UC Irvine

UC Irvine Previously Published Works

Title

Gamma ray bursts in their historic context

Permalink

<https://escholarship.org/uc/item/1c43m5hr>

Author

Trimble, V

Publication Date

2004

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Gamma Ray Bursts In Their Historic Context

Virginia Trimble

*Department of Physics and Astronomy, University of California, Irvine CA 92697 USA, and
Astronomy Department, University of Maryland, College Park MD 20742 USA*

Abstract. Gamma ray bursts remained essentially non-understood or misunderstood from their 1973 discovery (not, I will claim, “serendipitous”) to the first, 1997, redshift. This is by no means a record. The poster explored some of the examples of longer-standing puzzles and the after-dinner talk some of the details of the GRB case. The most striking feature of the GRB history is probably the unanimity with which “all we, like sheep, went astray,” which followed the epoch of “we have turned everyone to his own way.” Some of the reasons for this, the range of hypotheses, and how GRBs were presented to the astronomical and larger communities are discussed.

INTRODUCTION: THE PATTERN OF PUZZLING PHENOMENA

Some astronomical discoveries are understood as soon as they are made (pulsars as rapidly rotating, magnetized neutron stars, for instance) or even before (this is called prediction, black holes between 1800 and 1972, are a classic case). For those where understanding is considerably delayed, there seems to be a characteristic pattern, which also applies to the gamma ray bursts.

First, an observation is perceived as puzzling, sometimes by the person who made it or his immediate contemporaries; sometimes only considerably later (Tycho’s nova stella, for example). Second, one or more explanations are forthcoming, at most one of which can be right. Eventually the issue is resolved, and the solution holds down to the present. Then, with expert hindsight, it becomes possible to recognize what additional input was needed for the resolution. (Very occasionally the need is spotted in advance and the items sought.) The time between perception of a puzzle and basic resolution was, for the GRBs, 24 years (1973 discovery to 1997 redshift), and this seemed like a very long stretch to those of us who lived through it. The record is, however, considerably longer.

THE LONGEST-LASTING PUZZLES

Pulsating variable stars

Boulliau discovered the periodicity of Mira in 1667 and attributed it to rotation at that period and spots more numerous and darker than those on the sun. When, in 1782-84, Piggott and Goodricke found periods for Algol and Delta Cephei, they explained both as eclipses by a dark companion. And there the situation remained for more than 100

years, until it became clear that, at least for Delta Cephei, the period was short enough that one star would have to be inside the other. Plummer in 1911 and Shapley in 1914 suggested pulsation, with $P \simeq (G\rho)^{-1/2}$, and Eddington by 1926 [4] had worked out the kappa (opacity) mechanism, persuasively for the Cepheids and tentatively for the Miras. What was needed was simply the idea of a dynamical time scale for stars, which would have been possible for Kelvin and Helmholtz (or even Newton, but he did not have enough information about stellar densities), and the duration of the puzzle from Boulliau to Eddington was 259 years. This is the longest I am aware of.

Sources of Stellar Energy

Newton remarked that the sun had been bright for a long time and that there was probably a reason for it. John Herschel put forward friction or electrical discharges in 1833 (arguably the recognition of a puzzle). Gravitational potential energy, in the form of infall of meteoric material or gravitational contraction came from Julius Mayer in 1841 and James Waterston in 1843-53 respectively. Their papers were both rejected, and the core idea habitually therefore called the Kelvin-Helmholtz time scale. It is only about 30,000,000 years, too short for some geological and biological considerations of the earth. Perrine, Eddington, and others advocated “subatomic” energy early in the 20th century. Atkinson and Hourtermans put forward catalyzed cyclic nuclear fusion in 1929-31 (before the discovery of the neutron, so that their helium atom when it spun off had four protons and two electrons in its nucleus). The details came from Hans Bethe in 1939, and had required all sorts of advances in fundamental physics, including special relativity, mass spectroscopy, the Saha equation (to get the composition of the sun), and quantum mechanical tunneling. Duration: 106 years.

Coronal lines in Solar Eclipse Spectra

Charles A. Young found the 5303 Angstrom line in 1869. The explanation was, of course, the element coronium, by analogy with helium (for a solar photospheric feature) and nebulium (for the 3727 line and others in planetary nebulae). As the periodic table filled out, it became clear there was no room for the latter, and Ira Bowen tied the nebular lines to forbidden transitions of oxygen and nitrogen in 1927. It was not, however, until 1939-42 that Grottrian and Edlen were able to recognize 5303 as a transition of Fe XIV, for a duration of 70 years. Both a new idea (that the corona is hot, rather than that the gas responsible is lighter than hydrogen) and a great deal of 20th century physics were needed.

Advance of the Perihelion of Mercury

LeVerrier reported a value of 0.5 arcmin/century in 1859. The intra-Mercurial planet responsible, Vulcan, was supposedly seen the same year and many times thereafter.

The alternative of gravity deviating from $1/r^2$ was largely rejected as conflicting with lunar and Venusian data, but tentatively supported around 1900 by Simon Newcomb (the first president of the American Astronomical Society, who generally gets fairly bad press). The solution was, of course, to be found in exceedingly new physics, as noted immediately by Einstein in his 1915 publication on the general theory of relativity. Duration, 56 years.

The Solar Neutrinos

The construction of Ray Davis's chlorine detector was carried out in close collaboration with theorists, who said he would see a good many events per month. By 1968 it was clear that the observed rate was, at most, one third of the prediction (3 SNU vs. 9 SNU). Explanations blossomed, if not quite 1000. These included modest adjustments in nuclear physics, much odder physics (quark catalysis, multiplicative creation, Maxwell tail depletion, WIMP energy transport), inefficient extraction of Ar^{37} , and large or small changes in solar physics (complete mixing, a metal poor interior, rapid rotation or magnetic support, and central black hole or an iron core). Curiously, the right answer, neutrino oscillation (propounded by Bruno Pontecorvo) was in the inventory from the beginning. But it took several new experiments and several new neutrino detectors to make this clear. SAGE and GALLEX were supposed to look at the problem, but the key discoveries were really the detection of oscillation of atmospheric neutrinos (SuperKamiokande 1999) and detection of the rotated flavors in the solar flux (SNO 2001). The duration of the puzzle was a personal matter, with opinion gradually swinging in the direction of new weak interaction physics between about 1985 and 2001. I think I opted for oscillation in 1994, about the same time Hans Bethe did, which was not a coincidence.

The Also-Rans

These are the puzzles for which I suspect that the correct answer is not yet on our plates or anyhow not entirely recognized. Some examples are (a) stability of spiral arms (winding-up problem recognized in 1896, density waves from Lin and Shu in 1964, and on to the present), (b) acceleration of cosmic rays (recognized as extrasolar in 1911, Fermi mechanism 1949, GRB component 2003...), (c) nature of the dark matter (cluster velocities, Zwicky 1933, LSPs, axions, Wimpzillas and all 1974 to present), (d) heating of solar and stellar coronae (recognition of Fe XIV in 1939, heating by acoustic waves, MHD waves, microflares, electric currents, and so forth down to the present, for a duration of at least 64 years). And there are probably others.

Credits

Most of these topics are discussed in one or more of Hoskins [7, 8], Bahcall [1], Hufbauer [9], Eddington [4], and Russell, Dugan, & Stewart [13].

THE DISCOVERY OF “GAMMA-RAY BURSTS OF COSMIC ORIGIN”: SERENDIPITY AND MISCONCEPTIONS

Casual reviews and textbooks quite often announce that GRBs were discovered in 1969 or even 1967. In so far as science is shared knowledge, both are clearly wrong. The discovery must be placed in 1973, when Klebesadel, Strong, and Olson (1973) [11] submitted and published their *Astrophysical Journal Letter*. The 1969 date is that of the first burst reported in that letter and the 1967 one a burst recognized slightly later, for which, however, there is a little directional information, so that it could conceivably have come from the sun.

Others have argued that the events should have been announced earlier. The same statement was made in 1968 about pulsars, and I think both again are clearly wrong. The time to announce is when you are as sure as you can be, short of building a new array or launching a new satellite, that you have seen something interesting. Thus post-mature publication is rather rare, but premature publication fairly common. It is easy to see the difference by tracking the rates of citations to the “discovery papers” over the years. For GRBs and pulsars (and also the classic 1963 Schmidt paper on quasars, the 1965 announcement of the 3K microwave background, and many others), a dozen or so citations in the first year rise to many tens or hundreds in the next couple of years and then jaggedly taper off, with occasional secondary peaks when something new and exciting happens in the field. The GRB spike with optical identifications in 1997 is one example; another is a sudden recrudescence of citations to a 1965 paper by Gunn and Ostriker when the effect they had considered finally became detectable.

In contrast, a premature publication has an equally clear, but different signal. There is a similar enormous peak in the first couple of years, sometimes as high as that for appropriate publication, but the taper is almost as fast as the rise, and continues right on down to zero, apart from occasional mentions in comprehensive review papers. Examples include the announcements of a pulsar in the remnant of SN 1987A (whose period was an alias of the video camera scan rate) and of a planet orbiting the pulsar 1829-10 (with a six month period that was incorrect allowance for the eccentricity of our orbit).

Were GRB’s predicted? It is easy to say, and I have said earlier, that two kinds of events were predicted and two have been seen. Unfortunately, they were not the same kinds. The predictions were for shock break out in core collapse supernovae (Colgate [3]) and evaporation of primordial black holes (Hawking [6]). The latter counts as a prediction despite its date because the author does not seem to have been aware of the data (compare Einstein and the Michelson-Morley experiment, perhaps!). The two sorts seen were the classical GRBs and the soft gamma repeaters. I think that the quip about the two kinds not being the same is no longer entirely true. Indeed, there is still no

observational evidence for Hawking radiation from evaporating PBHs, but the current understanding, at least of events like GRB 030329 is not really so very different (apart from extreme asymmetry) from what Colgate suggested.

Was the discovery “serendipitous”? The answer here is, I conclude, yet another firm no. The fictional princes of Serendip (Sri Lanka) went about making fortunate discoveries by accident. A much better model is the remark of Louis Pasteur, in French, of course, that chance favors (only) the prepared mind. The discovery paper, and other discussions from the 1970s say firmly that the Los Alamos group were aware of the 1968 report by Cline, Holt, and Hones [2] of a burst of gamma rays from the solar flare of 1966, near the beginning of a new solar cycle. Thus the discoverers immediately recognize some solar flares as the cause of the near-simultaneous triggering of detectors on widely-separated satellites, and set out to make sure that there were no other such non-terrestrial triggers. Except that there were. Unexpected, but not serendipitous.

Two “alternative universes” come to mind. First, could GRBs have been discovered any earlier? That is, was anything up before the Vela and Kosmos series satellites carrying gamma ray sensors with very good time-tagging of photon arrivals? The answer is yes, and one (possible) GRB was recognized retrospectively in the data stream from a balloon that had been looking for positrons connected with aurorae on 26 September 1966. Gamma ray emission at 511 keV just after a positively charged particle passes through your outer layer is, of course, a positron signal. It is, however, a little hard to imagine this sort of record being recognized as anything cosmically interesting unless there had been simultaneous balloon flights more or less on opposite sides of the earth.

Second, would GRBs have been discovered fairly soon anyhow? To this it is a bit harder to say no, categorically. Confirmations of specific events came very quickly from a large fraction of the devices above the atmosphere at the time and soon thereafter, including OSO-5, 6, and 7, OGO 3 and 5, SAS-1, IMP-6, Apollo-16, Uhuru, TD-1, IMP-7, and 1972-076B, as well as the Kosmos series. None of these, of course, had been planned with any such entities as GRBs in mind. In particular, event 1972-2 appeared in detectors, anti-coincidence shields, and so forth in IMP-6, SAS-1, OGO-5, and OSO-6 as well as three Vela satellites. One is led to suspect that sooner or later people associated with these projects might by chance have found out about each others’ excess noise, in somewhat the same way that Princeton found out about the Penzias and Wilson detection of excess antenna temperature.

WE HAVE TURNED EVERY ONE TO HIS OWN WAY

I do not know whether GRBs resulted in more “prompt” models than any other astrophysics phenomenon in history. What is true is that the set is better documented. As early as the 1974 “Texas” symposium on relativistic astrophysics, Mal Ruderman remarked, first, that the only model not taken seemed to be anti-matter comets hitting white holes, second, that the only theorist who had not published a model was J.P. Ostriker (who still has not; I asked him, in the only original research done for this presentation; but he and R.D. Blandford have discussed an electromagnetic model), and, third, that he personally was betting on Black Hole ridden by Accretion to win, with Glitch to place.

The first 118 models, up to 1992, were tabulated by Nemiroff [12] and indeed include a remarkable range of compact objects, large and small, innovative physics, known and unknown, and so forth. Nemiroff's list is not quite complete even for its period. Colgate's prediction is there, but not Hawking's. Also missing are Jelley (1974) [10] who put forward a Leblanc-Wilson type collapse of a rotating magnetized star (that is, something rather similar to the current best buy) and Harris [5] who contemplated exhaust trails from interstellar space craft (accounting for the possible positron annihilation line in some spectra), looked for events in straight lines across the sky, and didn't find any. Slightly later came magnetar models and collisions of comets in the Kuiper belt.

Among my favorites within the list are directed stellar flares (the first beamed model), proposed by Brecher and Morrison, and Goblins (chunks of neutron star material released from gravitational confinement and exploding to white dwarf densities), the last published paper of Fritz Zwicky. Should one disparage this wild blooming of scenarios? It is easy to do so, in the form, for instance of a theorem said to be due to R.O. Redman, "A competent theorist can explain any set of observations using any theory," and a corollary actually due to Malcolm Longair, "In many cases he need not to even be competent." But the reality is (another quote, this time from Joe Weber) that theorists are cheap and telescopes are expensive, and so the fullest possible consideration of any datum is only just good sense.

Several curious sidelights appear in any extended list of models. One is the "incident of the gamma ray in the nighttime," when, in 1974, at that same Texas symposium, Stirling Colgate said, "I do not believe that hard X-ray pulses are not created in supernova events. We have just not been fortunate enough to observe the phenomenon." Most GRB2003 participants would probably say that, with GRB 0303029 = SN2003dh, we have now observed it. Another oddity is the distinction between "premature" considerations and "just in time" delivery. The Brecher and Morrison beaming was perhaps too early to make a mark, as was a 1993 paper by Rhoads and Paczynski that included a forecast of radio afterglows. In contrast, the 1997 prediction by Meszaros and Rees of optical afterglows slipped into print just before the photons were recorded. Only Stanton Peale's conclusion that the interior of Io should be very hot (because of tidal stressing by Jupiter) published days before the discovery of the volcanos strikes us as better timed.

ALL WE, LIKE SHEEP, WENT ASTRAY

The earliest models included a mix of galactic and extragalactic events, but, throughout the 1980s and lingering even into the BATSE era, there was truly remarkable consensus in favor of a picture where the bursts were something which happened on the surfaces of old, fairly nearby, neutron stars that had managed (somehow) to maintain magnetic fields near 10^{12} G and which the neutron stars would survive. How did this happen? There were, I think, something like seven contributing factors.

First was the 1976 discovery of X-ray bursts by Grindlay and his colleagues seen largely in the direction of the galactic center and correctly interpreted by Joss and by Woosley as nuclear explosions on neutron star surfaces. Indeed Woosley's model had the status of a prediction, because it was originally intended to apply to GRBs.

Next was the discovery of cyclotron resonance features in the X-ray spectrum of the intermediate mass X-ray binary Her X-1, recorded in a 1976 balloon flight coordinated by Truemper. The implied field was about the 10^{12} G expected from pulsar data. Third came an excessive distrust of any deviation of the N(S) relation from a $-3/2$ power law. The $3/2$ slope is an indicator of a uniform distribution in space, while flattening would mean we are seeing the edge (and steepening that the sources are commoner far away and/or long ago). Indeed a couple of early announcements of flattening were over-interpretations of data with awkward selection effects and such, but I believe that the White et al. (1983) balloon results [14], which actually included several events down to a fluence of 10^{-7} erg/cm² should have received wider credence that it did.

Fourth were spectral features, of two sorts, reported in the GRBs themselves. According to the Soviet detectors, the March 5th, 1979 burst (now a part of a long trail of a soft gamma repeater) had a redshifted positron annihilation feature near 400 keV (which a real detection of the line in the region of Sgr A* had conditioned us to believe). Less soft events in the next couple of years also seemed to have positron annihilation redshifted just about as much as you would expect for photons that had climbed out of a neutron star potential. These were gradually discounted beginning in 1984, when SMM did not record them.

The other sort of spectral feature was cyclotron resonance, like that of Her X-1, reported by the Soviet group, soon after from SMM and HEAO-1, and especially from Ginga data. The implied magnetic fields for the 10-70 keV resonances were again in the 10^{12} G range expected for neutron stars. This was apparently a defining datum for a large fraction of the GRB community. Indeed even I, in a 1990 review, listed as the best buy model "mergers of binary neutron stars with strong magnetic fields at cosmological distances." BATSE never really saw any such spectral features, and the number of GRBs who would still defend them is small, but not zero.

Fifth "distractor" was the reported optical flashes at GRB positions but at earlier times, for instance a 1928 blip on a Harvard plate where a 1978 GRB happened. Schaefer and Hudec each found some of these, and they meant that the source had to survive a burst and do it again, something like once per century. In retrospect, the inventory included plate flaws, flare stars, and possibly a few SNe in the GRB host galaxies.

Sixth was the "no host" problem, the absence of bright galaxies in the error boxes of some of the best localized, brightest GRBs. This also came from examination of archival plates, again by Schaefer, and has to be ascribed to a combination of the hosts (at least of long duration events) being genuinely rather faint and bad luck. Seventh and last were the neutron star runaways. Some pulsars are indeed relatively high velocity objects, probably kicked by asymmetrical supernovae and binary star disruption during their formation. For the Crab pulsar, the velocity is somewhat less than 200 km/sec, which won't have carried it far in its lifetime, but there were claims of 1000 km/sec and more for pulsars once in (but now on the edge or outside of) much older remnants. That sort of velocity would easily take a 1-10 Gyr neutron star into an extended galactic halo. And an extended galactic halo of sources could, just barely, account for the early CGRO burst census, in which they remained very isotropic on the sky but, at long last, we definitely saw the edge of the distribution.

Even at the time of the 1995 staging of the 75th anniversary celebration of the Great (Curtis-Shapley) Debate, organized by Robert Nemiroff, a vote of the participants

showed about equal numbers for “galactic” and “extragalactic.” The only effect of the debate itself (between D.Q. Lamb and B Paczynski) was a large increase in the faction of the audience who declined to vote. Not all minds were made up until (or perhaps after) the 1997 redshift measurement, though in September 1991, Martin Rees offered Bohdan Paczynski a 100:1 bet in favor of “galactic,” and then, after the fall of 1992 announcement of BATSE isotropy and turn-over in $N(S)$, opined, “We were both fools. I for offering the bet; Bohdan for not accepting it.”

SHARING GRBS WITH THE WORLD

I looked two sorts of places to see how this played out, first the proceedings of the biennial Texas Symposia, and, second semi-popular journals (Sky and Telescope and Scientific American). At Texas, GRBs did not exist in 1972, in 1974 got both a data review and the Ruderman theoretical presentation already mentioned. In 1976 and 1978 they again did not exist. The talks in 1980 and 1982 relied heavily on the spectral features, and 1984 saw another eclipse. Local neutron stars with strong fields were the only respectable models in 1986 and 1988, and Hartmann in 1990 again emphasized the cyclotron (Ginga) features as our definitive line of evidence.

The 1992 Texas GRB talk was given by Paczynski. Not that he had ever really deviated from extragalactic models, but it marked the sea change introduced by early CGRO data that he was the person asked to review the topic. Rees voted 50% for extragalactic, 30% for extended halo, and 20% for galactic disk and the $N(S)$ turnover still not being real.

On the popular front, both journals recognized the existence of the phenomenon (with articles written by or drawing directly on the expertise of the discoverers). Optical counterparts yielded renewed attention, and then the early BATSE data. S&T also reported a range of false alarms (correlations with QSOs, repeaters, etc), but both settled into the extragalactic camp by the end of 1997. Most striking to this reader, going back to the four Scientific American articles and the 16 S&T reports and reading them all at once, was how much more informative the early articles, written by active GRB astronomers, were than the later ones written by staff members. In one of the worst (S&T 1996) the issue of the cyclotron features is not even mentioned.

How have GRBs done in (non-technical) astronomy education? Not very well. Of 15 standard texts in multiple editions that I examined, only Jay Pasachoff’s regards them as interesting enough to mention the galactic/extragalactic issue before 1997 or to address how one might resolve it by optical counterparts with spectra. Since 1997, about half of the textbooks mention GRBs and say that the optical data settled a previously existing (but apparently not known to them!) controversy. Most of the books also do not score high on two other items from the history of astronomy (big bang vs. steady state and heliocentric vs. geocentric) where we have the opportunity to use the idea of falsifiability to indicate how science differs from “other ways of knowing.”

ACKNOWLEDGEMENTS

I am indebted to Ed Fenimore for the invitation to deliver the after dinner talk at GRB2003 and to Mark Galassi for efforts above and beyond the call of duty to figure out the logistics of dinner in two separated rooms followed by a talk with images. Special thanks go to an anonymous colleague who generously downloaded and printed out pictures of the astronomers and physicists who appear in Sect. 2, and some others, for display with poster, but who very sensibly declined any further involvement with the presentation or, indeed, with me. It is not possible to leave this topic without quoting Philip Morrison, who said (in another after dinner talk) that “it is hard to waste 10^8 dollars.” He had in mind both the GRB discovery and evidence for plate tectonics that came from a seismic array also intended to detect illicit tests for nuclear weapons. Only inflation and LIGO make me suspect this should now be 10^9 dollars.

REFERENCES

1. Bahcall, J.N., Neutrino Astrophysics. Cambridge University Press (1989).
2. Cline, T.R., Holt, S.S & Hones, E.W., JGR **73**, 434 (1968).
3. Colgate, S.A., Canadian J Phys. **46**, S476 (1968).
4. Eddington, A.S., The Internal Constitution of the Stars. Cambridge University Press (1926).
5. Harris, M.J., J. British Interplanetary Assoc. **43**, 551 (1990).
6. Hawking, S.W., Nature **248**, 30 (1974).
7. Hoskin, M.A. (Ed.) The Cambridge Illustrated History of Astronomy, Cambridge University Press (1997).
8. Hoskin, M.A. (Ed.) 1 The Cambridge Concise History of Astronomy, Cambridge University Press (1999).
9. Hufbauer, K., Exploring the Sun: Solar Science Since Galileo. Johns Hopkins University Press (1991).
10. Jelley, J.V., Nature **249**, 747 (1974).
11. Klebesadel, R.W., Strong, I.B., & Olsen, R.A., ApJ **182**, L85 (1973).
12. Nemiroff, R.J., Comments on Astrophysics **17**, 189 (1994).
13. Russell, H.N., Dugan, R.S. & Stewart, J.Q., Astronomy. Boston: Ginn & Co. (1926).
14. White, R.S. et al., Nature **271**, 635 (1983).