper source frame is thus $f_p^{(\mathrm{em})} = f_p^{(\mathrm{obs})}/\Gamma^{2\alpha}$. Substituting this relation and Eqn (3) into Eqn (2) we arrive at the proper optical thickness of the source

$$\tau_{\gamma\gamma}^{(\mathrm{em})} = \frac{\tau_{\gamma\gamma}^{(\mathrm{obs})}}{\Gamma^{4+2\alpha}}\,. \tag{4}$$

For the observed mean spectral photon indices $\alpha = 1.8-2$ [41] we get $\tau_{\gamma\gamma}^{(\mathrm{em})} < 1$ for $\Gamma \gtrsim 100$, so ultrarelativistic expansion of the source solves the compactness problem.

Note that similar bounds on the Lorentz-factor of the fireball expansion can be derived from demanding the free escape of high-energy gamma-quanta with $E_\gamma \gtrsim 1$ GeV. Indeed, as follows from Eqn (1), during a two photon interaction an electron–positron pair is not created if the angle between the photon wave vectors in the laboratory frame is $\chi \lesssim 2m_e c^2/\sqrt{E_t E_\gamma}$. For relativistic motion photon interaction is possible in principle for angles $\chi \lesssim 1/\Gamma$. Then for typical energies of target photons $E_t \sim 1$ MeV we obtain that only photons with energies

$$E_\gamma \lesssim 10^4\ (\mathrm{MeV})\left(\frac{\Gamma}{100}\right)^2\left(\frac{E_t}{1\ \mathrm{MeV}}\right) \tag{5}$$

can freely escape from the moving medium.

Therefore, the optically thin non-thermal spectra of GRBs and the observed high energy photons actually *demand* the presence of ultrarelativistic expansion of a photon – lepton plasma fireball [32, 33, 42]. This is a key feature in constructing cosmological models of GRBs. In particular, an important restriction on the baryon loading of the expanding fireball follows from here. The energy release $\Delta E_\gamma \gtrsim 10^{51}$ erg in gamma-rays during the short GRB time t_b much exceeds the Eddington luminosity $L_\gamma = \Delta E_\gamma/t_b \gg 10^{38}(M/M_\odot)$ erg s^{-1} for any reasonable burst durations and masses of possible astrophysical progenitors. This means that the radiation pressure will accelerate baryons up to limiting Lorentz-factors $\Gamma_{\lim} \sim (L_\gamma/L_{\rm Ed})^{1/3} \gg 1$ [8]. The most energetically economic assumption is that the kinetic energy of baryons is of the order of the energy of radiation and, moreover, it is the baryon kinetic energy conversion into radiation that gives rise to the observed gamma-ray emission[9]. As shown in Refs [43, 44], the presence of baryons in an optically thick photon–lepton fireball always leads to their acceleration and if there are too many of them (the fireball is heavily 'baryon contaminated'), the expansion velocity is non-relativistic $v \sim \sqrt{E/M_b} \ll c$ and the fireball remains optically thick. If the baryon contamination is low, the expansion can be ultrarelativistic with Lorentz-factor $\Gamma \sim E/(M_b c^2) \gg 1$, so that after reaching ultrarelativistic velocity the initially optically thick shell 'clears up'.

The assumption of the conversion of baryon kinetic energy into radiation underlies the model by Mészáros and Rees [5, 6] in which a GRB results from the interaction of a relativistic shock wave with the ambient interstellar (intergalactic) medium (the external shock wave model). It is also used in an alternative model of internal shocks [45, 7, 8], according to which a short GRB is generated by interaction of internal relativistic shock waves before the braking in the ambient medium has begun. Introducing the efficiency of baryon kinetic energy conversion into radiation $\epsilon_{b\gamma}$ we can write

$$M_b = \frac{E_\gamma}{\epsilon_{b\gamma} c^2 \Gamma} \approx \frac{5\times 10^{-6} M_\odot}{\epsilon_{b\gamma}} \left(\frac{E_\gamma}{10^{51}\ \text{erg}}\right)\left(\frac{\Gamma}{100}\right)^{-1}. \qquad (6)$$

The big energy release ($\sim 10^{51} - 10^{53}$ erg) can naturally be associated with low baryon contamination in some electrodynamical models (see Section 4).

3.2 Braking of the relativistic fireball in the ambient medium

At the initial stage the fireball expands adiabatically and the physical conditions change by the same laws as in the early Universe at the radiation-dominated stage [32, 44]. Let initially the rest mass of baryons in the fireball be much smaller than the radiation energy density $M_b \ll E_r(R_0) \sim T^4 R_0^3$, where T_0 is the initial temperature and R_0 is the initial radius. Energy conservation during the expansion in the laboratory frame implies $\Gamma(R)E(R) = E_0 = M_b c^2 + E_r(R_0) \sim T_0^4 R_0^3$. During adiabatic expansion $T_0 R_0 = T(R)R$ so we obtain the dependence of the Lorentz-factor increase at the stage of acceleration $\Gamma(R) \propto R$. Once the initial radiation energy has converted into the baryon kinetic energy the acceleration stops and the fireball expands with a constant Lorentz-factor $\eta = E_0/M_b c^2$ until the momentum is passed to the shocked ambient medium and the braking stage begins. In this respect the fireball expansion looks completely like supernova remnant evolution in the interstellar medium. However a specific of ultrarelativistic motion is that the momentum transfer occurs efficiently when the mass of gas swept up before the shock front is $\Delta M_{\rm ext} \approx M_b/\eta \ll M_b$ [10]. This occurs at a distance where the energy of the swept matter (in the laboratory frame) matches the initial energy of the fireball

$$\frac{4}{3}\pi R_{\rm dec}^3 (\eta n_0)(\eta m_p c^2) = E_0 \qquad (7)$$

(n_0 is the baryon density in the comoving frame so when passing to the laboratory frame an additional Lorentz-factor arises), so

$$R_{\rm dec} \approx 10^{17}\text{cm}\left(\frac{E_0}{10^{53}\ \text{erg}}\right)^{1/3}\left(\frac{n_0}{1\ \text{cm}^{-3}}\right)^{-1/3}\left(\frac{\eta}{100}\right)^{-2/3}. \qquad (8)$$

By that time the Lorentz-factor of the fireball decreases by about a factor of two. The corresponding dynamic time in the observer's frame $t_{\rm dec} \approx R_{\rm dec}/c\Gamma^2$ can be of the order of seconds for typical parameters, which may be of the same order as the duration of the GRB itself. However this external shock wave scenario meets difficulties in reproducing the complex structure of the GRB time profile [46], although it possibly can be applied to describe GRBs with smooth time profiles.

At present, another scenario is more popular in which a GRB is generated by interaction of internal shock waves that appear when relativistically moving shocks expelled from the central engine catch up with each other during a time period ΔT [7]. In this mechanism the time structure of the GRB directly reflects the impulse energy release by the central source generating relativistic shells which is the energy reservoir for the GRB. A fast shell with Lorentz-factor 2η catches up with a slower one with Lorentz-factor η, which has been expelled by the source a time interval δt before the fast one, at a distance $R_{\rm int} \sim 2\eta^2 c\delta t \sim 3\times 10^{14}(\text{cm})\delta t(\eta/100)^2$. Here a single gamma-ray pulse is generated. The characteristic time variability observed in the laboratory frame is then $t_{\rm var} \sim R_{\rm int}/2c\eta^2 \sim \delta t$ and is independent of the Lorentz-factor, and the observed GRB in the laboratory frame goes on for the time ΔT over which the central engine works. So the initial Lorentz-factor of relativistic ejection cannot be recovered from the analysis of the time profile of a GRB. This mechanism can reproduce very complex time profiles for GRBs [46, 47] and predicts the existence of an early optical afterglow of GRBs which was really observed from GRB990123 (see below).

3.3 Afterglows of GRBs

The kinetic energy transmitted to the ambient matter at the braking stage of the external shock wave is released as a soft afterglow of a GRB. On the front of a relativistic shock wave

[8] The power 1/3 appears due to the decrease of the energy of photons striking the moving electron and the decrease of the photon flux by Γ times. So one should take into account the factor $1/\Gamma^2$ in the expression for acceleration in the electron comoving frame (see Ref. [36], problem 4.4a).

[9] Another approach is to assume the domination of the baryon kinetic energy over the radiation energy $\Gamma M_b c^2 \gg \Delta E_\gamma^{\rm (obs)}$ (see, for example, Refs [37, 42]).Then the process becomes highly uneconomic energetically and demands a huge energy release, which is improbable.

[10] This relation, unobvious at first glance, can be derived straightforwardly from energy–momentum conservation written for a shell interacting with the ambient matter (see, e.g., [2] for more detail).

effective particle acceleration to high energies occurs and nonthermal electron distributions can be formed up to Lorentz-factors of $10^3 - 10^5$ [48]. Magnetic fields are always present in the swept gas so the stored kinetic energy is converted to radiation by the synchrotron mechanism, which is observed as soft X-ray and optical afterglows of GRBs.

Afterglows of GRBs generated by the interaction of a relativistic shock wave with the interstellar medium can be considered in full analogy with supernova remnants. In this problem the mechanism of the GRB itself is unimportant (like the supernova explosion mechanism). On time scales much larger than the energy release time, the time structure of the energy release is 'forgotten' and only what energy and momentum are passed to the ambient matter is important. This point enabled the observable effects of GRB afterglows to be calculated well before the first afterglow from GRB970228 had been discovered (see Ref. [11] and references therein)[11].

In the 'standard model' soft afterglows (X-ray, optical) of GRBs emerge due to synchrotron emission of electrons on the relativistic shock wave front braking in the interstellar medium. A power-law electron distribution $\propto \Gamma^p$ is formed in the shock front spanning the energy range $[\Gamma_{\min}, \Gamma_{\max}]$. The model has only two free parameters: (1) the ratio of the chaotic magnetic field energy to the total thermal energy of gas behind the shock front ϵ_B and (2) the fraction of the total thermal energy in electron chaotic motion ϵ_e (in other terms, the ratio of thermal energies of electron and proton components).

Using the relations between physical quantities on a relativistic shock wave front [50] and the equipartition parameters ϵ_B and ϵ_e, one can write down expressions for the hydrodynamic and magnetic parameters behind the front as functions of the time of observation t_s, Lorentz-factor Γ, and the ambient gas density n_1 [51]: the magnetic field strength

$$B = 4\Gamma\sqrt{2\pi\epsilon_B n_1 m_p c^2}\,, \tag{9}$$

electron Lorentz-factor

$$\gamma_e = 610\,\epsilon_e\Gamma\,, \tag{10}$$

and the maximum synchrotron frequency of a single electron (in the laboratory frame)

$$\nu_{\max} = \frac{eB}{m_e c}\gamma_e^2\Gamma \approx 10^{19}\ \mathrm{Hz}\left(\frac{\epsilon_e}{0.1}\right)^2\left(\frac{\epsilon_B}{0.1}\right)^{1/2} \times\left(\frac{\Gamma}{300}\right)^4\left(\frac{n_1}{1\ \mathrm{cm}^{-3}}\right)^{1/2}. \tag{11}$$

The synchrotron power emitted by one electron in the proper frame is

$$P_s = \frac{4}{3}\,\sigma_T c U_B \gamma_e^2\,, \tag{12}$$

where $U_B = B^2/8\pi$ is the magnetic field energy density, and σ_T is the Thompson cross-section. The cooling time of one electron in the proper frame is $t_c = \gamma_e m_e c^2/P_s$ and in the laboratory frame is Γ times shorter:

$$t_c = \frac{3m_e c}{4\sigma_T U_B \gamma_e \Gamma}\,. \tag{13}$$

From the last equation we can obtain the value of the Lorentz-factor γ_c to which the electron cools, and then substitute this expression into Eqn (11). Thus the characteristic frequency is obtained which is called the cooling frequency:

$$\nu_c \approx 10^{17}\ \mathrm{Hz}\left(\frac{\epsilon_B}{0.1}\right)^{-3/2}\left(\frac{\Gamma}{300}\right)^{-4}\left(\frac{n_1}{1\ \mathrm{cm}^{-3}}\right)^{-3/2}\left(\frac{t}{1\ \mathrm{s}}\right)^{-2}. \tag{14}$$

The synchrotron spectrum of electrons accelerated on the relativistic shock front will be determined by whether or not the synchrotron cooling of electrons can occur over the characteristic dynamical time, i.e. by the relation between frequencies $\nu_{\max}$ and ν_c. At the early stage of the fireball evolution (during the generation of GRB itself) for all typical parameters $\nu_c < \nu_{\max}$, the cooling is fast and all electrons cool instantaneously to $\gamma_c < \Gamma_{\min}$, so the spectrum consists of four power-law intervals with a maximum flux at ν_c and maximum energy release νF_ν at the frequency $\nu_{\max}$ (see review [2] for more detail).

At later stages (the afterglow formation) the electron cooling behind the external shock wave front is slow, $\Gamma_{\min} < \gamma_c$, only electrons with $\gamma_e > \gamma_c$ can cool down, and most electrons with $\gamma_e \sim \gamma_{\min}$ remain hot. The spectrum also consists of 4 power-law parts, however now the maximum of the emitted flux is at the frequency $\nu_{\max}$ while most energy is emitted at ν_c.

In both regimes the low-frequency end of the spectrum increases as $F_\nu \sim \nu^{5/2}$ or $F_\nu \sim \nu^2$ up to ν_a due to synchrotron self-absorption[12], then the growth becomes smoother $F_\nu \sim \nu^{1/3}$ reflecting the frequency dependence of a single electron synchrotron spectrum with energy $\gamma_e m_e c^2$. This growth continues up to the frequency ν_c (fast cooling) or $\nu_{\max}$ (slow cooling). At higher frequencies the spectrum decreases according to the cooling law of synchrotron emission $F_\nu \sim \nu^{-1/2}$ (fast cooling — up to the frequency $\nu_{\max}$) or $F_\nu \sim \nu^{-(p-1)/2}$ (slow cooling — up to the frequency ν_c). Beyond these frequencies the spectrum becomes steeper $F_\nu \sim \nu^{-p/2}$ [53].

With the shell expansion the spectrum shifts as a whole toward slower frequencies, so inside a fixed band of the detector the flux will decrease as a power-law function of time, with the power-law index changing when crossing the break frequencies ν_a, $\nu_{\max}$, ν_c.

The model assumes spherical symmetry and remains valid in the case of collimated emission into an opening angle θ until the Lorentz-factor of the expansion $\Gamma > 1/\theta$. After that the temporal behavior of the afterglow must change [54].

This simple model proved very successful in quantitative description of some GRB afterglows (GRB970228, GRB970508, etc.) [16]. Moreover, the time history of afterglows at different frequencies were used to put constraints on some physical parameters of GRBs and the ambient medium (n_0, ϵ_B, ϵ_e, and p [14, 55]). Their values were found to be quite

[11] Note that the effects of a powerful GRB with energy 10^{51} erg on the interstellar medium were first considered by Bisnovatyĭ-Kogan and Timokhin [49].

[12] The quadratic dependence on frequency appears when the self-absorption is produced by low-energy electrons which radiate efficiently at higher energies (see Refs [9, 52]).

reasonable ($\epsilon_e \sim 0.6$, $\epsilon_B \sim 0.1$, $n_1 \sim 5$ for GRB970508 [55, 56]). That the spherically-symmetric model well described the optical afterglow of GRB970508 on timescales as long as 200 days until its flux had become indiscernible against the background of the host galaxy has been used as an indication of the quasi-sphericity of this GRB. However, it is difficult to reproduce an increasing optical flux of the afterglow of GRB970508 at early phases in the framework of the standard model [56]. The slow-down of the X-ray afterglow of GRB981226 also cannot be described by the simple external shock wave model [57].

In addition to the possible asymmetry of the GRB itself, inhomogeneities of the ambient gas can lead to appreciable deviations from the standard model. As an alternative possibility of the afterglow effect we considered the heating of a close normal stellar atmosphere by a powerful GRB (the so-called 'mini-supernova' model [58]). Such a situation appears when a GRB occurs in a close binary system, for example as a result of the accretion-induced collapse of a white dwarf into a neutron star [59], or the recently proposed electrodynamical model [60]. The heating of the external layers of the star is produced by hard gamma-quanta or relativistic particles (in both cases the absorption coefficient is about $\kappa \sim 2$ g cm^{-2}). The external parts of the star begin rapidly expanding and gradually radiating the stored thermal energy. Specifically, we numerically calculated the effect of the deposition of a thermal energy $\Delta E = 10^{50}$ erg into the outer $10^{-3} M_{\odot}$ layers of a red giant with radius $4000 R_{\odot}$. The calculations were done using the non-relativistic hydrocode 'STELLA' with account of radiation transfer [61]. The optical light curve obtained well describes the first 20 days of the power-law fading of the afterglow of GRB970228 (Fig. 4). If a star with a smaller radius is illuminated by the same energy, the expansion velocity of the envelope can become relativistic. In this case the GRB 'induces' relativistic expansion of a small quantity of baryons ($\sim 10^{-5} M_{\odot}$) and the relativistic shock wave which forms in the interstellar medium will produce afterglow effects similar to those described above in the relativistic fireball model. This result demonstrates the possible diversity of optical afterglow formation mechanisms.

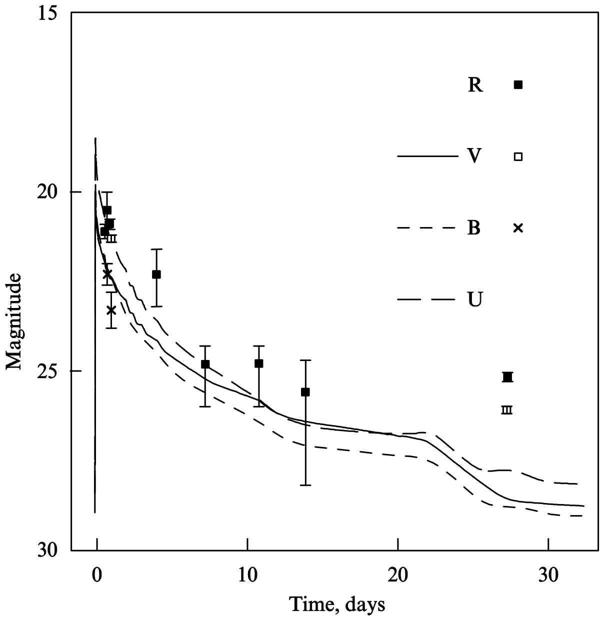

Figure 4. Optical U, B, V, R light curves of the afterglow GRB970228 in the mini-supernova model. In this model, a thermal energy of 10^{50} erg is deposited into the external $10^{-3} M_{\odot}$ layers of a red supergiant with radius $4000 R_{\odot}$. The assumed distance is 1 Gpc. The observed B, V, R-magnitudes are taken from Ref. [62]. This figure is from Ref. [58].

To date, 14 low-energy afterglows were observed from the 17 GRBs registered by Beppo-SAX. The lack of notable afterglows from some powerful GRBs (for example, from GRB970111) may indicate the unusual conditions in the GRB surroundings, for example, a very low density of the ambient medium. Such a situation is feasible in the model of coalescing neutron star binaries, which can fly out of the galactic plane to large distances over the $\sim 10^9$ years after their formation, or even are relics of early star formation in gas-poor elliptical galaxies. This effect was calculated in paper [63], in which the fraction of elliptical galaxies among the 'host galaxies' of GRBs is estimated to be about 10 – 15%. Thus the observed morphology of GRB host galaxies can be used as an independent test of various GRB models.

3.4 GRB990123 and the early optical afterglow of GRBs

The powerful GRB990123 (see below) proved to be a genuine 'rosetta stone' for the relativistic fireball model. The model became especially popular in 1997 shortly after the discovery of the optical afterglow from GRB970228 because the effect of afterglow and its main properties were predicted by this model before the first afterglow had been detected [11]. A new success of the model was the prediction of early optical afterglows from long GRBs with possible time overlapping between gamma- and optical emission, which was made by Sari and Piran literally two weeks before its detection in GRB990123 [64, 65]. Sari and Piran assume that the gamma-ray emission itself is generated by interacting internal shock waves at the fireball center. In long GRBs (with a duration of tens or hundreds of seconds), a relativistic wind from the central engine generates, during the GRB time t_{γ}, shells which catch up with each other and form a relatively extended quasi-homogeneous shell with a thickness $\Delta \sim t_{\gamma} c$ moving with a bulk Lorentz-factor Γ. The interaction of the forward shock from this shell with the interstellar medium produces late X-ray and optical afterglows. As was shown in Ref. [66], gamma-ray emission in such bursts from internal shocks (also forming the thick shell) and softer emission from the external shock (interacting with the ambient medium) can overlap in time.

The early optical afterglow appears in the reverse shock moving inside the shell Δ from the external shock front. This wave heats the gas and accelerated electrons in the shell which then cool down due to synchrotron emission and adiabatic expansion. In contrast to the external shock, which is later braked in the interstellar medium and gives rise to a prolonged optical afterglow, the reverse shock radiates over a relatively short time interval comparable with the duration of the GRB $t_A \sim \min(t_{\gamma}, \Delta/c)$. After reaching a maximum (when the reverse shock intersects the internal shell boundary) new electrons are not accelerated any more, hence no electron emission is produced beyond the frequency ν_c, and the cooling frequency itself rapidly decreases due to adiabatic expansion.

The energy of the reverse shock is comparable to that of the forward shock wave (i.e. is of the order of the energy of the GRB itself), however the temperature in the reverse shock is much lower, by a factor of $\Gamma_A^2/\Gamma_0 \sim \Gamma_0 \gg 1$ (the indices indicate to which moment the Lorentz-factor relates). The maximum of synchrotron emission of electrons in the observer's reference frame is proportional to the square of the energy of the chaotic electron motions $E = \gamma_e m_e c^2$ (i.e. to the temperature), the magnetic field strength in the shell B and its bulk Lorentz-factor (blue shift), $\nu_{\max} \sim \gamma_e^2 B\Gamma$. Therefore if the external wave front, by forming the shell, generates gamma-ray emission with MeV energies, the reverse shock wave mostly radiates at energies from 1 MeV$/\Gamma^2$ to a few keV, i.e. in the optical range.

The expected maximum of the early afterglow was predicted at a level of 7th stellar magnitude, and was actually observed at $8^m.95$; the time of the optical maximum was predicted to occur from 30 to 50 s after the GRB beginning and was actually observed 45 s after the beginning — not too bad a coincidence for such a crude model! Applying this model the observed early optical afterglow from GRB 990123 showed a good coincidence between the theory and observations [67]. The evaluated initial Lorentz-factor $\Gamma_0 \sim 200$ corresponds to values needed for free escape of hard gamma-ray emission (1) [39]. So there is strong confirmation of the simultaneous existence in this GRB of internal shocks and external shock wave braking in the ambient gas.

An independent analysis of the temporal properties of the observed gamma-ray emission [68] showed that the width of individual gamma-ray peaks does not change during the burst. If the gamma-ray emission were generated at the external front of the expanding envelope, the width of later peaks should be larger. Indeed, if $\Delta t'$ is the width of a pulse as measured by a clock moving with a Lorentz-factor $\Gamma \gg 1$ together with the shell, the observed duration of the pulse in the laboratory frame is

$$\Delta T = \Gamma(1 - \beta\cos\theta)\Delta t' \,, \qquad (15)$$

where θ is the angle between the observer and the velocity vector. Relativistic aberration makes it possible to observe only parts of the envelope lying inside the narrow angle $\theta \sim 1/\Gamma$. Substituting $\theta = 0$ and $\theta = 1/\Gamma$ into Eqn (15) we find that the individual gamma-ray pulse width at the end of the burst must be about two times larger than at the beginning: $\Delta T(\theta = 1/\Gamma) \simeq 1/\Gamma = 2\Delta T(\theta = 0)$. This is not the case (and not only in this burst!), so in the only kinematically available model of gamma-ray emission generation on the external shock front, the Lorentz-factor should remain constant over the total duration of the GRB and the emission should come from the region $\theta \ll 1/\Gamma$ [69]. However, in the case of GRB990123 the optical emission reaches maximum and starts power-law fading already during the GRB itself, which is indicative of the beginning of the external shock wave braking [$\Gamma(t) \propto t^{-3/8}$ for a Blandford – McKee self-similar solution]. So, based only on purely kinematic considerations, the conclusion can be made that in this burst the gamma-ray emission was generated in the central region irrespective of the specific model. However, in spite of the obvious success of the internal shock model [67], it would be premature to accept its universal character. For example, in this model the Lorentz-factors of individual shells must differ by a factor of 2, which would reflect on the widths of generated gamma-ray peaks, which is not observed.

At last, the huge energy release from this source 3×10^{54} erg (assuming spherical symmetry) aroused new interest in the problem of possible gravitational lensing of GRBs [29]. If this were the case, a less intensive GRB with a similar time profile would be observed from the same region [30] after a few weeks. No such burst has been detected.

4. Possible astrophysical sources of cosmological GRBs

As seen from above, the relativistic fireball model explains sufficiently well many observed features of both the GRB itself and especially its afterglow at low energies. In order to produce such a fireball, a small quantity of baryons ($\sim 10^{-5} M_\odot$) should be accelerated up to relativistic velocities with Lorentz-factors $\Gamma \gtrsim 100$. This requirement puts stringent constraints on the possible models. An energy release of the order of the solar rest mass energy does not violate any physical law. The main problem is in converting a notable fraction of this energy into the bulk kinetic energy of a small quantity of relativistic particles which then radiates away in the form of hard electromagnetic emission. In addition, the observed durations of GRBs of 10 – 100 s and the complex time profiles apparently reject an explosive instantaneous energy release (as in the external shock wave model), so realistic models of GRBs must provide a prolonged variable energy release (as in the scenario of internal shock waves). The last point is non-trivial: the size of the emitting region as evaluated from the observed variability time scale is of the order of $R_e \sim t_{\rm var} c \sim 10^7$ cm, while the energy release takes millions characteristic light times R_e/c [70].

Thus, the requirements on modern models of cosmological GRBs can be summarized as follows:

(1) electromagnetic energy release $\sim (0.1-1) M_\odot c^2$;

(2) duration of energy release 10 – 100 s, possibly more;

(3) characteristic variability time scale $\lesssim 10$ ms;

(4) formation rate in a typical galaxy of about one per $\sim 10^6$ years;

(5) small quantity of accelerated baryons $\sim (10^{-5} - 10^{-6}) M_\odot$ (additional requirement to form relativistic fireball).

Declining spherical symmetry decreases the energy required by a factor $\theta^2/4$, by increasing by the same factor the formation rate per galaxy. However, as we already mentioned, at present it is impossible to reliably estimate the degree of anisotropy of gamma-ray emission in most GRBs, and all known prolonged afterglows are satisfactorily described by the spherically-symmetric model.

A lot of possible astrophysical models have been suggested, and they can be subdivided into several classes depending on the source of energy which is converted into electromagnetic radiation. The first group comprises the models in which the energy available is the gravitational energy of the interaction of matter with a compact star (black hole). Essentially, to this group belongs the model of binary compact stars (double neutron star or neutron star and black hole) coalescing due to gravitational wave emission [71 – 73].

The total binding energy of a disk with mass $M_{\rm d}$ around a black hole lies within the range from 6 to 42% of the disk rest mass energy $M_{\rm d} c^2$ depending on the angular momentum of the black hole $J = aGM_{\rm bh}^2/c$. Introducing the coefficient η_γ of the gravitational to electromagnetic energy conversion we obtain an estimate of the available electromagnetic energy in these models:

$$\Delta E_\gamma \simeq 8 \times 10^{53}\,(\text{erg})\,\eta_\gamma \left(\frac{M_{\rm d}}{M_\odot}\right) \qquad (16)$$

(extreme Kerr black hole, $a = 1$);

$$\Delta E_\gamma \simeq 1.2\times 10^{53}\,(\mathrm{erg})\,\eta_\gamma\left(\frac{M_\mathrm{d}}{M_\odot}\right) \qquad (17)$$

(Schwarzschild black hole, $a=0$).

The conversion of the released gravitational energy to electromagnetic energy is possible, for example, due to the transfer of heat generated in the disk into the relativistic fireball through neutrino – antineutrino annihilation into electron – positron pairs $\nu\tilde{\nu}\to e^+e^-$. The mechanism is effective when the neutrinos can annihilate before crossing black hole horizon. Recent numerical calculations [74, 75] show that this mechanism can produce a fireball with energy $10^{51}-2\times10^{52}$ erg. An alternative conversion mechanism is related to magnetohydrodynamic processes in the disk [73, 76, 77], which result in transferring the thermal energy generated in the disk into the flux of relativistic particles.

Another class of models uses the rotational energy of a black hole as the energy reservoir. Examples are provided by the 'failed supernova' model [78] or 'hypernova' model [79]. In this case it is possible to extract up to 29% of the rest-mass energy of the black hole $M_\mathrm{bh}c^2$. It is astrophysically warranted that a black hole resulting from the coalescence of a binary neutron star or neutron star and black hole will be rapidly spinning. The black hole rotational energy is converted into the flux of relativistic particles during the magnetohydrodynamic interaction of the black hole with the surrounding disk of matter (Blandford – Znajek mechanism [80]). The extracted energy is

$$\Delta E_\gamma \sim \eta_\gamma f(a) M_\mathrm{bh} c^2\,, \qquad (18)$$

where η_γ is the MHD-efficiency, $a\equiv Jc/GM_\mathrm{bh}^2$ is the Kerr parameter which is equal to 1 for a maximally spinning black hole and the function

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

is small for small a $\left[f(a)\sim a^2/8\right]$ and has a sharp maximum $f(1)\approx 0.29$. The maximum energy of the GRB in this case can be as high as

$$\Delta E_\gamma^\mathrm{max}\sim 5\times10^{53}(\mathrm{erg})\,\eta_\gamma\left(\frac{M_\mathrm{bh}}{M_\odot}\right). \qquad (19)$$

Black holes formed from coalescing binary neutron stars have masses $2.5M_\odot$ [81] and rapid rotation $a\sim1$, so it can be expected that $\Delta E_\gamma^\mathrm{max}\sim10^{54}$ erg. If a black hole results from the core collapse of a massive star, its mass may be higher, $\sim 10M_\odot$ [82]. If the black hole does not rotate initially, the expected Kerr parameter after the coalescence of a neutron-star – black hole pair is $a\sim M_\mathrm{ns}/M_\mathrm{bh}<0.1-0.3$, so the large mass is compensated by a low efficiency of the Blandford – Znajek process. However, if a superstrong magnetic field is present in the disk, the efficiency of the energy release again reduces to estimates (16) or (17) with a maximum luminosity determined by an equation like Eqn (21) given below.

The energy 10^{54} ergs is also attainable in the model of a hypernova by Paczyński [79]. In this model, an extremely rotating black hole with a mass of about $10M_\odot$ is formed in the core collapse of a very massive star in a close binary system (the binarity is needed for the black hole to rapidly rotate). For high energy release efficiency it is assumed that a superstrong magnetic field 10^{15} G exists near the black hole. This field can be maintained by the surrounding gaseous disk. The maximum luminosity in the Blandford – Znajek mechanism can be estimated as (e.g., see Eqn (4.5) in Ref. [83])

$$L_\mathrm{max}^\mathrm{BZ}\sim10^{51}(\mathrm{erg\,s^{-1}})\left(\frac{B}{10^{15}\ \mathrm{G}}\right)^2\left(\frac{M_\mathrm{bh}}{10M_\odot}\right)^2. \qquad (20)$$

The black hole rotational energy is passed to the kinetic energy of the envelope by the magnetic field, with the outermost parts of the envelope expanding to relativistic velocities and producing a fireball. Since the lifetime of very massive stars is about several million years, hypernovae must be born at the sites of intense star formation and absent in galaxies with low star formation rates (for example, in ellipticals).

The model became popular after the host galaxies of some GRBs (GRB970508, GRB971214, GRB980703) were identified with galaxies with active star formation. For example, the lower limit to the star formation rate in the host galaxy of GRB980703 at redshift 0.966 is evaluated from spectroscopical observations to be about $7M_\odot$ per year (i.e. by an order of magnitude higher than in our Galaxy) [84].

The strong magnetic field associated with rapid rotation of the compact object is capable of producing electromagnetic luminosities up to

$$L_\mathrm{max}^\mathrm{em}\sim B^2R^3\omega\sim6\times10^{51}\ (\mathrm{erg\ s^{-1}})\times\left(\frac{B}{10^{15}\ \mathrm{G}}\right)^2\left(\frac{R}{10^6\ \mathrm{cm}}\right)^3\left(\frac{P}{1\ \mathrm{s}}\right)^{-1}. \qquad (21)$$

Here R is the characteristic radius, and P is the period of rotation. Using such estimates, a GRB origin during the formation of a millisecond pulsar was proposed [85] and the consequences of binary neutron star coalescence were analyzed [86]. The maximum energy available in such processes is about the binding energy of the neutron star $\sim0.15Mc^2\sim3\times10^{53}(M/M_\odot)$ erg. An interesting GRB mechanism close to that proposed by Usov [85] has recently been proposed by Spruit [60]. It is based on the recently discovered secular r-mode instability (Rossby waves) in rapidly rotating neutron stars emitting gravitational waves [88]. According to this model, a weakly magnetized neutron star (the surface field $B\sim10^7$ G) is spun-up by accretion in a binary system down to periods 1 – 3 ms. The secular instability starts and gravitational radiation brings away angular momentum which is supplied from the disk. The neutron star can start rotating differentially and in a time period of 1 month its internal azimuthal magnetic field can increase up to $\sim10^{17}$ G. When such a field floats up to the surface, the surface magnetic field can increase up to 10^{16} G over a time scale of the order of seconds, giving rise to a GRB with a duration of 1 – 100 s [85]. The attractiveness of electrodynamical models stems from their ability to produce fireballs with low baryon contamination.

Finally, S I Blinnikov [89] recently put forward a hypothesis that GRBs may be indicators of a physical process occurring in the 'mirror world' of elementary particles. According to this idea, mirror neutrinos (which are sterile for ordinary matter) are born during collapses or mergings of stars made of 'mirror' material, which interacts with ordinary matter only through gravitation. Due to neutrino oscillations some fraction of the sterile neutrinos are converted into ordinary ones. The annihilation or decay of these neutrinos generates an electron – positron plasma and a relativistic fireball with low baryon contamination, as required to produce a GRB.

If the energy released in GRB990123 was indeed $\sim 3 \times 10^{54}$ erg and the maximum luminosity was 6×10^{53} erg s^{-1}, all known mechanisms for generation of gamma-ray emission by *known* physical processes meet difficulties and must involve extreme conditions, such as large black hole masses, superstrong magnetic fields $B \gtrsim 10^{16}$ G, etc. The beaming factor of gamma-ray emission reduces the energetics required, but it is an additional free parameter. If the observed change in the law of fading of the afterglow of GRB990123 is explained by the jet [87], the beaming factor can be of order 1/100. Such a beaming factor does not contradict the possible high formation rate of binary neutron stars ($\sim 1/10^6$ yr^{-1} per Galaxy). However, it seems absolutely clear that in the present situation with GRBs we have met with the problem of explaining a huge electromagnetic energy release, and the solution of this enigma will undoubtedly have a significant effect on the development of high energy physics and the physics of elementary particles.

5. Conclusion

Observations of X-ray and optical afterglows of GRBs strongly proves the cosmological origin of most (all ?) GRBs. No concurrent hypothesis has been so far suggested to explain the afterglows in the model of GRB originating in the Galactic halo or inside the Galactic disk.

In GRBs located at distances of several Gigaparsecs an energy of $10^{51} - 3 \times 10^{54} \sim M_{\odot}c^2$ erg only in the form of hard electromagnetic emission is released over a time period of several seconds producing luminosities up to 6×10^{53} erg s^{-1}. This luminosity is about $2 \times 10^{-6} c^5/G$ (the 'fundamental luminosity' $c^5/G = Mc^2/(r_g/2c)$, where $r_g = 2GM/c^2$ is the gravitational radius for a mass M). This energy is several orders of magnitude larger than all known to date energies released in cataclysmic astrophysical processes.

Ultrarelativistic motion with Lorentz-factors $\Gamma > 100$ s appears with necessity in the cosmological GRB model. Such huge velocities have never been observed so far in astrophysics. In the most elaborated models for cosmic GRBs (the relativistic fireball model) the kinetic energy of the relativistically expanding photon – lepton plasma with a low addition of baryons ($\sim 10^{-6} M_{\odot}$) is converted into gamma-rays by collisions of internal shock waves with each other or during the interaction of the fireball with the ambient medium. The last process leads to braking of the fireball, and the bulk kinetic energy is converted to the thermal energy of electrons behind the shock front. This heat is then radiated away by the synchrotron mechanism and via inverse Compton scattering at lower energies thus forming the observed X-ray, optical, and radio afterglows. In fact, the observed afterglows were predicted by the relativistic fireball model.

In spite of the rapidly increasing information and detail of knowledge of GRB properties and their afterglows and, hence, growing difficulties in explaining all the known features of GRBs by one mechanism, relativistic fireballs remain an obligatory element of any reasonable cosmological GRB model. An elegant explanation is required for why such a low mass of baryons is accelerated up to large Lorentz-factors and what is the reason for their acceleration. From the observational point of view it is required to increase the statistics of GRB afterglows to check whether GRBs are observed in gas-poor elliptical galaxies with a low star formation rate. If yes, some models (for example, hypernovae) loose their universality.

During the generation of a cosmological GRB we meet physical conditions which by their characteristics are similar to those in the early Universe. Whatever such GRBs ultimately are, they must be associated with ultra high energy particle generation and their interaction with the ambient medium. Undoubtedly, studies of the physics of such processes in these natural laboratories will be one of the important fields in the astrophysics of the beginning of the XXI century.

The author acknowledges V L Ginzburg for encouragement in writing this paper, S I Blinnikov, V S Beskin and the staff of the relativistic astrophysics department of the Sternberg Astronomical Institute for discussions and useful notes.

This work is partially supported by the grants RFBR 98-02-16801, INTAS 96-0315, NTP 'Astronomiya' (projects 1.4.4.1 and 1.4.2.3), and 'Universities of Russia' No. 5559.

References

1. Luchkov B I, Mitrofanov I G, Rozental' I L *Usp. Fiz. Nauk* **167** 743 (1996) [*Phys. Usp.* **39** 695 (1996)]
2. Piran T *Phys. Rep.* **314** 575 (1999); preprint astro-ph/9810256
3. Kulkarni S et al. *Nature* (London) **395** 663 (1998)
4. Tavani M *Astrophys. J.* **497** 21 (1998)
5. Rees M J, Mészáros P *Monthly Not. RAS* **258** 41P (1992)
6. Mészáros P, Rees M J *Astrophys. J.* **405** 278 (1993)
7. Rees M J, Mészáros P *Astrophys. J.* **430** L93 (1994)
8. Paczyński B, Xu G *Astrophys. J.* **427** 709 (1994)
9. Katz J I *Astrophys. J.* **422** 248 (1994)
10. Sari R, Piran T *Astrophys. J.* **455** L143 (1995)
11. Mészáros P, Rees M J *Astrophys. J.* **476** 232 (1997)
12. Vietri M *Astrophys. J.* **478** L9 (1997)
13. Tavani M *Astrophys. J.* **438** L87 (1997)
14. Waxman E *Astrophys. J.* **489** L33 (1997)
15. Reichart D E *Astrophys. J.* **485** L99 (1997)
16. Wijers R A M J, Rees M J, Mészáros P *Monthly Not. RAS* **288** L51 (1997)
17. Kouveliotou C et al. *Astrophys. J.* **510** L115 (1999)
18. Hurley K et al. *Nature* (London) (1999) (in press); preprint astro-ph/9811443
19. Briggs M S et al. *Astrophys. J.* **459** 40 (1996)
20. Meegan C A et al. *Astrophys. J. SS* **106** 65 (1996)
21. Kommers J M et al. *Astrophys. J.* (1999) (in press)
22. Fenimore E E et al. *Nature* (London) **366** 40 (1993)
23. Usov V V , Chibisov G V *Astron. Zh.* **52** 192 (1975) [*Sov. Astron.* **19** 115 (1975)]
24. Shklovskij I S, Mitrofanov I G *Monthly Not. RAS* **212** 545 (1985)
25. Briggs M S et al., in *Gamma-Ray Bursts* (AIP Conf. Proc., Vol. 384, Eds C Kouveliotou, M S Briggs, G J Fishman) (Woodbury, N.Y.: AIP Press, 1996) p. 335
26. Owens A et al. *Astron. Astrophys.* **339** 370 (1998)
27. Akerlof C W et al. *GCN* 205 (1999); *IAUC* 7100 (1999); submitted to *Nature* (London)
28. Fruchter A et al. *GCN* 255 (1999)
29. Djorgovski S G et al. *GCN* 216 (1999)
30. Blandford R D, Helfand D J *Monthly Not. RAS* (1999) (in press); preprint astro-ph/9902004
31. Cavallo G, Rees M J *Monthly Not. RAS* **183** 359 (1978)
32. Goodman J *Astrophys. J.* **308** L47 (1986)
33. Paczyński B *Astrophys. J.* **308** L43 (1986)
34. Ruderman M *Ann. N.Y. Acad. Sci.* **202** 164 (1975)
35. Schmidt W K H *Nature* (London) **271** 525 (1978)
36. Rybicki G B, Lightman A P *Radiative Processes in Astrophysics* (London: John Whiley & Sons, 1979)
37. Rozental I L, Belousova I V *Astrophys. Space Sci.* **249** 117 (1997)

38. Blinnikov S I, Kozyreva A V, Panchenko I E *Astron. Zh.* (1999) (in press)
39. Fenimore E E, Epstein R I, Ho C *AAS* **97** 59 (1993)
40. Harding A K, Baring M G, in *Gamma-Ray Bursts* (AIP Conf. Proc., Vol. 307, Eds G Fishman et al.) (Woodbury, N.Y.: AIP Press, 1994) p. 520
41. Schaefer B E et al., in *Gamma-Ray Bursts* (AIP Conf. Proc., Vol. 307, Eds G J Fishman, J J Brainerd) (Woodbury, N.Y.: AIP Press, 1994)
42. Krolik J H, Pier E A *Astrophys. J.* **373** 277 (1991)
43. Paczyński B *Astrophys. J.* **363** 218 (1990)
44. Shemi A, Piran T *Astrophys. J.* **365** L55 (1990)
45. Narayan R, Paczyński B, Piran T *Astrophys. J.* **395** L83 (1992)
46. Sari R, Piran T *Astrophys. J.* **485** 270 (1998)
47. Panaitescu A, Mészáros P *Astrophys. J.* (in press); preprint astro-ph/9810258
48. Bykov A M, Mészáros P *Astrophys. J.* **461** L37
49. Bisnovatyĭ-Kogan G S, Timokhin A N *Astron. Zh.* **74** 483 (1997) [*Astron. Rep.* **41** 423 (1997)]
50. Blandford R, McKee C F *Phys. Fluids* **19** 1130 (1976)
51. Sari R, Piran T, Narayan R *Astrophys. J.* **497** L41 (1998)
52. Katz J I, Piran T *Astrophys. J.* **490** 772 (1997)
53. Mészáros P, Rees M J, Wijers R A M J *Astrophys. J.* **499** 301 (1998)
54. Rhoads J E *Astrophys. J.* **497** L1 (1997)
55. Wijers R A M J, Galama T J *Astrophys. J.* (in press); preprint astro-ph/9805341
56. Granot J, Piran T, Sari R, preprint astro-ph/9808007
57. Frontera F et al. *IAUC* 7078 (1998)
58. Blinnikov S I, Postnov K A *Monthly Not. RAS* **293** L29 (1998)
59. Dar A et al. *Astrophys. J.* **388** 164 (1992)
60. Spruit H C *Astron. Astrophys.* **341** L1 (1999)
61. Blinnikov S I et al. *Astrophys. J.* **496** 454 (1998)
62. Galama T et al. *Nature* (London) **387** 479 (1997)
63. Lipunov V M, Panchenko I E, Postnov K A, Prokhorov M E *Astron. Astrophys.* (1999) (in press)
64. Sari R, Piran T *Proc. Astron. Astrophys.* (1999) (in press); preprint astro-ph/9901105
65. Sari R, Piran T *Astrophys. J.* **520** 641 (1999); preprint astro-ph/9901338
66. Sari R *Astrophys. J.* **489** L37 (1997)
67. Sari R, Piran T *Astrophys. J. Lett.* **517** L109 (1999); preprint astro-ph/9902009
68. Fenimore E E, Ramirez-Ruiz E, Wu B *Astrophys. J. Lett.* (1999) (in press); preprint astro-ph/9902007; LA-UR-99-529
69. Ramirez-Ruiz E, Fenimore E E *Astron. Astrophys. Suppl.* (in press); preprint astro-ph/9812426
70. Katz J I, Piran T, Sari R *Phys. Rev. Lett.* **80** 1580 (1998)
71. Blinnikov S I et al. *Sov. Astron. Lett.* **10** 177 (1984)
72. Eichler D et al. *Nature* (London) **340** 126 (1989)
73. Paczyński B *Acta Astron.* **41** 257 (1991)
74. Ruffert M et al. *Astron. Astrophys.* **319** 122 (1997)
75. Popham R, Woosley S E, Fryer C *Astrophys. J.* (1999) (in press); preprint astro-ph/9807028
76. Mészáros P, Rees M J *Astrophys. J.* **482** L29 (1997)
77. Katz J I *Astrophys. J.* **490** 633 (1997)
78. Woosley S E *Astrophys. J.* **405** 273 (1993)
79. Paczyński B *Astrophys. J.* **499** L45 (1998)
80. Blandford R D, Znajek R L *Monthly Not. RAS* **179** 433 (1977)
81. Ruffert M, Janka H-T *Astron. Astrophys.* **338** 535 (1998)
82. Postnov K A, Prokhorov M E, in *Proc. Rencontres de Moriond* "*Gravitational Waves and Experimental Gravity*" (Ed. Tran Than Van) (Paris: Edition Frontières, 1999) (in press)
83. McDonald D A et al., in *Black Holes: the Membrane Paradigm* (Eds K S Thorne, R H Price, D A Macdonald) (New Haven: Yale Univ. Press, 1986)
84. Djorgovski S G et al. *Astrophys. J.* **508** L17 (1998)
85. Usov V V *Nature* (London) **357** 472 (1992)
86. Lipunova G V, Lipunov V M *Astron. Astrophys.* **329** L29 (1998)
87. Kulkarni S R et al. *Nature* (London) **398** 389 (1999); preprint astro-ph/9902272
88. Anderson N *Astrophys. J.* **502** 708 (1998)
89. Blinnikov S I, in *Proc. XXVII ITEP Winter School* (London: Gordon and Breach, 1999) (in press); preprint astro-ph/9902305