Mysterious energy machines in outer space

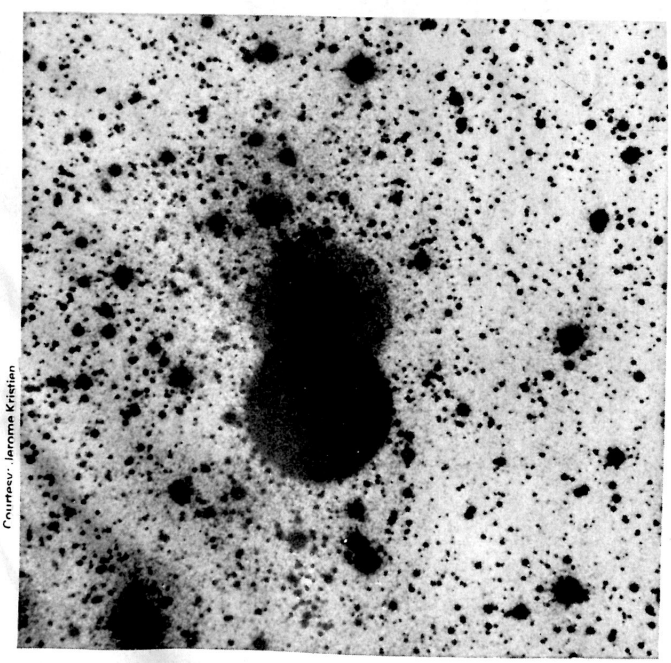
Jayant V. Narlikar

Space astronomy has provided evidence of cosmic sources which pour out enormous amounts of energy in short bursts. Where does all this energy come from? This question poses a challenge to theorists. Some of the possible processes are indeed stranger than the phenomena themselves

ASTRONOMY provides a laboratory much wider in scope than any that exists on the Earth. The theory of nuclear fusion in the 1930s and the present work on black hole astrophysics show that fundamental physics can get valuable inputs from cosmic observations. Recent observations from ground-based telescopes and space-based detectors have revealed the existence of many violent phenomena in the universe which pour out vast quantities of energy. These observations have posed challenges to the theorist to account for the sources of such violent events.

To human beings confined to the *terra firma*, the strongest source of energy appears to be our Sun. Radiating power at a rate of some four hundred million million million watts, the Sun generates so much energy that the amount of sunlight falling on the Earth could meet all foreseeable human needs of energy many thousand times over. Yet, in the cosmic setting, the Sun, an average star, is very low down on the energy ladder.

Though stars and galaxies held the centre stage in astronomy for the first three or four decades of this century, new astronomical techniques, which came in use after World War II, showed the existence of powerful sources of energy both within and outside our Galaxy. In the mid 1950s, extragalactic radio sources were discovered, with energy reservoirs estimated at 10^{58} to 10^{62} ergs. In the early 1960s, the remarkable quasistellar objects (QSOs, or quasars) which, while keeping a star-like appearance, can outshine an entire galaxy, were discovered. And he late 1960s and the early 1970s s w the detection of strong X-ray ources in binary or double star sy stems. Besides such steady emitters of energy, space astronomy has now presented evidence of cosmic sources which pour out energy in X-rays or gamma-rays in a series of short and violent bursts. These are esoteric phenomena indeed!



Courtesy: Jerome Kristian

The dark object at the centre is the giant star in the constellation Cygnus. Orbiting this star is a compact object of 8 solar masses that accretes matter from the star's outer layers and emits X-rays and radio waves. This object Cygnus X-1 is presently the best known candidate for a black hole. Also see Fig 1 on page 20

But where does all this energy come from? What are the processes which release such vast amounts of energy in so dramatic a fashion? Theoretical astrophysicists have been searching hard for answers. Indeed, some of the theoretical scenarios accounting for the sources of this energy that are presented here appear no less esoteric than the phenomena themselves. But remember, what appears unusual or impossible in the terrestrial context may be quite reasonable and plausible in the cosmic setting.

Stellar energy

Before we look into these esoteric phenomena of cosmic energy generation, consider first the now familiar example of how stars generate energy. Stars, like the Sun, shine because they produce energy deep in their interiors through the thermonuclear fusion of hydrogen into helium. The total mass of all the fusion products, however, falls short of the mass of the original hydrogen used up. This deficit in mass is released as energy. Using Einstein's famous mass-energy relation, $E=mc^2$, we find that of the total mass put into fusion, only a tiny fraction, about 0.007, is converted to useful energy. Even so, it is sufficient to keep the Sun shining for billions of years.

A star is able to control thermonuclear fusion in its core because of its strong self-gravity. The thermonuclear reactions in the core of the Sun cause high temperatures which generate strong outward pressures; these act against the Sun's own gravity, preventing it from contracting, and holding it in equilibrium. Remove the internal pressures in the Sun somehow, and it would collapse under its own gravitational force to a mere point in about 29 minutes! (A simple calculation based on Newtonian dynamics and gravitation will show this.)

The equation of stellar structure, originally set up in the 1920s by A. S. Eddington, the English astronomer, relates the gravitational force in the

star to the progressive changes of pressure from its centre outwards, the magnitude of pressure to density and temperature, and the fall of temperature outwards to the flow of energy from the interior to the surface. From these equations, stable models of stars emerge, with central temperatures high enough to start and sustain thermonuclear fusion. The key role, of course, is played by the controlling force of gravity. The large mass of an astronomical system makes gravity an important, often the most important, factor in determining its behaviour.

With this background, we shall now look at some of the cosmic energy sources. For purposes of comparison with the Sun, we will express masses in the unit of one solar mass, $M_\odot$=2 billion billion billion tons.

Accretion discs

Take first the binary stars, which emit X-rays. Why and how do they emit these X-rays?

Over three decades ago, Fred Hoyle, Raymond Lyttleton and Hermann Bondi had proposed the mechanism of how stars accrue interstellar matter. The basic idea was that the strong gravitational pull of a massive object would draw the surrounding matter on to the object. The idea found an echo in the 1960s and 1970s in the work of others—Zeldovich and Guseynov, Pringle and Rees and of Shakura and Sunyaev. This later work related to the binary star system and its eventual outcome was to provide an explanation why binaries emit X-rays.

The ideal situation for this model to work is when there are two stars going round each other, with one star, A, a giant or supergiant and the other star, B, a compact neutron star, as shown in figure 1. In a close binary system, A and B are not separated much more than the sum of their radii. Under these conditions, star B exerts strong tidal forces on Star A, pulling out material from its surface which goes round B and ultimately falls onto it. The dynamics of this process suggests that the infalling matter forms a disc round B. In a steady state, the disc preserves its structure and shape. However, it is fed with a supply of matter from A.

This material falling into B radiates energy as it winds its way inwards within the disc. Detailed calculations of the temperature of the disc show that in binary star systems, most of the emitted radiation would be in the range of frequencies commonly associated with soft X-rays.

Techniques of X-ray astronomy in the 1960s were, therefore, designed to detect soft X-rays. And among the early X-ray sources to be discovered was Hercules X-1, which has indeed been identified with the binary system containing the star HZ Hercules. But this compact companion B is not visible in optical observations, though its presence can be deduced from observations of A, and further, X-rays from B are periodically cut off when the star is eclipsed by A. The star B, which emits X-rays, is, therefore, assumed to be a neutron star which is a superdense mass mainly of neutrons formed when the nuclear fuel in a star eventually burns out and the star contracts under its own gravity; because of the strong gravity, only nutrinos and high-energy photons escape from the star, but not light, and hence it is not visible. Also, the mass of star B is estimated to be close to that of the Sun, which is well within the maximum permitted theoretical limit for the mass (about 3Mo) of a neutron star. In general, all binary stars emitting X-rays can now be explained on the assumption that one of the stars, star B, is a neutron star.

Except one. This is the binary X-ray source, Cygnus X-1; the mass of the compact companion star B in this system is estimated to be at least 5Mo, and therefore, it cannot be a neutron star. Indeed at present there is no known state of matter which, in a highly dense form, could generate large enough internal pressures to support such a massive star in equilibrium. The only alternative state possible for the massive star is that it contracts under its gravity to end up as a black hole. Hence it is usually concluded that the companion member in the Cygnus X-1 is a black hole.

To date, Cygnus X-1 is the only apparently clear-cut example of a binary containing a black hole. This circumstance has led some astronomers to doubt whether Cygnus X-1 could really be interpreted as a black hole. But then, if there is no black hole in the system, what is the source of its X-radiation?

The Penrose process

On a much larger scale than stars, astronomers have noticed highly collapsed objects in the nuclei of galaxies. For example, the nucleus of the galaxy M87 is believed to contain a massive dense object of 5 billion solar mass, probably a black hole. Accretion discs can also form around rotating massive objects which are sufficiently collapsed. It has been suggested that such discs may form around massive black holes and give rise to radiation that is observed in quasars, for example. Before we come to this picture, let us consider an interesting idea of extracting energy from rotating black holes, proposed by Roger Penrose in 1968.

A black hole is an object which is so highly collapsed and dense that the strong gravitational pull at its surface prevents even light from escaping. Since, commonly, light (or any other form of the electromagnetic wave) is our sole means of detecting an object, the black hole must remain invisible. Nevertheless, it can be detected by the gravitational influence it exerts on its surroundings.

Since a black hole, by definition, gobbles up all matter and radiation coming sufficiently close to it, matter falling into it from the surroundings tends to increase its gravitational mass. With this one-way process going on, it seems unlikely at first sight that energy could be extracted from a black hole. Yet this is precisely what the Penrose process is designed to achieve.

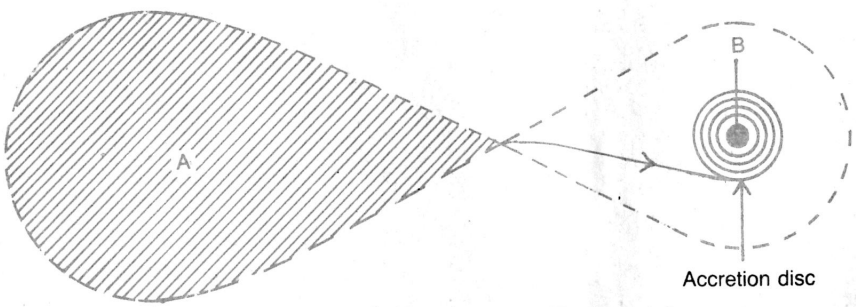


Fig 1: Schematic description of a close binary system. The star A is usually a giant or supergiant star while B is a compact star. Material flowing from A to B forms a disc round B. This disc emits radiation predominantly in the form of X-rays

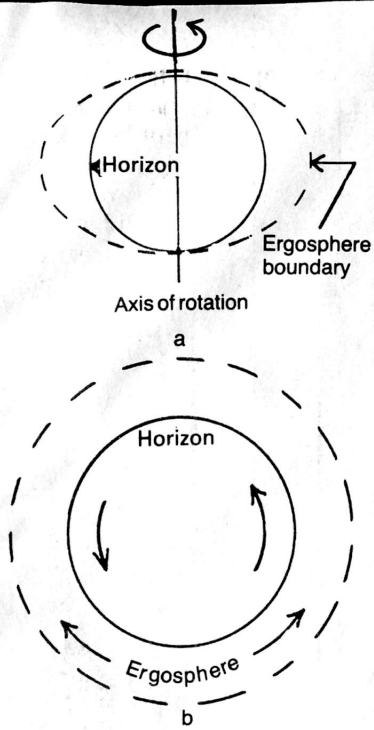


Fig 2 (a): The meridional section of a Kerr black hole. The inner (thick) circle is the horizon while the outer (dotted) curve is the static limit

Fig 2 (b): The equatorial section of the Kerr black hole shows two concentric circles, the inner (thick) one being the horizon and the outer (dotted) one being the static limit

To see how this comes about, let us take into account a crucial property of a *rotating* black hole, its surface area. The geometrical features of the space outside such a black hole are shown in figures 2(a) and 2(b).

Figure 2(a) shows the meridional section of space through the axis of rotation of the black hole. The inner (circular) curve is the section of the "horizon" which acts as an absolute barrier, for any physical signal from inside to go out. The outer curve is the so-called static limit which touches the horizon at the poles but is otherwise outside it. No physical observer can maintain a fixed position relative to distant stars if he is on or inside the static limit: he is carried along in the direction of rotation of the black hole no matter how hard he resists this tendency. Figure 2(b) shows the equatorial sections of the horizon and the static limit, the former lying inside the latter. The space between the horizon and the static limit is called the *ergosphere*. This is the region where it is possible to extract energy from the black hole.

Now, one of the fundamental rules governing the interaction of a black hole with ambient matter is that, whatever happens, the area of the horizon cannot decrease. This one-way behaviour is not unknown to physicists. In fact, in the last century physicists like Clausius, Kelvin and Boltzmann, while studying thermodynamics, had noted that all physical systems, left to themselves, increase their internal disorder. This disorder is measured by a quantity called the "entropy" of the system. The second law of thermodynamics tells us that the entropy of an isolated system can never decrease. The non-decrease in the area of a black hole horizon is by analogy called the second law of black hole physics. This law when applied to the Kerr black hole tells us that it is possible to extract energy if we *decrease* the angular momentum of the black hole also. This is the basis of the Penrose process.

The process itself is illustrated in figure 3. A piece of matter of mass m is hurled towards the ergosphere. While it goes round the black hole it is split into two parts, I and II. Part I falls into the black hole while part II emerges from the ergosphere with a total energy EII that is greater than mc^2. The energy of part II has been acquired through rotation in the ergosphere. The amount of energy and the angular momentum gained by it are at the expense of the black hole.

Clearly, the process works at its most efficient level if the black hole area is not allowed to increase. It can be shown that at the most about 30 per cent of the total rest mass energy of a black hole can be extracted this way. Although this is the highest efficiency factor known for any cosmic energy machine, black hole theoreticians have not been clever enough yet to think of an astrophysical scenario where this process can actually work. However, there is a scenario which is considered workable and has been advocated by many as the 'best buy' amongst the various cosmic energy machines.

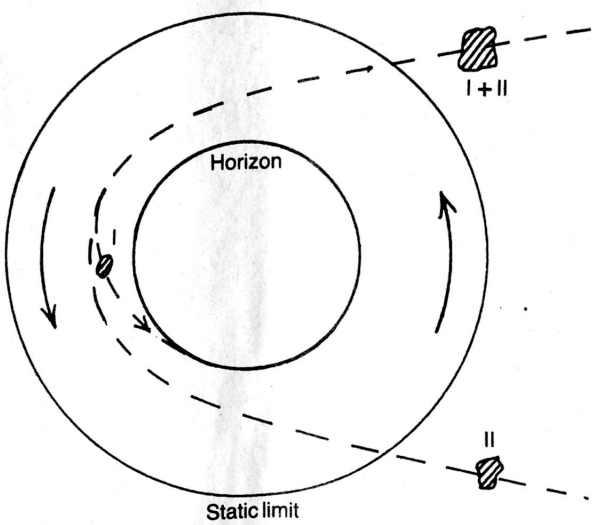


Fig 3: Penrose's scheme of extracting energy from a rotating black hole by diminishing the black hole's angular momentum. The final residual mass (II) which emerges has greater energy than the total mass (I + II) which went into the ergosphere

Radio galaxies and quasars

The discovery of radio sources outside our Galaxy and of quasars brought the energy problem to a head. The early identification of the radio source Cygnus A with a pair of colliding galaxies led to the belief that radio sources arose from such spectacular phenomena as collision of galaxies. However, this belief turned out to be false on several counts.

One of the reasons was theoretical. Galactic collisions are not improbable. Indeed, in a cluster of galaxies such close encounters are to be expected once in a while. The difficulty arises when we consider the amount of energy that can be released in a typical collision. Newtonian dynamics tell us how to calculate the energy of collision between two galaxies of known masses. For masses of the order of 10^{11} $M_\odot$ and of galactic dimensions, the collision energy comes out around 10^{59} erg.

Large though this number is, it is not large enough! In 1958, Geoffrey Burbidge made an estimate of the store of energy needed in a typical strong radio source. In making his estimate, Burbidge assumed that radio waves arise from fast electrons accelerated by ambient magnetic fields, a process known as synchrotron emission. Using the observed parameters of radio sources, Bur-

bidge calculated that the *minimum* energy needed in a typical radio source had to be as high as 10^{60} to 10^{62} erg.

Clearly, even if energy conversion to radio waves was assumed to be 100 per cent efficient, the collision process would be unable to deliver the requisite amount of energy. This theoretical reason was later substantiated by observations which clearly demonstrated that colliding galaxies were not involved on the site of radio emission, even in the classic case of Cygnus A.

Detailed studies of the largescale structure of radio sources reveal a picture shown schematically in figure 4. There we see two blobs of radio emitting regions well separated from each other with an optical source in between. The alignment is very good, suggesting a linear structure. In many cases, radio contours show jet-like features leading to hot spots (that is, radio bright compact regions) within the blobs. What do these jets contain? What astrophysical process is responsible for the highly collimated beams emanating from the centre and moving out to the periphery?

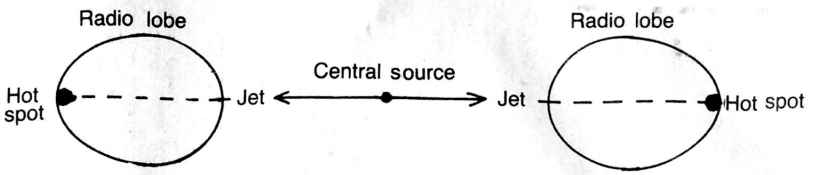


Fig 4: Schematic diagram of a double radio source. Jets from the centre seem to connect with hot spots at the ends

This is where the twin exhaust model of Roger Blandford and Martin Rees comes in. This theory is briefly illustrated in figure 5. There we have a supermassive black hole which is surrounded by a gas cloud, the entire system rotating about a fixed axis. It is assumed that in this system there is an energy source in the nuclear region which ejects plasma. In a spherically symmetric system, this plasma could come out spherically. However, in a rotating system the ejected particles come out asymmetrically, choosing the line of least resistance. This happens to be along the axis of rotation.

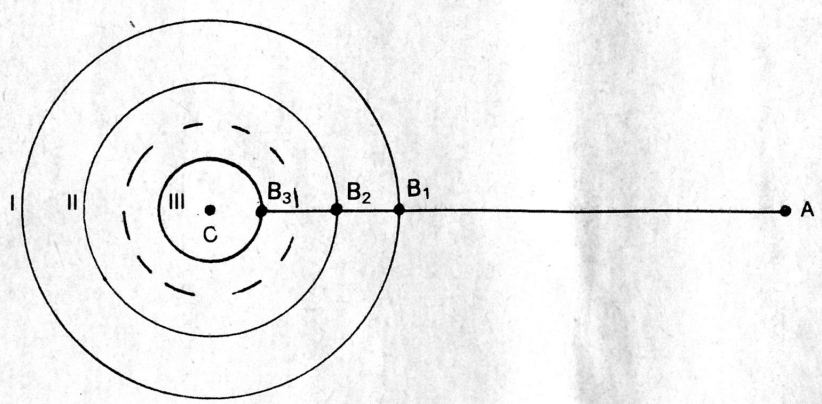
Blandford and Rees showed that the ejection process can be described in gas dynamics by a process which generates the de Laval nozzle where the gas comes out in a highly

White holes

WHAT is a white hole? How does it form? For this, first consider the more familiar example of how a black hole is formed through gravitational collapse. In the figure here, we see three stages, I, II, III in chronological order, during the collapse of a massive spherical body. Suppose there is an observer B_1, on the surface of the collapsing body; B_1, B_2, B_3 are his three positions at stages I, II, and III, respectively. A is another observer located on the radial direction B_3, B_2, B_1 away from the centre C of the body. We will assume that A is located sufficiently far away from the body to be effectively free from its gravitational influence.

Suppose that B emits light waves of a fixed wavelength towards A. Two effects join together to increase the wavelength of this light wave as measured by A; one is due to the gravitational attraction exercised by the body on light, while the other is the well-known Doppler effect arising as B moves away from A. If the surface of the spherical body falls freely towards C, both these effects increase so that the wavelength and the faintness of light received by A (from B) increase rapidly. General relativity tells us that the wavelength shoots to infinity and the light intensity drops to zero as the surface of the body crosses a critical sphere, shown in the figure by the broken circle. Beyond this stage, no communication is possible from B to A. Thus the light emitted at stage III will never make it to A. The critical sphere is known as the Schwarzschild sphere (after the physicist K. Schwarzschild, who first gave a rigorous solution of Einstein's equations outside a massive sphere).

In fact, the body becomes a black hole when its outer surface crosses the Schwarzschild sphere. Although invisible to A thereafter, the body has a finite future ahead of it until it shrinks to the point C. This final state of infinite density also implies that the geometry of space-time has broken down at C. Mathematicians call it the space-time singularity.



The time-symmetry inherent in relativity leads us to the conclusion, however, that the entire sequence of events described above can be reversed, just as a motion picture film is reversed to depict phenomena backwards in time. This reversal leads us to the white hole. Like the black hole, the white hole also is a solution of Einstein's equations.

But there is one remarkable difference noticed by the observer A in this reversed situation. If B is still sending signals to him, their wavelength may not necessarily increase by the time they reach A. The gravitational effect still operates to increase the wavelength, but the Doppler effect *works the other way*. For, the part of the surface of the body where B is located now approaches A instead of receding from him.

White hole calculations show that, for a small period in the beginning when the object erupts from C, the Doppler effect dominates so that light reaching A arrives with *reduced* wavelength and *increased* intensity. Thus in the figure, this effect is more marked in stage III (which is now *earlier*) than in stage II and is still less in stage I. Notice also that the Schwarzschild sphere no longer acts as a barrier to signals from B to A. The object is therefore visible to A right from its initial explosive beginning. What is more, the intensity of the object is more marked in the initial stages. Hence the name white hole in contrast to its invisible counterpart, the black hole.

J. V. N.

The C-field

THE steady state theory proposed by Fred Hoyle, Hermann Bondi and Tommy Gold in 1948 required matter to be created continuously as the universe expands in order that the overall density of matter in the universe remains "steady". Where does this matter come from? Does continuous creation violate the law of conservation of matter and energy?

Not necessarily! At first it looks tempting to suggest that matter is created at the expense of some energy reservoir. But then, that reservoir will also be depleted sooner or later. The C-field proposed by Hoyle has negative energy and need never be depleted through expansion and matter creation. While expansion weakens it, matter creation strengthens it and the two processes are held in balance in the steady state universe.

According to Einstein's general relativity, positive energy attracts while negative energy repels. Hence a reservoir of C-field acts as a source of repulsion, which grows stronger as it is compressed. This is precisely what happens in a massive oscillator.

J. V. N.

collimated jet, or rather to a couple of oppositely oriented collimated jets. Two such jets are necessary in a stable system in order to conserve momentum.

The supermassive black hole accretes matter from the surroundings in a disc perpendicular to the axis of rotation. It is estimated that not only radio emission but also the optical and X-ray emission in quasars may be traced back to such an accretion disc. The black hole needed to account for the observed emission has to have a mass in excess of 10^8 $M_\odot$. For such a black hole, the total energy is about 2×10^{62} erg, or more. If it is assumed that an efficient energy conversion mechanism exists which converts, say, 10 per cent of this energy to the observed radiation, we can explain the phenomena of quasars.

There are, however, several issues still to be resolved. A big problem, for example, is to understand why many radio sources, especially quasars, show only one jet.

Massive oscillators

In a classic paper in early 1963 which anticipated the role of supermassive objects and their gravitational energy reservoirs, Hoyle and William Fowler had suggested that gravitational collapse could lead to the formation of such objects. In the case of the Sun, we saw that the pressures generated by thermonuclear reactions can successfully withstand the gravitational contraction. But, generally, such pressures are ineffective if the object is much more massive than our Sun. For example, an object of 10^6 $M_\odot$ can barely survive as a stable star. A more massive object will inevitably collapse unless it breaks apart under some internal instabilities. What will happen to such objects if they manage to retain a coherent shape during collapse?

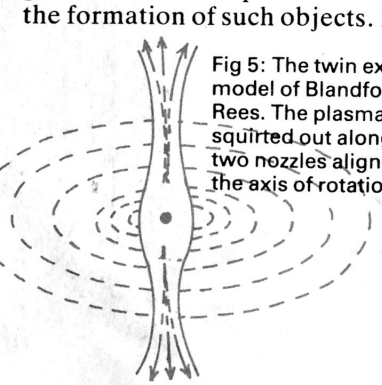


Fig 5: The twin exhaust model of Blandford and Rees. The plasma is squirted out along the two nozzles aligned with the axis of rotation

General relativity tells us that the end-product of gravitational collapse is a space-time singularity. Relativists believe that the collapse to a singularity is invariably preceded by the formation of a black hole.

The black hole, however, suffers from one defect: it cannot be observed directly. Its existence must be deduced indirectly. And indirect or circumstantial evidence is always subject to the criticism: Is this the only way to explain what is observed?

Now, if we reverse the collapse process, we get explosion. The Big Bang universe is an example of such an explosion, although it is somewhat unusual in the sense that the universe is not a subset of any larger system. Can we have finite or compact explosions which are the time-reversed versions of gravitational collapse? Since relativity is a time-symmetric theory, it permits such explosions and these are called *white holes* (see box on p. 22).

In fact, it was suggested, back in the 1960s, that white holes could exist as delayed explosions after the universal big bang. In 1975, Krishna Apparao and I had studied the astrophysical behaviour of white holes and shown that radiation or particles thrown out by a white hole would be 'blueshifted', that is, upgraded in energy. This effect lasts for a short (and early) period of a white hole's life. Nevertheless, it could be directly observed and could account for gamma-ray bursts or X-ray transients. White holes could exist in the stellar mass range or on the supermassive scale of exploding galactic nuclei such as those found in Seyfert galaxies.

On the other hand, there have been calculations (by D. M. Eardley) to show that the white hole as a delayed core of a big bang universe is highly unstable and too short-lived to be of any astrophysical interest! Besides, astrophysicists are unhappy about the existence of white holes because they have an origin in a space-time singularity, though this oft-quoted objection is conveniently forgotten in the case of the big bang universe which also originates in a singularity!

Both these objections to white holes can be eliminated—if the white hole does not begin in a singularity. Suppose a collapsing object encounters a hard repulsive force which makes it bounce when it is very dense. If the bounce radius is very small but finite, the object, in its expanding phase, would still behave like a white hole and fulfil its role of energising particles and radiation.

Such objects can be called *massive oscillators*. These objects oscillate because, when their radii are large, they begin to contract under their own gravity, and when the radii are small, they bounce. The kinetic energy stored in these oscillators is comparable to their rest mass energy. Hence these objects could be good candidates for cosmic energy machines.

What is the repulsive force which causes the bounce? One could use the scalar field of zero rest mass, called the C-field (see box). The existence of massive oscillators does not depend on any particular form of the repulsive force, but it does depend on the assumption that such a dominant short-range force exists in nature.

Prof. Narlikar is with the Astrophysics Group at the Tata Institute of Fundamental Research, Bombay.