

# Binary black hole mergers

Thomas W. Baumgarte and Stuart L. Shapiro

Solving the equations of general relativity presents unique challenges. Nowadays many of those have been met, and new numerical simulations are revealing surprising astrophysical phenomena.

**Thomas Baumgarte** is a professor of physics at Bowdoin College in Brunswick, Maine. **Stuart Shapiro** is a professor of physics and astronomy at the University of Illinois at Urbana-Champaign.

**After decades of effort**, numerical relativists can now simulate the inspiral and merger of two black holes orbiting each other. That computational triumph has come none too soon—physicists are on the verge of detecting gravitational waves for the first time, and at long last they know what to look for.

Black holes are strong-field objects whose properties are governed by Einstein's theory of gravitation—general relativity. A black hole is a region of spacetime that cannot communicate with the external universe. The boundary of that region defines the surface of the black hole, called the event horizon. Isolated black holes are remarkably simple. They are described by analytic solutions to Einstein's equations and are uniquely parameterized by just three quantities: their mass  $M$ , spin  $J$ , and charge  $Q$ . Since charged objects in space are rapidly neutralized by the surrounding plasma, one usually assumes  $Q = 0$  for real astrophysical black holes.

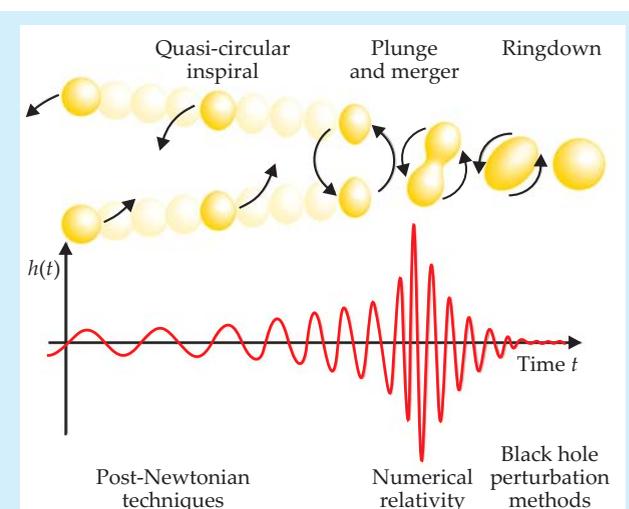
Stellar-mass black holes, which have masses from several to several tens of solar masses ( $M_{\odot}$ ), can form when massive stars exhaust their nuclear fuel and undergo collapse. They were first identified in binary x-ray sources in our galaxy, accreting gas from normal stellar companions. Spinning stellar-mass black holes accreting from disks of magnetized plasma may also trigger gamma-ray bursts (GRBs). During the early history of the universe, highly massive and supermassive black holes likely formed from smaller seed black holes and grew by a combination of mergers and gas accretion. The cores of nearly all nearby bulge galaxies, including our own Milky Way, harbor a supermassive black hole with a mass between  $10^6$  and  $10^9 M_{\odot}$ . Supermassive black holes are believed to be the central engines powering quasars and active galactic nuclei (AGNs), the most energetic sources of electromagnetic radiation currently known. Black holes, it seems, are making their presence felt all over the universe.

## Spacetime ripples

Binary black holes are among the most promising sources of gravitational radiation. General relativity describes gravitational waves as ripples on the background curvature of spacetime that propagate at the speed of light. In some ways they are like water waves traveling on an otherwise smooth sea. Unlike water waves, however, gravitational waves are not motions of material particles but ripples in the fabric of spacetime itself. According to general relativity, the orbit of a binary system decays in three phases, as shown in figure 1, due to the loss of energy and angular momentum carried away by gravitational waves. Radio observations of the Hulse–Taylor binary pulsar confirmed that such losses occur at the rates predicted by general relativity—a discovery for

which Russell Hulse and Joseph Taylor Jr were awarded the Nobel Prize in Physics in 1993. But gravitational waves have yet to be detected directly.

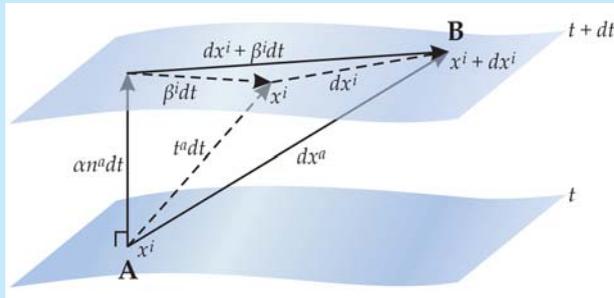
That should change with the Laser Interferometer Gravitational-Wave Observatory (LIGO) in the US, VIRGO in Italy, and similar ground-based detectors elsewhere, which can observe waves with frequencies of 10–1000 Hz. The target date for the Advanced LIGO–VIRGO network to become operational is 2015, at the completion of the latest upgrades (see PHYSICS TODAY, December 2010, page 31). Prime candidates for generating detectable radiation are binary black holes whose constituents each have masses of 10–50  $M_{\odot}$ . Since gravitational-radiation emission causes orbital eccentricity to decay, those binaries will be in tight, circular orbits when the dominant gravitational wave frequencies—twice the binaries' orbital frequencies—pass through the LIGO–VIRGO window. Thus the detectors will be able to measure gravitational waves generated in the last minutes of the binary inspiral; they will also observe radiation emitted during the merger and during the ringdown phase, in which the merged



**Figure 1. Coalescence of a compact binary.** The loss of energy and angular momentum via the emission of gravitational radiation drives compact-binary coalescence, which proceeds in three different phases. The strongest gravitational-wave signal, illustrated here as the gravitational-wave amplitude  $h$ , accompanies the late inspiral phase and the plunge and merger phase; for that part of the coalescence, post-Newtonian and perturbation methods break down, and numerical simulations must be employed. (Adapted from ref. 3.)

## Box 1. The 3 + 1 decomposition

In the flat spacetime of special relativity, the proper interval  $ds$  between two nearby points  $x^a = (t, x^1, x^2, x^3) = (t, x^i)$  and  $x^a + dx^a = (t + dt, x^i + dx^i)$  is given by the “spacetime Pythagorean theorem” as  $ds^2 = -dt^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 = \sum_{a,b} \eta_{ab} dx^a dx^b$ , where the only nonvanishing components of the Minkowski metric tensor  $\eta_{ab}$  are the diagonal ones:  $\eta_{ab} = \text{diag}(-1, 1, 1, 1)$ . The interval measures proper time when it is less than zero (that is,  $dt^2 = -ds^2$ ) and proper distance when it is greater than zero.



In general relativity, where spacetime is curved, the metric tensor takes a more general form  $g_{ab}$ , and the proper interval becomes  $ds^2 = \sum_{a,b} g_{ab} dx^a dx^b$ .

In a 3 + 1 decomposition, the spacetime is foliated into spatial slices of constant coordinate time  $t$ , as shown in the figure. The timelike unit vector  $n^a$  is normal to each slice, whereas the vector  $t^a$  connects points with the same spatial coordinates  $x^i$  on neighboring slices separated by  $dt$ . The lapse function  $\alpha$  determines the advance of proper time along  $n^a$  between neighboring slices, whereas the shift vector  $\beta^i$  determines the shift of spatial coordinates with respect to where they would be if they were determined simply by following  $n^a$ . The spacetime displacement vector  $dx^a$  connects the point **A**, with spatial coordinates  $x^i$  at time  $t$ , to the point **B**, with spatial coordinates  $x^i + dx^i$  at time  $t + dt$ .

The proper interval can again be computed from the spacetime Pythagorean theorem, which yields the 3 + 1 decomposition,

$$ds^2 = -\alpha^2 dt^2 + \sum_{i,j} \gamma_{ij} (dx^i + \beta^i dt)(dx^j + \beta^j dt).$$

Here  $\gamma_{ij}$  is the spatial metric that measures distances within each spatial slice.

black hole relaxes to its final, stationary equilibrium state. Physicists do not have a precise handle on the rate at which LIGO–VIRGO will detect black hole coalescences; estimates range from 0.4 to 1000 events per year.

The Laser Interferometer Space Antenna (LISA) is being considered as a space-based detector that will observe gravitational waves at frequencies of  $10^{-4}$  through  $10^{-1}$  Hz, lower than those seen by the ground-based interferometers. Such waves are emitted by coalescing binaries with at least one intermediate-mass or supermassive black hole. Models predict that LISA will detect 3–300 events per year arising from binaries with total masses in the range of  $10^3$ – $10^6 M_\odot$  and with redshifts as great as 15, in which case the gravitational radiation was emitted more than 13 billion years ago. Even now, tantalizing optical observations suggest that such binaries do indeed form in AGNs and quasars.

An understanding of how binary black holes coalesce is crucial for the construction of theoretical gravitational wave-

form templates to be used in so-called matched-filtering data analysis. Those templates will increase the likelihood of a detection and provide physical interpretations of detected signals. Moreover, a comparison of observed and theoretical gravitational waveforms can serve as a test of general relativity in the strong-field regime: As in high-energy physics, collisions provide a powerful means of probing the nature of an interaction.

But solving the binary-coalescence problem in general relativity has proven to be quite challenging. Analytic post-Newtonian expansions in  $v^2/c^2$ , where  $v$  is the black hole speed and  $c$  is the speed of light, determine the early inspiral phase, and black hole perturbation methods can treat the final ring-down of the merged remnant. But the late inspiral phase—as well as the plunge and merger phase, where the wave amplitude is largest—requires a numerical simulation, and that poses complications. Black holes contain physical spacetime singularities, regions where the gravitational tidal field (curvature) becomes infinite. It is crucial, but hardly easy, to choose a computational technique that avoids encountering those singularities. Moreover, most of the computational resources are usually spent resolving the strong-field region near the holes, yet the waves, which represent small perturbations on the background field, must be extracted in the weak, far-field region. So the numerical scheme must successfully cope with the problem of vast dynamic range. Finally, different formulations of Einstein’s equations behave very differently when implemented numerically, and we numerical relativists had to find suitable formulations that generate stable solutions.

## Numerical relativity

The usual form of Einstein’s equations elegantly unites space and time into a single entity, spacetime. But in order to follow the evolution of a system starting at some initial time, we need to undo that unification and split spacetime back into space and time.<sup>1–3</sup> That is, we need to cast Einstein’s equations into a form appropriate for solving a so-called initial-value or Cauchy problem. Carrying out such a 3 + 1 decomposition of Einstein’s equations results in a set of equations that constrain the gravitational fields at every instant of time, and a set of equations that evolve the fields in time. To obtain a complete solution, we first construct initial data, consistent with the constraint equations, that describe the solution at an initial instant of time. Then we determine the subsequent time development by solving the evolution equations.

The structure of the 3 + 1 decomposition is familiar from Maxwell’s equations, which similarly consist of a set of constraint equations for the electric (**E**) and magnetic (**B**) fields,

$$C_E \equiv \nabla \cdot \mathbf{E} - 4\pi\rho = 0, \quad C_B \equiv \nabla \cdot \mathbf{B} = 0, \quad (1)$$

and a set of evolution equations,

$$\frac{\partial}{\partial t} \mathbf{E} = \nabla \times \mathbf{B} - 4\pi\mathbf{j}, \quad \frac{\partial}{\partial t} \mathbf{B} = -\nabla \times \mathbf{E}, \quad (2)$$

where  $\rho$  and  $\mathbf{j}$  are, respectively, the charge and current densities. Here and throughout we have set  $c = 1$ .

It is always possible to write the magnetic field as a curl of a vector potential **A**, that is  $\mathbf{B} = \nabla \times \mathbf{A}$ , so that the constraint  $C_B = 0$  is satisfied identically and the evolution equations take the form

$$\frac{\partial}{\partial t} \mathbf{E} = -\nabla^2 \mathbf{A} + \nabla(\nabla \cdot \mathbf{A}) - 4\pi\mathbf{j}, \quad (3)$$

$$\frac{\partial}{\partial t} \mathbf{A} = -\mathbf{E} - \nabla\phi.$$

Here  $\phi$  is an arbitrary gauge function. Assuming charge conservation,  $\partial\rho/\partial t = -\nabla \cdot \mathbf{j}$ , equation 3 implies

$$\frac{\partial}{\partial t} C_E = 0. \quad (4)$$

If the constraints are satisfied initially, then they will be satisfied at all times.

A well-known 3 + 1 decomposition of Einstein's equations is the ADM formulation, named after Richard Arnowitt, Stanley Deser, and Charles Misner.<sup>4</sup> The resulting equations are similar to Maxwell's equations 1 and 3 and form a set of coupled, multidimensional partial differential equations. Instead of  $\mathbf{E}$  and  $\mathbf{A}$ , the fundamental variables are a spatial metric tensor  $\gamma_{ij}$  that measures distances within a slice of constant time  $t$ , and a tensor called the extrinsic curvature  $K_{ij}$  that is related to the time derivative of  $\gamma_{ij}$ —just as  $\mathbf{E}$  is related to the time derivative of  $\mathbf{A}$ . Instead of a gauge function  $\phi$ , a lapse function  $\alpha$  and a shift vector  $\beta^i$  appear in the ADM equations. The lapse and shift functions represent the coordinate freedom in general relativity. For a geometric interpretation, see box 1.

In electromagnetism, the leading-order radiation field scales with the time derivative of the source's dipole moment  $\mathbf{d}$ ; that is  $\mathbf{A} \propto \dot{\mathbf{d}}/r$ , where  $r$  is the distance to the source. In general relativity, the leading-order contribution to gravitational wave radiation scales with the second time derivative of the source's quadrupole moment. The amplitude  $h$  of that space-time perturbation measures the fractional change in the separation of two nearby, free test particles. The anticipated sensitivity of Advanced LIGO-VIRGO is a mind-boggling  $h = 10^{-23}$ .

Unlike Maxwell's equations, however, Einstein's equations are nonlinear, and so they introduce a new set of phenomena and challenges. Simulations have employed various strategies to solve the equations. In finite-difference applications, the spacetime continuum is represented as a discrete lattice or grid, and all partial derivatives are approximated as the differences between the values of functions on neighboring grid points. In spectral or pseudospectral techniques, all functions are expanded in terms of a complete set of basis functions—for example, Chebyshev polynomials—and the initial equations are recast as equations for the expansion coefficients.

Those methods alone enabled simulators to solve a number of important problems. One class is initial data problems, or solutions to the constraint equations. Various groups have

## Box 2. A reformulation of Maxwell's equations

Absent the term  $\nabla(\nabla \cdot \mathbf{A})$ , Maxwell's equations (see equation 3 in the text) could be combined into a wave equation for  $\mathbf{A}$ . One could go into the Coulomb gauge and demand  $\nabla \cdot \mathbf{A} = 0$ , but an alternative way to eliminate the offending term—one that also maintains gauge invariance—is to define  $\Gamma \equiv \nabla \cdot \mathbf{A}$  as a new independent variable. Maxwell's equations then take the form

$$\frac{\partial}{\partial t} \mathbf{E} = -\nabla^2 \mathbf{A} + \nabla \Gamma - 4\pi \mathbf{j}, \quad (2a)$$

$$\frac{\partial}{\partial t} \mathbf{A} = -\mathbf{E} - \nabla \phi.$$

The evolution of  $\Gamma$  is given by

$$\frac{\partial}{\partial t} \Gamma = -\nabla \cdot \mathbf{E} - \nabla^2 \phi = -4\pi \rho - \nabla^2 \phi, \quad (2b)$$

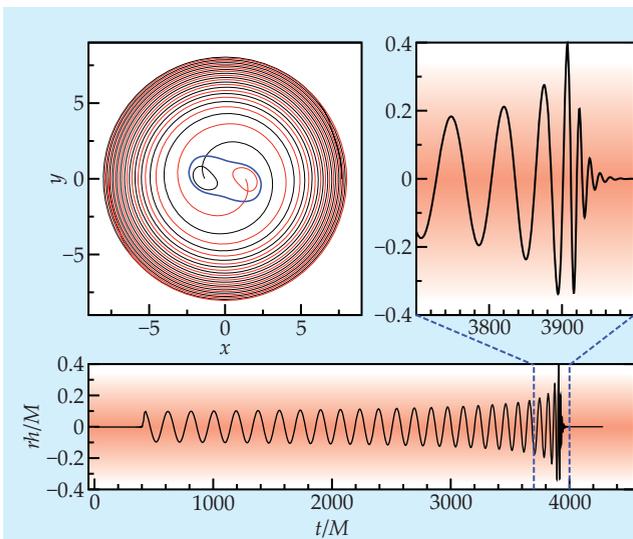
where we have used the constraint equation for  $\mathbf{E}$  in the final equality. We leave it as an exercise to show, using equations 2a and 2b, that any violations in the constraint  $C_E$  now satisfy the wave equation

$$\left( -\frac{\partial^2}{\partial t^2} + \nabla^2 \right) C_E = 0 \quad (2c)$$

instead of being constant in time as in text equation 4.

obtained, for example, solutions that provide snapshots of a black hole or neutron star binary in a quasi-stationary, circular orbit. (For a review, see reference 3.) Another class consists of spacetimes obeying spherical or axial symmetry, restrictions that reduce the number of dynamical degrees of freedom and simplify the equations.

The simulations of an axisymmetric head-on collision of two black holes represented a milestone achievement.<sup>5</sup> They revealed the geometric structure of the merging horizons and showed that only about 0.1% of the total mass of the black holes is radiated away in the collision as gravitational waves—that's much less than the upper limit of 29% allowed by Stephen Hawking's area theorem and much less than was later found to be emitted from black holes that merge following a quasi-circular inspiral. Another important discovery was the existence of critical phenomena in gravitational col-



**Figure 2. Black hole orbital trajectories.** The plot at the top left shows the centers of the individual horizons (red and black) of two equal-mass, rapidly spinning black holes undergoing binary inspiral and merger. Also shown, at the end of the inspiral, are the horizons of the individual black holes and (in blue) the newly formed common horizon. The plot at the bottom shows the real part of the dominant quadrupole mode of the emitted gravitational-wave amplitude  $h$ ; the blowup reveals details of the end of the merger and of the ringdown phase during which the merged black hole settles down to its final equilibrium state. Here  $r$  is the distance to the source. Distances and times are normalized by the initial total mass of the binary  $M$ , and Newton's gravitational constant and the speed of light have been set equal to 1. In those geometrized units, a solar mass is 4.95  $\mu\text{sec}$  or 1.48 km. (Adapted from ref. 16.)

lapse.<sup>6</sup> Other simulations tracked the collapse of rotating neutron stars and of clusters of collisionless matter, as well as the head-on collisions of those objects.

Given those successes and others, it seemed like a small step to relax the assumptions of spherical or axial symmetry and to simulate the realistic inspiral and merger of binary black holes in quasi-circular orbits. The advance of gravitational-wave detectors added urgency to the need for such simulations, so in the mid 1990s NSF funded a grand challenge project to accelerate code development. It soon became evident, however, that the relaxation of symmetry introduced a new set of unexpected difficulties. When simulations evolved all three spatial dimensions—even for weak gravitational fields in vacuum—they developed numerical instabilities and crashed after only short integration times. The outcome was even more dire for black holes.

### A series of advances

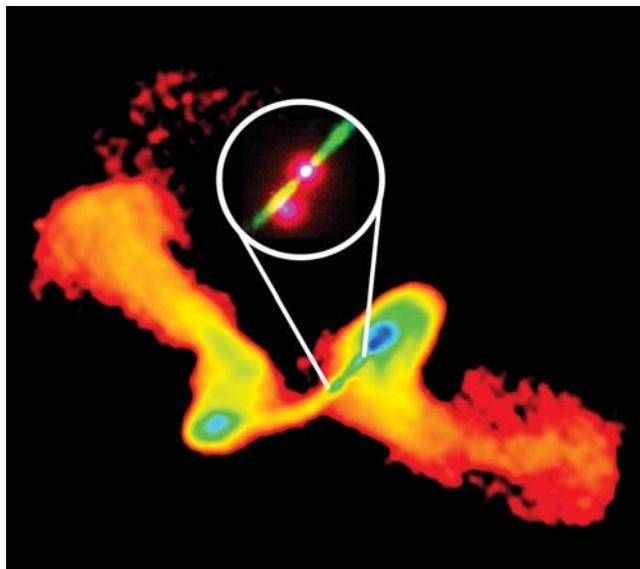
Thanks to progress on a number of fronts, numerical relativists have overcome the problems encountered in the early simulations. The first important step involved the reformulation of Einstein's equations.

Starting with the ADM equations, for example, one can generate an equivalent set of equations by introducing new variables that absorb derivatives of some of the original variables and by adding multiples of the constraints to the evolution equations. It turns out that some formulations behave much better than others when implemented numerically. The fact that reformulating the equations should have any effect at all may seem puzzling, since all manipulations involve mathematical identities—a solution to one set of equations is automatically a solution to any other. But one way to understand how a reformulation can help is to recognize that adding multiples of the constraint equations, for example, affects the mathematical character of the equations. In particular, some formulations satisfy criteria that guarantee stable or otherwise well-behaved solutions, while others do not.<sup>1</sup> Perhaps more intuitively, numerical simulations do not yield exact solutions: Round-off errors will lead to the constraints not being satisfied. Those constraint violations behave differently in different formulations and can grow unstably.

Maxwell's equations readily illustrate how changing the formulation can alter the behavior of constraint violations. Equation 4 demonstrates that in the standard formulation, a nonzero constraint, once it has developed during a numerical simulation, will persist. In box 2 we reformulate Maxwell's equations so that constraint violations satisfy a wave equation instead of equation 4 and can propagate away.

Reformulations of Einstein's equations work in a similar way. Successful applications include the Baumgarte-Shapiro-Shibata-Nakamura formulation<sup>7</sup> and the generalized harmonic formulation.<sup>8</sup> With BSSN it soon proved possible to simulate coalescing binary neutron stars,<sup>9</sup> which, at least initially, do not contain any singularities. Simulating black holes, however, necessarily requires careful handling of their interior spacetime singularities. One approach invokes black hole excision, whereby the interior of a black hole is removed from the computational mesh. That surgery is justified physically since, by definition, the black hole interior cannot affect the exterior. An alternative approach exploits so-called moving-puncture coordinates, imposed by adopting special criteria for the lapse and shift functions defined in box 1. The spatial slices simulated with those coordinates never reach the spacetime singularity, so there is no need for black hole excision.

Other computational advances have contributed to the



**Figure 3. Evidence of a spin flip?** The radio galaxy NGC 326 displays X-shaped radio jets that appear to have changed direction suddenly. One possible explanation of the sudden change is that a merger of black holes led to a flip in the spin of the black hole emitting the jets. (Image from National Radio Astronomy Observatory/Associated Universities Inc, observers Matteo Murgia and colleagues.) The inset (courtesy of the Space Telescope Science Institute) shows the innermost jets, formed from the most recent emission.

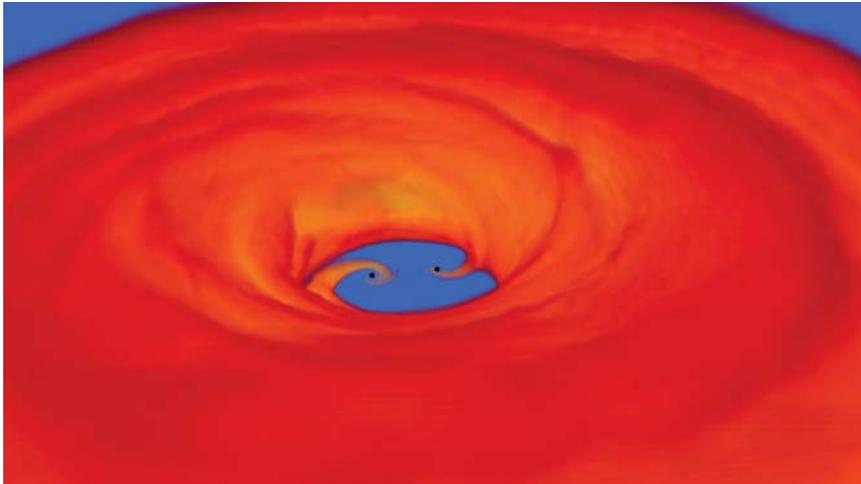
successful simulation of black hole mergers. Recent numerical work on parallel computers employs high-order approximation methods and adaptive mesh refinement (that is, nested grids with adjustable resolution) to meet the demands inherent in the problem.

After decades of effort and anticipation, the combination of the above techniques enabled the first successful simulations of binary black hole inspiral and merger, first by Frans Pretorius and shortly after by Manuela Campanelli and colleagues and by John Baker and colleagues.<sup>10</sup> The announcement of the simulations generated great excitement and opened the door to the exploration of physical and astrophysical consequences of black hole mergers.

### Waves, recoils, and spin flips

Among the first results from simulations of binary black hole mergers were gravitational waveforms, such as shown in figure 2, that could be used in building theoretical templates for gravitational-wave searches. Those real waveforms bear a somewhat surprising similarity to earlier, qualitative sketches, at least for binaries with a high degree of symmetry. Some genuine surprises emerged, though, for asymmetric binaries with unequal masses, misaligned spins, or both.

One surprise surfaced in connection with the recoil of a merged binary. The recoil itself is to be expected when the binary is asymmetric. To understand why, consider a spinning, S-shaped lawn sprinkler with its two curved arms. Each arm emits water, and hence linear momentum, in a tangential direction; that's what makes the sprinkler spin. If both arms emit water at the same rate, the loss of linear momentum from each of the two arms cancels, and the sprinkler feels no net force. Suppose, though, that one arm emits water at a greater rate than the other. The loss of linear momentum no longer cancels, and a net force and an acceleration of the center of mass are the results. If such a lawn sprinkler were



**Figure 4. Coalescence in gas.** Shown is a snapshot from a simulation of equal-mass, nonspinning black holes immersed in a circumbinary gaseous disk. Blue regions are empty; lighter shades of warm colors indicate higher density. Tidal torques from the binary carve out a hollow in the disk, but some gas spirals into the hollow and accretes onto the inspiraling black holes. Precursor electromagnetic radiation from the gas precedes the merger and peak gravitational-wave signal, and aftermath radiation is generated as the disk fills in the hollow and gas is steadily accreted onto the black hole remnant. (Simulation by Brian Farris, Yuk Tung Liu, and Stuart Shapiro.)

placed into outer space, its center of mass would describe a circle. Now imagine slowly increasing the flow of water. That would lead to the center of mass describing an outward spiral. Finally, if the water were suddenly shut off, the sprinkler would coast in some direction in its orbital plane.

A similar effect exists for binary black holes. Each companion in the binary is constantly accelerated and emits linear momentum in the form of gravitational waves. For symmetric binaries, the emissions from the two companions cancel each other, but for asymmetric binaries they do not. Moreover, as the binary orbit decreases, the black holes' speed, acceleration, and emission of linear momentum increase. Therefore, during the inspiral of an asymmetric binary the center of mass describes an outward spiral. The process ends when the binary merges, leaving the remnant with a recoil motion.

Recoil was first simulated for nonspinning black holes with unequal masses.<sup>11</sup> Consistent with earlier post-Newtonian estimates, those simulations found a maximum recoil speed of about 175 km/s, achieved for binaries whose black holes had a mass ratio of about 0.36. The surprise is that for spinning black holes, the recoil speeds can be much larger: For equal-mass binaries in which the black hole spins have large and antialigned components in the orbital plane, recoil speeds can exceed 4000 km/s.<sup>12</sup> That high value has important astrophysical consequences because it exceeds the escape speed from even the largest galaxies. According to the popular model of hierarchical structure formation, large galaxies form by the successive mergers of smaller ones. As they merge, the supermassive black holes they host also merge. If those supermassive black hole mergers were to routinely result in recoil speeds exceeding the escape speeds from the remnant galaxies, the number of bulge galaxies containing supermassive black holes would be smaller than is inferred from observations. That issue already has triggered a number of cosmological investigations. The key question is the likelihood of mergers that result in very large recoil speeds. Even for a random spin distribution, the likelihood is probably small. Moreover, several studies suggest that interactions with ambient gas tend to align the black hole spins with the binary's orbital angular momentum and thereby reduce the resulting recoil speed.<sup>13</sup>

Black hole spin affects binary mergers in other interesting ways. If one neglects the angular momentum emitted by gravitational radiation, the spin of a merger remnant is the sum of the two progenitor black hole spins and the binary's orbital angular momentum just prior to merger. The remnant spin may therefore point in a direction different from those

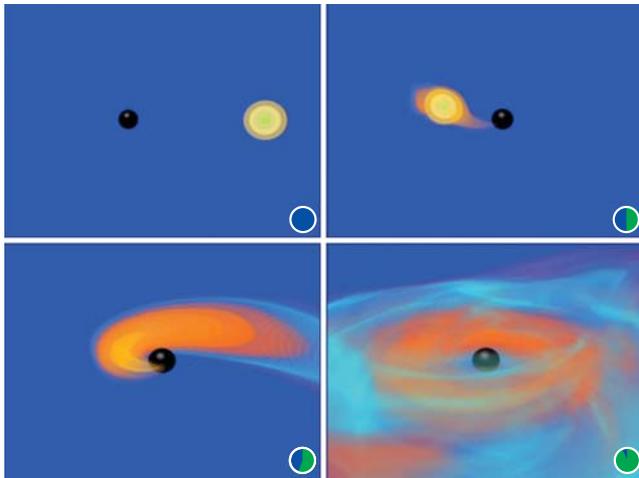
of its progenitors. Such a change in spin may have observable consequences, since current theory holds that radio jets from radio galaxies are emitted along the spin axis of black holes at the centers of those galaxies. Indeed, some radio jets display an X-shaped pattern, such as shown in figure 3, which suggests that the spin of the black hole powering the jets flipped at some point in the past.

### Immediate challenges

Simulations have far to go before they completely survey different black hole mass ratios and spins. Exploring that multidimensional parameter space remains a daunting task, but it is of crucial importance if physicists are to assemble the waveform templates needed for matched-filtering data analysis. New international collaborations aim to extend the quality and scope of the simulations and to facilitate the incorporation of theoretical waveforms into data analysis—all in time for the first gravitational-wave detections. The NINJA (Numerical Injection Analysis) team is assembling waveforms from various groups and using them to test search algorithms. The NRAR (Numerical-Relativity and Analytical-Relativity) collaboration is stitching together post-Newtonian and numerical waveforms to provide the best-calibrated template families covering the largest parameter space of binary masses and spins.

To appreciate some of the challenges, consider the mass ratio. Current computational resources lack the capability to simulate binaries with mass ratios smaller than about  $10^{-2}$ . Extreme-mass-ratio inspirals with mass ratios less than about  $10^{-5}$  can be handled analytically by applying black hole perturbation theory, which employs gravitational radiation-reaction forces to drive the inspiral. Such systems, which include stellar-mass black holes in orbit about supermassive black holes, may form in galaxy cores and provide sources of low-frequency gravitational waves for LISA. But neither post-Newtonian methods nor perturbation theory provide sufficiently accurate waveform templates for the intermediate-mass-ratio inspirals that may arise when intermediate-mass black holes spiral into supermassive black holes.

A comprehensive survey of black hole mergers will also inform cosmologists' understanding of structure formation. The coalescence of massive and supermassive black holes is triggered by the merger of their host galaxies. Galaxy mergers can drive the hierarchical buildup of large-scale structure in the early universe. Meanwhile, the shock heating and accretion of ambient gas by a coalescing binary can generate appreciable electromagnetic radiation. That situation raises



**Figure 5. A neutron star–black hole encounter.** The panels show four instants of the tidal disruption of a neutron star by a black hole in a binary system. The black hole has a mass three times that of the neutron star; neither object is spinning. The bulk of the matter is captured by the black hole, with only about 4% remaining in a disk. Color coding indicates varying densities; green regions are densest, blue regions are least dense. Clocks in the lower right-hand corners of the panels track the simulation's progress. (Simulation by Zachariah Etienne and colleagues; see ref. 15.)

the exciting prospect of a near-simultaneous detection of electromagnetic and gravitational waves from a binary merger. The observation of electromagnetic radiation could confirm putative binary mergers, and the radiation itself could serve as a probe of the gas in galaxy cores and the physics of accretion.

With the binary coalescence problem in vacuum under control, numerical relativists are already turning to environmental issues as the next challenge. Meeting that challenge requires solving Einstein's equations together with both the equations of relativistic magnetohydrodynamics to evolve the magnetized plasma and the equations of radiation transport to follow the emitted photons. Both black hole recoil and mass loss due to gravitational waves have important effects on the accreting gas. Preliminary relativistic simulations provide rough estimates of electromagnetic precursor and afterglow radiation from mergers in large, gaseous clouds or circumbinary disks;<sup>14</sup> see figure 4. Multiwave instruments such as the Large Synoptic Survey Telescope now being planned should be able to monitor that transient emission.

Binary neutron stars and binaries comprising a black hole and a neutron star also show promise as sources of gravitational waves. In fact, relativistic simulations of binary neutron stars achieved success before those of binary black holes,<sup>9</sup> in part due to the absence of initial spacetime singularities. Not so with black hole–neutron star binaries, which pose the combined computational challenges of relativistic matter and singularities.<sup>15</sup>

When two neutron stars coalesce, the possibility exists that the merged stars will have a mass exceeding the maximum mass allowed for a single neutron star, about  $2\text{--}3 M_{\odot}$ . When that happens, the fate of the system depends on its mass. High-mass systems undergo prompt collapse to black holes; low-mass systems form differentially rotating (that is, not rigidly rotating) hypermassive stars whose delayed collapse after many periods is triggered by magnetic fields, gravitational-wave emission, or both. A neutron star coalesc-

ing with a stellar-mass black hole is tidally disrupted prior to merger, as shown in figure 5; a neutron star plunging into a supermassive black hole is swallowed whole.

Both neutron-star and black hole–neutron star binaries are candidate sources of short duration, high-energy GRBs. The disk that forms from the debris around the spinning remnant gives rise to magnetic fields and outflows collimated along the rotation axis; those can power a GRB. Alternative GRB models posit hypermassive neutron stars, among other possibilities. Relativistic simulations model the various scenarios and explore the exciting possibility of a simultaneous detection of gravitational waves and a GRB from the same cosmic source.

Many of the computational hurdles of simulating black hole mergers have been overcome. But work has just begun to extract gravitational waveforms and electromagnetic signatures in advance of their detection and to determine how likely are the events that generate that radiation and what are their cosmological impacts. In the years ahead, the job will be aided by advances in computer hardware, improved simulations, and highly anticipated observational data.

*We gratefully acknowledge NSF and NASA for their research support.*

*The online version of this article includes videos of the simulations that produced figures 4 and 5.*

## References

1. M. Alcubierre, *Introduction to 3 + 1 Numerical Relativity*, Oxford U. Press, New York (2008).
2. C. Bona, C. Palenzuela-Luque, C. Bona-Casas, *Elements of Numerical Relativity and Relativistic Hydrodynamics: From Einstein's Equations to Astrophysical Simulations*, Springer, New York (2009).
3. T. W. Baumgarte, S. L. Shapiro, *Numerical Relativity: Solving Einstein's Equations on the Computer*, Cambridge U. Press, New York (2010).
4. R. Arnowitt, S. Deser, C. W. Misner, in *Gravitation: An Introduction to Current Research*, L. Witten, ed., Wiley, New York (1962), p. 227. See also J. W. York Jr, in *Sources of Gravitational Radiation*, L. L. Smarr, ed., Cambridge U. Press, New York (1979), p. 83.
5. S. G. Hahn, R. W. Lindquist, *Ann. Phys.* **29**, 304 (1964); L. Smarr et al., *Phys. Rev. D* **14**, 2443 (1976); P. Anninos et al., *Phys. Rev. Lett.* **71**, 2851 (1993); R. A. Matzner et al., *Science* **270**, 941 (1995).
6. M. W. Choptuik, *Phys. Rev. Lett.* **70**, 9 (1993).
7. T. Nakamura, K. Oohara, Y. Kojima, *Prog. Theor. Phys. Suppl.* **90**, 1 (1987); M. Shibata, T. Nakamura, *Phys. Rev. D* **52**, 5428 (1995); T. W. Baumgarte, S. L. Shapiro, *Phys. Rev. D* **59**, 024007 (1999).
8. H. Friedrich, *Commun. Math. Phys.* **100**, 525 (1985); D. Garfinkle, *Phys. Rev. D* **65**, 044029 (2002).
9. M. Shibata, *Phys. Rev. D* **60**, 104052 (1999).
10. F. Pretorius, *Phys. Rev. Lett.* **95**, 121101 (2005); M. Campanelli et al., *Phys. Rev. Lett.* **96**, 111101 (2006); J. G. Baker et al., *Phys. Rev. Lett.* **96**, 111102 (2006).
11. J. G. Baker et al., *Astrophys. J. Lett.* **653**, L93 (2006); F. Hermann et al., *Class. Quantum Grav.* **24**, S33 (2007); J. A. González et al., *Phys. Rev. Lett.* **98**, 091101 (2007).
12. J. A. González et al., *Phys. Rev. Lett.* **98**, 231101 (2007); M. Campanelli et al., *Phys. Rev. Lett.* **98**, 231102 (2007); C. O. Lousto, Y. Zlochower, <http://arxiv.org/abs/1108.2009>.
13. See, for example, T. Bogdanović, C. S. Reynolds, M. C. Miller, *Astrophys. J. Lett.* **661**, L147 (2007); M. Volonteri, K. Gültekin, M. Dotti, *Mon. Not. R. Astron. Soc.* **404**, 2143 (2010).
14. M. Megevand et al., *Phys. Rev. D* **80**, 024012 (2009); B. D. Farris, Y. T. Liu, S. L. Shapiro, *Phys. Rev. D* **81**, 084008 (2010); T. Bode et al., *Astrophys. J.* **715**, 1117 (2010); C. Palenzuela, L. Lehner, S. L. Liebling, *Science* **329**, 927 (2010); O. Zanotti et al., *Astron. Astrophys.* **523**, A8 (2010).
15. M. Shibata, K. Taniguchi, *Phys. Rev. D* **77**, 084015 (2008); Z. B. Etienne et al., *Phys. Rev. D* **79**, 044024 (2009); F. Foucart et al., *Phys. Rev. D* **83**, 024005 (2011); B. C. Stephens, W. E. East, F. Pretorius, *Astrophys. J. Lett.* **737**, L5 (2011).
16. G. Lovelace, M. A. Scheel, B. Szilágyi, *Phys. Rev. D* **83**, 024010 (2011). ■