

# Some comments on triggers, energetics and beaming

M.J. Rees

Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK  
e-mail: mjr@ast.cam.ac.uk

Received January 26; accepted February 28, 1999

**Abstract.** There is now a consensus that gamma-ray bursts involve extraordinary power outputs, and highly relativistic dynamics. The trigger is probably a binary merger or collapse involving compact objects. The most plausible progenitors, ranging from NS-NS mergers to various hypernova-like scenarios, eventually lead to the formation of a black hole with a debris torus around it, the extractable energy being up to  $10^{54}$  ergs. The various modes of energy extraction from such systems are discussed. Some key theoretical issues are highlighted, along with the types of observation that would discriminate among the various models.

**Key words:** gamma-rays: bursts

---

## 1. Introduction

Through the impetus of the BeppoSAX satellite, there is now general agreement that the bursts (or at least a substantial subset of them) are ultraenergetic phenomena, posing a severe theoretical challenge and – irrespective of which detailed model proves correct – involving some fascinating and extreme physical processes. There is now a stronger motivation to develop models in fuller detail. This paper outlines some of the key issues. I shall focus on the burst itself, since the afterglow is being considered by several other speakers.

## 2. What is the trigger?

The photon luminosity, for the few-second duration of a typical burst, is of course colossal: it exceeds by many thousands the most extreme output from any active galactic nucleus (thought to involve supermassive black holes), and is 14 orders of magnitude above the Eddington limit for a stellar-mass object. The total energy, however, is not out of line with some other phenomena encountered in astrophysics – indeed it is reminiscent of the energy released

in the core of a supernova, the big difference being that the primary sudden event (with a timescale of seconds) is not smothered by a stellar envelope, as in a supernova, but manifests itself in hard radiation that escapes more promptly.

Unless they are beamed into less than one percent of the solid angle, the triggers for GRBs are thousands of times rarer than supernovae. The most widely favoured and conventional possibility is coalescence of binary neutron stars (see, for example, Narayan et al. 1992). Systems such as the famous binary pulsar will eventually coalesce, when gravitational radiation drives them together. When a neutron star (NS) binary coalesces, the rapidly-spinning merged system would be too massive (for most presumed equations of state) to form a single NS; on the other hand, the total angular momentum is probably too large to be swallowed immediately by a black hole. The expected outcome, after a few milliseconds, would therefore be a spinning black hole (BH), orbited by a torus of neutron-density matter.

Other types of progenitor have been suggested – e.g. a NS-BH merger, where the neutron star is tidally disrupted before being swallowed by the hole; the merger of a white dwarf with a black hole [WD-BH]; or a category labeled as hypernovae or collapsars, where the collapsing core is too massive to become a neutron star, but has too much angular momentum to collapse quietly into a black hole (as in a so called “failed supernova”). The details of the latter model are addressed by Woosley in his contribution. The simple point that I wish to stress, however, is that a BH plus debris torus is a common ingredient of all these models; moreover the overall energetics of these various progenitors differ by at most an order of magnitude, the spread reflecting the differing spin energy in the hole and the different masses left behind in an orbiting torus. (There has been some confusion on this point in recent literature, through failure to appreciate that the dominant energy from a NS-NS event comes after a black hole forms, rather than during the precursor stage that Narayan et al. 1992, discussed.) How might such a system

generate relativistic outflow or a release of electromagnetic energy?

### 3. Energy from a black hole and debris torus?

Two large reservoirs of energy are in principle available: the binding energy of the orbiting debris, and the spin energy of the black hole. The first can provide up to 42% of the rest mass energy of the torus, for a maximally rotating black hole: the second can provide up to 29% (for a maximal spin rate) of the mass of the black hole itself. How can the energy be transformed into outflowing relativistic plasma after such a coalescence event? There seem to be two options. The first is that some of the energy released as thermal neutrinos is reconverted, via collisions outside the dense core, into electron-positron pairs or photons. The rate of this process depends on the square of the neutrino luminosity. The second option is that strong magnetic fields anchored in the dense matter convert the rotational energy of the system into a Poynting-dominated outflow, rather as in pulsars. Let us consider these two options in turn.

(i) Neutrinos could give rise to a relativistic pair-dominant wind if they converted into pairs in a region of low baryon density (e.g. along the rotation axis, away from the equatorial plane of the torus). The  $\nu\bar{\nu} \rightarrow e^+e^-$  process can tap the thermal energy of the torus produced by viscous dissipation. For this mechanism to be efficient, the neutrinos must escape before being advected into the hole; on the other hand, the efficiency of conversion into pairs (which scales with the square of the neutrino density) is low if the neutrino production is too gradual. Typical estimates suggest a limit of  $\lesssim 10^{51}$  erg (Ruffert 1997; Ruffert et al. 1997; Ruffert & Janka 1998; Popham et al. 1998), except perhaps in the “collapsar” or failed SN Ib case where Popham et al. (1998) estimate  $10^{52.3}$  ergs for optimum parameters. If the pair-dominated plasma were collimated into a solid angle  $\Omega_j$  then of course the apparent “isotropized” energy would be larger by a factor  $(4\pi/\Omega_j)$ , but unless  $\Omega_j$  is  $\lesssim 10^{-2} - 10^{-3}$  this may fail to satisfy the apparent isotropized energy of  $10^{53.5}$  ergs implied by a redshift  $z = 3.4$  for GRB 971214.

(ii) An alternative way to tap the torus energy is via magnetic fields threading the torus (Paczynski 1991; Narayan et al. 1992; Mészáros & Rees 1997b; Katz & Piran 1997). Even before the BH forms, a NS-NS merging system might lead to winding up of the fields and dissipation in the last stages before the merger (Mészáros & Rees 1992; Vietri 1997a).

The above mechanisms tap the rotational energy available in the debris torus. However, a hole formed from a coalescing compact binary is guaranteed to be rapidly spinning, and, being more massive, could contain a larger reservoir of energy than the torus; this energy, extractable in principle through MHD coupling to the rotation of

the hole by the Blandford & Znajek (1977) ( $B - Z$ ) effect, could be even larger than that contained in the orbiting debris (Mészáros & Rees 1997b; Paczyński 1998). Collectively, any such MHD outflows have been referred to as Poynting jets.

Simple scaling from the familiar results of pulsar theory tells us that fields of order  $10^{15}$  G, are needed to carry away the rotational or gravitational energy in the time scales of tens of seconds (Usov 1994; Thompson 1994). If the magnetic fields do not thread the BH, then a Poynting outflow can at most carry the gravitational binding energy of the torus. For a maximally rotating and for a non-rotating BH this is 0.42 and 0.06 of the torus rest mass, respectively. The torus mass in a NS-NS merger is  $M_t \sim 0.1 M_\odot$  (Ruffert & Janka 1998), and for an NS-BH or WD-BH merger it may be  $M_t \sim 1 M_\odot$  (Paczynski 1998; Fryer & Woosley 1998). The extractable energy could amount to several times  $10^{53} \epsilon(M_t/M_\odot)$  ergs, where  $\epsilon$  is the efficiency in converting gravitational into MHD jet energy. Tori masses even higher than  $\sim 1 M_\odot$  may occur in scenarios involving massive supernovae. Conditions for the efficient escape of a high- $\Gamma$  jet may, however, be less propitious if the “engine” is surrounded by an extensive envelope.

If magnetic fields of comparable strength thread the BH, its rotational energy offers an extra (and even larger) source of energy that can in principle be extracted via the  $B - Z$  mechanism (Mészáros & Rees 1997b). For a maximally rotating BH, this is  $0.29 M_{\text{bh}} c^2$  ergs, multiplied, of course, by some efficiency factor. A near-maximally rotating black hole is guaranteed in a NS-NS merger. The central BH will have a mass of about  $2.5 M_\odot$ ; the NS-BH merger and hypernova models may not produce quite such rapidly-spinning holes, but the hole masses are larger, so the expected rotational energy should be comparable. Spinning holes can thus power a jet of up to  $\sim 1.5 \cdot 10^{54}$  ergs. Even allowing for low total efficiency (say 30%), a system powered by the torus binding energy would only require a modest beaming of the  $\gamma$ -rays by a factor  $(4\pi/\Omega_j) \sim 20$ , or no beaming if the jet is powered by the  $B - Z$  mechanism, to produce the equivalent of an isotropic energy of  $10^{53.5}$  ergs.

### 4. The gamma-ray emission mechanism

Well-known arguments connected with opacity, variability timescales and so forth require highly relativistic outflow. Best-guess numbers are Lorentz factors  $\Gamma$  in the range  $10^2$  to  $10^3$ , allowing rapidly-variable emission to occur at radii in the range  $10^{14}$  to  $10^{16}$  cm. The entrained baryonic mass would need to be below  $10^{-4} M_\odot$  to allow these high relativistic expansion speeds.

Because the emitting region must be several powers of ten larger than the compact object that acts as trigger, there is a further physical requirement: the original

energy outflowing in a magnetised wind would, after expansion, be transformed into bulk kinetic energy (with associated internal cooling). This energy cannot be efficiently radiated as gamma rays unless it is re-randomised. This requires relativistic shocks. Impact on an external medium would randomise half of the initial energy merely by reducing the expansion Lorentz factor by a factor of 2. Alternatively, there may be internal shocks within the outflow: for instance, if the Lorentz factor in an outflowing wind varied by a factor more than 2, then the shocks that developed when fast material overtakes slower material would be internally relativistic (Rees & Mészáros 1994).

In an unsteady outflow, if  $\Gamma$  were to vary by a factor of 2 on a timescale  $\delta t$ , internal shocks would develop at a distance  $\Gamma^2 c \delta t$ , and randomise most of the energy. For instance, if  $\Gamma$  ranged between 500 and 2000, on a timescale of 1 second, efficient dissipation would occur at  $3 \times 10^{16}$  cm s.

There is a general consensus that the longer complex bursts must involve internal shocks, though simple sharp pulses could arise from an external shock interaction (the latter would in effect be the precursor of the afterglow). An external shock moving into a smooth medium would obviously give a burst with a simple time-profile. A blobby external medium could give features, but only if the covering factor of blobs is low, implying modest efficiency. (This issue is discussed by Dermer in his contribution to this meeting).

Even if the bursts were caused by a completely standardised set of objects, their appearance would be likely to depend drastically on orientation relative to the line of sight. Along any given line of sight, the time-structure would be determined partly by the advance of jet material into the external medium, but probably even more by internal shocks within the jet, which themselves depend on the evolution of the torus, from its formation to its eventual swallowing or dispersal.

The radiation processes for the gamma rays are probably no more than synchrotron radiation. This would imply the presence of magnetic fields where the shocks occur. If the outflow from the central trigger is Poynting-dominated, then a field of  $10^{15}$  G at (say)  $10^7$  cm would imply a comoving field of  $10^7(\Gamma/100)^{-1}$  G out at  $10^{13}$  cm – strong enough to ensure rapid cooling of shocked relativistic electrons. (Note, conversely, that even if magnetic fields were not important near the central trigger, they must be present, with about the same amount of flux that Poynting-dominated models require, at the location of the actual gamma-ray emission.)

We are a long way from modelling what triggers gamma ray bursts. If we had a precise description of the dynamics, along with the baryon content, magnetic field, and Lorentz factor of the outflow, we could maybe predict the gross time-structure. But we could not predict the intensity or spectrum of the emitted radiation – still less answer key questions about the emission in

other wavebands – without also having an adequate theory for particle acceleration in relativistic shocks. We need the answers from plasma physicists to the following poorly-understood questions: (i) Do ultra-relativistic shocks yield power laws? The answer probably depends on the ion/positron ratio, and on the relative orientation of the shock front and the magnetic field (e.g. Gallant et al. 1992). (ii) In ion-electron plasmas, what fraction of the energy goes into the electrons? (iii) Even if the shocked particles establish a power law, there must be a low-energy break in the spectrum at an energy that is in itself relativistic. But will this energy, for the electrons, be  $\Gamma m_p c^2$ , or (or even, if the positive charges are heavy ions like Fe,  $\Gamma m_{\text{Fe}} c^2$ )? (iv) Can ions be accelerated up to the theoretical maximum where the gyroradius becomes the scale of the system? If so, the burst events could be the origin of the highest energy cosmic rays. (v) Do magnetic fields get amplified in shocks? This is relevant to the magnetic field in the swept-up external matter outside the contact discontinuity, and determines how sharp the external shock actually is.

## 5. Intrinsic time scales

A question which has remained largely unanswered so far is what determines the characteristic duration of bursts, which can extend to tens, or even hundreds, of seconds. This is of course very long in comparison with the dynamical or orbital time scale for the “triggers”, which is measured in milliseconds. While bursts lasting hundreds of seconds can easily be derived from a very short, impulsive energy input, this is generally unable to account for a large fraction of bursts which show complicated light curves. This hints at the desirability for a “central engine” lasting much longer than a typical dynamical time scale.

Observationally (Kouveliotou et al. 1993) the short ( $\lesssim 2$  s) and long ( $\gtrsim 2$  s) bursts appear to represent two distinct subclasses, and one early proposal to explain this was that accretion induced collapse (AIC) of a white dwarf (WD) into a NS plus debris might be a candidate for the long bursts, while NS-NS mergers could provide the short ones (Katz & Canel 1996). As indicated by Ruffert et al. (1997),  $\nu\bar{\nu}$  annihilation will generally tend to produce short bursts  $\lesssim 1$  s in NS-NS systems, requiring collimation by  $10^{-1} - 10^{-2}$ , while Popham et al. (1998) argued that in collapsars and WD/He-BH systems longer  $\nu\bar{\nu}$  bursts may be possible.

An acceptable model requires that the surrounding torus should not completely drain into the hole, or be otherwise dispersed, on too short a time scale. There have been some discussions in the literature of possible “runaway instabilities” in relativistic tori (Nishida et al. 1996; Abramowicz et al. 1998; Daigne & Mochkovitch 1997): these are analogous to the runaway Roche lobe overflow predicted, under some conditions, in binary systems.

These instabilities can be virulent in a torus where the specific angular momentum is uniform throughout, but are inhibited by a spread in angular momentum. In a torus that was massive and/or thin enough to be self-gravitating, bar-mode gravitational instabilities could lead to further redistribution of angular momentum and/or to energy loss by gravitational radiation within only a few orbits. Whether a torus of given mass is dynamically unstable depends on its thickness and stratification, which in turn depends on internal viscous dissipation and neutrino cooling.

The disruption of a neutron star (or any analogous process) is almost certain to lead to a situation where violent instabilities redistribute mass and angular momentum within a few dynamical time scales (i.e. in much less than a second). A key issue for gamma ray burst models is the nature of the surviving debris after these violent processes are over: what is the maximum mass of a remnant disc/torus which is immune to very violent instabilities, and which can therefore in principle survive for long enough to power the bursts? It is the mass of this residual torus – i.e. what is left after violent instabilities on a dynamical timescale have done their work – that is the relevant  $M_t$  in the above expressions (in Sect. 3) for the extractable energy of the torus.

If the trigger is to liberate its energy over a period 10 – 100 s via Poynting flux – either through a relativistic wind “spun off” the torus or via the  $B - Z$  mechanism – the required field is a few times  $10^{15}$  G. A weaker field would extract inadequate power; on the other hand, if the large-scale field were even stronger, then the energy would be dumped too fast to account for the longer complex bursts. It is not obvious why the fields cannot become even higher. Note that the virial limit is  $B_v \sim 10^{17}$  G.

Kluźniak & Ruderman (1998) note that, starting with  $10^{12}$  G, it only takes of order a second for simple winding to amplify the field to  $10^{15}$  G; amplification in a newly-formed torus could well occur more rapidly, for instance via convective instabilities, as in a newly formed neutron star (cf. Thompson & Duncan 1993; Thompson 1994). Kluźniak and Ruderman suggest, however, that the amplification may be self-limiting because magnetic stresses would then be strong enough for flares to break out. A magnetic field configuration capable of powering the bursts is likely to have a large scale structure. Flares and instabilities occurring on the characteristic (millisecond) dynamical time scale would cause substantial irregularity or intermittency in the overall outflow that would manifest itself in internal shocks. There is thus no problem in principle in accounting for sporadic large-amplitude variability, on all time scales down to a millisecond, even in the most long-lived bursts. Note also that it only takes a residual torus (or even a cold disk) of  $10^{-3} M_\odot$  to confine a field of  $10^{15}$  G, which can extract energy from the black hole via the  $B - Z$  mechanism. Even if the evolution time scale for the bulk of the debris torus were no more

than a second, enough may remain to catalyse the extraction of energy from the hole at rate adequate to power a long-lived burst.

## 6. Isotropic or beamed outflows?

Computer simulations of compact object mergers and black hole formation can address the fate of the bulk of the matter, but there are some key questions that they cannot yet tackle. In particular, high resolution of the outer layers is needed because even a tiny mass fraction of baryons loading down the outflow severely limits the attainable Lorentz factor – for instance a Poynting flux of  $10^{53}$  ergs could not accelerate an outflow to  $\Gamma > 100$  if it had to drag more than  $\sim 10^{-4}$  solar masses of baryons with it. Further 2D numerical simulations of the merger and collapse scenarios are under way largely using Newtonian dynamics, and the numerical difficulties are daunting. There may well be a broad spread of Lorentz factors in the outflow – close to the rotation axis  $\Gamma$  may be very high; at larger angles away from the axis, there may be an increasing degree of entrainment, with a corresponding decrease in  $\Gamma$ . Even if the outflow is not narrowly collimated, some beaming is expected because energy would be channeled preferentially along the rotation axis. Moreover, we would expect baryon contamination to be lowest near the axis, because angular momentum flings material away from the axis, and any gravitationally-bound material with low angular momentum falls into the hole. In hypernovae, the envelope is rotating only slowly and thus would not initially have a marked centrifugal funnel; even  $10^{53}$  ergs would not suffice to blow out more than a narrow cone of the original envelope with a Lorentz factor of more than 100. So in these models the gamma rays would be restricted to a narrow beam, even though outflow with a more moderate Lorentz factor (relevant to the afterglow) could be spread over a wider range of angles. A wide variety of burst phenomenology could be attributable to a standard type of event being viewed from different orientations.

Two further effects render the computational task of simulating jets even more challenging. The first stems from the likelihood that any entrained matter would be a mixture of protons and neutrons (neutrons, being unconstrained by magnetic fields, could also drift into a jet from the denser walls at its boundary). If a streaming velocity builds up between ions and neutrons (i.e. if they have different Lorentz factors in the outflow) then interactions can lead to dissipation even in a steady jet where there are no shocks (Derishev et al. 1999). A second possibility (Mészáros & Rees 1998b,c) is that entrained ions in a relativistic jet could become concentrated in dense filaments confined by the magnetic field. As already mentioned, the comoving field strength, even out at  $10^{13}$  cm, is of order  $10^6$  G. Trapped filaments of iron-rich thermal, with density up to  $10^{19}$  cm $^{-3}$  and with kT of order a keV, could be

confined by such fields. Such filaments must of course have a small volume-filling factor: otherwise they would load down the jet too much. However, in these strong fields the gyroradii would be so small that filaments could survive against thermal conduction and other diffusion processes even if their dimensions (transverse to the field) were less than 100 cm. Such thin filaments can provide a large covering factor even while filling a tiny fraction of the volume. If they were moving relativistically outwards, they could contribute ultra-blueshift spectral features – for instance, K-edges of Fe could be shifted up to hundreds of keV.

## 7. Brief comments on the afterglows

The discovery of afterglows has not only has extended observations to longer time scales and other wavebands, making the identification of counterparts possible, but also provided confirmation for much of the earlier work on the fireball shock model of GRB, in which the  $\gamma$ -ray emission arises at radii of  $10^{13} - 10^{15}$  cm (Rees & Mészáros 1992, 1994; Mészáros & Rees 1993; Paczyński & Xu 1994; Katz 1994; Sari & Piran 1995). In particular, this model led to the prediction of the quantitative nature of the signatures of afterglows, in substantial agreement with subsequent observations (Mészáros & Rees 1997a; Costa et al. 1997; Vietri 1997a; Tavani 1997; Waxman 1997; Reichart 1997; Wijers et al. 1997).

Astrophysicists understand supernova remnants reasonably well, despite continuing uncertainty about the initiating explosion; likewise, we may hope to understand the afterglows of gamma ray bursts, despite the uncertainties about the “trigger” that I have already emphasised. The simplest hypothesis is that the afterglow is due to a relativistic expanding blast wave. The complex time-structure of some bursts suggests that the central trigger may continue for up to 100 seconds. However, at much later times all memory of the initial time-structure would be lost: essentially all that matters is how much energy and momentum has been injected, its distribution in angle, and the mass fractions in shells with different Lorentz factors.

The simplest spherical afterglow model – where a relativistic blast wave decelerates as it runs into ambient matter, leading to a radiative output with a calculable spectrum, and a characteristic power law decay – has been remarkably successful at explaining the gross features of the GRB 970228, GRB 970508 and other afterglows (e.g. Wijers et al. 1997). The gamma-rays we receive come only from material whose motion is directed within one degree of our line of sight. They therefore provide no information about the ejecta in other directions: the outflow could be isotropic, or concentrated in a cone of any angle substantially larger than one degree (provided that the line of sight lay inside the cone). At observer times of more than a week, the blast wave would however be decelerated to a moderate Lorentz factor, irrespective of

the initial value. The beaming and aberration effects are thereafter less extreme, so we observe afterglow emission not just from material moving almost directly towards us, but from a wider range of angles.

The afterglow is thus a probe for the geometry of the ejecta – at late stages, if the outflow is beamed, we expect a spherically-symmetric assumption to be inadequate; the deviations from the predictions of such a model would then tell us about the ejection in directions away from our line of sight. It is quite possible, for instance, that there is relativistic outflow with lower  $\Gamma$  (heavier loading of baryons) in other directions (e.g. Wijers et al. 1997); this slower matter could even carry most of the energy (Paczyński 1998). An argument for a broad beaming angle in the energy outflow is that, if the energy were channeled into a solid angle  $\Omega_j$  then, (Rhoads 1997), one expects a faster decay of  $\Gamma$  after it drops below  $\Omega_j^{-1/2}$ . A simple calculation using the usual scaling laws leads then to a steepening of the flux power law in time. The lack of such an observed downturn in the light curve has been interpreted as further supporting the sphericity of the entire fireball. There are several important caveats, however. The first one is that the above argument assumes a simple, impulsive energy input (lasting  $\lesssim$  than the observed  $\gamma$ -ray pulse duration), characterized by a single energy and bulk Lorentz factor value. Estimates for the time needed to reach the non-relativistic regime, or  $\Gamma < \Omega_j^{-1/2} \lesssim$  few, could then be under a month (Vietri 1997b; Huang et al. 1998), especially if an initial radiative regime with  $\Gamma \propto r^{-3}$  prevails. (It is however unclear whether, even when electron radiative time scales are shorter than the expansion time, such a regime applies, as it would require strong electron-proton coupling (Mészáros et al. 1998). Waxman et al. (1998) have also argued on observational grounds that the longer lasting  $\Gamma \propto r^{-3/2}$  (adiabatic regime) is more appropriate.) Furthermore, even the simplest reasonable departures from this ideal model (e.g. a substantial amount of energy and momentum ejected behind the main shell with a lower Lorentz factor) would drastically extend the afterglow lifetime in the relativistic regime, by providing a late “energy refreshment” to the blast wave on time scales comparable to the afterglow time scale (Mészáros & Rees 1998a). Anisotropy in the burst outflow and emission affects the light curve at the time when the inverse of the bulk Lorentz factor equals the opening angle of the outflow. If the critical Lorentz factor is less than 3 or so (i.e. the opening angle exceeds  $20^\circ$ ) such a transition might be masked by the transition from ultrarelativistic to mildly relativistic flow, so quite generically it would be difficult to limit the late-time afterglow opening angle in this way if it exceeds  $20^\circ$ . Since some afterglows are unbroken power laws for over 100 days (e.g. GRB 970228), if the energy input were indeed just a simple impulsive shell the opening angle of the late-time afterglow at long wavelengths is probably greater than  $1/3$ , i.e.  $\Omega_{\text{opt}} \gtrsim 0.4$ .

However, even this still means that the energy estimates from the afterglow assuming isotropy could be 30 times too high.

The beaming angle for the gamma ray emission could be far smaller, and is much harder to constrain directly. The ratio of  $\Omega_\gamma/\Omega_x$  has been considered by Grindlay (1998) using data from Ariel V and HEAO-A1/A2 surveys, who did not find evidence for a significant difference between the deduced gamma-ray and X-ray rates, and concluded that higher sensitivity surveys would be needed to provide significant constraints. More promising for the immediate future, the ratio  $\Omega_\gamma/\Omega_{\text{opt}}$  can also be investigated observationally (see also Rhoads 1997). The rate of GRB with peak fluxes above  $1 \text{ ph cm}^{-2} \text{ s}^{-1}$  as determined by BATSE is about 300/yr, i.e.  $0.01/\text{sq deg/yr}$ . According to Wijers et al. (1998) this flux corresponds to a redshift of 3. If the gamma rays were much more narrowly beamed than the optical afterglow there should be many “homeless” afterglows, i.e. ones without a GRB preceding them. The transient sky at faint magnitudes is poorly known, but there are two major efforts under way to find supernovae down to about  $R = 23$  (Garnavich et al. 1998; Perlmutter et al. 1998). These searches have by now covered a few tens of “square degree years” of exposure and would be sensitive to afterglows of the brightness levels thus far observed. It therefore appears that the afterglow rate is not more than a few times  $0.1/\text{sq deg/yr}$ . Since the magnitude limit of these searches allows detection of optical counterparts of GRB brighter than  $1 \text{ ph cm}^{-2} \text{ s}^{-1}$  it is fair to conclude that the ratio of homeless afterglows to GRB is unlikely to exceed  $\sim 20$ . It then follows that  $\Omega_\gamma > 0.05 \Omega_{\text{opt}}$ , which combined with our limit to  $\Omega_{\text{opt}}$  yields  $\Omega_\gamma > 0.02$ . The true rate of events that give rise to GRB is therefore at most 600 times the observed GRB rate, and the opening angle of the ultrarelativistic, gamma-ray emitting material is no less than  $5^\circ$ . Combined with the most energetic bursts, this begins to pose a problem for the neutrino annihilation type of GRB energy source.

Obviously, the above calculation is only sketchy and should be taken as an order of magnitude estimate at present. However, it should improve as more afterglows are detected and the modelling gets more precise.

## 8. Conclusions and prospects

Simple blast-wave models seem able to accommodate the present data on afterglows. However we can at present only infer the energy per solid angle; there are reasons to suspect that the afterglow is not too narrowly beamed; on the other hand the constraints on the angle-integrated  $\gamma$ -ray energy are not strong – these could be concentrated in a high-Lorentz-factor beam only a few degrees across. As regards the trigger, there are two key questions. First, does it involve a black hole orbited by a dense torus (which

I’ve advocated as a “best buy”)? Second, if so, can we decide between the various alternative ways of forming it: NS-NS, NS-BH or hypernova?

The locations should help to settle the second question. This is because a hypernova would be expected to lie in a region of recent star formation; on the other hand, a neutron star binary could take hundreds of millions of years to spiral together, and could by then (especially if given a kick velocity on formation) have moved many kiloparsecs from its point of origin (Bloom et al. 1998). There is also already tentative evidence that some detected afterglows arise in relatively dense gaseous environments – e.g. by evidence for dust in GRB 970508 (Reichart 1998) and the absence of an optical afterglow and strong soft X-ray absorption in GRB 970828 (Groot et al. 1997; Murakami et al. 1997). On the other hand, fits to the observational data on GRB 970508 and GRB 971214 suggest external densities in the range of  $0.04 - 0.4 \text{ cm}^{-3}$ , which would be more typical of a tenuous interstellar medium (Wijers & Galama 1998).

We must also remain aware of other possibilities. For instance, we may be wrong in supposing that the central object becomes dormant after the gamma-ray burst itself. It could be that the accretion-induced collapse of a white dwarf, or (for some equations of state) the merger of two neutron stars, could give rise to a rapidly-spinning pulsar, temporarily stabilised by rapid rotation. The afterglow could then, at least in part, be due to a pulsar’s continuing power output (cf. Usov 1994). It could also be that mergers of unequal mass neutron stars, or neutron stars with other compact companions, lead to the delayed formation of a black hole. Such events might also lead to repeating episodes of accretion and orbit separation, or to the eventual explosion of a neutron star which has dropped below the critical mass, all of which would provide a longer time scale, episodic energy output.

And there could be more subclasses of classical GRB than just short ones and long ones. For instance, GRB with no high energy pulses (NHE) appear to have a different (but still isotropic) spatial distribution from those with high energy (HE) pulses (Pendleton et al. 1996). Some caution is needed in interpreting this, since selection effects could lead to a bias against detecting HE emission in dim bursts (Norris 1998). Then, there is the apparent coincidence of GRB 980425 with the SN Ib/Ic 1998bw (Galama et al. 1998). Much progress has been made in understanding how gamma-rays can arise in fireballs produced by brief events depositing a large amount of energy in a small volume, and in deriving the generic properties of the long wavelength afterglows that follow from this. There still remain a number of mysteries, especially concerning the identity of their progenitors, the nature of the triggering mechanism, the transport of the energy and the time scales involved.

Gamma-ray bursts, even if we do not understand them, may still be useful as powerful beacons for probing the

high redshift ( $z > 5$ ) universe. Even if their total energy is reduced by beaming to a “modest”  $\sim 10^{52} - 10^{52.5}$  ergs in photons, they are the most extreme phenomena that we know about in high energy astrophysics. The modeling of the burst itself - the trigger, the formation of the ultrarelativistic outflow, and the radiation processes - is a formidable challenge to theorists and to computational techniques. It is, also, a formidable challenge for observers, in their quest for detecting minute details in extremely faint and distant sources. And if the class of models that we have advocated here turns out to be irrelevant, the explanation of gamma-ray bursts will surely turn out to be even more remarkable and fascinating.

*Acknowledgements.* I am especially grateful to Peter Mészáros and Ralph Wijers for extended collaboration, and to Josh Bloom and Stan Woosley for discussions. This research has been supported by the Royal Society.

## References

- Abramowicz M.A., Karas V., Lanza A., 1998, A&A (in press) (astro-ph/9712245)
- Blandford R.D., Znajek R.L., 1977, MNRAS 179, 433
- Bloom J., Sigurdsson S., Pols O., 1998, MNRAS (in press) (astro-ph/9805222)
- Costa E., et al., 1997, Nat 387, 783
- Daigne F., Mochkovitch R., 1997, MNRAS 285, L15
- Derishev E.V., Kocharovskiy V.V., Kocharovskiy V.I.V., 1999 (preprint)
- Fryer C., Woosley S., 1998 (submitted to ApJL) (astro-ph/9804167)
- Galama T., et al., 1998, Nat (in press) (astro-ph/9806175)
- Gallant Y.A., Hoshino M., Langdon A.B., Arons J., Max C.E., 1992, ApJ 391, 73
- Garnavich P., et al., 1998, ApJ 493, L53
- Grindlay J., 1998 (preprint)
- Groot P., et al., 1997, in Gamma-Ray Bursts, Meegan C., Preece R. & Kosshut T. (eds.). AIP: New York, p. 557
- Huang Y.F., Dai Z.G., Lu T., 1998, A&A (in press) (astro-ph/9807061)
- Katz J., Canel L.M., 1996, ApJ 471, 915
- Katz J., 1994, ApJ 422, 248
- Katz J., Piran T., 1997, ApJ 490, 772
- Kluźniak W., Ruderman M., 1998 (preprint)
- Kouveliotou C., et al., 1993, ApJ 413, L101
- Mesz P., Rees M.J., 1992, ApJ 397, 570
- Mesz P., Rees M.J., 1993, ApJ 405, 278
- Mesz P., Rees M.J., 1997a, ApJ 476, 232
- Mesz P., Rees M.J., 1997b, ApJ 482, L29
- Mesz P., Rees M.J., 1998a, ApJL 496, L1
- Mesz P., Rees M.J., 1998b, ApJL 502, L105
- Mesz P., Rees M.J., 1998c, MNRAS 299, L10
- Mesz P., Rees M.J., Wijers R.A.M.J., 1998, ApJ 499, 301
- Murakami T., et al., 1997, in Gamma-Ray Bursts, Meegan C. (ed.)
- Narayan R., Paczyński B., Piran T., 1992, ApJL 395, L83
- Nishida S., Lanza A., Eriguchi Y., Abramawicz M.A., 1996, MNRAS 278, L41
- Norris J., 1998 (private communication)
- Paczyński B., 1991, Acta. Astron. 41, 257
- Paczyński B., 1998, ApJ 494, L45
- Paczyński B., Xu G., 1994, ApJ 427, 708
- Pendleton G., et al., 1996, ApJ 464, 606
- Perlmutter S., et al., 1998, Nat 391, 51
- Popham R., Woosley S., Fryer C., 1998, ApJ (submitted) (astro-ph/9807028)
- Rees M.J., Mesz P., 1992, MNRAS 258, 41P
- Rees M.J., Mesz P., 1994, ApJ 430, L93
- Reichart D., 1997, ApJ 485, L57
- Reichart D., 1998, ApJ 495, L99
- Rhoads J.E., 1997, ApJ 487, L1
- Ruffert M., 1997, A&A 317, 793
- Ruffert M., Janka H.-T., Takahashi K., Schaefer G., 1997, A&A 319, 122
- Ruffert M., Janka H.-T., 1998, A&A (in press)
- Sari R., Piran T., 1995, ApJ 455, L143
- Tavani M., 1997, ApJ 483, L87
- Thompson C., 1994, MNRAS 270, 480
- Thompson C., Duncan R.C., 1993, ApJ 408, 194
- Usov V.V., 1994, MNRAS 267, 1035
- Vietri M., 1997a, ApJ 478, L9
- Vietri M., 1997b, ApJ 488, L105
- Waxman E., 1997, ApJ 489, L33
- Waxman E., Kulkarni S., Frail D., 1998, ApJ 497, 288
- Wijers R.A.M.J., Bloom J.S., Bagla J.S., Natarajan P., 1998, MNRAS 294, L13
- Wijers R.A.M.J., Rees M.J., Mesz P., 1997, MNRAS 288, L51
- Wijers R.A.M.J., Galama T., 1998, ApJ (submitted) (astro-ph/9805341)
- Woosley S., Eastman R., Schmidt B., 1998, ApJ (submitted) (astro-ph/9806299)