VARIABILITY IN GAMMA-RAY BURSTS: A CLUE

RE'EM SARI AND TSVI PIRAN

Racah Institute of Physics, Hebrew University, P.O. Box 91904, Jerusalem, Israel Received 1997 January 6; accepted 1997 March 13

ABSTRACT

We show that external shocks cannot produce a variable gamma-ray burst, unless the burst is produced by an extremely narrow jet (angular opening of $\leq 10^{-4}$) or if only a small fraction of the shell emits the radiation and the process is very inefficient. Internal shocks can produce the observed complex temporal structure provided that the source itself is variable. In this case, the observed temporal structure reflects the activity of the "inner engine" that drives the bursts. This sets direct constraints on it. Subject headings: gamma rays: bursts — hydrodynamics — relativity — shock waves

1. INTRODUCTION

Five years of BATSE's observations with perfect isotropy and paucity of weak bursts show that the origin of gammaray bursts (GRBs) is probably cosmological. Therefore, given the measured flux, GRBs involve immense amounts of energy, $\sim 10^{51}$ ergs. The "compactness problem" then shows that the observed gamma rays must be emitted by a medium with highly relativistic velocities having Lorentz factor $\gamma \ge 100$ (Fenimore, Epstein, & Ho 1993; Woods & Loeb 1995; Piran 1996). While the energy source varies from one model to another (merging of binary neutron stars, failed supernovae, or collapse of magnetic stars) and is relatively speculative, all models of cosmological GRBs involve a relativistic moving shell which converts its (kinetic or magnetic) energy to radiation at a large radius. In all these models the observed radiation does not emerge directly from the "inner engine" that drives the shell, which remains hidden.

Most bursts are highly variable with a variability scale significantly smaller than the overall duration. Following Fenimore, Madras, & Nayakshin (1996), we use kinematic considerations to constrain different GRB models. We show that the overall duration, T, reflects directly the length of time that the "inner engine" operates, and the observed temporal variability reflects variability in the "inner engine." The only exceptions to this conclusion occur if the engine produces an extremely narrow jet or if GRBs are extremely inefficient. These considerations also limit the emission radius, R_e —the place where the energy of the shell is converted to radiation-to be significantly smaller than was previously thought. The maximal emission radius is quite close to the minimal radius at which a GRB can be produced without becoming optically thick. This is also the place where internal shocks between different parts of the shell that move with different Lorentz factors would naturally take place. Thus, the observed temporal structure is consistent with the "internal shock" model for conversion of the kinetic energy to radiation (Rees & Mészáros 1994; Narayan, Paczyński, & Piran 1992; Sari & Piran 1997). In the "external shock" scenario the radiation arises from deceleration of the shell due to the interaction with an external medium such as the ISM (Mészáros & Rees 1992, 1993; Katz 1994; Sari & Piran 1995). Sufficiently small radii, as required by the temporal limitations, are impossible in this case.

In § 2 we discuss the angular spreading problem, which is the key to our discussion. We show in § 3 that in the framework of models in which the duration of the burst is given by the radius of emission, all solutions to the angular spreading problem result in extremely narrow jets or an extremely low efficiency. In § 4 we discuss models in which the total duration of the burst corresponds directly to the time that the "inner engine" operates. The internal shock scenario fits this picture. We show that the hydrodynamic version of the external shock scenario (and most likely all other versions) is incompatible with these limits.

2. ANGULAR SPREADING

Special relativistic effects determine the observed duration of the burst from a relativistic shell. Consider an infinitely thin relativistic shell with a Lorentz factor γ_e (the subscript e is for the emitting region) and an angular width larger than γ_e^{-1} . Because of relativistic beaming an observer can see only a region of size γ_e^{-1} . Therefore a shell with an angular size larger than γ_e^{-1} can be considered as spherical. Let R_e be a typical radius characterizing the emitting region (in the observer frame) such that most of the emission takes place while the shell is moving from $R_e - \Delta R_e/2$ to $R_e + \Delta R_e/2$. The observed duration between the first photon (emitted at $R_e - \Delta R_e/2$) and the last one (emitted at $R_e + \Delta R_e/2$) is

$$T_{\rm radial} \cong \Delta R_e / 2\gamma_e^2 c$$
 . (1)

Because of radiation beaming an observer sees up to solid angle of γ_e^{-1} from the line of sight. Two photons emitted at the same time and radius R_e , one on the line of sight and the other at an angle of γ_e^{-1} away, travel different distances to the observer. The difference, $R/2\gamma_e^2$ leads to a delay in the arrival time by (Ruderman 1975; Katz 1994)

$$T_{\text{angular}} \cong R_e / 2\gamma_e^2 c . \tag{2}$$

Fenimore et al. (1996) have shown that the observed pulse will have a fast rise and a slow decay with FWHM $\sim 0.22 R_e / \gamma_e^2 c.$

Comparison of equations (1) and (2) using $\Delta R_e \leq R_e$ reveals that $T_{\text{angular}} \ge T_{\text{radial}}$. As long as the shell is spherical on an angular scale larger than γ_e^{-1} , any temporal structure that could have risen due to irregularities in the radial structure of the shell or the material that it encounters will be spread on a time given by $T_{angular}$. Thus $T_{angular}$ is a lower limit for the observed temporal variability: $\delta T \ge T_{angular}$. If the shell has a finite thickness, Δ (measured in the

observer's rest frame), then the duration of the burst must

be longer than Δ/c . We therefore have

$$T = \begin{cases} T_{\text{angular}} = R_e/2c\gamma_e^2 & \text{if } \Delta < R_e/\gamma_e^2 \text{ (type I)}, \\ \Delta/c & \text{otherwise (type II)}. \end{cases}$$
(3)

It is convenient to classify different GRB models as type I or type II according to whether the first or the second situation obtains.

In type I models, the burst's duration is determined by the emission radius, and it is independent of Δ . These models include the standard "external shock model," its magnetic version, i.e., the relativistic magnetic wind model (Usov 1994) in which a magnetic Poynting flux runs into the ISM, or the scattering of starlight by a relativistic shell (Shemi 1994; Shaviv & Dar 1995).

In type II models, the burst's duration is determined by the thickness of the shell. These models include the "internal shock model" or its magnetic-dominated version given by Thompson (1994).

The majority of GRBs have a complex temporal structure (e.g., Fishman & Meegan 1995; Meegan et al. 1996) with typical variations on a timescale, δT , significantly smaller than the total duration T. We define the ratio $\mathscr{V} \equiv T/\delta T$, which is a measure of the variability and an upper limit for the number of peaks (\mathscr{V} is larger than the number of peaks in bursts that contain "quiet" periods without emission). Figure 1 presents a burst of duration $T \sim 75$ s and typical peaks of width $\delta T \lesssim 1$ s, thus $\mathscr{V} \sim 100$. We adopt these as canonical numbers for this paper.

Consider a type I model, where $\Delta/c < T_{\text{angular}}$ and $T = T_{\text{angular}}$. Angular spreading means that any variability in the emission on a timescale smaller than T_{angular} is erased unless the spherical symmetry is broken within angular size smaller than γ_e^{-1} . Thus a burst produced from such a shell, in a spherical geometry, must be a smooth single-humped burst with $\mathscr{V} = 1$ and no temporal structure on a timescale $\delta T \ll T$. Put in other words, a shell of a type I model, and with angular width larger than γ_e^{-1} cannot produce a variable burst with $\mathscr{V} \ge 1$. This is the angular spreading problem. Fenimore et al. (1996) called this the "curvature effect."

On the other hand, a type II model contains a thick shell, $\Delta > R_e/\gamma_e^2$, and can produce a variable burst. The variability timescale is again limited by $\delta T > T_{angular}$; however, $T_{angular}$ can be shorter than the total duration T. The temporal variability can now reflect radial inhomogeneity of the



FIG. 1.—Part of burst 2553 (at E > 25 keV) of duration $T_{90} = 75$ s. The variability is on a timescale $\lesssim 1$ s. The variability parameter for this burst is $\mathscr{V} \sim 100$.

shell. Since the width, Δ , is determined by the time that the "inner engine" operates, and radial inhomogeneities in the shell reflect its variability, we find that both the total duration and the variability timescale reflect those of the source. This is a remarkable conclusion in view of the fact that the fireball hides the "inner engine" and that it was believed that we would not be able to obtain any direct information on it.

3. ANGULAR VARIABILITY AND OTHER CAVEATS

Thin shells, with $\Delta < R_e/\gamma_e^2$, can produce variable bursts only if the opening angle of the emitting region is sufficiently small—that is, spherical symmetry is broken on scales significantly narrower than γ_e^{-1} . Otherwise, the angular spreading will erase any variability on short timescales.

We begin by estimating the maximal size of an emitting region that can produce temporal structure of the order of $\delta T = T/\mathscr{V}$. Imagine two points (r_1, θ_1) and (r_2, θ_2) , r being the distance from the origin and θ the angle from the line of sight, that emit radiation at times t_1 and t_2 , respectively. In principle, one can carefully choose the emission points (r_1, θ_1) and (r_2, θ_2) to produce an arbitrarily narrow pulse. For example, one can arrange that the emitting regions are located on the ellipsoid which is the locus of points from which photons reach the observer at the same time. However these ellipsoids are different for different observers, and what looks shorter for a specific observer will look longer to most other observers. The same is true if we vary the emission times, t_1 and t_2 . Therefore, we assume that $r_1 = r_2 = R_e$ and $t_1 = t_2$. Consequently, quite generally the difference in the arrival time between two photons will be

$$\delta T \approx \frac{R_e(\theta_2^2 - \theta_1^2)}{2c} = \frac{R_e \bar{\theta} | \theta_2 - \theta_1 |}{c} = \frac{R_e \bar{\theta} \delta \theta}{c}, \quad (4)$$

where we have used θ_1 , $\theta_2 \ll 1$, $\bar{\theta} \equiv (\theta_1 + \theta_2)/2$, and $\delta \theta \equiv |\theta_2 - \theta_1|$.

Since an observer sees emitting regions up to an angle γ_e^{-1} away from the line of sight, $\bar{\theta} \sim \gamma_e^{-1}$, and the size of the emitting region $r_s = R_e |\theta_2 - \theta_1|$ is limited by

$$r_s \le \gamma_e c \,\delta T \,. \tag{5}$$

The corresponding angular size is

$$\delta\theta \le \frac{\gamma_e \, c \, \delta T}{R_e} = \frac{1}{\mathscr{V}\gamma_e} \,. \tag{6}$$

Note that Fenimore et al. (1996) considered only emitting regions that are directly on the line of sight for which $\bar{\theta} \sim |\theta_2 - \theta_1|$ and obtained the limit $r_s = \gamma_e c(T \,\delta T)^{1/2}$, which is larger than our estimate in equation (5). However, only a small fraction of the emitting regions will be exactly on the line of sight. Most of the emitting regions will have $\bar{\theta} \sim \gamma_e^{-1}$.

The above discussion suggests that one can produce GRBs with $T \approx T_{\text{radial}} \approx R_e/c\gamma_e^2$ and $\delta T = T/\mathscr{V}$ if the emitting regions have angular size smaller than $1/\mathscr{V}\gamma_e$. The first idea that comes to mind is a narrow jet. However, for a typical burst the maximal opening angle is smaller than 10^{-4} ! Hydrodynamic acceleration can produce jets with angular width γ_e^{-1} or larger. The jets require another acceleration mechanism. Additionally, either rapid modulation of the jet or inhomogeneities in the ISM are required to produce the temporal variability.

The second possibility is that the shell is relatively "wide" (wider than γ_e^{-1}) but the emitting regions are narrow. This may occur if either the ISM or the shell itself is very irregular. An example of this situation is described schematically in Figure 2. However, the emitting regions will have a small covering factor, and this situation is extremely inefficient. The area of the observed part of the shell is $\pi R_e^2/\gamma_e^2$. To comply with the temporal constraint, the total area of the emitting regions is $\sqrt[4]{\pi r_s^2}$. The ratio between the two, i.e., the fraction of the shell that emits the radiation, is

$$\frac{\mathscr{V}\pi r_s^2}{\pi R_a^2/\gamma^2} \le \frac{1}{4\mathscr{V}} \ll 1 , \qquad (7)$$

where we have used the definition of type I models, $R_e = 2c\gamma_e^2 T$, and equation (5) for the maximal size of the emitting objects. This sets an upper limit for the efficiency which is less than 1%.

To obtain a high efficiency, i.e., a covering factor of order unity, with emitting regions of size r_s , we must have $\sim 4\mathscr{V}^2$ emitting regions. But a sum of $4\mathscr{V}^2$ peaks each of width $1/\mathscr{V}$ of the total duration does not produce a complex time structure. Instead it produces a smooth time profile with small variations, of order $1/2\mathscr{V} \ll 1$, in the amplitude.

The problem of type I models, with shells that are spherical on an angular size of more than γ_e^{-1} , is fundamental. It does not depend on the nature of the emitting regions: ISM clouds, starlight, or fragments of the shell. This is the case, for example, in the models of Shaviv & Dar (1995), who consider interaction of a smooth shell with external fragmented medium. This low efficiency poses a serious energy crisis for most (if not all) cosmological models. Recall the huge amount of energy, 10^{51} ergs, observed. It is difficult to imagine sources that can emit considerably larger amounts of energy, of the order of 10^{53} ergs or more, in the form of a relativistic shell. If such sources exist, it is not clear what



FIG. 2.—An attempt to produce variability in type I models by breaking the spherical symmetry. A shell with angular size γ^{-1} is drawn (the angular size is highly exaggerated). The spherical symmetry in this example is broken by the presence of bubbles in the ISM. The relative angular size of the shell and the bubbles is drawn to scale assuming that a burst with $\mathscr{V} = 15$ is to be produced. Consequently $\mathscr{V} = 15$ bubbles are drawn (more bubbles will add up to a smooth profile) with radial position between R and 2R. The fraction of the shell that will impact these bubbles is small, leading to high inefficiency. As \mathscr{V} increases, the efficiency problem becomes more severe, $\sim \mathscr{V}^{-1}$.

will happen to the rest of the kinetic energy of the shell. Note that a flux of 10^{53} ergs per 10^{6} yr pre-Galaxy, of 100 GeV cosmic rays, is comparable to the observed cosmic-ray background at that energy.

4. TYPE II MODELS

The simplest solution to the angular spreading problem is if the emission radius is sufficiently small that angular spreading does not erase temporal structure with timescale δT , i.e., $R_e \leq 2\gamma_e^2 c \, \delta T$. This can take place in type II models in which the overall duration is $T = \Delta/c$, longer than $T_{\text{angular}} = R_e/2\gamma_e^2 c$.

In this class of models one needs multiple shells to account for the observed temporal structure. Each shell produces an observed peak of duration δT , and the whole complex of shells (whose width is Δ) produces a burst that lasts a time $T = \Delta/c$. The observed temporal structure will be longer between the temporal structure of the "inner engine" and the angular spreading time.

Thus, in contrast to previous worries (Piran 1996; Mészáros 1995), we find that there is some direct information that we have on the "inner engine" of GRBs. It must be capable of producing the observed complicated temporal structure. This severely constrains numerous models.

Type II behavior arises naturally in the internal shock model (Rees & Mészáros 1994; Narayan, Paczyński, & Piran 1992), where the shells are created with variable Lorentz factors and therefore collide with one another and convert a considerable fraction of their kinetic energy into internal energy which is radiated. Sari & Piran (1997) have recently given both an upper limit (above which external shocks occur before the internal shocks) and a lower limit (below which the flow is optically thick) for the Lorentz factor of the internal shocks:

$$100 \le \gamma_e \le 1200 \left(\frac{\delta T}{1 \text{ s}}\right)^{-1/2} \left(\frac{T}{100 \text{ s}}\right)^{1/8} l_{18}^{3/8} , \qquad (8)$$

where $l \equiv (E/c^2 n_1)^{1/3} \sim 10^{18}$ cm. The corresponding radius, R_e , is given by

$$3 \times 10^{14} \left(\frac{\delta T}{1 \text{ s}} \right) \le R_e \le 4 \times 10^{16} \left(\frac{T}{100 \text{ s}} \right)^{1/4} l_{18}^{3/4} .$$
 (9)

The lower limits might be higher for low values of δT ; the full expressions are in Sari & Piran (1997).

It is worthwhile to recall that internal shocks can extract at most half of the shell's energy, and a relativistic shell with kinetic energy and Lorentz factor comparable to the original one is left. If the shell is surrounded by ISM and collisionless shocks occur, the relativistic shell will dissipate by "external shocks" as well, which predicts an additional smooth burst, with comparable energy. The additional burst, whose timescale and spectrum depend on model parameters, was not yet observed. An alternative is that the shell continues to move freely and eventually contributes low-energy cosmic rays of about $10^{2}-10^{4}$ GeV, depending on the Lorentz factor of the shell. This flux is about 10^{-2} of the observed flux at 10^{2} GeV. It is almost comparable to the observed flux if all particles are above 10^{4} GeV.

While internal shocks are naturally type II, it is interesting to ask whether external shocks could give rise to type II behavior. This would have been possible if we could have set the parameters of the external shock model to satisfy



FIG. 3.—The external shock problem. The deceleration radius R_e and the Lorentz factor of the shocked shell γ_e as functions of the initial Lorentz factor γ , for a shell of fixed width $\Delta = 3 \times 10^{12}$ cm. For low values of γ , the shocked material moves with Lorentz factor $\gamma_e \sim \gamma$. However, as γ increases the reverse shock becomes relativistic, reducing significantly the Lorentz factor $\gamma_e < \gamma$. This phenomenon prevents the "external shock model" from being type II.

 $R_e \leq 2\gamma_e^2 c \, \delta T$. For thin shells the deceleration radius is given by (Mészáros & Rees 1992):

$$R_e = l\gamma^{-2/3} , \qquad (10)$$

and the observed duration is therefore $T = R/\gamma^2 = l\gamma^{-8/3}$. The deceleration is gradual and the Lorentz factor of the emitting region γ_e is similar to the original Lorentz factor of the shell γ . It seems that with an arbitrary large Lorentz factor γ we can obtain a small enough deceleration radius R_e , as required by type II. However, Sari & Piran (1995) have shown that equation (10) is valid only for thin shells satisfying $\Delta > l\gamma^{-8/3}$. As γ increases above a critical value $\gamma \ge \gamma_c = (l/\Delta)^{3/8}$, the shell can no longer be considered thin. In this situation the reverse shock penetrating the shell becomes ultrarelativistic and the shocked matter moves with Lorentz factor $\gamma_e = \gamma_c < \gamma$ independent of the initial Lorentz factor of the shell γ . The deceleration radius is now given by $R_e = \Delta^{1/4} l^{3/4}$, and it is also independent of the initial Lorentz factor of the shell. The behavior of the deceleration is not given by $R_e = \Delta^{1/4} l^{3/4}$. eration radius R_e and observed duration as function of the shell Lorentz factor γ is given in Figure 3 for a shell of thickness $\Delta = 3 \times 10^{12}$ cm.

5. CONCLUSIONS

Relativistic motion is essential in GRBs. Relativistic kinematic arguments (Fenimore et al. 1996) strongly limit GRB models. The duration of a GRB is determined either by the width of the emitting shell, $T = \Delta/c$, or by the emission radius, $T_{\text{angular}} = R_e/\gamma_e^2$. Models in which $T_{\text{angular}} > \Delta/c$ (type I) can produce only a single-hump smooth burst. They cannot produce a variable burst. The standard "external shock" model is the classical example of a type I model. Therefore, this conclusion rules out this scenario. An exception to this conclusion is if the "inner engine" emits an extremely narrow jet (with angular width smaller than $1/\psi \gamma_e \sim 10^{-4}$). Such a jet cannot be produced by a standard fireball in which the matter is accelerated by thermal pressure, and it requires another acceleration mechanism. Alternatively, a variable burst can be produced by an irregular shell with angular fluctuations of large amplitude and small size, or with a highly irregular ISM. In this case the process is extremely inefficient owing to a low covering factor, and the total energy needed for a GRB is \mathscr{V} times larger than the observed energy (of the order of 10⁵³ ergs).

Type II models do not suffer the angular spreading problem and can produce the observed temporal structure. The "internal shock" model is the classical example for this type. Note that it is impossible to change the parameters of an "external shock" model so that it will become type II. In this case the temporal structure reflects the activity of the "inner engine." The overall duration is the time it operates, while the variability reflects the variability of the source. This is good news, since we now have direct information on the "inner engine." It is also bad news, since only a few known models can produce the observed highly variable temporal structure observed in GRBs.

We would like to thank Jonathan Katz and Ramesh Narayan for helpful discussions. This research was supported in part by a grant from the US-Israel BSF and by NASA grant NAG 5-3516.

REFERENCES

Note added in proof.—The additional burst predicted here for the "internal shock scenario" has been observed by now as the continuing X-ray, visible, and radio emission seen in GRB 970228 and GRB 970508.