

Articles

Building Black Holes: Supercomputer Cinema

STUART L. SHAPIRO AND SAUL A. TEUKOLSKY

A new computer code can solve Einstein's equations of general relativity for the dynamical evolution of a relativistic star cluster. The cluster may contain a large number of stars that move in a strong gravitational field at speeds approaching the speed of light. Unstable star clusters undergo catastrophic collapse to black holes. The collapse of an unstable cluster to a supermassive black hole at the center of a galaxy may explain the origin of quasars and active galactic nuclei. By means of a supercomputer simulation and color graphics, the whole process can be viewed in real time on a movie screen.

In a galaxy far, far away, long, long ago, 100 million neutron stars swirl around the center at velocities near the speed of light. Suddenly the delicate balance between their orbital motion and the inward pull of their mutual gravitational attraction is upset. Stars begin to rush toward the center. An avalanche ensues. Out of this catastrophic collapse, a supermassive black hole arises. Gaseous debris accretes onto the black hole, radiating profusely before being swallowed. A quasar is born.

QUASARS ARE THE MOST ENERGETIC OBJECTS IN THE universe, with a power output up to 10^{48} erg s^{-1} , which is 10^{15} times the luminosity of the sun. Most astrophysicists believe that quasars and active galactic nuclei are powered by supermassive black holes, with masses in the range of 10^6 to 10^8 solar masses (1).

How are these supermassive black holes formed? The scenario depicted above is one speculation. But a speculation is not a calculation. Only a detailed quantitative investigation can demonstrate that such an idea is theoretically viable.

How might we observe such a catastrophic event? Direct astronomical observations are out of the question. The collapse is such a short-lived phenomenon that catching a black hole in the act of formation is unlikely.

Supercomputers can provide answers to both of these questions. By performing a supercomputer simulation we can demonstrate that the above scenario is indeed a viable mechanism for the formation of supermassive black holes. Moreover, with the aid of the supercomputer, we can "observe" the collapse in real time on a movie screen.

At the Cornell National Supercomputer Facility, we have recently solved Einstein's equations of general relativity for the dynamical evolution of a relativistic star cluster (2, 3). Besides the obvious application to astrophysics, the study of relativistic star clusters has a wider computational significance. The problem falls into two broad categories of research now being hotly pursued by computational physicists working in many different areas: nonlinear dynamics and field theory on a space-time lattice. Interest in these subjects is growing rapidly because of new computer hardware and new numerical algorithms. It is significant that techniques developed for solving problems arising in one area are commonly useful for solving problems arising in a different area—when expressed in computa-

The authors are professors in the Departments of Astronomy and Physics and are members of the Center for Radiophysics and Space Research at Cornell University, Ithaca, NY 14853.

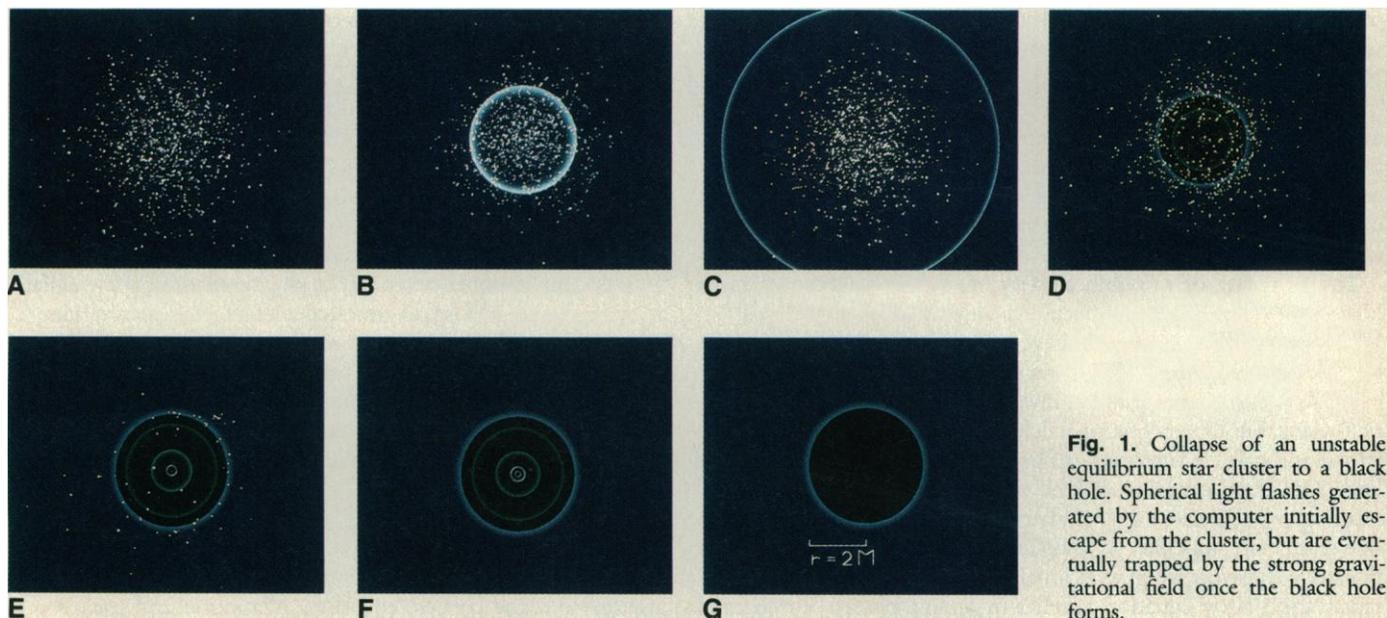


Fig. 1. Collapse of an unstable equilibrium star cluster to a black hole. Spherical light flashes generated by the computer initially escape from the cluster, but are eventually trapped by the strong gravitational field once the black hole forms.

Fig. 2. Schematic illustration showing how clocks in the strong gravitational field of a black hole at the origin run slower than clocks far away.

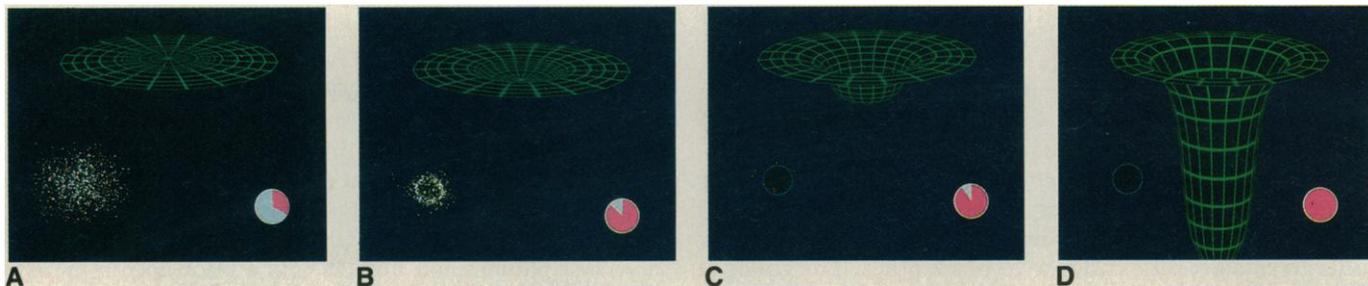
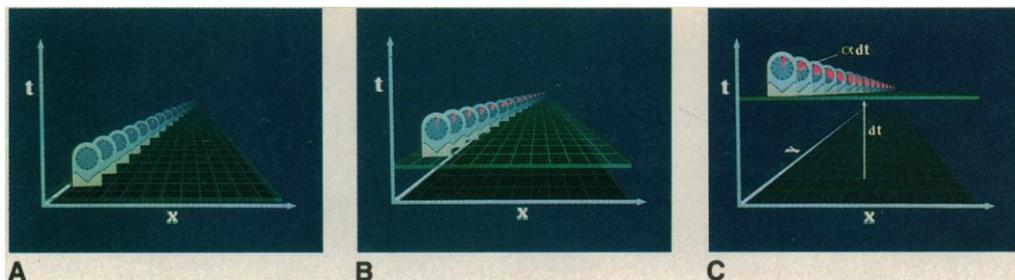


Fig. 3. “Collapse of the lapse.”

tional terms, seemingly different problems in many cases share a common numerical structure. For example, some of the algorithms employed for a relativistic star cluster have also been used for long-term weather forecasting, plasma fusion studies, and the design of airplane wings.

Our work addresses several long-standing issues in stellar dynamics. A major unsolved problem concerns the stability of relativistic star clusters in dynamical equilibrium. The pioneering treatments of the stability of relativistic star clusters (4) were restricted to linear perturbation theory. The more general study of stability by numerical means involves setting up the equilibrium cluster as initial data in a time-dependent simulation. If the cluster is stable, it does not evolve away from its initial configuration. We have used our numerical model in this way to resolve issues of cluster stability.

The stability of a single gaseous star can be diagnosed from its binding energy. Consider a sequence of such stars in equilibrium, parametrized by central density. The binding energy of each star along the sequence is a function of the central density. It is a theorem that all stars are dynamically stable up to the first maximum of the binding energy. Beyond this point all members of the sequence are dynamically unstable. This theorem holds both in Newtonian gravity and in general relativity (5).

No such simple binding energy criterion has been established for an equilibrium sequence of collisionless star clusters [Ipser (6) has presented a sufficient, but not necessary, criterion]. Our stability analysis shows that the binding energy does indeed provide a diagnostic for stability—a numerical demonstration awaiting a formal mathematical proof.

The final fate of unstable clusters, and their collapse to black holes, has previously been discussed only in qualitative terms (7). With our computer model, we have been able to track the complete nonlinear evolution of unstable configurations.

We have restricted our initial investigation to spherically symmetric clusters, but the gravitational field of the system can be arbitrarily strong and particle velocities can be arbitrarily close to the speed of light. Our computational scheme combines the tools of numerical relativity (8) with those of N -body particle simulations (9). Stars in a star cluster are modeled as a collisionless gas of particles, which interact exclusively by gravitational attraction. Consequently, they satisfy the Vlasov equation familiar in plasma physics for charged

particles experiencing Coulomb interactions. We solve this equation in full general relativity by a mean-field particle simulation scheme.

Our model can accurately follow the collapse of an unstable cluster to a black hole. A black hole is a region where gravity is so strong that nothing, not even light, can escape. The formation of a black hole is inevitably accompanied by the appearance of space-time singularities—regions inside the black hole where the gravitational tidal force and matter density become infinite. Because these infinities are always expected to occur inside black holes, they do not influence the evolution of the causally disconnected regions outside the hole [“cosmic censorship” (10)]. However, because of these infinities, tracking the formation of a black hole numerically is technically difficult. Yet only by solving such problems can we study the full nonlinear nature of relativistic gravity.

Physical Picture

To understand how one simulates the evolution of a star cluster in general relativity, it is useful to consider first how the problem can be tackled in Newtonian physics. Stars in the cluster move according to Newton’s Laws of Motion

$$\frac{d\mathbf{x}}{dt} = \mathbf{u} \quad (1)$$

$$\frac{d\mathbf{u}}{dt} = -\nabla\Phi \quad (2)$$

where \mathbf{x} and \mathbf{u} are the position and velocity vectors of each star and Φ is the gravitational potential. In the simulation, these equations are integrated for a small timestep for a representative sample of cluster stars, typically 10,000 or so. From the new star positions, we compute the mass density, ρ , at the new time by dividing the volume of the cluster into small bins and counting the number of stars in each bin. The density then serves as the source term for the gravitational field equation, which in Newtonian physics is Poisson’s equation for Φ ,

$$\nabla^2\Phi = 4\pi G\rho \quad (3)$$

where G is the gravitational constant. The new potential is then inserted into the particle equations of motion and the process is

Fig. 4. Collapse of the inner region of a star cluster as seen by two sets of observers.



repeated for the next small timestep. This approach to evolving a self-gravitating collisionless system is known as a mean-field particle simulation scheme.

Solving the relativistic problem is similar to the Newtonian approach just described. The Newtonian particle equations of motion are replaced by the geodesic equations of general relativity, whereas Poisson's equation is replaced by Einstein's equations for the gravitational field. The restriction to spherical symmetry leads to a considerable reduction in the number of dynamical phase-space degrees of freedom that have to be tracked in the simulation. Instead of the full six degrees of freedom, three components of position and three components of velocity for each particle, an orbit is uniquely specified by the radial coordinate and the radial and transverse components of the velocity. Conservation of angular momentum further reduces this to two dynamical degrees of freedom. In addition, the relativistic gravitational field variables at each instant are functions only of radius.

Solving Einstein's equations for the gravitational field is not trivial. The usual form of these equations is

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \quad (4)$$

where $G_{\mu\nu}$ is the Einstein tensor and $T_{\mu\nu}$ is the stress-energy tensor. This form intertwines space and time, and hence is not suitable for performing a computer calculation of time evolution. We need to recast the equations in the form of an initial-value problem: given the state of the system over all space at one instant of time, t , we require equations that determine the state of the system over all space at the next instant, $t + dt$. The new state can then be used as initial data to continue the integration to the next instant, and so on.

The required splitting of space and time in general relativity is given by the ADM (Arnowitt-Deser-Misner) or 3 + 1 decomposition of Einstein's equations (11). In this decomposition, the original form of Einstein's equations (Eq. 4) is split into two kinds of equations: constraint equations and evolution equations. The constraint equations contain no time derivatives and relate field variables at a given instant of time. The evolution equations are first-order differential equations in time that propagate the initial data to the next instant of time.

This kind of decomposition is not unique to general relativity. Indeed, Maxwell's equations of electromagnetism are already in this form: the constraint equations of electromagnetism are

$$\nabla \cdot \mathbf{E} = 4\pi\rho_e \quad (5)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (6)$$

while the evolution equations are

$$\frac{\partial \mathbf{E}}{\partial t} = c\nabla \times \mathbf{B} - 4\pi\mathbf{J} \quad (7)$$

$$\frac{\partial \mathbf{B}}{\partial t} = -c\nabla \times \mathbf{E} \quad (8)$$

Here \mathbf{E} is the electric field, \mathbf{B} is the magnetic field, ρ_e is the electric

charge density, \mathbf{J} is the current density, and c is the speed of light. In a numerical simulation of Maxwell's equations, one would start with \mathbf{E} and \mathbf{B} fields that satisfy the constraints, Eqs. 5 and 6. The subsequent evolution of these fields is governed by Eqs. 7 and 8. The whole procedure is self-consistent: the constraint equations are always guaranteed to hold if they are satisfied initially and if the fields are evolved according to Eqs. 7 and 8.

Note that the Newtonian limit of the ADM decomposition yields one constraint equation, which is just Poisson's equation, Eq. 3. There are no evolution equations—Newtonian gravitation is not a truly dynamical field.

In full general relativity, the ADM equations are similar to Maxwell's equations, but more complicated. There are more fields and equations because general relativity describes a tensor field, whereas Maxwell's equations describe vector fields. More important, the field equations in general relativity are nonlinear. This makes the equations more difficult to solve, and yields solutions exhibiting more complicated behavior.

The Movie

This investigation provides an example of a computational problem that requires extensive use of graphical display to visualize the dynamical behavior uncovered by the simulation. Imagine trying to make sense of a listing of the positions and velocities of 10,000 stars at 1,000 successive instants of time! Instead, we have taken the output from selected cases and had the computer produce a color movie that depicts the collapse of unstable clusters to black holes.

The equations were solved on the Cornell National Supercomputer Facility (an assembly of Floating Point Systems FPS 264 array processors hosted by an IBM 3090-600). The computer tracked the motion of a representative sample of stars in the cluster. The positions of these stars at successive times were then displayed on color film. The completed film lasts 8 minutes and was made on a Cray XMP at Digital Productions in Los Angeles.

The movie consists of six scenes. Highlights from scene 1 are shown in Fig. 1, A to G. Initially, the stars orbit about their mutual center at speeds close to the speed of light. The stars obey a Maxwell-Boltzmann energy distribution. The cluster is initially in equilibrium—the motion of the stars exactly counterbalances the inward pull of their combined gravity. If Newton's theory of gravity were correct, the cluster would remain in this equilibrium state forever. However, according to general relativity, gravity is actually stronger than predicted by Newton's theory, and the cluster is unstable to catastrophic collapse. As time advances, the stellar orbits spiral inward toward the center. At the very center the concentration of mass becomes so great that a black hole forms. During the simulation, the computer sends out spherical flashes of light from the center of the cluster. The propagation of these flashes indicates when the black hole forms. These flashes are depicted by the bright outward-moving shells in Fig. 1, B to G. Before the black hole forms,

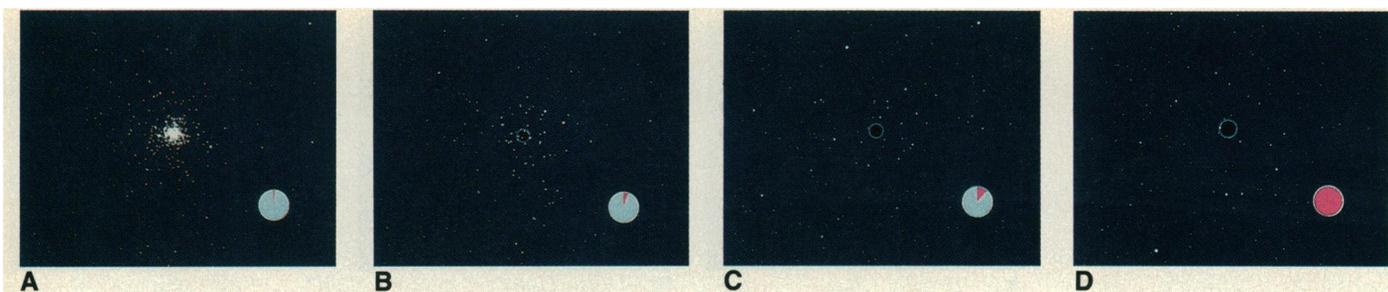


Fig. 5. Collapse of the inner region of an extended star cluster to a black hole. This may be the mechanism by which a supermassive black hole forms in a galactic nucleus to power a quasar.

the light rays have no problem traveling outward forever, escaping from the cluster entirely. However, once the black hole forms, the light rays are permanently trapped within (Fig. 1, D to F). In this example, the black hole radius (the horizon) continues to grow until all of the stars are consumed. The outer boundary of the black hole has been silhouetted with a light blue halo for easier visualization. The numerical value of the final radius, R , of the black hole is accurately given by

$$R = \frac{2GM}{c^2} \quad (9)$$

where M is the total mass of the cluster. (G and c have been set equal to unity in Fig. 1G.) This agrees with the theoretical value, and is called the Schwarzschild radius.

Scene 2 is a cartoon illustrating that the strong gravitational field that develops near a black hole slows down time (Fig. 2). The black hole is at the origin of the coordinates, $x = y = 0$. The z -axis is suppressed for clarity. Coordinate time, t , is plotted vertically (space-time diagram). The clocks measure proper time, τ , which agrees with coordinate time far away where gravity is weak. Clocks that are far from the black hole advance more rapidly than clocks close to it. The lapse function, α , gives the ratio of time measured by a local clock, $d\tau$, to time measured very far away, dt (the gravitational red shift). The pink region of the clock measures the proper time elapsed, $d\tau = \alpha dt$.

Scene 3 depicts the "collapse of the lapse" (12) (Fig. 3). Plotted in green is $\log \alpha$ as a function of radius from the origin for the collapse shown in scene 1. The clock in the lower right-hand corner measures

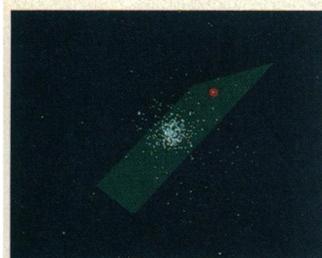
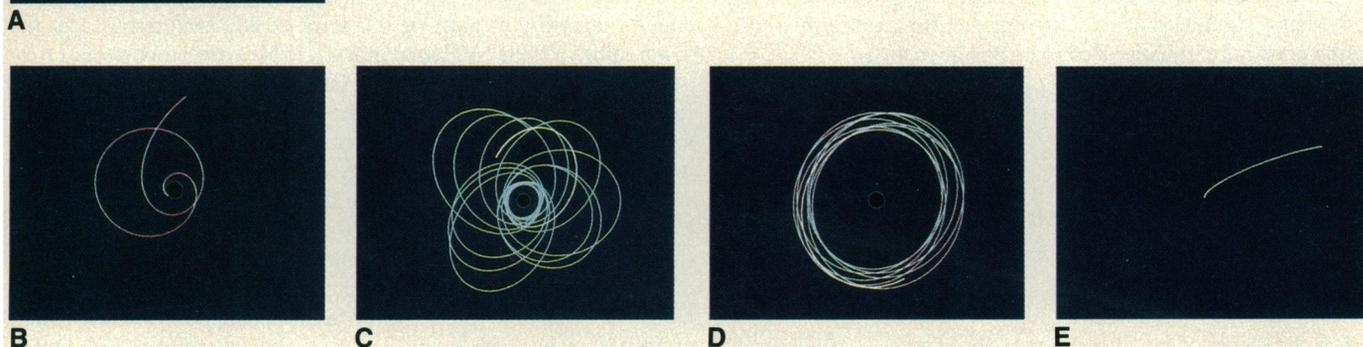


Fig. 6. Selected orbits for the collapse depicted in Fig. 5.



the advance of coordinate time during the collapse. The cluster configuration at the corresponding time is shown in the lower left-hand corner. At $t = 0$ the gravitational field of the cluster is not very strong, $\alpha \approx 1$, and the plot (Fig. 3A) is a flat surface. The function α plunges to zero ($\log \alpha \rightarrow -\infty$) in the center of the cluster when the black hole forms (Fig. 3D).

One of the goals of numerical relativity is to study black hole formation on the computer without encountering the space-time singularity that appears in a finite proper time at the center of the black hole. Reaching the singularity causes overflows and underflows in the simulation because of the infinities that accompany the singularity. In this case the simulation will crash before the evolution is complete, that is, before the fate of regions far outside the black hole can be ascertained.

This problem is avoided in general relativity by exploiting a gauge freedom. This freedom is analogous to the gauge freedom that is well known in electromagnetism. In relativity, we are free to label events in space-time with quite arbitrary sets of smooth coordinates. Only proper times and distances have physical significance, not the values of the space-time coordinates. Consider the march of the simulation from one value of coordinate time to the next. Because the lapse measures the advance of proper time relative to coordinate time, it is obviously desirable to choose a time coordinate in which the lapse falls to zero whenever the gravitational field starts to get strong. In this way, it is possible to postpone the formation of the singularity to large values of coordinate time.

There are an infinite number of ways to exploit this gauge freedom in practice. Geometric considerations have led theorists to propose two particular choices for a suitable time coordinate, maximal time-slicing and polar time-slicing. Both have the property that $\alpha \rightarrow 0$ in a strong field region, but which is better in a simulation?

Scene 4 is a comparison of the Maxwell-Boltzmann collapse of scene 1 as seen by two sets of observers who use the two different choices of time coordinate (Fig. 4). The frames show the stellar positions at the center of the cluster at the same coordinate times. Clocks carried by the maximal observers, which measure proper time, will advance farther than those carried by the polar observers.

Hence the collapse proceeds farther for the maximal observers, and the stars get closer to the singularity at the center of the black hole. In Fig. 4B the black hole has already formed for maximal time slicing. The surface of the cluster falls well inside the Schwarzschild radius by late times (Fig. 4C). By contrast, for polar observers, the black hole forms later and even at late times the surface barely reaches the Schwarzschild radius. Hence, we conclude that polar slicing avoids the singularity to a greater extent than maximal slicing. This is in agreement with theoretical predictions (13).

As described earlier, most astrophysicists believe that quasars (and active galactic nuclei) are powered by supermassive black holes, with masses exceeding 10^6 to 10^8 solar masses. The existence of supermassive black holes may not be restricted to such exotic objects. Our nearest neighbor, the Andromeda galaxy, probably has a black hole of this size at its center (14), as might other nearby galaxies. Even our own galaxy may contain a 10^6 solar mass black hole at the center (15). Where could such supermassive black holes come from?

Scene 5 demonstrates how such a black hole can form in a galaxy whose core consists largely of compact stars (Fig. 5). Compact stars—neutron stars and black holes of a few solar masses—are end-point products of the evolution of normal stars (5). They proliferate even ordinary galaxies like our own. It may take up to 10^{10} years for a galaxy to reach the stage at which its core is sufficiently relativistic for catastrophic collapse. At the onset of collapse a galaxy will consist of a tiny relativistic core embedded in a larger Newtonian “halo” of stars (3, 16). The ensuing implosion takes only a few minutes and is depicted in Fig. 5, A to D.

In these frames, we zoom in to view the central relativistic core of the cluster. The initial core (Fig. 5A) contains less than 0.5% of the total mass of the cluster. When the collapse is complete, the black hole contains a full 5% of the cluster mass (Fig. 5D). This is more than ten times the mass of the original core. The black hole thus swallows stars from a wider region than just the relativistic core. The bulk of the stars are not captured, however, but continue to orbit about the black hole in a new, stable, equilibrium configuration. This is in contrast to the case shown in scene 1, where the relativistic core in the initial cluster contained a large fraction of the total mass. The final equilibrium state shown here—a supermassive black hole embedded in an extended Newtonian star cluster—is just what may be required to explain the observations of quasars and the central regions of galaxies.

This particular case (scene 5) was difficult computationally because of its large dynamic range: the ratio of the central density to the mean density was 10^{13} . The calculation was carried out with 7000 particles, 350 radial grid points, and 4000 timesteps, and required 10 hours of CPU time on an FPS 264. The graphics required an equivalent amount of CPU time. Special choices of coordinates were needed to make the computation feasible at all.

Scene 6 illustrates four classes of orbits of stars near the black hole of scene 5. In spherical symmetry, each orbit is confined to a plane as in Newtonian theory. A star on a typical capture orbit (highlighted in red in Fig. 6A) will have its orbital plane (pale blue) pass through the center of the cluster. The star first swings around the center on a nearly elliptical trajectory (Fig. 6B). By the second pass of the star, the hole has grown sufficiently that it swallows the star. As a result, the mass and size of the black hole (see Eq. 9) both increase, which enables the hole to swallow stars that escaped capture on their initial passes. In this way the hole can consume stars that originate from outside the relativistic core of the cluster, which collapses almost immediately.

Figure 6C shows a star that orbits as close as it can to the black hole (about two Schwarzschild radii) without actually being captured. Notice the very large perihelion precession (rotation of the ellipse on each orbit). The tiny perihelion precession of the planet

Mercury in the relatively weak gravitational field of the Sun—one of the famous experimental tests of general relativity—pales in comparison.

As in Newtonian theory, a particle in a spherically symmetric gravitational field feels only the total amount of mass inside its radius, and its motion is not affected by the internal spherical distribution of the mass. In general relativity, this result is known as Birkhoff's theorem. Figure 6D shows a star on an almost circular orbit, which is unperturbed by the formation of the black hole—a clear demonstration of Birkhoff's theorem.

Figure 6E shows a star that starts far from the center, so that the black hole is a small dot on this scale. The unfortunate star happens to be on a very radial orbit and hence plunges directly into the black hole. The reach of the black hole out into the halo is greatest for such low angular momentum stars. However, most of the mass of the black hole comes from stars like that of frame A of scene 6, which originate within a few core radii.

Future Work

Most objects in astronomy are not spherically symmetrical and exhibit rotation. Handling these complications requires a considerable increase in computer resources, even for an axisymmetric configuration. For example, the number of phase-space dynamical degrees of freedom increases from two to four. Moreover, the gravitational field variables at each instant of time now depend on two spatial coordinates. Most interestingly, when spherical symmetry is broken in general relativity, the gravitational field becomes truly dynamical—gravitational waves are emitted. A fraction of the cluster's total mass-energy can thus be radiated away in the form of gravitational waves during the collapse. Tackling these more complex issues is the next challenge.

REFERENCES AND NOTES

1. See, for example, M. C. Begelman, R. D. Blandford, M. J. Rees, *Rev. Mod. Phys.* **56**, 255 (1984); M. J. Rees, *Annu. Rev. Astr. Astrophys.* **22**, 471 (1984).
2. S. L. Shapiro and S. A. Teukolsky, *Astrophys. J.* **298**, 34 (1985); *ibid.*, p. 58; *ibid.* **307**, 575 (1986).
3. ———, *ibid.* **292**, L41 (1985).
4. J. R. Ipser and K. S. Thorne, *ibid.* **154**, 251 (1968); J. R. Ipser, *ibid.* **158**, 17 (1969); E. D. Fackerell, *ibid.* **160**, 859 (1970).
5. S. L. Shapiro and S. A. Teukolsky, *Black Holes, White Dwarfs, and Neutron Stars: The Physics of Compact Objects* (Wiley, New York, 1983).
6. J. R. Ipser, *Astrophys. J.* **238**, 1101 (1980).
7. Ya. B. Zel'dovich and M. A. Podurets, *Astron. Zh.* **42**, 963 (1965) [English translation: *Sov. Astron.-A. J.* **9**, 742].
8. See, for example, J. M. Centrella, Ed., *Dynamical Spacetimes and Numerical Relativity* (Cambridge Univ. Press, Cambridge, 1986).
9. See, for example, J. A. Sellwood, *Annu. Rev. Astr. Astrophys.* **25**, 151 (1987).
10. R. Penrose, *Riv. Nuovo Cimento* **1**, 252 (1969). Specifically, Penrose has conjectured that gravitational collapse does not produce “naked singularities,” that is, singularities that are not clothed by black hole event horizons.
11. R. Arnowitt, S. Deser, C. W. Misner, in *Gravitation: An Introduction to Current Research*, L. Witten, Ed. (Wiley, New York, 1962), pp. 227–265.
12. L. Smarr and J. W. York, *Phys. Rev.* **D17**, 1945 (1978); *ibid.*, p. 2529.
13. D. M. Eardley, paper presented at Numerical Astrophysics: A Symposium in Honor of James R. Wilson, University of Illinois, October 1982; J. M. Bardeen and T. Piran, *Phys. Rep.* **96**, 205 (1983).
14. J. Kormendy, *Astrophys. J.* **325**, 128 (1988); A. Dressler and D. O. Richstone, *ibid.* **324**, 701 (1988).
15. J. H. Lacy, F. Baas, C. H. Townes, T. R. Geballe, *Astrophys. J.* **227**, L17 (1979).
16. G. D. Quinlan and S. L. Shapiro, *ibid.* **321**, 191 (1987).
17. This work has been supported in part by National Science Foundation grants AST 87-14475 and PHY 86-03284 at Cornell University. Computations were performed on the Cornell National Supercomputer Facility, a resource of the Center for Theory and Simulation in Science and Engineering (Cornell Theory Center). The Center receives major funding from the National Science Foundation and IBM Corporation, with additional support from New York State and members of the Cornell Research Institute. The original movie was produced in the Department of Astronomy at Cornell. We are indebted to C. Kochanek, A. Liu, and N. Silva for assistance with the graphics. The final version of the movie was produced at Digital Productions, Los Angeles, with support from the National Science Foundation. The work at Digital Productions was directed by S. Fangmeier and C. Upton, to whom we are grateful.