Astrophysical Modeling: Gamma-Ray Bursts

As our final class, we'll take a look at one of the most exciting and controversial fields in all astrophysics: gamma-ray bursts. These events may have a higher peak photon luminosity than any other events in the universe, and their physics is therefore extreme enough to have motivated a number of exotic suggestions. Their spectra and brightness also may make them great backlighting for the universe, meaning that absorption lines in their spectra can tell us about the composition and evolution of the $z \sim 5 - 20$ universe. The history of this field is also an object lesson in how new evidence can shift opinions dramatically. At one time or another, a substantial majority of people in the field have believed (with a fair degree of certainty) that gamma-ray bursts are (1) in the Galactic disk, (2) standard candles, (3) caused by merger and inspiral of two compact objects, and (4) the product of a special type of supernova. Current opinion favors (4), but new evidence can always change this. To me, the study of gamma-ray bursts encapsulates much of what makes the scientific process unique. It is a subject filled with rancor and conflict, but the emergence of new data has had its say in a way not available with pure philosophy.

Brief summary of properties

Let's first summarize briefly what gamma-ray bursts are. Indeed, it is somewhat difficult, because unlike many of the sources and phenomena we've discussed, gamma-ray bursts are rather heterogeneous in their properties. The first two properties come from their name: the emission is primarily in gamma rays (with a spectral νF_ν peak in the hundreds of keV), and the events have a limited duration (from milliseconds to about a thousand seconds, as seen so far). There seems to be a broad bimodal distribution of durations, one peak being less than a second and the other being at 10-20 seconds. Unlike X-ray bursts, the profile of the flux with time is not universal. Many bursts have a "FRED" profile (fast rise, exponential decay), but others are more spiky, or have some emission, a long quiescent period, and then have more emission. It appears at this time that the distribution of locations of bursts on the sky is consistent with isotropic, although occasional evidence for weak clustering in a subset of bursts is reported. There is also no definitive evidence that any burst has repeated, although some events are consistent with a repetition of up to four events. The positional uncertainties for almost all bursts are large (at least several degrees), which is why statements about isotropy and repetition are difficult to make. The flux observed at Earth has an extremely broad range between different burst, from a maximum of about 10^{-3} erg cm^{-2} s^{-1} to the flux limits of detectors, down to 10^{-8} erg cm^{-2} s^{-1}. All bursts that have been localized enough for pointed follow-up have X-ray afterglows lasting days (before they are too weak to detect), and about half have detectable optical afterglows. The spectrum and the time development of the bursts are adequately described by power laws with a few breaks in them. Redshifts (or at least lower limits to the redshift) have been obtained for a number of bursts, clearly indicating that many, perhaps all, bursts are at large cosmological distances.

History of detection

Gamma-ray bursts were first discovered as a byproduct of the Cold War. In the late 1960s there was a concern that the Soviets might test nuclear weapons in space. The US decided that it needed to be able to detect the gamma-ray emission that would result, and it therefore launched the Vela series of satellites. They were alarmed when, starting in 1968, the satellites detected gamma-ray flashes from space! The spatial resolution of the satellites was poor, but eventually it was determined that the flashes came from outside the solar system, so in 1973 the flashes were reported publicly.

In 1979 there was an apparent breakthrough in the study of gamma-ray bursts. On March 5, 1979, nine separate satellites detected a remarkably strong burst (impressive enough that this is simply known as the "March 5 event"). Many of these satellites were far enough away from the Earth that it was possible to localize the direction of this event by timing; an aid to this localization was that the event had an extremely sharp onset. This event came from the N49 supernova remnant in the Large Magellanic Cloud, and later was even more specifically determined to come from an X-ray hot spot in the cloud. This was exciting, because this was the first time that a GRB had been identified with a quiescent source. Moreover, this source repeated; 16 more bursts were seen over the following months. However, it is now thought that this event was the first identified member of a separate class, soft gamma-ray repeaters. At the time, though, this mislead people for a long time, because it appeared that this was clear evidence for a Galactic source of the bursts, and it was so clearly established that it appeared to be a fixed point in the data.

In the 1980s, other bits of evidence appeared to support the local origin of the bursts. Data from the Japanese satellite *Ginga* for several bursts suggested the existence of cyclotron absorption-like features in three bursts, one that appeared very secure. This also seemed to argue strongly for a relatively local origin. The point is that without any persistent sources or direct evidence of distance, a given flux is not informative about the distance (in the dark, a light could be a nearby firefly or a distant airplane). However, the argument was that if the distance was cosmological, the luminosity would be so high as to prevent the formation of lines near a compact object. At the end of the 1980s, virtually the entire community (with the notable exception of Bohdan Paczynski) was sure that gamma-ray bursts mostly came from neutron stars in the disk of the Galaxy.

In 1991, the Compton Gamma-Ray Observatory was launched, as one of NASA's Great Observatories program. The Burst and Transient Source Experiment (BATSE) was particularly well-suited for detection of GRBs, since it had a low flux limit and all-sky coverage. It also had better angular resolution than previous instruments, although even for bright bursts the location was no better than two degrees and for dim bursts it was 30 degrees or worse. Prior instruments had detected no deviation from isotropy, but it was expected that with BATSE's much more sensitive detectors that a bias towards the Galactic plane and center would be seen. It was not. However, what was seen was a rollover at low fluxes compared to what would be expected in a Euclidean universe with a constant density of sources.

This radically changed the way that most people thought. The combination of isotropy with a deficit of dim sources is exactly what is seen in cosmological populations of all types. The expanding universe means that beyond a redshift of roughly unity, there is less volume to play with, so if there was a constant comoving density of sources then there would appear to be fewer sources at large distances. When this result from BATSE was established firmly, therefore, most people switched over to thinking that GRBs were cosmological. As a result, previous evidence in favor of a more local origin was discounted (e.g., it was now felt that the *Ginga* "cyclotron lines" were statistically insignificant). There were also two sources discovered that were similar to the source of the March 5 event, which were given their own separate class, soft gamma-ray repeaters. However, Don Lamb and colleagues pointed out that a population of high-velocity neutron stars in the halo of our Galaxy could also explain the observed isotropy and falloff. The majority of the community didn't agree with this, although a debate held in 1995 in Washington, D.C. helped convince people that the case wasn't open and shut in favor of a cosmological origin. The main problem was one of physics versus astronomy. Isotropy and a rollover in the brightness distribution has, historically, suggested a distant cosmological origin. On the other hand, it's a lot easier to figure out energy sources on the scale of 10^{38-42} erg than on the scale of 10^{51} erg of gamma rays. For the next breakthrough, a smoking gun was needed.

The main problem was that now that the March 5 event was considered separate, no quiescent counterparts of GRBs existed. This divorced the field of GRBs from the rest of astronomy, and made further progress difficult. One problem was that most GRB localizations were with BATSE, which could only do a couple of degrees at best, and in that kind of area an unlimited number of sources exist. A second problem was that the interplanetary network (IPN), with which relative timing could do much better localization, was down to two satellites (BATSE and Ulysses) after 1992, so only a long, thin arc could be established. Into this mix, in 1996, came the Italian-Dutch satellite BeppoSAX. Initially the goals with this satellite had nothing to do with GRBs. However, starting in 1997, it was able to localize $\sim$10 bursts per year to accuracies of a few arcminutes. This is a small enough area that optical and X-ray pointed observations could be brought to bear quickly. People then looked for initially bright sources that faded... and found them. In 1999, a rapid pointing even found an optical source that reached 8th magnitude just seconds after the GRB. The optical observations localize the source to a fraction of an arcsecond, and has allowed spectra to be taken that prove the sources are at high redshifts (more than $z = 3$ in some cases). This settles the question, although to be cautious one must point out that so far only the long-duration class of GRBs has had follow-up observations of this sort. One can at least say that some GRBs, perhaps all, are cosmological. Recent observations with ASCA and Chandra suggest that some may have iron lines, although these are weak and debatable.

The flux distribution of bursts themselves is nicely fit by a model in which the rate of bursts per volume is unchanged throughout the history of the universe, and in which the bursts have a constant rest-frame luminosity (hence, in which they are non-evolving standard candles). This was the standard model for a while, but BeppoSAX observations and their follow-up have shown that the rate per volume was much higher in the old days; in fact, it appears roughly consistent with the star formation rate as a function of redshift, so the new standard model is that the bursts happen in star-forming regions. Note, however, that statistics are poor, and although it is clear that the GRB rate is now less than it was, it is not clear that it is nailed to the star formation rate.

Fireballs and afterglows

In the mid-1990s a number of researchers realized that regardless of what the energy source was, the release of 10^{51} erg or more in a few seconds or less would produce an expanding fireball, and that the interactions of this fireball with the surrounding medium would yield potentially robust signatures, including afterglows. The observation of these afterglows is therefore something of a confirmation of the models, although the models have enough parameters (due to legitimate uncertainties!) and the observations are featureless enough that the association wasn't instantly convincing. The initial models had a very quick release of energy (the standard picture was of merging compact objects, taking milliseconds), and the gamma-rays as well as the afterglow were attributed to shocks and interaction with a clumpy interstellar medium. Ed Fenimore and colleagues showed, however, that this would not produce the observed properties of GRBs, so now it is thought that the gamma-ray bursts reflect the rest-frame duration of the event, and that internal shocks in the outgoing fireball account for the gamma-ray emission. The general success and robustness of the fireball model seems at this time to be a theoretical fixed point in a very uncertain situation.

The Central Engine

The most interesting question related to GRBs is what powers them. The acceptance that they originate from significant redshifts represents a major shift in what has to be considered. When they were thought to originate from ~ 1 kpc, this was easily accounted for by any number of processes on neutron stars. But at $z = 1$, a 10^{-5} erg cm^{-1} s^{-1} burst means an isotropic luminosity of 10^{51} erg s^{-1}, which is another story entirely! In fact, this luminosity is much greater than the peak luminosity of supernovae. What, then, could do this?

In much of 1990s, attention focused on mergers of compact objects. The merger of two orbiting neutron stars releases some 10^{53} erg, and therefore would be enough to power a burst. However, there are some potentially major problems with this. First, the timescale of merger is milliseconds, not tens of seconds as is usually seen (and is the rest-frame time, from Fenimore's work). Second, if there are too many baryons in the fireball, then the energy all goes to them and the resulting Lorentz factor is too low. This would mean that afterglows would take years instead of days, and the peak in emission would be at much lower energies than observed. Therefore, there has to be some way to have the burst occur in a "clean" direction that has a deficit of baryons. There were suggestions that neutrinos and antineutrinos would annihilate and produce leptons far enough away that the requisite cleanliness would be achieved, but this turns out to be too inefficient. In addition, the merger times of binaries can be quite long (billions of years, as is the case for three NS-NS binaries in our Galaxy). One would therefore expect cases in which the binary, having received kicks from the two supernovae, had traveled great distances from their host galaxy, and had delayed merger enough to occur long after star formation had dropped off. One would not expect such a strong evolution as is seen, although there are lots of uncertainties about that as well.

The new standard model therefore involves a special type of supernova. The idea is that a massive star evolves quickly, so it doesn't have time to leave its birthplace or explode long after the starburst. The type that produce GRBs are sometimes called "hypernovae". The idea is that the formation of a rapidly rotating $\sim 10\,M_\odot$ black hole in such a supernova will establish a jet and a preferred axis that may clean away baryons. The total energy is perfectly adequate, as well. However, the progenitor stars have a good ten solar masses of baryons just waiting to slow down the fireball, and no one has come close to explaining how they end up so clean, at least in a quantitative way. It is, however, thought that a jet is necessary, because the isotropic energy release inferred is otherwise up to 3×10^{54} erg, which is the rest-mass energy of a neutron star!

Where to go from here

Looking back on the history of GRBs, several models or pictures have held sway for many years, but have ultimately been discarded. The current fashion seems to be in favor of hypernovae, but caution is required! There are many, many questions that have yet to be answered. How are the short bursts formed (hypernovae would have timescales of tens of seconds)? How are the high Lorentz factors achieved? What makes a hypernova different from a regular Type Ic supernova, which it is supposed to resemble (the rates are such that only a tiny fraction, 10^{-4} or less, of such supernovae produce gamma-ray bursts)? There is a lot of heat and not much light about these questions at the present.

Even without such detailed knowledge, however, it may be that GRBs will yield valuable information in other ways. It has been pointed out that their energy spectra and time development are such that their observed flux does not decrease rapidly with increasing redshift. These therefore could be ideal for probing the structure of the early universe through absorption-line systems, due to their brightness.

It is remarkable that so little is known about these sources after 30 years. However, HETE-II and the planned launch of Swift will greatly increase the number of localized GRBs. Combined with fast follow-up in other wavebands, statistics will improve dramatically. This will also increase the range of redshift observed (greater z because of better sensitivity and fast follow-up, and lower z because of greater sky coverage). I think one can expect that in five years a much greater knowledge will have been obtained. Full understanding, however, may have to wait for computers and MHD modeling of the bursts to catch up. After all, supernovae have been known for seven decades, and the models still have difficulty reproducing them! On the positive side, there is a lot of room for discovery and exploration of the bursts, both observationally and theoretically.