Review

Numerical relativity of compact binaries in the 21st century

Matthew D Duez and Yosef Zlochower

1 Department of Physics and Astronomy, Washington State University, Pullman, WA 99164, United States of America
2 Center for Computational Relativity and Gravitation and School of Mathematical Sciences, Rochester Institute of Technology, 85 Lomb Memorial Drive, Rochester, NY 14623, United States of America

E-mail: m.duez@wsu.edu

Received 11 January 2016, revised 13 August 2018
Accepted for publication 17 August 2018
Published 7 November 2018

Corresponding Editor Dr Beverley Berger

Abstract
We review the dramatic progress in the simulations of compact objects and compact-object binaries that has taken place in the first two decades of the twenty-first century. This includes simulations of the inspirals and violent mergers of binaries containing black holes and neutron stars, as well as simulations of black-hole formation through failed supernovae and high-mass neutron star–neutron star mergers. Modeling such events requires numerical integration of the field equations of general relativity in three spatial dimensions, coupled, in the case of neutron-star containing binaries, with increasingly sophisticated treatment of fluids, electromagnetic fields, and neutrino radiation. However, it was not until 2005 that accurate long-term evolutions of binaries containing black holes were even possible (Pretorius 2005 Phys. Rev. Lett. 95 121101, Campanelli et al 2006 Phys. Rev. Lett. 96 111101, Baker et al 2006 Phys. Rev. Lett. 96 111102). Since then, there has been an explosion of new results and insights into the physics of strongly-gravitating system. Particular emphasis has been placed on understanding the gravitational wave and electromagnetic signatures from these extreme events. And with the recent dramatic discoveries of gravitational waves from merging black holes by the Laser Interferometric Gravitational Wave Observatory and Virgo, and the subsequent discovery of both electromagnetic and gravitational wave signals from a merging neutron star–neutron star binary, numerical relativity became an indispensable tool for the new field of multimessenger astronomy.

Keywords: black holes, numerical relativity, relativistic binaries, compact objects

(Some figures may appear in colour only in the online journal)
2.5. Methods for evolving the fluid equations............. 11
2.5.1. Conservative formulation................................. 11
2.5.2. Excision and punctures in the presence of matter. ................................................................. 12
2.5.3. More physics: equations of state, neutrinos, magnetic fields. ................................................. 13
3. Black hole–black hole binary simulations............. 14
3.1. Early efforts.......................................................... 14
3.2. Post-breakthrough results................................... 14
3.2.1. Recoils. .......................................................... 14
3.2.2. Modeling the remnant properties............... 16
3.2.3. Numerical relativity at the extremes............... 17
3.2.4. Waveform modeling......................................... 18
3.2.5. Semi-analytic waveform models.................. 20
4. Relativistic stars and disks..................................... 21
4.1. Black hole–black hole binaries with accretion... 21
4.2. Relativistic stars.................................................. 22
4.2.1. Radial stability and collapse outcome........... 23
4.2.2. Magnetohydrodynamic evolution.................. 23
4.2.3. Nonaxisymmetric mode instability............... 24
4.2.4. Stability of self-gravitating black hole accretion disks......................................................... 24
5. Black hole formation.............................................. 25
5.1. Population I/II core collapse and collapsars....... 25
5.2. Massive star collapse in the early universe.......... 26
6. Non-vacuum compact binaries............................... 27
6.1. White dwarf—compact object binaries............... 28
6.2. Black hole–neutron star binaries...................... 28
6.2.1. Expectations before numerical relativity simulations............................................................. 28
6.2.2. Inspiral and merger in numerical relativity: parameter space exploration and gravitational waves. ................................................................. 29
6.2.3. Post-merger in numerical relativity: neutrinos, ejecta, and MHD........................................ 30
6.3. Neutron star–neutron star mergers.................... 31
6.3.1. Pre-breakthrough simulations...................... 31
6.3.2. Post-merger evolution: neutrinos, ejecta, MHD.......................................................... 32
6.3.3. General relativistic effects: prompt collapse threshold and post-merger gravitational waves................................................................. 32
6.3.4. Longer term evolution of remnants: subgrid scale modeling..................................................... 33
7. Comparison to observations.................................... 34
7.1. Gravitational wave astronomy............................ 34
7.2. GW170817: The age of multimessenger astronomy begins...................................................... 34
8. Conclusion........................................................... 35
Acknowledgments...................................................... 36
References............................................................. 36

1. Introduction

With the discovery of gravitational waves from merging black hole–black hole binaries by the Laser Interferometric Gravitational Wave Observatory (LIGO) in 2015 [4–6], the subsequent observations of other black hole mergers by LIGO [7, 8] and later by LIGO and Virgo [9], and the simultaneous observation of gravitational waves and electromagnetic spectra from the merger of a neutron star–neutron star binary by LIGO, Virgo, and a large team of astronomers in 2017 [10–12], the new field of gravitational wave and multimessenger astronomy was born.

Fundamental to this new science is the ability to infer the dynamics of the sources based on the observed signals, something that can only be accomplished using detailed theoretical predictions based on numerical simulations of the nonlinear Einstein field equations of general relativity both in vacuum and coupled to the equations of magnetohydrodynamics. Indeed, the body of techniques that emerged based on efforts to solve this system, known as numerical relativity, was designed largely with such gravitational wave source modeling in mind, but it has also been turned to other astrophysical phenomena involving strongly-curved, dynamical spacetime.

Only rather exotic phenomena involve sufficiently strong spacetime curvature to require numerical relativity. Newtonian gravity clearly works quite well for main sequence stars, planets, and the like. As is well-known, relativity becomes important when speeds approach the speed of light $c$, so a reasonable guess would be to expect important relativistic effects as the escape velocity approaches $c$. Then an object of mass $M$ and radius $R$ will require relativistic treatment if $R$ is close to the gravitational radius $r_G \equiv 2GM/c^2$, the radius of a nonspinning black hole of mass $M$. The same condition can be stated in terms of the dimensionless compactness $C \equiv GM/r_G^2$. Strong-gravity objects have high compactness (order unity being the standard of ‘high’). Black holes ($C \sim 1$) and neutron stars ($C \sim 0.1$) are compact objects by this definition. White dwarfs ($C \sim 10^{-4}$) are a marginal case—relativity plays a large role in their stability condition but not their equilibrium structure—and are usually also classified as compact.

Formulating the integration of the Einstein equations so that evolutions are stable and the coordinates evolve sensibly turned out to be a difficult task. There was some worry that numerical relativity might not be ready when the advanced gravitational wave detectors needed it. Finally, in 2005 [1–3], the first stable black hole–black hole binary merger simulations were carried out. There followed a race to produce accurate waveforms for gravitational wave observation efforts, which were already underway.

In this review, we describe how numerical relativity has come to be a robust tool for studying strong-gravity systems. We also review some of the major accomplishments of numerical relativity to date. To provide an appropriate scope, we focus on applications to compact binaries and black hole formation, processes where general relativity is essential and whose astrophysical importance is clear.

The article is organized as follows. In the rest of the introduction, we provide background on general relativity, black holes, and relativistic stars. In section 2, we cover methods for evolving the Einstein field equations and coupled matter sources. Particular attention is given to the historical ‘breakthrough’ discoveries that enabled stable evolutions of
multiple-black-hole spacetimes. The next sections review simulation results, covering the period before the breakthroughs and after. Section 3 is devoted to black hole–black hole binary mergers. The following three sections describe simulations of phenomena with matter, especially neutron stars. Finally, in section 7, we return to our original motivation and consider what has been learned by the confrontation of numerical relativity predictions with actual LIGO-Virgo observations.

1.1. The field equations of general relativity

The theory of special relativity introduced the notion of spacetime. In that theory, spacetime is a geometrically flat 4-dimensional (4D) manifold. General relativity extends this notion to non-flat manifolds. In general relativity, the Newtonian notion of a gravitational force is replaced by geodesic motion in a curved spacetime. Unless acted on by other (non-gravitational) forces, objects whose size is much smaller than the local spacetime radius of curvature travel along geodesics.

Throughout this paper, we will use geometric units. In these units, the speed of light, c, and Newton’s constant G, are taken to be 1. A consequence of this is that distances, time intervals, masses, and energy all have the same units. By convention, the unit of each of these is denoted by an arbitrary mass $M^3$.

In the section below, we will provide an extremely brief overview of the field equations of general relativity. For a comprehensive overview, we suggest consulting [13] for a very accessible introduction to general relativity, and [14, 15], [16] for a more advanced treatment. The material below was synthesized from these references.

The geometry of a spacetime can be entirely described by a line element $d\mathbf{s}^2$. In Minkowski spacetime, in Cartesian coordinates, the line element takes the form

$$d\mathbf{s}^2 = -dt^2 + dx^2 + dy^2 + dz^2. \quad (1)$$

Note that $d\mathbf{s}^2$ is not necessarily positive. For timelike paths, the proper time along the path is given by the integral of $d\tau = \sqrt{-ds^2}$. Equation (1) can be written as

$$d\mathbf{s}^2 = \eta_{\mu\nu}dx^\mu dx^\nu, \quad (2)$$

where $x^{\mu} = (t, x, y, z)$ and the components of the symmetric tensor $\eta_{\mu\nu}$ are given by diag($-1, 1, 1, 1$). Here we used two standard conventions, the timelike coordinate is listed first and repeated Greek indices are summed over. By convention, the index of the timelike coordinate is 0, while the spatial coordinates have indices 1, 2, 3. In arbitrary coordinates, the line element becomes

$$d\mathbf{s}^2 = g_{\mu\nu}dx^\mu dx^\nu, \quad (3)$$

where $y^{\mu}$ is some new set of coordinates and

$$g_{\mu\nu} = \left( \frac{\partial x^\alpha}{\partial y^\mu} \right) \left( \frac{\partial x^\beta}{\partial y^\nu} \right) \eta_{\alpha\beta}. \quad (4)$$

3 For example, solar mass intervals of distance and time are about 1.5 km and $5 \times 10^{-6}$ s, respectively.

Here, we will not make a distinction between the components of a tensor and a tensor itself. For our purposes a tensor $L$ of type $(p, q)$ is a set of $p \times q$ functions, denoted by

$$L^{\alpha_1\alpha_2\cdots\alpha_p}_{\beta_1\beta_2\cdots\beta_q}, \quad (5)$$

which, under a change of coordinates from some coordinate system $x$ to another $y$, transform as

$$L^{\alpha_1\alpha_2\cdots\alpha_p}_{\beta_1\beta_2\cdots\beta_q}(x) = \left( \frac{\partial x^{\alpha_1}}{\partial y^\mu} \right) \left( \frac{\partial x^{\alpha_2}}{\partial y^\nu} \right) \cdots \left( \frac{\partial x^{\alpha_p}}{\partial y^{\nu_p}} \right) \times \left( \frac{\partial y^\mu}{\partial y^{\delta_1}} \right) \left( \frac{\partial y^\nu}{\partial y^{\delta_2}} \right) \cdots \left( \frac{\partial y^{\nu_p}}{\partial y^{\delta_q}} \right) L^{\mu_1\mu_2\cdots\mu_p}_{\nu_1\nu_2\cdots\nu_q}(x). \quad (6)$$

Thus the metric $g_{\mu\nu}$ is a tensor. Associated to $g_{\mu\nu}$ is its matrix inverse $g^{\mu\nu}$ (i.e. $g^{\mu\nu}g_{\nu\sigma} = \delta^\mu_\sigma$, where $\delta$ indicates the usual Kronecker delta function).

The metric, $g_{\mu\nu}$, in special relativity is intrinsically flat. By this, we mean any of several equivalent statements: there is a coordinate transformation such that the new metric is everywhere identical to $\eta_{\mu\nu}$, as discussed above, parallel geodesics will remain parallel, and the intrinsic curvature of the metric, as measured by the Riemann curvature tensor, vanishes everywhere. We briefly describe how geodesics and the Riemann curvature tensor are calculated.

A geodesic is the generalization of a straight line in Euclidean space. In Cartesian coordinates, the tangent vector to a straight line is a constant. This can be expressed as

$$t^\mu = \frac{d}{d\lambda} x^\mu(\lambda), \quad (7)$$

$$\frac{d}{d\lambda} t^\mu(\lambda) = 0, \quad (8)$$

where $t^\mu$ is the tangent vector to the line and $x^\mu(\lambda)$ are the Cartesian coordinates of each point of the line. Equation (8) can be re-written as

$$t^\mu \frac{\partial x^\nu}{\partial \lambda} = 0. \quad (9)$$

However, even in flat space, this equation does not hold true in arbitrary coordinates. To fix this, we replace the ordinary derivative $\partial / \partial x^\nu$ with the covariant derivative $\nabla_\nu$. The equation for a geodesic in arbitrary coordinates, as well as flat and non-flat metrics, is

$$\frac{d}{d\lambda} x^\mu(\lambda) = t^\mu(\lambda), \quad (10)$$

$$t^\mu \nabla_\nu t^\nu = 0. \quad (11)$$

Finally, $\nabla_\nu$, the covariant derivative associated with the metric $g_{\mu\nu}$, is defined by its action on arbitrary tensors $L^{\mu_1\mu_2\cdots}_{\nu_1\nu_2\cdots}$, which is given by

$$\nabla_\alpha L^{\mu_1\mu_2\cdots}_{\nu_1\nu_2\cdots} = \partial_\alpha L^{\mu_1\mu_2\cdots}_{\nu_1\nu_2\cdots} + \Gamma^\mu_{\alpha\beta} L_{\beta\nu_1\nu_2\cdots} + \Gamma^{\nu_1}_{\alpha\beta} L^{\mu_1\mu_2\cdots}_{\beta\nu_2\cdots} + \cdots \quad (12)$$

$$- \Gamma^{\beta}_{\alpha\nu} L^{\mu_1\mu_2\cdots}_{\beta\nu_2\cdots} - \Gamma^{\nu}_{\alpha\beta} L^{\mu_1\mu_2\cdots}_{\beta\nu_2\cdots} - \cdots ,$$

where
\[ \Gamma^\sigma_{\mu\nu} = \frac{1}{2} g^{\sigma\alpha} (\partial_\mu g_{\alpha\nu} + \partial_\nu g_{\alpha\mu} - \partial_\alpha g_{\mu\nu}) , \] (13)

and \( \partial_\alpha \) is shorthand for \( \partial/\partial x^\alpha \). The components of \( \Gamma^\sigma_{\mu\nu} \) are collectively known as the Christoffel symbols. Unlike the metric and the tensors constructed from the Christoffel symbols below, the components of the Christoffel symbols do not transform according to equation (6).

The Riemann curvature tensor is constructed from the metric and has the following form:

\[ R_{\mu\nu\rho\sigma} = \partial_\sigma \Gamma^\alpha_{\mu\nu} - \partial_\nu \Gamma^\alpha_{\rho\sigma} + \Gamma^\alpha_{\mu\beta} \Gamma^\beta_{\rho\sigma} - \Gamma^\alpha_{\rho\beta} \Gamma^\beta_{\mu\sigma} . \] (14)

The Riemann curvature tensor can be further split into a trace-free part, known as the Weyl tensor \( C_{\mu\nu\rho\sigma} \), and the Ricci tensor

\[ R_{\mu\nu} = R_{\mu\nu\rho\sigma} \Gamma^\rho_{\sigma} . \] (15)

Finally, the Einstein tensor \( G_{\mu\nu} \), is given by

\[ G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \frac{1}{2} g_{\mu\nu} \rho. \] (16)

Regardless of coordinates, the Riemann curvature tensor is identically zero in special relativity. General relativity extends the notion of spacetime to include non-flat metrics, where the Riemann curvature tensor \( R_{\mu\nu\rho\sigma} \) is non-vanishing.

In general relativity, the curvature of spacetime possesses its own dynamics, and indeed one of the most important phenomena treated by numerical relativity, the merger of two black holes, is a vacuum problem. That is, \( T_{\mu\nu} = 0 \). Numerical relativity is also used to study phenomena involving matter flows in strongly-curved dynamical spacetimes. Two problems where such relativistic effects should be particularly important are compact object mergers involving neutron stars and the formation of black holes by stellar collapse.

Matter and energy constitute the stress-energy tensor \( T_{\mu\nu} \) that is the source term in Einstein equations. In general, \( T_{\mu\nu} \) will be a sum of stress tensors for the gas, electromagnetic, and neutrino fields. The energy and momentum conservation equations

\[ \nabla_\mu T^{\mu\nu} = 0 \] (18)

provide evolution equations for the matter.

In the non-vacuum systems we will consider, the particles—including nucleons, nuclei, electrons, positrons, and photons, but not necessarily neutrinos—form a nearly perfect fluid, meaning the mean free path is very small compared to the system’s scale, making the collection a fluid, and viscosity and heat transport are small enough to be ignored. This fluid will have a stress tensor

\[ T_{\mu\nu}^{\text{gas}} = (\rho_0 + u + P) u_\mu u_\nu + P g_{\mu\nu} , \] (19)

where \( \rho_0, u, P \), and \( u_\mu \) are the rest mass density, internal energy, pressure, and 4-velocity. (Note \( \rho_0 \) must be distinguished from the total energy density \( \rho = \rho_0 + u \).) The 4-velocity has only three independent components, with the Lorentz factor \( \gamma \equiv u^0 \) given by the normalization condition \( u \cdot u = -1 \):

\[ \gamma^2 = 1 + u^0. \] (20)

Equations (18) and (19) must be supplemented by the rest mass conservation equation

\[ \nabla_\mu (\rho_0 u^\mu) = 0 \] (21)

and also an equation of state (EoS)

\[ P = P(\rho_0, T, X_i) \] (22)

\[ u = u(\rho_0, T, X_i) , \] (23)

where \( T \) is the temperature and \( X_i \) are composition variables.

For a detailed exposition of relativistic hydrodynamics, including its numerical treatment, see the book by Rezzolla and Zanotti [634].

1.3. Black holes

Perhaps one of the most interesting predictions of general relativity is the existence of black holes. Black holes are regions in spacetime where the curvature is sufficiently strong that light, and therefore any physical signal, cannot escape. Astrophysically, black holes form as stellar objects collapse. Despite all the microphysics that goes into the dynamics of stellar objects, once equilibrated, a black hole can be completely described by two parameters: its mass and spin.\(^6\) In geometric units the magnitude of the spin angular momentum \( S \) is bounded by the mass \( m \), where \( S < m^2 \). Typically one defines a specific spin \( a \), where \( a = S/m \) and a dimensionless spin \( \chi \), where \( \chi = S/m^2 \).

If the black hole is non-spinning, it is known as a Schwarzschild black hole [17, 18], and if \( S \) is non-zero, it is known as a Kerr black hole [19]. In both cases the black hole spacetimes are named after their discoverers.

A black hole has no material surface, of course, but there is a boundary separating the region from which it is impossible ever to escape (the black hole interior) to the outside universe. This boundary is called the event horizon. The event horizon is a null surface, i.e. one must move at the speed of light to stay on it. In order to determine if a true event horizon exists, one needs to know the entire future of the spacetime (otherwise there is always the possibility that an observer can escape

\(^4\) Readers familiar with differential geometry will notice that in equation 13 we are limiting ourselves to coordinate (holonomic) bases, which are sufficient for numerical relativity.

\(^5\) Note that while the components of \( R_{\mu\nu\rho\sigma} \) transform according to equation (6), the components of \( \Gamma^\sigma_{\mu\nu} \) do not.

\(^6\) More correctly, an equilibrated black hole is completely described by its mass, spin, and charge. However, astrophysical black holes are expected to have effectively zero charge because accretion from the interstellar medium should rapidly discharge them.
the supposed black hole at a later time). In practice, numerical relativists find event horizons by evolving a cluster of null geodesics or a null surface backwards in time starting at the very end of their simulations. (For a review, see [20].) We will see that some methods of numerically handling black hole interiors (excision methods) require some knowledge of the horizon location during the simulation. For these purposes, numerical relativists use the apparent horizon. Apparent horizons are two-dimensional (2D) surfaces that may exist at each time in a numerical simulation. They are defined to be surfaces from which outward-pointing null rays do not expand. This very unusual situation can only occur in the vicinity of a black hole, but finding such surfaces only requires information about the metric and extrinsic curvature at a given time. For a stationary black hole, the apparent horizon will coincide with the event horizon; in a dynamical spacetime, it will be inside the event horizon.

Black holes can form binaries, as demonstrated by LIGO’s recent detection of gravitational waves [4, 5]. In such a case, the black holes are not truly in equilibrium, but each can still be described reasonably well as Kerr or Schwarzschild black holes, at least when the binary components are well separated. Such a binary can then be described by several intrinsic parameters, such as the mass ratio of the two black holes, the spin magnitudes and orientations of the two black holes, and the orbital eccentricity.

1.4. Relativistic stars

Much astrophysical thinking is guided by idealized equilibria, such as the spherically symmetric star and the thin accretion disk, and this remains true in the study of compact object systems. Neutron stars are extremely compact objects, containing a little over a solar mass ($\rho \approx 4$ km) within a radius of $\sim 10$ km, and so must be studied using general relativity; they will be our primary type of relativistic star.

For spherical neutron stars of an assumed barotropic equation of state $P = P(\rho)$, the Tolman–Oppenheimer–Volkoff (TOV) [21, 22] equations of hydrostatic equilibrium yield a sequence of equilibria, one for each central density $\rho_c$. At a critical central density $\rho_{\text{crit}}$, the mass reaches a maximum value $M_{\text{max,TOV}} = M(\rho_{\text{crit}})$, and only the configurations on the ascending ($dM/d\rho > 0$) side, which generally turns out to be $\rho_c < \rho_{\text{crit}}$, are stable. A star on the unstable side will either collapse to a black hole or undergo large radial oscillations about the lower density configuration of the same mass. (See figure 1.)

Can neutron stars exist with $M > M_{\text{max,TOV}}$? If the neutron star is spinning, this provides some additional support against gravity. Codes exist for generating the resulting 2D (axisymmetric) equilibria (e.g., [24, 25]). The matter is usually taken to be a perfect fluid with purely azimuthal flows [$u^r = u^\phi = 0$, $u^\rho = u^\Omega$], and a rotation law for $\Omega$ must be specified. Rotation might be uniform ($\Omega =$constant) or differential ($\Omega$ varies through the star). Viscosity (or similar angular momentum transport mechanisms) will tend to produce uniform rotation, but differentially rotating equilibria can persist on timescales shorter than that of viscosity.

\[ \Omega = \text{constant} \]

\[ \Omega \text{ or similar angular momentum transport mechanisms} \]

\[ \text{differential rotation, providing rotational support while} \]

\[ \text{viscosity (or similar angular momentum transport mechanisms)} \]

\[ \text{will tend to produce uniform rotation, but differentially} \]

\[ \text{rotating equilibria can persist on timescales shorter than that} \]

\[ \text{of viscosity.} \]

\[ \text{Rotation in equilibrium stars is constrained, though, by the} \]

\[ \text{mass-shedding limit, at which fluid on the equator is in geodesic} \]

\[ \text{('Keplerian') orbit. Faster rotation at the equator would} \]

\[ \text{centrifugally eject mass. This restricts the degree of possible} \]

\[ \text{uniform rotation particularly severely, so that the maximum} \]

\[ \text{mass only increases by around 20%, the most massive} \]

\[ \text{configurations being found close to (but not exactly on) the mass-} \]

\[ \text{shedding limit} \]

\[ \text{Stars with masses between} \]

\[ M_{\text{max,TOV}} \]

\[ \text{and this higher limit are called supramassive. A sufficient} \]

\[ \text{condition for instability of uniformly rotating stars (on the} \]

\[ \text{secular timescale on which uniform rotation is maintained) can} \]

\[ \text{be determined via locating the turning point in the constant} \]

\[ \text{angular momentum sequence} \]

\[ \text{similar to TOV sequences. (The actual instability onset occurs} \]

\[ \text{slightly on the 'stable'} \]

\[ \text{side} \]

\[ \text{The point of onset of dynamical instability can be} \]

\[ \text{determined by numerical simulations.} \]

\[ \text{Stars with mass above the supramassive limit are called} \]

\[ \text{hypermassive. Such equilibria can exist with the help of} \]

\[ \text{differential rotation, providing rotational support while} \]

\[ \text{evading the mass-shedding limit by keeping the rotation rate} \]

\[ \text{sub-Keplerian near the equator. Numerical relativity confirms} \]

\[ \text{that such stars can persist stably for multiple dynamical times-} \]

\[ \text{cales, as we discuss in section 4.2.} \]

\[ \text{One might also look to thermal support—hot nuclear} \]

\[ \text{matter—to increase the maximum mass, effectively changing} \]
the equation of state to give more pressure support. Effects of thermal support have been studied for uniformly rotating neutron stars by Goussard et al [30] and for uniformly and differentially rotating neutron stars by Kaplan et al [31]. The latter suggest an approximate turning point method for assessing stability which has been numerically confirmed by Weih et al [32] and used by Bauswein and Stergioulas to explain some numerical relativity findings on the threshold mass for prompt collapse of a neutron star–neutron star binary merger remnant to a black hole [33].

The distinctions introduced above between normal, supramassive, and hypermassive neutron stars have played a large role in the interpretation of neutron star–neutron star binary merger simulations. When two neutron stars merge, the resulting object will either collapse to a black hole or settle to a dynamical equilibrium state in roughly a dynamical timescale $\sim 10^{-1}$ ms. An equilibrium remnant could be described as a type of relativistic star. If two 1.4 $M_\odot$ neutron stars merge, this remnant could easily have a mass in excess of $M_{\text{max}}^{\text{TOV}}$. However, the remnant will also be spinning rapidly and differentially, and it will have acquired a great deal of heat from the merger shock. Thus, normal, supramassive, and hypermassive remnants are all possible, depending on the stars’ masses and the unknown value of $M_{\text{max}}^{\text{TOV}}$. However, while differentially rotating stars are in equilibrium on a dynamical timescale, they evolve on the secular timescales of effects that transport angular momentum ($\sim 10$ ms). Similarly, the equilibrium will be adjusted by loss of thermal support on the neutrino cooling timescale ($\sim s$). Except in the unlikely event that it sheds enough mass to drop below the supramassive limit, a hypermassive remnant will ultimately collapse on one of these timescales. Thus, the main outline of the post-merger evolution seems to depend on one parameter, the mass of the binary, and one EoS-related number, the neutron star maximum mass.

1.5. Posing the problem: recasting the field equations as an initial value problem

Returning to the Einstein equations themselves, similar to how the 4-vector $A^\mu$ in electromagnetism is not unique due to gauge freedom, the metric that satisfies equation (17) (and any relevant boundary conditions) is not unique. In general relativity, the gauge freedom comes in the form of the freedom to choose coordinates arbitrarily. In order to get a unique solution, we need to impose gauge conditions. Many, but not all, formulations of the Einstein equations for numerical simulations use what is known as a 3 + 1 decomposition [34] (see also recent texts on numerical relativity [35–37]). In a 3 + 1 decomposition, the coordinates are constructed by using a family of non-intersecting, spatial hypersurfaces. The basic setup is illustrated in figure 2. The spacetime is split into spatial hypersurfaces labeled by a coordinate $t$. On each spatial slice, coordinates $x^\nu$ are specified (we use the convention that Latin indices take on the values 1, 2, or 3 of the space-like dimensions). A point labeled $x_0^\nu$ on one spatial slice and another with the same label on a different slice may be skewed with respect to the unit normal direction $n^\nu$ (which must be timelike). Here, the two points are shifted with respect to each other by a spatial vector $\beta^\nu$. If a particle moves from spatial slice $t_0$ to $t_0 + dt$ along the normal, it will experience a proper time interval of $\alpha dt$, where $\alpha$ is known as the lapse function. The lapse function and shift vector $\beta^\nu$ are freely specifiable as functions of both the spatial and time coordinates. The choice of lapse function determines how far neighboring spatial hypersurfaces are from each other.

On each $t = \text{const}$ hypersurface, there is an induced metric $\gamma_{ij}$. The induced metric is simply the spatial components of the spacetime metric $g_{ij} = g_{\nu\mu} n^\nu n^\mu$. On the other hand, the 4D metric can be specified by providing the lapse function, shift vector, and spatial metric. The 4-metric has the form

$$g_{\mu\nu} = \begin{pmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & \gamma_{11} & \gamma_{12} & \gamma_{13} \\ g_{20} & \gamma_{21} & \gamma_{22} & \gamma_{23} \\ g_{30} & \gamma_{31} & \gamma_{32} & \gamma_{33} \end{pmatrix},$$

where $g_{00} = -\alpha^2 + \gamma_{ij}\beta^i\beta^j$ and $\beta_i = \gamma_{ij}\beta^j$. While the surface normal has components $n^\mu = (1, -\beta^1, -\beta^2, -\beta^3)/\alpha$ and $n_\mu = -\alpha\nabla_\mu t$.

The 3-metric $\gamma_{ij}$ and its matrix inverse $\gamma^{ij}$ can be used to define a covariant derivative, Christoffel symbols, Riemann tensor, and Ricci tensor. The formulas for these are nearly identical to those presented in section 1.1, with the exception that indices only take on the values 1, 2, and 3. We will distinguish these 3-dimensional (3D) tensor (and tensor-like objects) from their 4D counterparts by either using different symbols (e.g. using the symbol $D_i$ to indicate the 3D covariant derivative), or by prepending a superscript 3 surrounded by parentheses (e.g. $(3)R_{ij}$).

Using the surface normal $n^\mu$, we can define a spatial tensor $K_{ij}$ known as the extrinsic curvature, by projecting the tensor

$$K_{ij} = g^{\mu\nu} \nabla_\mu n_\nu.$$
\[
(\nabla_{\nu}n_{\nu} + \nabla_{\nu}n_{\mu})/2 \text{ onto the slice. The resulting tensor is related to the time derivative of } \gamma_{ij} \text{ by } D_t \gamma_{ij} = -2\alpha K_{ij} + D_\nu \beta^\nu + D_j \beta_i,
\]
where \( D_j \) is the covariant derivative associated with \( \gamma_{ij} \).

Finally, with these choices, the ten Einstein equations become six evolution equations for \( \gamma_{ij} \) and four constraint equations. This is analogous to the way the Maxwell equations split into evolution equations for \( \vec{E} \) and \( \vec{B} \) and two constraint equations for \( \text{div}\vec{E} \) and \( \text{div}\vec{B} \).

The resulting field equations, usually known as the Arnowitt–Deser–Misner [34] (ADM) equations, but are actually a reformulation of the standard ADM equations by York [38], are given by (see, e.g. [35–37])

\[
\begin{align*}
\partial_t \gamma_{ij} &= -2\alpha K_{ij} + D_\nu \beta^\nu + D_j \beta_i, \\
\partial_t K_{ij} &= -D_j D_\nu \alpha + \alpha (R_{ij} - 2K_{ik}K^k_j)
- 8\pi \alpha (S_{ij} - \frac{1}{2} \gamma_{ij} (S - \rho))
+ \beta^j D_i K_j + K_{ik} D_j \beta^k + K_j D_i \beta^k,
\end{align*}
\]

where \((R_{ij}) \) and \((S_{ij}) \) are the Ricci curvature tensor and Ricci scalar associated with \( \gamma_{ij} \) and the source terms are given by

\[
\rho = n_\nu n_{\mu} T^{\mu \nu},
S' = -\gamma^\nu n^\mu T_{\mu \nu},
S_j = T_{ij},
S = \gamma^\nu S_{\nu j}.
\]

Equations (24) and (25) form the evolution equations, while equations (26) and (27) are the Hamiltonian and momentum constraint equations, respectively. In equations (24)–(31) Latin indices are raised and lowered with the spatial metric \( \gamma_{ij} \), i.e. \( \beta^j = \gamma^{ij} \beta^i \) and \( \beta^\nu = \gamma^\mu \beta_\mu \). The tensor \( \gamma^{ij} \) is the matrix inverse of \( \gamma_{ij} \). Greek indices are raised and lowered with the full metric \( g_{\mu \nu} \) and its inverse \( g^{\mu \nu} \).

2. Numerical relativity formalisms and techniques

There are two major classes of techniques used for numerical simulations of the Einstein equations. These are finite-difference methods, typically coupled to adaptive mesh refinement techniques, and pseudospectral methods. To understand how these methods work, we will consider some toy problems. Most of the finite difference codes are based on modifications to the ADM system (see section 2.3). These equations are in a form with mixed second and first derivatives. Basically, the system is such that only first time derivatives occur, but first and second spatial derivatives occur. (Of course, auxiliary evolution variables can be introduced so that the system only has first spatial derivatives, but at the cost of introducing additional constraints.) A good toy problem to illustrate how such equations are evolved is thus

\[
\partial_t \Pi - \beta^j \partial_j \Pi = -\nabla^2 \Phi,
\]

\[
\partial_t \Phi - \beta^j \partial_j \Phi = \Pi,
\]

where the spatial coordinates will be denoted by \( x, y, z \). We will explore numerical techniques for solving equation (32) in the next section.

2.1. Finite differencing

To solve equation (32), we consider a discrete grid labeled with 3 integer indices \((i, j, k)\), where the values of \( x, y, z \) at a point \((i, j, k)\) are given by \((x_0 + i dx, y_0 + j dy, z_0 + k dz)\). Furthermore, we denote the values of a function \( f(x, y, z) \) on this grid by \( f_{i,j,k} \). We then approximate spatial derivatives using these points. For example,

\[
\frac{\partial^2 f}{\partial x^2} = \frac{(f_{i+1,j,k} + f_{i-1,j,k} - 2f_{i,j,k})}{dx^2} + O(dx^2),
\]

is an approximation to \( \frac{\partial^2 f}{\partial x^2} \) using a three point stencil. By using more points in the stencil, this derivative can be made more accurate in the sense that the error will scale with higher powers of \( dx \). Modern numerical relativity codes tend to use sixth-order to eighth-order finite differencing [39–41]. With these approximations, equation (32) becomes

\[
\begin{align*}
\partial_t \Pi_{i,j,k} &= \sum_{l,m,n} (C_{i,j,k}^{l,m,n} \Pi_{l,m,n} + D_{i,j,k}^{l,m,n} \Phi_{l,m,n}), \\
\partial_t \Phi_{i,j,k} &= \sum_{l,m,n} (E_{i,j,k}^{l,m,n} \Pi_{l,m,n} + F_{i,j,k}^{l,m,n} \Phi_{l,m,n}),
\end{align*}
\]

where the coefficients \( C, D, E, F \) and the values of \((l, m, n)\) in the sums are determined by the finite difference stencil used. Thus, the partial differential equation (32) becomes a set of coupled ordinary differential equations for \( \Pi_{i,j,k} \) and \( \Phi_{i,j,k} \). These equations are then typically solved using standard Runge–Kutta techniques [42].

A major drawback of the above technique is that it is extremely computationally wasteful. For convenience here, we will assume that \( dx = dy = dz = h \). The smaller the grid spacing \( h \), the smaller the error, but the maximum value of \( h \) one could use and still have an acceptably accurate solution generally varies quite strongly over space (e.g. many points are needed to resolve black holes, but few needed far away). Furthermore, the regions where high resolution (i.e. small values of \( h \)) are needed tend to be quite small. Thus, if one used a uniform grid capable of resolving the entire space, essentially all the calculation time would be spent evolving the overresolved regions. This problem can be ameliorated to some extent by choosing special coordinates [43–45] that concentrate gridpoints in certain regions. The state-of-the-art technique for overcoming this inefficiency is the use of adaptive meshes [46–50]. In an adaptive mesh code, a coarse grid covers the entire computational domain, with increasingly finer grids placed in location where high resolution is required (see figure 3).

Evolutions also involve a discretization in time, and the smaller the timestep \( \delta t \), the more steps are needed to cover a given time interval, and the more expensive the simulation. Explicit time integration methods are subject to the
Courant–Friedrichs–Lewy stability condition [51], which limits $dt$ on a mesh to be less than around $h/v_s$, where $v_s$ is the maximum signal speed, which for spacetime evolution is the speed of light. The effect on $dt$ is a price to be paid for smaller $h$.

### 2.2. Pseudospectral methods

The other major techniques used in black-hole simulations fall under the category of pseudospectral methods [52–56]. In spectral methods the evolved fields are expressed in terms of a finite sum of basis functions. An example of this would be to describe a field on a sphere in terms of an expansion in spherical harmonics. These methods have the advantage that if the fields are smooth then the error in truncating the expansion converges to zero exponentially with the number of basis functions used in the expansion. In pseudospectral methods, values of functions are stored at special gridpoints, the colocation points, corresponding to Gaussian quadrature points of the basis functions [57]. Codes can then transform between spectral and colocation-point representations via Gaussian quadrature. This is especially useful for computing products and other pointwise operations which are much simpler using gridpoints. In pseudospectral form, spectral methods can be thought of as a particular limit of finite differencing, the limit that uses the entire domain as its stencil so as to make the highest-order derivative operator [58]. The order of this operator will then increase with the number of colocation points, giving the method faster convergence than a fixed polynomial order. The payoff is that differentiations and interpolations become more expensive much more quickly than fixed-stencil finite difference methods as resolution is increased. Also, exponential convergence is lost for functions that are only smooth to finite order. Perhaps the best known numerical relativity code that uses pseudospectral techniques is the Spectral Einstein Code [52–54], or SpEC, used by the SXS (simulating extreme spacetimes) collaboration.

The choice of finite difference versus spectral methods affects which method of handling black holes is easier to implement. A black hole interior presents a major challenge to any numerical technique because of the curvature singularity it harbors (see, e.g. Wald [59] for a discussion on the inevitability of forming curvature singularities). Fortunately, the singularity is concealed behind an event horizon. The region inside the horizon cannot affect the external solution, so numerical simulations need not evolve it accurately. They only need to keep it from causing the simulation to crash. One way to do this is to simply not evolve a region inside the horizon, i.e. to excise this region. Spectral methods can do this naturally, because even near the inner edge of the grid, no points are needed from the other side of the excision boundary to take derivatives. A boundary condition physically should not be needed (because no information flows out of a region where all characteristic speeds go inward), and none is required. The other method, the puncture method, described in detail below, involves allowing singularities in the computational domain. In the appropriate gauge, these singularities are sufficiently benign that finite difference methods can handle them. It would be more difficult to evolve a puncture stably with a spectral code [56].

### 2.3. Making the problem well posed

As will be discussed later, the ADM equations by themselves proved to be unstable for many strong-field problems, including black-hole mergers. With all the difficulties encountered trying to implement the ADM equations in the 1990s, emphasis changed to developing new 3 + 1 systems and analyzing their well-posedness [60–62, 62–73]. Informally speaking, a hyperbolic system of equations is well posed if the solution depends continuously on the initial and boundary data. Ill-posed systems can have solutions that grow without bound even for very small evolution times.

To understand how reformulation of the basic evolution equations can make or destroy well-posedness, consider a simple vector wave equation

\[
\begin{align*}
\frac{\partial \vec{E}}{\partial t} &= \vec{\nabla} \times \vec{B}, \\
\frac{\partial \vec{B}}{\partial t} &= -\vec{\nabla} \times \vec{E},
\end{align*}
\]  

subject to

\[
\begin{align*}
C_E &= \vec{\nabla} \cdot \vec{E} = 0, \\
C_B &= \vec{\nabla} \cdot \vec{B} = 0.
\end{align*}
\]

This system is well posed in the sense that the solution $(\vec{E}(t), \vec{B}(t))$ depends continuously in the initial data. Any constraint violation (failure of $C_E$ or $C_B$ to be zero) will be preserved (not grow or decrease) by the evolution system.

This system can be transformed into two separate identical second-order equations for $\vec{E}$ and $\vec{B}$ of the form

\[\text{Figure 3. Schematic of how mesh refinement works in one dimension. Shown are the values of a function at discrete points along the x axis. A coarse grid (black) covers the entire domain and progressively finer grids (blue and red) cover the parts of the domain where the function varies rapidly.}\]
The latter system is not quite equivalent to the original system in that constraint violations now grow linearly in time. A better system is obtained by noting that 
\[ \nabla \times \nabla \times \vec{A} = -\nabla^2 \vec{A} + \nabla (\nabla \cdot \vec{A}). \]
Using this, and the assumption that \( \nabla \cdot \vec{A} = 0 \), we get
\[ \Box \vec{A} = 0, \tag{39} \]
where
\[ \Box = \frac{\partial^2}{\partial t^2} - \nabla^2. \tag{40} \]
If we solve equation (39) with a small divergence, the norm of the divergence will remain bounded. More generally, if the constraint equations (37) are satisfied, then any solution to equation (39) is also a solution to
\[ \Box \vec{A} + \kappa \nabla (\nabla \cdot \vec{A}). \tag{41} \]
If \( \kappa \) is chosen larger than one, small violations of the \( \nabla \cdot \vec{A} = 0 \) constraint can blow up arbitrarily quickly. Thus, with seemingly inconsequential changes, we can turn a system from one with a minor blowup in the constraints to either one with a catastrophic blowup in the constraints, or no blowup at all. However, with the addition of an auxiliary field, we can do even better. Consider the system [74]
\begin{align*}
\frac{\partial E}{\partial t} &= \nabla \times \vec{B} + \nabla \psi_E, \\
\frac{\partial \vec{B}}{\partial t} &= -\nabla \times \vec{E} + \nabla \psi_B, \\
\frac{\partial \psi_E}{\partial t} &= -\nabla \cdot \vec{E} - \psi_E, \\
\frac{\partial \psi_B}{\partial t} &= -\nabla \cdot \vec{B} - \psi_B. \tag{42}
\end{align*}
For this system, the constraints satisfy
\[ \Box C_E - \left( \frac{\partial C_E}{\partial t} \right) = 0, \]
\[ \Box C_B - \left( \frac{\partial C_B}{\partial t} \right) = 0. \tag{43} \]
Solutions to equations (43) decay exponentially in time. Thus if numerical effects introduce constraint violations, these will be damped away. The main message here is that even for a physically motivated evolution system, the addition or removal of terms nominally equal to zero can make the difference between a solvable and an insolvable system. Furthermore, for a constrained system, like the equations of electromagnetism and general relativity, enlarging the system of equations by adding new fields can suppress unphysical constraint violations.

The standard ADM formulation is now known to be ill-posed in the nonlinear regime [68, 69]. On the other hand, many well-posed formulations have been proposed but turned out not to be an immediate panacea for unstable black hole simulations. Some of the most influential of these formulations are the Bona-Masso family [75–78], the NOKBSSN family [79–81], Z4 family [82–89], the Kidder–Scheel–Teukolsky family [90, 91], and the generalization-harmonic family [73, 92, 93].

One of the key improvements in numerical simulations prior to the breakthroughs of 2005 was the introduction of the so-called NOKBSSN formulation of the Einstein equations in 3 + 1. This system, which is named after its developers Nakamura, Oohara, Kojima, Shibata, Baumgarte, and Shapiro [79–81], modifies the standard ADM equations in several crucial ways. Firstly, the spatial metric \( \gamma_i \) is split into an overall conformal factor \( e^\phi \) and a conformal metric \( \tilde{\gamma}_{ij} \), where \( e^{2\phi} \tilde{\gamma}_{ij} = \gamma_{ij} \) and the determinant of \( \tilde{\gamma}_{ij} \) is unity. This conformal metric has its corresponding Christoffel symbols \( (3)^{\Gamma}_{ij} \) and Ricci tensor \( (3)^R_{ij} \). Second, the three combinations \( \gamma^i_{\flat} (3)^{\Gamma}_{ij} = (3)^{R}_{ij} (k = 1, 2, 3) \), as well as the \( K = \gamma^i_{\flat} K_{ij} \) are promoted to evolved variables. Third, the remaining evolved extrinsic curvature variables are trace-free conformal extrinsic curvature variables \( \tilde{A}_E = e^{-4\phi} [K_{ij} - (1/3)K\tilde{g}^{ij}] \). Finally, the momentum constraint equations are used to modify the evolution equations for \( (3)^{\Gamma}_{ik} \), which introduces a constraint damping quality to the system. These changes, in conjunction with a particular choice of gauge conditions, namely the use of certain Bona-Masso [75] type lapse conditions (known as 1+log slicing) and \( \Gamma \)-driver shift conditions [94] led to the first genuinely stable, fully nonlinear implementations of the Einstein equations for systems without symmetries, at least for non-black-hole spacetimes. The factoring out of the conformal factor \( \phi \) proved to be particularly advantageous for collisions of black holes [95, 96]. However, the state of the art for black-hole evolutions in the early 2000s only allowed for head-on collisions and grazing collisions [95, 96], where the black holes merge well before completing one orbit. The NOKBSSN system also proved to be stable using higher-order finite-differencing methods [44] for head-on collisions, as well.

A key technique used in many of these early evolutions was the fixed-puncture formalism [97, 98]. In this formalism, a two-sheeted Einstein–Rosen bridge associated with a single black hole is mapped into a single sheet with a singularity at the center. As shown in figure 4, the standard Schwarzschild spacetime can be recast as a puncture by taking a spatial slice that passes through the bifurcation sphere\(^{10}\). On either side of the sphere, the slice extends infinitely far. Next we introduce a coordinate \( \mathcal{R} \), which is related to the usual Schwarzschild coordinate \( r \) by
\[ r = \mathcal{R} \left( 1 + \frac{M}{2\mathcal{R}} \right)^2. \tag{44} \]
The spatial metric then takes in the form.

\(^{10}\)In the extended Schwarzschild spacetime, also called the Kruskal extension, there is both a black hole and a white hole. The bifurcation sphere is the point where the two horizons meet (see figure 4).
Figure 4. (Left) A spacetime diagram of a Schwarzschild black hole with the \((\theta, \phi)\) angular coordinates suppressed. Each point represents a sphere of radius \(4\pi r^2\). Note that there are two curves corresponding to each value of \(r\). Radially ingoing and outgoing light rays travel along \(45^\circ\) lines in this diagram. The event horizon(s) correspond to the \(r = 2M\) diagonal curves. The dotted line represents a spatial slice with a two-sheeted topology. (Right) The spatial surface corresponding to the dotted line shown with one spatial dimension \((z)\) suppressed. Points on the dotted line correspond to circles here. The horizons correspond to the thick circle and the \(r = 5M\) and \(r = 10M\) curves each map to one circle inside the horizon and one outside. The central point maps to the \(r = \infty\) of Region III.

\[
\text{ds}^2 = \psi^4 \left( d\mathcal{R}^2 + \mathcal{R}^2 d\Omega^2 \right),
\]

where \(\psi = 1 + M/2\mathcal{R}\). Here \(\mathcal{R} = 0\) and \(\mathcal{R} = \infty\) both correspond to \(r = \infty\) and the metric is singular at \(\mathcal{R} = 0\). This singularity is not the curvature blow-up singularity at the center of the black hole, that singularity is in the future of this slice, rather the singularity is entirely gauge and results of the black hole, that singularity is in the future of the central point.

2.4. Breakthroughs in numerical relativity

Perhaps one of the most significant breakthroughs in numerical relativity occurred when Pretorius, who had previously developed an adaptive-mesh-refinement (AMR) code based on the generalized harmonic system [93], included constraint damping techniques [60] developed for the \(Z4\) system by Gundlach et al [88]. The key development there was that the general relativistic action can be extended to include terms that vanish when the constraints are satisfied, but act to damp these constraint violations when they are nonzero. When applied to the generalized harmonic system used by Pretorius, the constraint violations for a black hole–black hole binary remained bounded and Pretorius was thus able to perform the first successful evolution of an orbiting binary [1]. Pretorius presented his initial results at the Banff International Research Station Workshop on Numerical Relativity in April 200511.

The actual system Pretorius evolved was the Einstein equations coupled to a scalar field. The initial data consisted of two scalar boosted stars with supercritical densities. The stars collapsed into two black holes that then orbited and merged. Figure 5 is a reproduction of figure 3 in of Pretorius’ paper [1]. It shows the very first merger waveform from an orbiting black hole–black hole binary ever published.

Just four months after Pretorius submitted his groundbreaking paper, a new breakthrough was announced that became known as the moving punctures approach [2, 3]. This new method was developed independently by the groups then at the University of Texas at Brownsville (UTB) and NASA

11 http://bh0.physics.ubc.ca/BIRS05/
Goddard Space Flight Center (GSFC) and first demonstrated publicly in the Numerical Relativity 2005: Compact Binaries workshop at NASA GSFC\cite{12}. Notably, the moving punctures approach allowed groups worldwide to evolve black hole–black hole binaries. It is based on an extension of the standard NOKBSSN system, with several important changes. (1) The singular conformal factor is replaced by a non-singular function $\chi = e^{-4\phi}$ (although evolutions with $\phi$ itself are also used \cite{3}) which is evolved fully numerically. (2) The gauge conditions explicitly allow the punctures to move. Previously, the shift condition was chosen so that the punctures could not move. (3) The standard $1 + \log$ lapse condition

$$\partial_t \alpha = -2\alpha K,$$

would require $K$ to be singular at the puncture in order for the lapse to change from a zero value when the puncture is on a given point to a non-zero value when the puncture has passed. This singular behavior is removed by changing the lapse condition to an advection equation

$$\partial_t \alpha - \vec{\nabla} \cdot \alpha = -2\alpha K.$$  

Finally, within the evolution equations for $\tilde{\Gamma}^i$ there is a singular term on the puncture location proportional to the lapse. By choosing an initial lapse that is identically zero on the puncture, this singularity is also removed. With these changes, the gauge naturally evolves such that the black holes orbit each other and inspiral in coordinate space. The pathologies seen with the fixed-puncture approach vanished with these new dynamic punctures. Furthermore, because the moving punctures approach was similar to existing codes, groups around the world were able to rapidly develop their own versions.

These include the BAM code \cite{49} developed at Jena, the Maya–Krane code developed at Penn State \cite{107}, as well as the original codes, LazEv developed at The University of Texas at Brownsville, and Hahndol, developed at GSFC. More recently, the publicly available EinsteinToolkit \cite{41, 108} code included an open-source implementation, known as McLachlan. LazEv, Maya–Krane, and McLachlan all used the Cactus Computational Toolkit \cite{109}, originally developed at the Albert Einstein Institute in Golm, Germany.

There were several significant differences between Pretorius’ techniques and the moving punctures approach. Unlike in the moving punctures approach, the system Pretorius developed used excision to handle the black hole singularities, compactified the computational domain to include spatial infinity, and of course, used the generalized harmonic system with constraint damping. This system was sufficiently unlike the other techniques used by numerical relativists at the time that it was only slowly adopted.

Figure 6 shows reproductions from the breakthrough papers using the moving punctures approach. The waveforms, with various analyses, and the horizons are shown.

Soon after the announcement of the moving punctures breakthrough, simulations were reported where the binary completed more than one full orbit \cite{110} and then multiple orbits \cite{111}. The latter, in particular, compared the merger waveforms from simulations starting at various separations and found that the merger waveform was insensitive to initial conditions. This was followed shortly afterwards by the discovery of the orbital hangup effect \cite{112} for spinning binaries. This effect either delays or accelerates the merger depending on whether the spins of the two black holes are (partially) aligned or counteraligned with the orbital angular momentum, and proved to be important for parameters estimation of LIGO sources \cite{5}. Some of the important discoveries that proceeded from these simulations will be described below in section 3.

2.5. Methods for evolving the fluid equations

2.5.1. Conservative formulation. Many of the early numerical relativity hydrodynamic simulations used a formulation introduced by Wilson in 1972 \cite{113}. In Wilson’s scheme, the evolution variables are

$$\rho_{\ast} = W\sqrt{\gamma}\rho_0, E = \rho_{\ast}c, S_i = \rho_{\ast}h u_i.$$  

a variable for rest-mass density, internal energy density, and momentum density, respectively. Often, there are numerical advantages to evolving entropy rather than internal energy, so some codes (e.g. \cite{114, 115}) specializing to Gamma-law EoS evolved a variable $e_{\ast} = W\sqrt{\gamma}(\rho_0 e)^{\gamma}$. The equations can be finite differenced in a conservative form: fluxes are calculated at cell interfaces; the same flux added to one grid cell is removed from its neighbor, and no truncation error accrues to the total rest mass. However, the variables evolved (in particular $E$) are not those that are physically conserved, so an explicit artificial viscosity must be added to correctly account for shocks.

Shock handling is accommodated more naturally if one evolves the physically conserved variables, meaning that one

---

Figure 5. A reproduction of the waveform shown in figure 3 of [1] courtesy of the author. This was from the first fully nonlinear numerical simulation of the last orbit, merger, and ringdown of a black hole–black hole binary. It was generated by Pretorius using the generalized harmonic coordinate approach. The curves show the waveforms as calculated at various radii and translated in time. Reprinted figure with permission from [1], Copyright (2005) by the American Physical Society.
should evolve the total energy density rather than internal energy density. The resulting equations can be solved using established high-resolution shock capturing methods. This is the path followed in what has come to be called the Valencia formulation [23, 116, 117], which all current numerical relativity hydrodynamics codes essentially follow. The evolution equations take conservative form

$$\partial_t U + \nabla \cdot F = S$$

(49)

where the conservative variables $U$ are

$$U = (\rho^*, W \sqrt{\frac{\rho}{\rho^*}}, X, \rho^*, \tau)$$

(50)

$$\tau = \sqrt{\gamma a^2 T^{00}} - \rho^*, S_i = \sqrt{\gamma a T_i^0}.$$

(51)

Conservative shock-capturing hydrodynamics codes in numerical relativity have achieved at best 3rd-order convergence [118].

After computing $U$ at a new timestep, it remains to recover the original (‘primitive’) variables such as $\rho_0$ and $u_i$. It turns out to be sufficient to recover $W$ and $T$, but this will involve some sort of root-finding process.

2.5.2. Excision and punctures in the presence of matter. If the black hole interior is removed via excision, matter must be able to flow into the black hole and ‘disappear’ in a stable way. For problems involving collapse to a black hole, one must maintain accuracy inside the collapsing object until an apparent horizon is located, after which a region inside this horizon can be excised. Scheel et al [119] did this in 1D for spherical collisionless matter. Next, Brandt et al [120] introduced a code for evolving nonvacuum black hole spacetimes using an isometry inner boundary condition at the apparent horizon. Techniques for matter excision using horizon penetrating coordinates were introduced roughly simultaneously by Duez et al [121] and Baiotti et al [122]. Optimal methods for one-sided differencing of the fluid equations near the excision boundary are investigated by Hawke et al [123], although these initial simulations were less sensitive to this than to the
gauge choices needed to keep the coordinates horizon-penetrating. Excision is still the method used for simulations in the generalized harmonic formulation (e.g. SpEC). A particularly elegant grid structure for hydrodynamic excision is provided by the cubed-sphere arrangement [124].

In NOKBSSN, it is much simpler to use moving puncture gauges and avoid explicit excision. One might worry that material inflow into the puncture would cause numerical problems, but fortunately this turns out not to be the case. Numerical experiments showed that puncture simulations can handle stellar collapse to a black hole [125] and spherical accretion into a black hole [126] with no code changes except a small extra dissipation in the metric evolution (in [125]) and a means of resetting fluid variables near the puncture where conservative to primitive variable recovery fails (in [126]). Shortly after this realization, Shibata and Uryu carried out the first NOKBSSN black hole–neutron star merger simulations [127].

2.5.3. More physics: equations of state, neutrinos, magnetic fields. Information about the properties of the matter enters through the equation of state. An extremely simple but nevertheless useful equation of state is the polytropic law

\[ P = \kappa \rho_0^\Gamma \left(1 - \frac{\Gamma}{2}\right) \rho_0^\epsilon, \]

where \( \kappa \) and \( \Gamma \) are constants. Higher \( \Gamma \) means stiffer EoS. A notable feature of this EoS is that it is barotropic; there is no explicit temperature dependence, meaning the matter must be degenerate or the temperature itself be a function of density (as, for example, in an isentropic gas). In many cases, we may wish to allow an initially polytropic gas to pick up added thermal pressure and internal energy via shock heating. In this case, one uses the more general Gamma-law EoS

\[ P = \left(1 - \frac{\Gamma}{2}\right) \rho_0^\epsilon, \]

where \( \epsilon \) is now given not by the polytropic law but by the energy density evolution equation. For adiabatic evolution, the polytropic law should be maintained. One gets a surprising amount of mileage out of this simple EoS family. Nonrelativistic ideal degenerate Fermi gases have \( \Gamma = 5/3 \); relativistic ideal degenerate Fermi gases have \( \Gamma = 4/3 \); stars with both radiation and gas pressure with a constant fraction of the total from each have \( \Gamma = 4/3 \). Neutron stars are not polytropes, but much of the early numerical relativity work involving neutron stars modeled them as \( \Gamma = 2 \) polytropes.

Ultimately, an accurate treatment of dense matter is needed. For a general astrophysical gas, there may be many composition variables \( X_i \), each in need of its own evolution equation. However, when dealing with high densities and temperatures above \( \sim \)MeV, the matter can be assumed to be in a nuclear statistical equilibrium, in which case there is only one composition variable, the proton fraction or electron lepton number fraction \( Y_e = n_p/(n_p + n_e) \). This variable evolves due to charged-current weak nuclear interactions, which do not always have time to equilibrate. Thus, for our equation of state, we are left with functions of three variables, e.g. \( P(\rho, T, Y_e) \).

Unfortunately, they are unknown functions for densities much above nuclear saturation, so numerical relativity simulations of neutron stars must explore the range of equations of state consistent with known nuclear and astrophysical constraints. For the problem most relevant for gravitational waves, compact binary inspirals, the situation simplifies. The nuclear matter is very degenerate and in beta equilibrium, so the EoS is effectively one-dimensional (1D): \( P = P(\rho_0) \).

These 1D EoS are conveniently parameterized as piecewise-polytropes, for which the density is divided into intervals, and each interval has its own polytrope law. For example, in the \( i \)th interval, covering the density range \( \rho_{i-1} < \rho < \rho_i \), the pressure is \( P = \kappa \rho^{\Gamma_i} \). The polytropic indices \( \Gamma_i \) and transition densities \( \rho_i \) are free parameters. \( \Gamma_0 \) and \( \kappa_0 \) covers the low-density range where the pressure, dominated by relativistic electrons, is known. The other \( \kappa_i \) are set by requiring \( P \) to be continuous at \( \rho_i \). Fortunately, only a few free parameters needed to adequately cover the range of plausible EoS [128].

At the end of inspiral, tidal disruption breaks beta equilibrium, as the matter decompresses faster than charged weak interactions can adjust \( Y_e \). Also, the occurrence of shocks heats the matter so that it is no longer degenerate. A number of studies add a Gamma-law thermal piece to the pressure to allow shocks to heat the matter. Any cold EoS can by thus augmented as follows:

\[ \epsilon = \epsilon_{\text{cold}}(\rho_0) + \epsilon_{\text{th}} \]

\[ P = P_{\text{cold}}(\rho_0) + (\Gamma_{\text{th}} - 1)\rho_0^\epsilon_{\text{th}} \]

where now \( \epsilon_{\text{th}} \) comes from the energy density evolution. The thermal Gamma law may not capture important parts of the true 3D EoS. This has been tested in the context of neutron star–neutron star binary mergers by Bauswein et al [129]. They find that the thermal Gamma law approximation can alter the post-merger gravitational wave frequency by 2%–8%, post-merger torus mass by 30%, and delay time to collapse to a black hole by up to a factor of 2. In addition, only 3D EoS provide the physical temperature information needed for neutrino calculations.

The evolution of the lepton number and \( Y_e \) are given by weak nuclear processes such as electron and positron capture, which emit neutrinos that travel some distance, as well as the reverse absorption processes. Also, neutrino cooling is the dominant source of cooling in most simulations with neutron stars, and neutrino absorption above the neutrinosphere can be an important driver of winds. Newtonian simulations, especially in the supernova context, have long concerned themselves with these effects, and around 2010 they began to be incorporated into numerical relativity simulations. At first, neutrino emission effects were approximated by local sink terms for the energy and lepton number (‘neutrino leakage’) [130–134]. Effective emission rates differ in optically thick and optically thin regions; the neutrino optical depth can be computed by an inexpensive iterative procedure [134, 135]. The current state-of-the-art for numerical relativity is neutrino transport in an energy-integrated moment closure approximation [136–139], which is impressive progress in so short a time but still far from a full solution to the 6D Boltzmann equation.

A final major piece of realistic matter numerical relativity simulations is the electromagnetic field evolution. Neutron star interiors have plenty of free charges and very high electrical conductivity, so the magnetohydrodynamic (MHD) approximation is valid in most regions. Thus, we must add the Maxwell stress tensor for the electromagnetic field \( T_{\text{EM}}^{\mu\nu} \) to the total...
stress tensor \(T_{\mu\nu}\), and so magnetic terms appear in \(\tau\) and \(S_i\). The evolution of the magnetic field \(B^i\) is given by the induction equation—in words, that magnetic field lines are attached to (‘frozen into’) fluid elements. The electric field is set by the MHD condition that the electric field in the conducting fluid is constrained to vanish \(\nabla \cdot B = 0\). At all times, the magnetic field should satisfy the constraint \(\nabla \cdot E = 0\). In practice, magnetic monopoles are avoided by constrained transport (staggering magnetic and electric variables so that the change of \(\nabla \cdot B\) exactly vanishes [140]), by evolving a vector potential \([141]\) (which, with appropriate staggering, is equivalent to constrained transport \([142, 143]\)), or by divergence cleaning (extending Maxwell’s equations so that monopoles damp and propagate off the grid \([74, 144, 145]\)).

An interesting limit of the MHD equations occurs in magnetospheres, where the Maxwell piece of \(T_{\mu\nu}\) dominates, so that \(\tau\) and \(S_i\) become essentially the electromagnetic energy density and Poynting flux, respectively. Magnetospheres differ from vacuum electromagnetism because enough free charges remain to prevent electric potential differences along field lines \((E \cdot B = 0)\). These conditions are expected to obtain in the region around neutron stars and the polar jet region around accreting black holes. Specialized codes have been developed to evolve the relativistic force-free equations, evolving either the electric and magnetic fields \([146–148]\) or the magnetic field and Poynting flux \([149]\). Lehner et al. \([150]\) introduce a scheme for evolving the full MHD equations in high-density regions and the force-free equations in low-density regions. This scheme was successfully used to study the collapse of magnetized neutron stars, but for future applications a single set of equations able to handle both fluid and field-dominated regimes was desirable. This was done by Palenzuela \([151]\) in the context of a resistive MHD code. Resistivity and the force-free limit might sound like different issues, but in fact the inhibition of flow by charged particles across field lines in a magnetosphere can be modeled as an anisotropic resistivity \([146]\), so by allowing sufficiently general Ohm’s laws, Palenzuela’s code can both handle resistivity inside stars and impose the force-free limit outside. Also, Paschalidis et al. \([152]\) have adjusted their MHD code to extend to the force-free limit, showing that MHD and force-free \((B^i, S_i)\) evolution just differ in the primitive variable recovery. These codes have been used to study magnetosphere interaction in the late inspiral of neutron star–neutron star \([153–155]\) and black hole–neutron star binaries \([156]\), which has been suggested as a mechanism to create precursor signals to short duration gamma ray bursts.

3. **Black hole–black hole binary simulations**

There have been several recent reviews of the history of numerical relativity \([35, 36, 157–160]\). Here we will briefly cover some of the major highlights.

3.1. **Early efforts**

Attempts at numerical simulations of black hole–black hole binaries date back to the 1960s with the pioneering work of Hahn and Lindquist \([161]\), who were able to simulate two initially stationary black holes for a short time. Later, with faster computers and improved algorithms, Smarr et al. \([162–164]\) able to simulate head-on collisions through merger. It was not until 1993 that computers were powerful enough to calculate accurate waveforms from such mergers \([165]\).

In the 1990s the National Science Foundation of the United States supported a large collaboration, the Binary Black Hole Grand Challenge Alliance, with the goal of advancing numerical relativity to the point where evolutions of orbiting black holes became feasible. There were several important developments enabled by the grand challenge. These include the full 3D evolutions of boosted single black holes using excision \([166]\), perturbative techniques to extract gravitational waveforms from numerical simulations \([167]\), stable evolutions of single black-hole spacetimes using characteristic techniques \([168]\), and the development of toolsets for parallel simulations. The alliance \([169]\), and independently Brügmann \([170]\), were able to evolve grazing collisions of black hole–black hole binaries in 3D. However, these simulations crashed after a short evolution time \((\lesssim 50 M)\).

As mentioned above, the Lazarus approach \([43, 101–105]\) could be used to extend these simulations and generate waveforms from closely separated binaries.

These grazing collisions were performed using the ADM system of equations evolved as a standard Cauchy problem. Using characteristic evolution techniques, the PITT Null code \([168, 171, 172]\) (developed at the University of Pittsburgh) could evolve highly distorted spacetimes for arbitrary lengths of time. For these long evolutions, the PITT code used a coordinate system based on outgoing null (i.e. lightlike) geodesics. In a two black-hole spacetime, these geodesics would form caustics in the vicinity of the two black holes, making the associated coordinate system singular. Thus the PITT code could not evolve the interior of a black hole–black hole binary. However, this code has since proven to be very useful for evolving the exterior region in which gravitational waves propagate away from the central system. An interior Cauchy evolution an exterior characteristic evolution can be combined to produce highly-accurate waveforms in a technique known as Cauchy-characteristic extraction \([172–178]\).

3.2. **Post-breakthrough results**

With the breakthroughs of 2005, there was rapid progress in our understanding of the physics of black hole–black hole binaries.

3.2.1. **Recoils.** One of the most remarkable results that came from these simulations is that the merger remnant can recoil at thousands of kilometers per second. Determining just how fast the remnant can recoil took several years and required many hundreds of individual simulations. Perhaps the most straightforward way to conceptualize why the emitted power due to an inspiral can have (instantaneously) a preferred direction is to consider the case of unequal-mass black holes. The asymmetry of the system leads to a small excess of radiation along the direction of the linear momentum of the smaller black hole. For perfectly circular
orbits, this effect would average out to zero over an orbit, but since the binary will also be inspiraling, the cancellation will not be exact and a net recoil will be generated. The net recoil only becomes significant during the fast plunge phase.

Initial measurements of the recoil concentrated on analyzing individual configurations [179], or several configurations, but at very close separations [180]. The study of recoils started in earnest with Gonzalez et al [181]. Theirs was the first to do what was previously unheard of, a large number of relatively long-term accurate simulations. In the case of [181], they performed over 30 individual simulations and determined that the recoil very nearly obeys the simple formula $V = 16A\eta^2\sqrt{(1 - 4\eta)(1 + B\eta)}$, where $A = 750 \text{ km s}^{-1}$ and $B = -0.93$, $\eta = q/(1 + q)^2$ is the symmetric mass ratio, and $q = m_1/m_2$ is the usual mass ratio. Based on the formula provided by Gonzalez et al, the maximum recoil generated by unequal mass binaries is $175 \pm 11 \text{ km s}^{-1}$. As we will see, this contribution to the recoil can be vastly swamped by contributions due to the spins of the black holes themselves. Figure 7 shows the results of their study.

Soon after, other groups showed that the maximum recoil for spinning binaries, where the spins are aligned and antialiigned with the angular momentum, is much larger. In [182] and [183], it was shown that the maximum recoil for an equal mass, spinning binary with one black hole spin aligned with the orbital angular momentum and other antialiigned is $\sim 475 \text{ km s}^{-1}$. A still larger recoil of $V_{\text{max}} \sim 525 \text{ km s}^{-1}$ for a mass ratio of $q \approx 0.62$ was found in [185] when they extended the analysis of aligned/counteraligned spin binaries to unequal masses.

The recoils induced by unequal masses and aligned/antialiigned spins is always in the orbital plane of the binary (which, by symmetry, does not precess). [186] performed a set of simulations that showed that the out-of-plane recoil, which is induced by spins lying in the orbital plane, can be much larger. These superkicks [186–190] were found to be up to $4000 \text{ km s}^{-1}$ when the spins were exactly in the orbital plane. The superkick configuration is quite interesting, not only because of the large recoil, but also because of the direction of the recoil. Figure 8 shows the basic setup. The spins are antialigned with each other and in the orbital plane. Such a system will not precess and the orbital angular momentum will always point in the $z$ direction. Furthermore, the system has $\pi$-rotation symmetry about the $z$-axis. This means the recoil cannot lie in the orbital plane. What actually happens in this case is the binary bobs up and down long the orbital axis at ever increasing speeds until it merges. This bobbing is controlled by the orientation of the spins, as shown in figure 9. The net effect is quite unexpected. The magnitude and direction of the recoil depends sinusoidally on the polar orientation of the spins (see figure 10). Originally, it was thought that these in-plane spins maximized the recoil, however, it was later found out in [191, 193, 194] that, due to the hangup and other nonlinear-in-spin effects [112], having partially misaligned spins actually leads to a substantially larger recoil (up to $5000 \text{ km s}^{-1}$). The basic setup of this hangup-kick configuration is very similar to the superkick, with the exception that the out-of-plane components of the spins are aligned. For small spins, the recoil depends sinusoidally on the polar orientation (i.e. $V \propto \sin \theta$). However, for larger spins, the recoil is substantially larger for smaller angles, as shown in figure 11.

Of critical importance for modeling the superkick is the dependence on mass ratio. It is perhaps surprising that even though the central supermassive black holes in galaxies with a bulge can range in mass from under a million to tens of billions of solar masses, roughly $\sim 93\%$ [195–197] of galactic mergers are expected to produce supermassive black hole mergers with mass ratios in the range $1/10 < q < 1$. If the fall-off of the recoil with mass ratio is steep, then even a 10:1 binary may have a negligible recoil. Based on post-Newtonian theory, the group formerly at UTB, now at the Rochester Institute of Technology (RIT), argued that this dependence should vary as $q^2$ [186, 198]. This was first put to the test in [199], where the authors found the recoils fall faster than $q^2$, at least in certain symmetric configurations. A follow-up series of papers by the RIT group [200, 201] modeled the spin and mass ratio dependence for more generic configurations and found a leading $q^2$ dependence on the recoil for these more generic configurations.

Modeling the recoil for generic configurations is complicated by the need to perform hundreds of simulations. Even for a fixed mass ratio and zero eccentricity, the recoil depends on six spin degrees of freedom. In order to tackle this problem, the dimensionality of the free parameters needs to be reduced. The basic procedure developed by the RIT group in a series of papers [191, 200, 201] starts with a family of simulations related by a rotation of the azimuthal direction of the spins at the initial separation (the spins of both black holes are rotated the same amount). This leads to a 1-parameter family of configurations with (typically) a nearly sinusoidal dependence of the recoil. The amplitude of this sinusoidal dependence is then measured as a function of the polar spin orientations. This procedure works well when the azimuthal
The modeling of recoils from binaries with spins aligned and counteraligned with the orbital angular momentum began soon after the breakthroughs in numerical relativity. The first such simulations were performed in [182] and [183], with the first systematic studies of the recoil from such binaries in [207] and [208]. And the process of generating empirical models for the remnant masses, spins, and recoils from such mergers were first performed in [198, 209–212]. In addition, these types of binaries have been studied for their use in waveform modeling with examples in the SXS [213, 214], Georgia Tech [215, 216] and RIT [217, 218] catalogs. More recently, because of the apparent lack of observed highly-recoiling AGN, the RIT group began a systematic study of the spin and mass ratio dependence of the recoil for aligned/counter aligned binaries in [185, 219].

3.2.2. Modeling the remnant properties. One of the important tasks required in order to make the wealth of information from numerical simulations useful for astrophysics was to model the radiated energy-momentum and the corresponding final mass, spin, and recoil of the remnant black hole from black hole–black hole binary mergers in terms of the initial parameters of the binary. Developing these models required thousands of computationally expensive simulations.

In developing these models, two different techniques were initially used, but current models now combine aspects of both. The first technique used post-Newtonian theory [181, 183, 186, 188, 189, 198–200, 207, 210], or other physical arguments [212, 220, 221] to determine the functional form and free parameters of an approximate model for the relevant quantity, and the other used ad-hoc expansions [112, 209, 211, 222]. The work of [209, 211] pioneered the technique of using symmetry arguments to limit the degrees of freedom in the models. Their construction only assumed that the remnant can be described by the spin vectors of each black hole and the mass ratio. The model then must obey the following two symmetries. If \( F(\vec{x}_1, \vec{x}_2, q) \) is a formula for the remnant, mass, spin vector, or recoil, then \( F(\vec{x}_1, \vec{x}_2, q) = F(\vec{x}_2, \vec{x}_1, 1/q) \), i.e. the physical outcome of a merger cannot depend on the labels (1, 2) of the two black holes. Second, if \( F \) must transform appropriately under parity. One, however, need not use the variables \( (\vec{x}_1, \vec{x}_2, q) \). Inspired by post-Newtonian expressions, the RIT group has made extensive use of the variables \( (\Delta, \delta M) \) as well the
variables $\eta$ and $\vec{S}_0$ [185, 190–191, 200, 201, 219, 223–225]. These are defined as

$$\vec{S}_1 = m^2_1 \vec{\chi}_1,$$

$$\vec{S}_2 = m^2_2 \vec{\chi}_2,$$

$$\vec{S} = (\vec{S}_1 + \vec{S}_2)/m^2,$$

$$\Delta = (\vec{S}_2/m_2 - \vec{S}_1/m_1)/m,$$

$$\delta m = (m_1 - m_2)/m,$$

$$\vec{S}_0 = \vec{S} + (1/2)\delta m \vec{\Delta}.$$  

Any expansion in terms of one set of variables can be reexpressed in terms of another. However, since the goal is to model the remnant with accuracy, one wants to use variables that minimize the number of free parameters required to fit the known data.

State of the art models for remnant properties now combine results from simulations of many different groups (see [184, 185, 201, 219, 226–228]) with the goal of reducing systematic biases (which may arise from different groups concentrating on different regions of parameters space).

3.2.3. Numerical relativity at the extremes. State of the art numerical relativity codes now routinely evolve binaries with mass ratios as small as $q \lesssim 1/10$ [216, 229–233], moderately-to-highly precessing systems [186, 191, 200, 201, 234–242], and binaries with moderate spins. However, much smaller mass ratios, and spins close to 1 are still quite challenging. Prior to the work of [243] it was not even possible to construct initial data for binaries with spins larger than $\sim 0.93$ [244]. This limitation was due to the use of conformally flat initial data. Conformal flatness of the spatial metric is a convenient assumption because the Einstein constraint system take on particularly simple forms. Indeed, using the puncture approach, the momentum constraints can be solved exactly using the Bowen-York ansatz [245]. There were several attempts to generate data for highly-spinning black hole–black hole binaries, while still preserving conformal flatness [246, 247], but these introduced negligible improvements. Lovelace et al [243] were able to overcome these limitations by choosing the initial data to be a superposition of conformally Kerr black holes in the Kerr–Schild gauge. Using these new data, they were soon able to evolve binaries with spins as large 0.97 [248], and later spins as high as 0.994 [249].

While spins of 0.92 may seem reasonably close to 1, the scale is misleading. The amount of rotational energy in a black hole with spin 0.9 is only 52% of the maximum. Furthermore, particle limit and perturbative calculations show even more extreme differences between spins of 1 and spins only slightly smaller. For example, Yang et al [250] studied an analog to turbulence in black-hole perturbation theory. For spins close to 1, there is an inverse energy cascade from higher azimuthal ($m$) modes to lower ones for $\ell$ modes that obey $\epsilon = |1 - \chi| \lesssim \ell^{-2}$. This give hints that a more informative measure of the spin is actually $1/\epsilon$. Similarly, analysis of Kerr geodesic [251, 252] and particle-limit calculations of recoils [253, 254] indicate that the dynamics of nearly-extremal-spin black holes cannot

---

13 Initial data are said to be conformally flat if the spatial metric associated with the data is proportional to the flat space metric.
be elucidated with any degree of certainty using lower spin simulations.

Recently, the group at RIT also introduced their version of highly-spinning initial data, also based on the superposition of two Kerr black holes [255, 256], but this time in a puncture gauge. The main differences between the two approaches is how easily the latter can be incorporated into moving-punctures code. They compared their results to the SXS results for both spins of $\chi = 0.95$ [255] and $\chi = 0.99$ [257], and found very good agreement.

The other type of extreme simulation concerns small mass ratios. Because current numerical relativity codes use explicit algorithms to evolve the spacetime, the Courant–Friedrichs–Lewy condition, which determines how large a timestep can be relative to the spatial discretization, severely limits the run speed when one of the black holes is much smaller than the other. Basically, the number of timesteps required near the smaller black hole is set by the size of that black hole, not by its dynamics. This, coupled to the fact that the inspiral for a small-mass-ratio binary is much slower than for a similar mass one, means that such simulations are extraordinarily expensive. To date, the smallest quasi-circular inspiral evolved so far had $q = 1/100$ [231, 232].

One promising method to overcome these limitations is to use semi-implicit techniques [258, 259], but these have not yet been shown to work for small-mass-ratio binaries.

Finally, in [260] another extreme was explored: that of binaries at far separations. There the authors used fully non-linear numerical relativity to model several orbits of binaries separated at $D = 20M$, $D = 50M$, and $D = 100M$ and compared the orbital dynamics to post-Newtonian and Newtonian predictions. Very good agreement between post-Newtonian and numerical relativity predictions for the orbital frequency was found for the $D \geq 50M$ cases. The longest simulations to merger published to date was in [261], where a binary was evolved for 175 orbits. For reference, an equal-mass binary at a separation of $D = 100M$ will complete over 2000 orbits before merging and requires about $8.2 \times 10^5 M$ of evolution time. (If the mass of each black hole in the binary is $3M_\odot$, then the merger from $D = 100M$ would take about 40 min.)

### 3.2.4. Waveform modeling.

One of the major goals of numerical relativity is to produce accurate waveforms for gravitational wave data analysis. The actual process of obtaining the waveform from a numerical simulation can be involved. See [262] for a review of modern techniques for extracting the gravitational waveform from a numerical simulation.

In figure 13, we show an example waveform from a recent black hole–black hole binary simulation [239, 263]. In order to use physical units, we consider a binary that has a total mass of $20M_\odot$. The plot shows the waveform from the last 48 orbits for an equal-mass binary in a precessing configuration. There are two distinct phases of the waveform. The longest phase is due to the slow inspiral and eventual plunge. In the figure, this corresponds to the start of the waveform until about 1.92s. Small oscillations in the amplitude are apparent. These are due to precession of the orbital plane. The overall ramp up of both the amplitude and frequency is due to the inspiral and its associated increase in the orbital frequency. Following a brief transition between 1.915s and 1.925s, the waveform changes to a damped sinusoid. This phase is due to the rapid equilibration of the now single black hole.

Remarkably, on September 14, 2015 the twin LIGO observatories detected the gravitational waveform from the inspiral and merger of two black holes [4, 5]. The resulting waveform is shown in figure 14.

In this section we will review the history of fully non-linear numerical simulations of black hole mergers to generate and verify the waveforms. The current generation of numerical relativity codes calculate the gravitational waveform of a merger simulations by calculating the Regge–Wheler–Zerilli perturbations (which can be related to the strain $h$), the Bondi News function $N$, or Weyl scalar $\psi_4$ (as well as combinations of the above). With the exception of the calculation of $N$ using Cauchy-characteristic extraction [174, 175, 178, 264, 265], the waveform is calculated at a series of finite radii and extrapolated to $r = \infty$ along an outgoing null (lightlike) paths using some form of either polynomial extrapolation, or perturbative expansion [266]. In a suitable gauge, $\psi_4 = \bar{N} = \bar{h}$, where an overbar denotes complex conjugation and a dot represents a time derivative. Note that gravitational wave detectors measure $h$ directly, while the emitted power is directly related to $N$. The points at $r = \infty$ along outgoing null rays are collectively known as future null infinity. While these points are formally outside the spacetime, they are quite useful for defining gravitational radiation.

For an isolated source\footnote{More precisely, for an asymptotically flat spacetime.}, the Bondi News function $N$ is directly related to the radiated power-per unit solid angle in the gravitational radiation by...
\[ \left( \frac{dP}{d\Omega} \right) = \frac{1}{4\pi} (N\bar{N}). \] (60)

\( N \) itself is defined on future null infinity, and the only gauge freedom in \( N \) is associated with the supertranslation freedom \( N(\tau, x^A) \rightarrow N(\tau + \sigma(x^A), x^A + \tau\omega^A(x^A)) \), where \( \tau \) is an affine parameter (a generalization of proper time along light-like curves), \( x^A \) denotes angular coordinates, and \( \sigma(x^A) \) and \( \omega^A \) are constant functions of angle.

However, as standard Cauchy codes cannot include future null infinity (but see [267–269] for an approach which may allow evolutions that include future null infinity), calculations of \( N \) involve a matching procedure, where data from a Cauchy code is used as boundary data for a characteristic evolution. This matching requires the specification of unknown data from the edge of the Cauchy domain to null infinity. This induces spurious radiation [178], which can be controlled by moving the matching procedure to farther radii.

In Chu et al [233] the authors compared their extrapolations of the Regge–Wheeler–Zerilli perturbations to the News calculation. They found that errors due to both gauge effects and extraction at finite radii lead to mismatches of \( a \sim 5 \times 10^{-4} \). It is interesting to note that this mismatch may be geometrical in nature and arising from the difficulty in defining gravitational radiation at a finite distance from the source.

The main technique for extracting radiation involves the calculation of the Weyl scalar \( \psi_4 \). Calculations of \( \psi_4 \) have the advantage that there is a simple, well defined procedure for calculating the Newman–Penrose scalars \( \psi_4 \) in a class of tetrads where \( \psi_4 \) represents the outgoing radiation (and \( \psi_0 \) the ingoing radiation). This class of tetrads, known as quasi-Kinnersley tetrads [105, 270–272] is unique up to an overall phase factor and normalization.

Since there are no analytically known waveforms from the mergers of black holes, in order to test the correctness of numerically derived waveforms one needs to both carefully audit the codes and compare results generated from different code bases. The first such comparison [273] was performed early on with waveforms generated by the inspiral of an equal-mass, low-spin binary obtained using the LazEv code [2, 44] developed at Brownsville and Rochester Institute of Technology, the Hahndol code [48, 274, 275] developed at NASA-GSFC, and Pretorius’ original code [1, 93]. For that test, each group evolved similar binaries, but at slightly different initial configurations. The results from this first comparison are shown in figure 15. Later on, as more groups developed their own codes, more large-scale comparisons were performed. For the Samurai [276] project, comparisons were made between the SpEC [277, 278] code developed by the SXS collaboration, the Hahndol code, the MayaKranc code [279] developed at Penn State / Georgia Tech, the CCATI code [207] developed at the Albert Einstein Institute, and the BAM code [39, 49] developed at the University of Jena. One of the major differences between the Samurai project and [273] was the use of simulated LIGO noise data to determine if the differences between the waveforms generated by the various codes is, in practice, detectable.

---

\( \text{Figure 13. The waveform (real and imaginary components of the (2,2) mode of the rescaled strain) from a precessing binary with total mass } 20M_\odot [239, 263]. \) To obtain the physical strain, the values plotted need to be rescaled by a factor of \( 9.7 \times 10^{-16}/D \), where \( D \) is the distance to the binary in units of kiloparsecs.
Such comparisons took on more urgency with the detection of gravitational waves in 2015 [4, 5]. In the process of verifying that detection two groups, the SXS collaboration and the group at RIT, generated waveforms as part of their work for the LIGO scientific collaboration. Despite the two codes sharing no common routines, and using different initial data generation techniques, different evolutions techniques, and different waveform extraction techniques, the dominant ($\ell = 2, m = 2$) modes produced by the two codes agreed to better than 99.9% [280].

One of the main goals of waveform modeling is to infer the properties of the source based on the observed waveforms. One of the first systematic studies of how sensitive the waveform is to the parameters of the binaries in the non-precessing case was performed in [635], where they studies the effects of spins on waveform detectability.

3.2.5. Semi-analytic waveform models. Because fully nonlinear numerical simulations of black hole mergers are computationally expensive, many semi-analytic and empirical approaches have been developed in order to model the waveform of black hole mergers using information from previously modeled simulations [214, 227, 228, 230, 233, 235, 237, 238, 248, 261, 276–278, 280–298, 299–316]. As demonstrated in [317], the systematic errors associated with waveforms models used by the LIGO Scientific Collaboration led to negligible errors in the inferred parameters of the source of GW150914.

One conceptually straightforward method for constructing a waveform model would be to interpolate known numerical waveforms as a function of the source parameters. The actual details of how this interpolation is performed is quite sophisticated and requires first representing waveforms in a reduced basis. The resulting models [242, 310, 316, 318] have been shown to be accurate, at least under the conditions tested [242].

Another class of models, usually referred to as phenomenological models [281, 286, 292, 309, 311, 319], are based on an expansion of the waveform in Fourier space. The amplitude and phase of the waveform are expanded as algebraic functions of frequency. The coefficients in these expressions are then modeled as functions of the binary’s parameters by fitting to existing waveforms.

Yet another method for modeling the waveform from a binary, when the two black holes are still far apart, is based on a series expansion in the black-hole separation and velocity,
known as the post-Newtonian expansion. For a review of post-Newtonian theory, see [320].

Since post-Newtonian theory is expected to be accurate when the binary separation is large, and become increasingly more inaccurate during the inspiral, a natural question is to determine where the post-Newtonian waveforms differ substantially from the numerical ones. The first direct comparisons of the waveform predictions from numerical relativity and post-Newtonian theory were performed by Buonanno, Cook, and Pretorius [321], shortly thereafter by the NASA-GSFC group [319, 322], with the group at Jena following soon after that [283, 284, 323]. These early comparisons were between post-Newtonian and numerical relativity predictions of the waveforms for non-precessing systems. The first comparison of precessing waveforms was done in [335]. All of these comparisons were for relatively short waveforms and in a regime where post-Newtonian theory is not particularly accurate. The longest comparison to date between post-Newtonian and numerical relativity waveforms was a 175 orbit evolution performed in [261]. Other studies of much more separated binaries (where the evolutions were not taken to merger) showed good agreement in the dynamics between post-Newtonian and numerical relativity for several orbits with separations of 100M and 50M for an equal-mass, nonspinning binary [260].

While current post-Newtonian waveforms are not particularly accurate during the late-inspiral phase, there are models inspired by post-Newtonian theory that reproduce numerical waveforms with greater accuracy. The Effective One Body formalism [324–328] recasts the problem of the evolution of the binary as an effective field theory for a single particle. This formalism contains free parameters which can be modeled using numerical relativity simulations. The resulting formalism EOB-NR can reproduce gravitational waveforms, at least for non-precessing binaries, with great accuracy. Furthermore, these waveforms are produced at a fraction of the cost of the original numerical simulations [288, 289, 291, 303, 329–334].

4. Relativistic stars and disks
4.1. Black hole–black hole binaries with accretion
Astronomers expect that the environment of supermassive black hole–black hole binary mergers will often be gas-rich; therefore, there is hope for an electromagnetic counterpart to the (low-frequency) gravitational wave signal. In a thin accretion disk around a single black hole, angular momentum flows outward (an effect of MHD turbulence), causing gas to slowly spiral inward, releasing energy radiatively as it falls deeper into the gravitational potential. (Thus, the more compact the object, the more efficiently accretion onto it can release energy.) A black hole–black hole binary near merger might be accompanied by gas orbiting the binary itself, forming a ‘circumbinary disk’. Early 1D studies of circumbinary disks predicted that gravitational torques from the binary would clear out a region of radius about twice the orbital separation (for binaries with mass ratio around unity), suggesting that accretion onto the black holes would be mostly frustrated. As the binary inspirals, eventually the inspiral timescale becomes smaller than the disk’s viscous timescale, presumably causing the disk to decouple from the binary, its inner edge unable to keep up with the shrinking binary. Newtonian 2D (vertically summed) [335, 336] and 3D [337–339] simulations confirm the evacuation of the region around the binary, but find that gas is efficiently carried in narrow accretion streams from the inner disk to the black holes. Meanwhile, perturbations to the disk caused by the merger itself, with the associated mass loss and kick of the central system due to gravitational waves, have been investigated by artificially reducing the mass and adding linear momentum to the central object around an equilibrium disk [340–344]. All of this suggests that high luminosity can be maintained after decoupling and through merger.

Numerical relativity studies of fluids near black hole–black hole binaries began shortly after the moving puncture revolution. It is not clear that this had to be the case. Some of the most advanced recent works consider disks around inspiraling binaries without dynamically evolved spacetimes. For example, Noble et al [345] used a 2.5 post-Newtonian-order approximation to the binary spacetime further than 10M from the binary combined with a 3.5 post-Newtonian approximation to the binary orbital evolution, while Gold et al [346] simply rotated their conformal thin sandwich initial data. Nevertheless, the first numerical relativity treatments did include spacetime evolutions through merger. These early studies by Bode et al [347–349] and Farris et al [350] considered binaries immersed in low-angular momentum gas (advection-dominated / Bondi-like inflow) and calculated electromagnetic luminosity from bremsstrahlung and synchrotron emission.

Clearly, magnetic field effects might have important effects on these inflows. Prior to MHD simulations, Palenzuela et al [351] and Moesta et al [352] performed force-free simulations of the effect of a black hole–black hole binary merger on nearby magnetic field lines (presumed to be anchored to a circumbinary disk outside the computational domain). An interesting finding of these simulations, shown in figure 16, is the appearance of dual jets by a sort of binary system generalization of the Blandford–Znajek effect; in this case energy is extracted from the orbital motion of the binary, rather than the spin energy of a black hole (the latter being the classic Blandford–Znajek effect). (It should be noted, though, that while multiple studies confirm the presence of dual jets, they also show that the emission is predominantly quadrupolar [352, 636].)

Newtonian MHD simulations by Shi et al [353] found that magnetohydrodynamic disks (as opposed to previous disks with alpha viscosity) accrete more rapidly and experience stronger tidal torques. Soon afterward, numerical relativity MHD simulations were carried out for the low-angular momentum plasma case by Giacomazzo et al [354] and for the circumbinary disk case by Farris et al [355]. Using numerical relativity MHD and a post-Newtonian black hole–black hole

17 The viscous timescale is the timescale on which angular momentum transport moves gas inward. The name comes from the common practice of modeling this transport process with a viscosity.
binary metric, Noble et al. [345] showed that neither binary torques nor decoupling reduce the overall accretion rate by a large factor. Gold et al. [346, 356] have carried out numerical relativity MHD simulations varying the binary mass ratio between 1:1 and 1:10, confirming that magnetic fields boost accretion rate, increase shock heating, and produce dual jets merging into a single jet at large distances. After decoupling but prior to merger, the jets coalesce into a single jet. The merger leads to a one-time boost in the jet’s magnetic field strength and outflow velocity, which the authors hope can provide a signature of black hole–black hole binary merger (as opposed to a single black hole disk flare).

Computational cost has forced nearly all simulations to date either to excise the inner region containing the binary or, for those cases that do track flow into the separate black holes, evolve for less than a viscous timescale. To observe viscously settled accretion flows, Farris et al. [357] performed long-term pre-decoupling 2D Newtonian evolutions resolving the inner region. They find that individual ‘mini-disks’ form around each black hole. Bowen et al. [358] have studied these mini-disks in their post-Newtonian spacetime, inserting disks around each hole and evolving them to an overall steady state. The discovery of these ‘mini-disks’ illustrates the possibility of further surprises as future simulations incorporate long evolutions as well as radiation transport with associated thermal effects.

4.2. Relativistic stars

Numerical relativity is the main tool for studying rapidly rotating relativistic stars, where it is used to test the stability of equilibrium configurations and the nonlinear evolution driven by instabilities. Because the outcome of these instabilities often involve black holes, relativistic stars will be a notable part of our story. For a full treatment of this topic, see the Living Review article by Paschalidis and Stergioulas [359, 360], the book on this subject by Friedman and Stergioulas [361], and also chapter 14 of Baumgarte and Shapiro [35].

Collapsing star simulations did not have to wait for the black hole problem to be solved. Simulations up to the turn of the century could use singularity avoiding slicings such as maximal slicing and its approximates to follow collapse a short while past apparent horizon formation before grid stretching effects (increasing distortion of slices needed to keep them from intersecting the singularity) destroyed the run’s accuracy. Evolutions to late times after collapse, and
evolutions of systems like black hole—neutron star binaries, which have a black hole throughout, had to wait until general black hole spacetimes could be stably evolved.

4.2.1. Radial stability and collapse outcome. Supramassive and hypermassive stars can be created using 2D stellar equilibrium codes (see [360]). Knowing that these equilibria exist, we next consider whether they are dynamically stable, and if not whether the instabilities are of a kind to destroy the equilibrium on a dynamical timescale or if they just introduce some small-scale ‘churning’ with effects on a secular timescale. Stability concerns the behavior of initially small perturbations. Numerical error provides perturbations on its own, but it is resolution-dependent, so stability studies often seed perturbations. A popular method for studying radial stability in stars is pressure depletion, a slight reduction of pressure below the equilibrium requirement. This can be done in a way that preserves the constraints and respects the equation of state by simply holding $\tau$ and $\delta_i$ fixed and slightly increasing $\rho_*$ [362]. To study the stability of nonaxisymmetric modes, these modes can be seeded by nonaxisymmetric perturbations of the density.

Because no black hole is involved (at least until after instability has clearly manifested itself), numerical relativity could begin addressing these questions even before the breakthroughs numerical relativity in 2005 that allowed for the evolutions of orbiting black hole—black hole binaries. Already in 2000, simulations by Shibata et al. [363] showed that the dynamical instability point for uniformly rotating supramassive $n = 1$ polytropes nearly coincides with the secular instability point. Despite the rapid rotation, when pressure-depleted unstable stars collapse to black holes, they leave almost no disk. A follow-up study of supramassive polytropes with various polytropic index $n < 2$ near their mass-shedding limit also found almost no disk mass around the post-collapse Kerr black holes [364]. These simulations used singularity avoiding slicings and could not evolve long after horizon formation. Subsequent studies using excision after apparent horizon location confirmed the no-disk result for uniformly rotating stars with stiff polytropic EoS [121, 122]. Based on angular momentum distribution at the maximum mass configuration, it is expected that this result carries over to realistic neutron-star EoS [365]. (In section 5, we will see that the story is quite different for non-compact stars with soft EoS.) After moving puncture gauge conditions were discovered, it was possible to evolve to a final black hole state and confirm that the metric matches the spinning puncture form [366].

Numerical relativity provides a straightforward way to test the dynamical stability of hypermassive neutron stars: just evolve for several dynamical times (perhaps with some initial perturbation). By this test, Baumgarte et al. [29] demonstrated the stability of a model with mass around $1.6M_{\odot}$ (the usual case for $\rho_*$ otherwise $\sim A_0/\Omega_0$, where $A_0 = B/\sqrt{\rho_0}$ is the Alfvén speed, and growth timescale $\sim \Omega^{-1}$. For realistic B fields, resolving $\lambda_{MRI}$ can be a serious computational challenge. Numerical relativity MHD studies sometimes avoid this problem by using magnetar-strength seed fields. The MRI has been identified in relativistic disk (e.g. [376–379]) and star [380] simulations. The effect of the MRI is to initiate turbulence and thus on a secular timescale to dissipate energy as heat and transport angular momentum outward. The MRI’s role in driving disk accretion is the subject of a vast amount of work; the distinctive role of numerical relativity has been to study its effect on differentially rotating neutron stars.

Axisymmetric simulations of magnetized hypermassive neutron stars were undertaken by Duez et al. [367] using a $\Gamma = 2$ EoS and initial angular velocity about three times higher at the center than at the equator. These simulations were notable for the narrative of this article as being among the first astrophysically interesting numerical relativity simulations to prolong evolution past collapse using excision. (Shortly afterward, both groups in the collaboration—University of Illinois at Urbana-Champaign (UIUC) and Kyoto University—switched to moving punctures.) A combination of magnetic wind/braking and MRI turbulence transports angular momentum outward, causing the envelope to expand and the core to contract. After about $10^5$ ms, the core undergoes collapse on a dynamical timescale. The collapse leaves

4.2.2. Magnetohydrodynamic evolution. These hypermassive neutron stars, although dynamically stable, are presumably driven to collapse on a secular timescale by processes that transport angular momentum outward, robbing the core of its rotational support. The ultimate source of angular momentum transport is most likely turbulence driven by the magnetorotational instability. To study secular evolution from first principles requires magnetohydrodynamic (MHD) simulations.

Such simulations must specify an initial state for the magnetic field. One injects a small magnetic field into the equilibrium state, usually not chosen to be an MHD equilibrium but thought of as a ‘seed’ of the more physical magnetic field that will grow through shear and turbulence. MHD simulations suggest that the magnetic field in a star with random initial state tends to settle to a helical mixture of poloidal and toroidal field [369]. Equilibrium fields can be constructed in relativity (e.g. [370–373]), but simpler seed fields are more often used, e.g. poloidal fields moving along isodensity contours constructed from an azimuthal vector potential

$$A_{\phi} = \begin{cases} A_{\phi} \varpi^2 \left( \rho_0 - \rho_{\text{seed}} \right)^n & \text{for} \rho_0 > \rho_{\text{seed}} \\ 0 & \text{otherwise} \end{cases},$$

where $\varpi$ is the cylindrical radius; $A_{\phi}$, $n$, and $\rho_{\text{seed}}$ are freely specifiable constants. The expectation was that, for magnetorotationally unstable systems, no memory of the seed field would long survive. Black hole-torus simulations give some support to this assumption for the interior of the torus but find that the appearance and strength of polar jets is very sensitive to seed field geometry [374].

Differential rotation with $d\Omega/dr < 0$ (the usual case for differential rotation) is unstable to the magnetorotational instability (MRI) [375]. The fastest growing MRI mode has size $\lambda_{MRI} \sim v_A/\Omega$, where $v_A \sim B/\sqrt{\rho_0}$ is the Alfvén speed, and growth timescale $\sim \Omega^{-1}$. For realistic B fields, resolving $\lambda_{MRI}$ can be a serious computational challenge. Numerical relativity MHD studies sometimes avoid this problem by using magnetar-strength seed fields. The MRI has been identified in relativistic disk (e.g. [376–379]) and star [380] simulations. The effect of the MRI is to initiate turbulence and thus on a secular timescale to dissipate energy as heat and transport angular momentum outward. The MRI’s role in driving disk accretion is the subject of a vast amount of work; the distinctive role of numerical relativity has been to study its effect on differentially rotating neutron stars.
a massive torus surrounding the black hole, a promising setup for a short-duration gamma ray burst [381]. Snapshots from this evolution are shown in figure 17. The same collaboration then performed similar studies for other differentially rotating neutron star initial configurations [382]. Use of a more realistic EoS produces a qualitatively similar outcome, but collapse is averted (unsurprisingly) if the star is not hypermassive. A non-hypermassive star with angular momentum too high for a uniformly rotating star settles to an equilibrium uniformly rotating star plus torus configuration.

4.2.3. Nonaxisymmetric mode instability. Nonaxisymmetric modes, which have the form $\delta \propto e^{i(M\phi - \omega t)}$, are interesting as gravitational wave sources. Since the background equilibrium is often differentially rotating, there is a clear conceptual difference between the fluid’s rotation and the perturbation mode, which rotates with constant pattern speed $\Omega_p = \omega/M$ everywhere. In fact, rotating perfect fluid stars are generically unstable because of gravitational waves via the Chandrasekhar–Friedman–Schutz instability, although in most realistic cases this is suppressed by viscosity or other effects. (See [360] and references therein.)

It is well-known that for rotating stars, the fundamental $L = M = 2$ (bar) mode becomes unstable for sufficiently high $T/|W|$; around 0.14 for a secular instability and around 0.27 for a dynamical instability. The unstable bars grow to nonlinear amplitude and lead to the shedding of high angular-momentum material. It is thus hard to imagine $T/|W| > 0.27$ stars persisting in nature. The dynamical bar mode instability (often called the ‘high $T/|W|$ instability’) has been confirmed in numerical relativity simulations of differentially rotating stars [383–385].

Numerical relativity investigations also found unexpected unstable growth of low (but nonzero) $M$ modes in strongly differentially rotating or toroidal stars at $T/|W|$ well below the dynamical bar mode threshold [386–389]. The instability has been seen in numerical relativity rotating stellar core collapse simulations [390] and so may be an important gravitational wave source from a galactic supernova. As the mechanism was not at first understood, the instability was called the ‘low $T/|W|$ instability’ or the ‘one-armed spiral instability’, names that sometimes persist. We now know that it is caused by a corotation resonance [391]. The corotation radius $r_c$ of a mode is the radius where $\Omega(r_c) = \Omega_p$. The mode has positive energy for $r > r_c$ and negative energy for $r < r_c$, and thus energy can be transferred outward at $r_c$ to strengthen the mode on both sides. The corotation instability takes many crossings to grow, so the mode must be trapped in a region containing $r_c$. In stars, a minimum of the vortensity can act like a trapping potential [392], and this can be produced by a toroidal density structure or extreme differential rotation. Subsequent numerical simulations are consistent with this model [393, 394]. Bar mode growth from both the low-$T/|W|$ [394] and the high-$T/|W|$ [395] instabilities has also been simulated in the presence of magnetic fields, where magnetic tension fails to suppress the instabilities for any realistic field strength.

4.2.4. Stability of self-gravitating black hole accretion disks. Given the ability to evolve matter in dynamical black hole spacetimes, we can carry out a similar analysis to that of relativistic stars, this time for self-gravitating tori around black holes. Once again we require axisymmetric constraint-satisfying equilibrium initial data, with a central black hole introduced either by a horizon inner boundary condition [396] or a puncture [397]. Then we just watch perturbations evolve.

Self-gravity can lead to instabilities in black hole-torus systems. Even in Newtonian physics, a disk will break apart if
it violates the Toomre stability criterion. A potential axisymmetric dynamical instability that has received much attention from numerical relativists is the runaway instability [398]. In this scenario, an equilibrium torus filling its Roche lobe will be unstable because a small amount of accretion into the black hole increases the black hole’s mass, pushing the Roche lobe into the torus. The mass transfer into the black hole is then unstable and destroys the torus in a dynamical timescale. Analysis of stationary configurations suggests that the disk’s self-gravity enhances the instability [399], but disks are stabilized by a positive radial angular momentum gradient in the disk and by black hole spin [400]. The first relativistic simulations, by Daigle and Font [401, 402], neglected torus self-gravity, treating evolution of the metric by allowing the mass and spin of the Kerr black hole to increase by accretion. These confirmed that even rather small angular momentum gradients prevent the instability. However, only live-metric simulations could properly include torus self-gravity. These were first undertaken by Montero, Font, and Shibata [403], and they did not find a runaway stability in any of their models. However, simulations by Korobkin et al [404] did demonstrate the existence of the instability in disks particularly prone to it.

There are also nonaxisymmetric instabilities present in black hole-torus systems analogous to those found in rotating stars. Before they were identified in stars, corotation instabilities were already known to exist in nearly-constant angular momentum disks (the Papaloizou–Pringle instability [405, 406]) and in disks with vortensity maxima (the Rossby wave instability [407]). Self-gravity is not an essential feature of the instability, and in fact self-gravity tends to suppress the Papaloizou–Pringle instability [408].

Torus self-gravity can trigger nonaxisymmetric instabilities not present in nonself-gravitating disks. These instabilities have been studied systematically by 3D simulations in Newtonian physics [409] and numerical relativity [410]. The latter study, by Korobkin et al., constitutes one of the first notable applications of cubed-sphere multipatch technology to numerical relativity. Simulations show that moderate self-gravity triggers ‘intermediate mode’ instability [408], spontaneous elliptic deformations of the disk that, in fact, can be considered the disk analogue of the high-$T/|W|$ instability [411]. An interesting effect of self-gravity on the $m = 1$ Papaloizou–Pringle mode is momentum transfer between the disk and black hole, leading to an outspinning motion of the black hole [410]. The nonlinear development of this instability was explored in numerical relativity by Kiuchi et al [412], who suggest it may be a significant source of gravitational waves (e.g. from a GRB central engine or the aftermath of a supermassive star collapse).

5. Black hole formation

In section 4.2, we considered the collapse of uniformly rotating compact, stiff stars, finding that they tend to collapse to Kerr black holes with no significant leftover material to form an accretion disk. Much more common astrophysically is the formation of black holes from non-compact stars with soft EoS around $n = 3$ (the marginal stability limit for Newtonian polytropes). Scenarios can be divided by the mass of the progenitor star. Population I and II stars with masses $\sim 10^2\, M_\odot$ form iron cores of mass $\sim M_\odot$, where $n = 3$ comes from the dominance of relativistic degenerate electrons to the pressure. The first generation of stars, the metal-free Population III stars, may have had masses $\sim 10^5\, M_\odot$. If metal-free gas is unable to cool and does not fragment into Pop III stars, these massive stars are radiation-pressure dominated and (because of convection) isentropic, leading them to also take the form of $n = 3$ polytropes. Like black hole spacetimes, polytrope systems can be scaled to any mass. However, this scale invariance is broken when one takes into account EoS stiffening and nuclear reactions, which depend on the actual density and not just the dimensionless compaction. Numerical relativity is needed to determine the collapse outcome: the mass and spin of the black hole and the properties of any accompanying disk. Even before these simulations could be carried to post-collapse equilibrium, it was possible to guess from a trick introduced by Shapiro and Shibata [413] that these collapses would be more likely to form massive disks. The post-collapse disk is roughly the matter with high enough initial angular momentum to orbit outside the innermost stable circular orbit (ISCO)\(^{18}\) of the black hole to be formed, which can be of order 10% of a supermassive star progenitor.

5.1. Population III core collapse and collapsars

Stellar mass black holes are thought to originate in the core collapse of massive stars for cases where, for some reason, the process of permanently expelling the gas around a protoneutron star fails. (The case of successful supernova explosion and neutron star formation has been the subject of much numerical work, the discussion of which would take us too far afield.) If the progenitor has sufficient angular momentum, the newly formed black hole may be surrounded by an accretion torus. This is the explanation of long-duration gamma ray bursts in the collapsar model of Woosley and MacFadyen [414–416]. Formation of a black hole-torus system may occur in several ways. The inner core may originally collapse to a protoneutron star, but a supernova may fail to occur (the shock stalls and does not sufficiently re-energize), so that the star eventually collapses under its accumulating mass; this is a Type I collapsar [416]. A mild explosion may occur, but enough material falls back onto the star to trigger collapse—a Type II collapsar [417]. Finally, the inner core might collapse directly to a black hole—a Type III collapsar [418].

The first 1D numerical relativity stellar collapse simulations were carried out in 1966 by May and White [419] using 18 In general relativity, the effective potential associated with orbits (i.e. timelike geodesics) around a black hole is similar to the effective potential in Newtonian gravity, but with additional attractive terms proportional to $1/r^3$. Because of this term, there is a region from the black-hole horizon to about three times the Schwarzschild radius of the black hole where there are no stable circular orbit. At the boundary of this region is the innermost stable circular orbit, or ISCO. A particle following an inspiraling quasi-circular orbit will plunge into the black hole once it crosses this ISCO.
the formulation of Misner and Sharp [420]: a Lagrangian method with, however, a slicing that does not avoid singularities. Thus, simulations could not be continued long after black hole formation, but this was enough to determine that a black hole in fact forms, rather than a neutron star. This problem can be overcome using a retarded time coordinate, which avoids the black hole interior [421, 422]. The same basic methods have continued to be used for subsequent spherical collapse simulations with increasingly sophisticated microphysics and neutrino transport [423–426], including detailed studies of the neutrino signals from failed supernovae solving the Boltzmann transport equation [427–429]. 1D general relativity simulations find that black hole formation tends to happen for high progenitor mass (which, due to stellar winds, may be much lower than the zero age main sequence mass, so black hole formation is more likely for low-metallicity stars, which suffer less mass loss). Prompt collapse may only occur for very high mass, low metallicity (perhaps only Population III stars) [426, 430, 431]. 2D simulations are needed to study rotating collapse. Some such simulations were carried out beginning with Nakamura [432] but lacked realistic initial conditions and EoS. In 2005, on the eve of the numerical relativity revolution in black hole treatment, Sekiguchi and Shibata [433] attempted greater realism using a set of two-component piecewise polytrope EoS. The paper limited itself to the criterion for prompt black hole formation because subsequent evolution could not be followed.

While waiting for numerical relativity, 2D post-collapse simulations were being used to study the evolution of the post-black hole formation torus and the possible initiation of a gamma rays burst. Lacking the true post-collapse configuration, these first simulations had to insert a black hole into a collapsing flow by hand. For Newtonian simulations, the black hole is a Newtonian or pseudo-Newtonian point-mass addition to the gravitational potential and an inner absorbing boundary [416, 417, 434–436]. For relativistic MHD simulations, a fixed Kerr metric was used [437–440]. A key input parameter is the initial angular momentum. To produce a promising torus, this is usually chosen to be large enough for circular orbit well outside the nascent black hole’s ISCO but small enough that the disk is compact and can lose energy efficiently by neutrinos. However, Lee and Ramirez-Ruiz [441] find promising behavior even for somewhat lower angular momenta; shocked gas on the equator forms a dwarf disk which accretes rapidly due to general relativistic effects even without magnetic fields or viscosity. Follow-up simulations by Lopez-Camara et al [436] suggest that low-j collapsars might differ from high-j collapsars by the former not producing an accompanying supernova. Hydrodynamic simulations, such as the original study by MacFadyen and Woosley [416] add an alpha viscosity and tend to find that the polar regions free-fall into the black hole while inside an accretion shock a thick torus forms and viscous heat-driven outflows are launched. Subsequent simulations with MHD for both high-j [434] and low-j [438, 442] cases differ primarily in the quick appearance of magnetically-driven polar jets.

Clearly, numerical relativity simulations were needed which include the collapse of a realistic rotating stellar core, self-consistent black hole formation, and evolution long past black hole formation to study the dynamics of the torus. With the ability to stably form and evolve black holes in numerical relativity, this became possible. The first such simulation was carried out by Sekiguchi and Shibata [443]. They evolved a high-entropy core from collapse through a second past black hole formation using a finite temperature equation of state and neutrino leakage. A range of initial j were used; low-j cores produced geometrically thin shocked disks, while high-j cores produce thick tori. The first 3D numerical relativity collapse simulations were performed by Ott et al [444]. Octant symmetry and eleven levels of adaptive mesh refinement made it possible to follow the collapse in 3D, but the simulation was only followed for ~0.1 s after black hole formation. Core collapse is followed by a bounce, but the accretion shock stalls, and the protoneutron star collapses to a black hole. Several key quantities in these simulations are plotted in figure 18. The dimensionless spin peaks at 0.75 for the most rapidly rotating case and then rapidly decreases as lower-j material is accreted. The collapse-bounce-collapse sequence of events leads to a distinct gravitational wave signal. In addition to this collapse waveform, gravitational waves may be produced by inhominogeneities in the collapsing matter, self-gravitational instabilities in the torus, and anisotropic neutrino radiation. (See [445, 446] and references therein.)

A difficulty for 3D collapsar simulations is the long (multi-second) timescale on which rapid accretion occurs. The inclusion of MHD and neutrino transport will make modeling the hyperaccretion phase even more challenging.

5.2. Massive star collapse in the early universe

Black hole formation from massive stars in the early universe is interesting primarily for explaining the ‘seeds’ from which supermassive black holes grew, but also as sources of electromagnetic and gravitational wave signals.

1D simulations of Population III stars indicate that stars with mass less than around $260 M_\odot$ end their lives in pair-instability supernovae, while more massive stars collapse directly to black holes [418, 431]. Nakazato et al [447] study the effects of neutrino emission on Pop III star collapse using a relativistic Boltzmann transport code. Effects of rotation are indirectly addressed by the 2D Newtonian simulations of Ohkubo et al [448], who, in a manner reminiscent of early collapsar simulations, insert a point mass by hand into the collapsing star. Rather than adding rotation and viscosity, the effects of a disk are modeled on larger scales by injecting a jet through the inner boundary. The jet drives an explosion, and nucleosynthesis outputs are calculated.

Supermassive stars began to be simulated in the 1970s. At the time, interest was primarily driven by the prospect of explaining active galactic nuclei in terms of these objects, which do radiate at their Eddington luminosity. These stars are radiation-pressure dominated and presumed isentropic, so they are nearly $n = 3$ polytropes. (Gas pressure makes $n$ slightly less than 3, but this deviation decreases with increasing mass.) Although they are not compact, the instability
of $>10^5 M_\odot$ stars is triggered by general relativity. These earliest simulations were 1D, assuming spherical symmetry. Appenzeller and Fricke [449, 450], using post-Newtonian gravity but including nuclear reactions, found prompt collapse to a black hole if $M > 10^6 M_\odot$; for lower masses nuclear burning of hydrogen explodes the star. Later simulations found that nonzero metallicity (albeit high given the context: $Z \sim 10^{-2}$) catalyzes CNO burning and triggers explosion [451]. Spherically symmetric full numerical relativity simulations of high-mass supermassive stars found prompt collapse to a black hole [452, 453].

In fact, this is another system for which we expect rotation to be extremely important. As a supermassive star cools and shrinks, it probably spins up to the mass-shedding limit, thereafter following a mass-shedding sequence as it cools till it hits a radial instability. Equilibrium sequences of rapidly rotating supermassive stars in general relativity have been produced by Baumgarte and Shapiro [454] for $n = 3$ and by Shibata et al [455] for $2.94 \leq n \leq 3$. Sajo et al [456] evolved an $n = 3$ supermassive star from it is critical point using 3D post-Newtonian physics, finding that the collapsing star remains axisymmetric (i.e. nonaxisymmetric modes do not have time to grow even though $T/W$ passes the critical value for bar formation) and that nearly all the mass falls into the black hole despite its rotation. Around the same time, 2D numerical relativity simulations were carried out by Shibata and Shapiro [457], tracking the collapse until about 60% of the mass was inside the apparent horizon. Because they could not evolve long past black hole formation, the final disk mass remained uncertain, but the authors estimated from the angular momentum distribution that it should be around 10% of the star’s mass.

The Illinois group returned to this problem in 2007 with black hole excision to determine the post-collapse state [458]. Their simulations vindicated the earlier angular-momentum based predictions of a massive disk (several percent of the total original rest mass) and a black hole with spin about 70% of the Kerr limit. This work also included magnetic fields, which do not affect the collapse but do affect the disk evolution. Insertion of a dynamically unimportant dipole field into the pre-collapse star leads to jet formation after collapse, with enough Poynting luminosity to potentially power an ultra-long GRB detectable at high redshifts [459]. Montero et al [460] carried out simulations with detailed microphysics and hydrogen and helium burning, finding that a metallicity of $10^{-3}$ is needed to cause thermonuclear explosion rather than black hole formation for mass-shedding supermassive stars of mass $\sim 5 \times 10^6 M_\odot$. Finally, Shibata and collaborators have included the deviation of $\Gamma$ from 4/3 [461] and nuclear reactions [462], finding collapse outcomes similar to the $\Gamma = 4/3$ studies [458].

Initial data sets for the above simulations assume the star is able to maintain uniform rotation and entropy. Collapse of differentially rotating supermassive stars has been simulated in 2D by Montero et al [460] and in 3D by Sajo and Hawke [463]. The latter study monitors quasi-periodic gravitational waves coming from the post-collapse system even after the hole’s quasinormal ringing damps. When a nearly extremal black hole is formed, the gravitational wave grows to fairly high amplitude after collapse for reasons which remain mysterious. Zink et al [464] studied the collapse of differentially rotating toroidal supermassive stars, finding that in this case the star is subject to strong nonaxisymmetric instabilities. These lead to the fragmentation of the star into self-gravitating, collapsing parts, in some cases leading to the formation of a supermassive black hole–black hole binary system [465].

The above all assume that supermassive stars are able to thermally relax to isentropy, which should be true if they are convective. For an alternative scenario, leading to a stellar mass black hole surrounded by a much more massive envelope, see Begelman [466, 467].

### 6. Non-vacuum compact binaries

Non-vacuum compact object binaries (that is, binaries made of two compact objects, at least one of which is not a black hole) inspiral due to gravitational radiation just like black hole–black hole binaries. Tidal deformation of the star(s) constitute an additional time-varying quadrupole, subtly affecting the inspiral and associated gravitational waveform. For most of the inspiral, this small tidal effect can be adequately modeled using post-Newtonian theory, and to good approximation the size and structure of the star(s) only affect the waveform via their dimensionless tidal deformability parameters $\Lambda = (2/3) k_2 (c^2 R/GM)^2$, where $k_2$ is the apsidal constant while $R$ and $M$ are the star’s radius and mass, respectively. $\Lambda$ can be thought of as a measure of a star’s response to an
external tidal field. Thus, one may hope to use gravitational waveforms from compact neutron star binaries to constrain the \( \Lambda \), and hence the EoS, of neutron stars [647]. An important application of numerical relativity, not discussed in this review, is to test--and if necessary improve--these models of tidal effects on waveforms during inspiral. Interested readers are referred to a sample of the papers on the topic [637–641].

Material effects become dramatic at the end of inspiral, which will involve a collision or, more often, a tidal disruption. The latter happens when the tidal force on a star from its companion exceeds the star’s own self gravity. This can be illustrated by a simple Newtonian order-of-magnitude calculation. Suppose the disrupting star has mass \( M \), radius \( R \) and is at a separation \( d \) from its companion of mass \( m \). (In all cases we consider, this companion will be more massive and more compact, often a black hole.) Then the self-gravitation and tidal accelerations are \( mR^2 \) and \( mR/d^3 \), respectively, and they match when

\[
\frac{d}{m} \sim \frac{(R/M)(m/M)}{2^{3/2}} .
\]

Tidal disruption is likely marked by a sharp decrease in the gravitational wave amplitude as the binary loses its quadrupolar shape. The subsequent fate of the disrupted star’s mass must be determined by simulations. Most attention has been given to the two observationally important possibilities of gas forming an accretion disk around the remaining binary object and gas being ejected from the system.

6.1. White dwarf—compact object binaries

Neutron star-white dwarf mergers and stellar mass black hole–white dwarf mergers are not easily amenable to numerical relativity because of the disparity of length scales between the two objects. Also, white dwarfs are not in nuclear statistical equilibrium. Usually, isotope abundances in a white dwarf can be considered fixed, but a merger event may trigger nuclear reactions, which would then provide a new energy reservoir and must be explicitly tracked. Because the white dwarf disruption happens on scales much larger than the neutron star or low-mass black hole, one might ask if a Newtonian white dwarf plus point mass treatment is sufficient. Such calculations have been done in SPH [468], indicating that the disrupted white dwarf shears into an accretion disk, a possible setup for a long-duration gamma ray burst.

Paschalidis et al [469, 470] have attempted to use numerical relativity to study white dwarf-neutron star mergers. To make simulations feasible, the white dwarf is replaced by a ‘pseudo-white dwarf’, only ten times bigger than the neutron star rather than 500. The merger outcome is a Thorne–Zytkow-like object which will cool to a hypermassive star and eventually collapse to a black hole. The authors acknowledge that simulations with more complete microphysics are still needed. 1D disk calculations by Metzger and collaborators [471, 472] indicate that, at least for systems with mass ratio not close to one, heating from nuclear reactions may unbind most accreting matter, and they judge it unlikely that enough mass accumulates on a neutron star to trigger collapse in those cases.

For Intermediate mass black hole-white dwarf tidal disruptions, the disparity of length scales is removed and tidal disruption happens in the strong gravity regime. In this case, we would have in mind nearly parabolic encounters in dwarf galaxies or globular clusters, rather than quasicircular inspiral and merger. Rosswog et al [473] carried out Newtonian SPH simulations, including nuclear burning, of such events, looking especially at cases where tidal compression triggers explosive nuclear burning. Because the code was Newtonian, the black hole had to be approximated by a Paczynski–Witt potential. The effects of black hole spin could only be studied with numerical relativity simulations, which were carried out by Haas et al [474], although without nuclear reaction effects. Both sets of simulations predict a residual accretion disk and accompanying soft x-ray flare lasting about a year.

6.2. Black hole–neutron star binaries

Because of their potential as gravitational wave sources and short gamma ray burst progenitors, most non-vacuum numerical relativity work has focused on black hole–neutron star binary and neutron star–neutron star binary mergers. In addition, these mergers are of interest as possible sources of short-duration gamma ray bursts, r-process nucleosynthesis, and kilonovae. (For reviews of multimessenger astronomy, see [642, 643].). In this section, we review black hole–neutron star binary merger simulations, a key application of numerical relativity dynamical black hole-handling technology. For a fuller treatment, see the Living Review by Shibata and Taniguchi [475].

6.2.1. Expectations before numerical relativity simulations. Black hole–neutron star binaries were historically the last to be simulated in numerical relativity, but simple arguments and Newtonian simulations gave some idea what to expect.

If a neutron star disrupts inside the ISCO of its companion black hole, no massive disk or ejecta is expected. What’s more, the gravitational wave in these cases should be nearly indistinguishable from that of a black hole–black hole binary system with the same masses. Defining \( d_{\text{ISCO}} = \kappa M_{\text{BH}} \) and using equation (62), we conclude that tidal disruption is likely for binaries with

\[
\frac{R_{\text{NS}}}{M_{\text{NS}}} > \kappa \left( \frac{M_{\text{BH}}}{M_{\text{NS}}} \right)^{2/3} .
\]

That is, disruption is favored by low neutron-star compactness \( C \equiv M_{\text{NS}}/R_{\text{NS}} \), low mass ratio \( q \equiv M_{\text{BH}}/M_{\text{NS}} \), and high dimensionless black-hole spin \( \chi \) (to reduce \( \kappa \)).

What happens when the neutron star fills its Roche lobe and mass transfer begins? The question was first addressed in Newtonian simulations by Lee and Kluzniak [476–478], with the black hole treated as a point mass and the neutron star treated as a polytrope. These simulations found that mass transfer is stable for stiff EoS but unstable for soft EoS [477, 478], a difference that carried over to rival nuclear theory-based EoS as studied by Janka et al [479] and Rosswog et al [480]. In the case of unstable mass transfer, the neutron star is
destroyed in a single mass-transfer event. Stable mass transfer, on the other hand, is episodic, yielding an unmistakably different gravitational wave signal. Replacing the Newtonian point mass with a Paczynski–Wiita potential makes mass transfer less stable, so that tidal disruption happens in one pass even for stiff realistic EoS [481, 482]. Newtonian simulations also found massive ejection of unbound matter during mergers. SPH simulations around Kerr black holes supported expectations that prograde black hole spin is favorable to disk formation [483].

6.2.2. Inspiral and merger in numerical relativity: parameter space exploration and gravitational waves. The first general relativistic simulations of black hole systems came soon after the moving puncture revolution. A head-on collision with a neutron star falling into a black hole was successfully modeled (using excision) by Löffler et al [644], and soon after Shibata and Uryu carried out binary merger simulations starting from roughly circular orbit (using moving punctures) [127]. Simulations by the UIUC, SXS, and LSU/BYU/LIU groups quickly followed [484–486]. The former two of groups used the NOKBSSN formalism with moving punctures; the latter two used the generalized harmonic formulation with explicit excision. These early simulations used simple, polytropic EoS, and in some cases improper treatment of low-density material led to underestimates in the disk and ejecta masses. They all found complete neutron-star disruption in a single mass transfer event. When tidal disruption occurs outside the ISCO, the neutron star deforms into a tidal stream, with inner material streaming toward the black hole and outer material streaming outward. Of the matter falling toward the black hole, most will fall into the black hole. Soon after flow into the black hole commences, material with sufficient angular momentum wraps around the black hole, causing the tidal stream to crash into itself. The resulting shock heats the gas, which begins setting into an accretion disk very close (tens of km) to the black hole. Of the nuclear matter expanding outward, some is bound and eventually falls back onto the disk, while the rest (the ‘dynamical ejecta’) is unbound and escapes permanently.

Using all the numerical relativity results available at the time, Foucart devised an analytic fit to the post-merger disk mass (defined as the rest mass of bound material outside the black hole 10 ms after merger; recall that mass is continuously falling onto the disk and into the horizon) as a function of $C$, $q$, and $\chi$ which confirms the expectation that, all else being equal, lower $C$, lower $q$, or higher $\chi$ increases disk mass [487]. Using their own set of simulations, Kawaguchi et al [488] devised a similar analytic fitting formula for the mass and asymptotic speed of unbound ejecta.

Using numerical relativity simulations, Shibata et al [489] found gravitational waves from black hole–neutron star binary mergers to always fall into one of three categories, illustrated in figure 19. When the neutron star falls into the black hole before being disrupted, the wave is similar to black hole–black hole binary waves. When the neutron star disrupts well outside the ISCO, the waveform cuts off at this point during the inspiral, and no merger or ringdown wave is seen. Disruption close to the ISCO gives a case with intermediate features: inspiral and merger waves but reduced ringdown wave. Information about the neutron star is contained in the gravitational wave cutoff frequency, Shibata et al emphasize that this cutoff frequency is not identical to the gravitational wave frequency at tidal disruption, but rather is somewhat higher. Presumably, this is because the star persists for a time as a clump of matter as it inspirals past the tidal disruption radius.

BH-polytrope simulations also found a strong dependence of the post-merger disk mass on the black-hole spin. Etienne et al [491] found that much more massive disks could be formed for neutron stars disrupted by black holes with prograde spin, which is perhaps to be expected, since such black holes have smaller ISCOs. An extreme case—mass ratio of 3 and prograde black hole spin at 97% of the Kerr limit—was simulated by Lovelace et al [492]; not even half of the rest mass is promptly accreted in this case. Retrograde spin, on the other hand, makes disruption outside the ISCO less likely and disk masses lower. The black hole spin orientation has been varied by Foucart et al [493, 494] and by Kawaguchi et al [495], with the general findings that large spin misalignments remove the increase in disk mass seen in prograde spins and lead to disks initially misaligned with the black hole spin. With high enough prograde black-hole spin, Foucart et al [494] were able to observe tidal disruption even in systems with mass ratios in the 5–7 range [494] where astrophysical black hole–neutron star binary systems are thought most likely to lie [496].

Most contemporary black hole–neutron star binary simulations use more realistic EoS. Because the matter only heats up as the gravitational wave signal turns off, waveform studies have sensibly concentrated on piecewise-polytropic parameter studies of EoS effects [490, 495, 497]. In a truly impressive effort, Kyutoku and collaborators carried out 134 merger simulations, varying both EoS and binary parameters.
[490, 497, 498]. The equation of state was modeled as a two-piece piecewise polytrope, which, since the low-density EoS is known, has two free parameters. They choose to systematically vary the high-density Γ and P1, the pressure at a fiducial density ρfidu = 10^{14.7} g cm^{-3}. The reason for using P1 rather than the transition density is because P1 is found to closely correlate with the neutron star radius and tidal deformability, making it a good candidate for an equation of state parameter that can be measured by gravitational wave signatures such as the cutoff frequency. These simulations have been used to calibrate analytic black hole– neutron star binary waveform models, valid in the range 2 < q < 5, covering the full inspiral and merger [498–501]. Fisher matrix and Bayesian analysis of the analytic model shows that LIGO detections can hope to significantly constrain the tidal deformability and P1, especially given dozens of realistic detections [498, 501].

Using piecewise polytrope EoS, East et al [502] simulated eccentric black hole– neutron star binary encounters. Here could finally be seen cases of episodic mass transfer in general relativity, as well as instances of the ‘zoom-whirl’ phenomenon observed in black hole– black hole binary simulations. Such events may well occur at interesting rates in dense stellar environments such as globular clusters [503]. If so, the richer dynamical possibilities of eccentric merger deserve more attention.

6.2.3. Post-merger in numerical relativity: neutrinos, ejecta, and MHD. Post-merger evolution requires finite-temperature EoS, which have been employed in simulations by the SXS collaboration [132, 139, 504, 505] and by Kyutoku et al [506]. These simulations fail to find episodic mass transfer even for stiff EoS. It is found that, when β equilibrium violation is allowed during tidal disruption (on these timescales, the lepton number advects), tidal streams are narrower, and ejecta velocities lower than β-equilibrium EoS would predict [505, 507]. Lepton number equilibrium in the post-merger disk is re-established about 10 ms after merger under the action of neutrino cooling, making it a good candidate for an equation of state parameter that can be measured by gravitational wave signatures such as the cutoff frequency. These simulations have been used to calibrate analytic black hole– neutron star binary waveform models, valid in the range 2 < q < 5, covering the full inspiral and merger [498–501]. Fisher matrix and Bayesian analysis of the analytic model shows that LIGO detections can hope to significantly constrain the tidal deformability and P1, especially given dozens of realistic detections [498, 501].

Using piecewise polytrope EoS, East et al [502] simulated eccentric black hole– neutron star binary encounters. Here could finally be seen cases of episodic mass transfer in general relativity, as well as instances of the ‘zoom-whirl’ phenomenon observed in black hole– black hole binary simulations. Such events may well occur at interesting rates in dense stellar environments such as globular clusters [503]. If so, the richer dynamical possibilities of eccentric merger deserve more attention.

6.2.3. Post-merger in numerical relativity: neutrinos, ejecta, and MHD. Post-merger evolution requires finite-temperature EoS, which have been employed in simulations by the SXS collaboration [132, 139, 504, 505] and by Kyutoku et al [506]. These simulations fail to find episodic mass transfer even for stiff EoS. It is found that, when β equilibrium violation is allowed during tidal disruption (on these timescales, the lepton number advects), tidal streams are narrower, and ejecta velocities lower than β-equilibrium EoS would predict [505, 507]. Lepton number equilibrium in the post-merger disk is re-established about 10 ms after merger under the action of intense neutrino emission (Lν ∼ 10^{53} erg s^{-1}, preferentially in electron antineutrinos), during which time the disk’s average Ye rises to about 0.1. (For stiff EoS, Ye might be as high as 0.2 [506].) As the disk cools, the equilibrium Ye decreases and a re-neutronization is seen in the disk.

Whether the neutrino emission is sufficient to power a gamma-ray burst remains uncertain. Just et al [508] simulate black hole tori using an energy-dependent M1 neutrino transport scheme, and find that, at least for favorable cases, neutrino-antineutrino annihilation can power a relativistic outflow sufficient for a low-energy short duration GRB. (black hole– neutron star binary mergers are more promising in this regard than neutron star– neutron star binary mergers, because the latter have more baryon loading from dynamical ejecta on the poles.)

After some early false negatives, tidal ejection of unbound matter was robustly identified in numerical relativity black hole– neutron star mergers [494, 509, 510]. In relativistic codes, unbound matter is usually identified as that with specific orbital energy \( e = -u_t - 1 > 0 \). The mass of dynamical ejecta for cases with tidal disruption is often large \( \sim 10^{-2} - 10^{-3} M_{\odot} \) and highly asymmetric—concentrated on one side in the orbital plane—with sufficient momentum to impart a kick of \( 10^5 \text{ km s}^{-1} \) on the remnant black hole in some cases [494, 509]. Additional matter is ejected into weakly bound orbit and will fall back onto the central remnant later. This fallback material was studied using Newtonian SPH by Rosswog [511] and then in numerical relativity by Chawla et al [486]. From the distribution of fallback times, these studies predict a late-time fallback accretion rate following a \( t^{-5/3} \) power law.

The possible importance of black hole– neutron star dynamical ejecta for r-process nucleosynthesis was pointed out by Lattimer and Schramm in 1976 [512]. Newtonian simulation confirm that black hole– neutron star binary mergers produce large ejecta masses of neutron-rich material that should undergo r-process nucleosynthesis and produce an optical/near-IR transient [513]. Indeed, so efficient are these mergers in producing r-process elements that Bauswein et al [514], use their own ejecta predictions from SPH conformally flat general relativistic simulations and the galactic abundance of r-process material to constrain the rate of black hole– neutron star mergers.

The SXS simulations confirm expectations that ejecta is very neutron rich, so as it de compressed it is expected to undergo r-process nucleosynthesis and produce the second and third r-process peaks. This is indeed what Roberts et al [515] find tracking nuclear reactions in the ejecta, although a weak first peak can be seeded by neutrino irradiation by the central black hole– disk system. Higher Ye, outflow may be provided by winds from the accretion disk. Recently, Fernandez et al [516] have produced models including both effects. Using outgoing ejecta and disk profiles from the SXS merger simulations as initial data, the disk was evolved to late times using 2D Newtonian hydrodynamics with an alpha viscosity to model angular momentum transport. As the disk evolves, neutrino cooling decreases to the point of insignificance, and by \( \sim 200 \text{ ms} \) the disk has reached an advective state (viscous heating balanced by advection of hot material inward rather than by radiative cooling) with strong convection and mass outflow. In most cases, the dynamical ejecta dominates, but if the disk outflow and dynamical ejecta should happen to have comparable masses, a solar-like distribution of r-process elements would follow.

Radioactive decay of r-process ejection might power a detectable signal, most likely in the near infrared, called a kilonova or macronova [517]. Tanaka et al [518] use ejecta from numerical relativity black hole– neutron star merger simulations as input to a (photon) radiation transfer code to predict light curves. The simulations did not include nucleosynthesis, but assumed the solar r-process pattern. They find that black hole– neutron star kilonovae can often appear as bright or brighter than neutron star kilonovae, and then in numerical relativity by Chawla et al [486]. From the distribution of fallback times, these studies predict a late-time fallback accretion rate following a \( t^{-5/3} \) power law.

The possible importance of black hole– neutron star dynamical ejecta for r-process nucleosynthesis was pointed out by Lattimer and Schramm in 1976 [512]. Newtonian simulation confirm that black hole– neutron star binary mergers produce large ejecta masses of neutron-rich material that should undergo r-process nucleosynthesis and produce an optical/near-IR transient [513]. Indeed, so efficient are these mergers in producing r-process elements that Bauswein et al [514], use their own ejecta predictions from SPH conformally flat general relativistic simulations and the galactic abundance of r-process material to constrain the rate of black hole– neutron star mergers.

The SXS simulations confirm expectations that ejecta is very neutron rich, so as it decompressed it is expected to undergo r-process nucleosynthesis and produce the second and third r-process peaks. This is indeed what Roberts et al [515] find tracking nuclear reactions in the ejecta, although a weak first peak can be seeded by neutrino irradiation by the central black hole– disk system. Higher Ye, outflow may be provided by winds from the accretion disk. Recently, Fernandez et al [516] have produced models including both effects. Using outgoing ejecta and disk profiles from the SXS merger simulations as initial data, the disk was evolved to late times using 2D Newtonian hydrodynamics with an alpha viscosity to model angular momentum transport. As the disk evolves, neutrino cooling decreases to the point of insignificance, and by \( \sim 200 \text{ ms} \) the disk has reached an advective state (viscous heating balanced by advection of hot material inward rather than by radiative cooling) with strong convection and mass outflow. In most cases, the dynamical ejecta dominates, but if the disk outflow and dynamical ejecta should happen to have comparable masses, a solar-like distribution of r-process elements would follow.

Radioactive decay of r-process ejection might power a detectable signal, most likely in the near infrared, called a kilonova or macronova [517]. Tanaka et al [518] use ejecta from numerical relativity black hole– neutron star merger simulations as input to a (photon) radiation transfer code to predict light curves. The simulations did not include nucleosynthesis, but assumed the solar r-process pattern. They find that black hole– neutron star kilonovae can often appear as bright or brighter than neutron star– neutron star kilonovae, because the former can produce more ejecta, and that the former will tend to be bluer than the latter. The same group applied these models to the purported kilonova associated with GRB 130603B, showing the observed near-infrared excess is consistent with either a soft EoS neutron star– neutron star binary merger or a stiff EoS black hole– neutron star merger [519]. Kawaguchi
et al [488] returned to black hole–neutron star kilonovae with analytic models for the heating and radiative diffusion but a large suite of merger simulations. Lanthamide-free disk wind might create a bluer signal [520], but in the models studied by Fernandez et al, this was all obscured by the dynamical ejecta [516].

The most dramatic, and least-understood, post-merger processes are magnetically driven. Early simulations with confined poloidal fields found that the field in the post-merger disk quickly wound into a toroidally dominated configuration, with no observable jets [486, 521] even when the MRI could be resolved [522]. Later simulations by Paschalidis et al [523] found that jets can more easily emerge from a field initially extending outside the neutron star. Such magnetospheric fields might also trigger observable signals that precede a GRB from the merger [156, 524]. In fact, even confined seed fields may be more promising than they originally seemed. Extremely high-resolution studies (Δx ≈ 100 m) by Kiuchi et al [525] find winds driven from inner disk heating, whose strength increases with resolution, with convergence not yet achieved, can pin magnetic flux to the black hole with associated Blandford–Znajek jets. Post-merger disk evolution turns out to be a difficult, multi-scale problem, and crucial properties like the rates of wind outflow, magnetic energy outflow, and neutrino annihilation energy deposit remain poorly constrained. For the time being, the long-term evolution of the disk is being investigated using Newtonian alpha-viscosity models (e.g. [526, 527]).

6.3. Neutron star–neutron star mergers

6.3.1. Pre-breakthrough simulations. The story of neutron star–neutron star binary merger simulations differs from that of other major numerical relativity problems in that many neutron star–neutron star binaries evolve well past merger without encountering black hole formation, so full numerical relativity merger simulations began earlier for neutron star–neutron star binary systems, and the ability to evolve space-times with black holes had a less dramatic effect. Neutron star–neutron star binary simulations have recently received a thorough review by Baiotti and Rezzolla [528]. They are also the subject of a Living Review by Faber and Rasio [529]. Techniques for constructing initial data are described in the review by Tichy [530]. Readers can find in these sources a more detailed presentation of the large subject of compact neutron star–neutron star binaries.

Simulations of neutron star–neutron star binary mergers were first undertaken in Newtonian physics with mostly simple polytropic equations of state. Nakamura and Oohara in a series of papers performed the first neutron star–neutron star binary merger simulations using finite differencing [531–534]. SPH simulations were performed by Rasio and Shapiro [535, 536]. Further simulations of both types followed [537–539]. Gravitational waves had to be studied in the quadrupole approximation, and the possibility of black hole formation could not be addressed. However, these Newtonian simulations showed some features that would be confirmed by numerical relativity simulations: the stars merge into a massive remnant rotating rapidly and differentially, and for stiff EoS the remnant is subject to bar mode deformations, leading to a sustained post-merger gravitational wave signal.

Post-Newtonian simulations were the next logical step. The first such simulation was carried out using smoothed particle hydrodynamics by Ayal et al [540]. Unfortunately the 1 post-Newtonian-order terms are not always small compared to Newtonian terms, indicating that truncating at this level is not a valid approximation, and post-Newtonian studies sometimes resorted to artificially reducing the post-Newtonian terms (e.g. [541]). The next advance was to the conformally flat approximation to general relativity. Here surprises seemed to arise when Wilson, Mathews, and Marronetti [542, 543] reported the neutron stars in their simulations collapsing individually to black holes before merging. However, they used an EoS with fairly low neutron star maximum mass and were found to have an error in one of their equations [544]. Pre-merger collapse is no longer considered likely, but the conformal flatness approximation has turned out to be a useful and reliable tool for neutron star–neutron star binary modeling (e.g. [545–547]).

Parallel to efforts toward incorporating relativity were efforts to include realistic microphysics in Newtonian simulations. Grid-based simulations by Ruffert et al [548–550] and SPH simulations by Rosswog et al [551–553] began the use of finite-temperature equations of state and inclusion of neutrino effects (in a leakage approximation) for neutron star–neutron star binaries. These simulations highlighted the potential of neutron star–neutron star binary mergers as GRB central engines. The study of neutron star–neutron star binary mergers with finite-temperature EoS was continued in conformally flat gravity by Oechslin et al [547]. Neutrino absorption effects were studied in Newtonian physics using flux-limited diffusion by Dessart et al [554] and through ray tracing and phenomenological extensions to leakage by Perego et al [555]. Using these very different methods, both groups find strong neutrino-driven winds ejected from the merged remnant, winds that could play important roles in generating r-process elements and kilonovae and in baryon loading the environment of a potential GRB.

Newtonian simulations were also able to investigate the ejecta and its potential for r-process nucleosynthesis before the first numerical relativity simulations [556–558], and later numerical relativity simulations have found results reasonably close to the earlier Newtonian studies.

Neutron star–neutron star mergers in full numerical relativity were first carried out by Shibata and Uryu at the turn of the century [114, 559, 560]. These initial simulations modeled the neutron stars as Γ = 2 polytropes. Their most significant discovery was that the remnant does collapse to a black hole, but only if its mass exceeds a certain threshold. Less massive systems form dynamically stable differentially rotating neutron star remnants (which in some cases are hypermassive). Other groups added AMR [50] and high-resolution shock-capturing techniques [561]. Binary polytrope simulations with other sophisticated numerical relativity codes followed (BAM [562], WhiskyTHC [118], SpEC [362]). All found similar results. As the next step in microphysical realism, Shibata and
collaborators tried cold nuclear-theory based EoS (augmented with a Gamma-law thermal component to capture shock heating) [563], confirming the possibility of nonaxisymmetric post-merger structure. Since then, the piecewise-polytrope parameterization has often been used to systematically vary the cold EoS, or even just as a cheap way to approximate a given cold EoS [564–567].

With the introduction of moving punctures, Baiotti et al [561] Kiuchi et al [568] revisited neutron star–neutron star binary mergers with various EoS, following high-mass cases beyond black hole formation to measure the post-collapse black hole and disk properties. For the APR EoS, equal-mass binaries with prompt black-hole formation leave very little torus mass (∼10⁻⁴M⊙), but about 10⁻²M⊙ disks remain for binaries with mass ratio around 0.8. Rezzolla et al [569, 570] have produced analytic fitting formulae relating post-collapse torus mass to the pre-merger binary parameters.

6.3.2. Post-merger evolution: neutrinos, ejecta, MHD. Now the main challenge was not metric evolution but microphysics. After two neutron stars merge, thermal, neutrino, and magnetic effects become important. Numerical relativity simulations with finite-temperature EoS and neutrino leakage were carried out by Sekiguchi et al [131, 571] for the stiff Shen EoS and a hyperonic EoS. A survey of neutron star–neutron star binary mergers with a large number of finite-temperature EoS has since been performed (in conformally flat general relativity) by Bauswein et al [572]. Kastaun and Galeazzi [574] carried out a further set of mergers with finite-temperature EoS (LS220 and SHT) and studied the structure of the hypermassive remnants in detail. Contrary to the widespread presumption that hypermassive stars ‘cheat’ the mass-shedding limit by having rapidly rotating cores with strong centrifugal support and more slowly rotating envelopes, the authors find that their hypermassive remnants have slowly rotating (pressure-supported) cores and extended, quasi-Keplerian envelopes. Simulations of low-mass neutron star–neutron star binary mergers (i.e. with non-hypermassive remnants) by Kastaun et al [575] and Foucart et al [576] show similar features. The composition of low-density regions and outflows is strongly affected by neutrino absorption. To capture these effects, numerical relativity simulations with energy-integrated M1 neutrino transport have been carried out by Wanajo et al [138], Foucart et al [576, 577], and Sekiguchi et al [578].

A number of studies have focused particularly on the dynamical ejecta [565, 577–579], which, like black hole–neutron star ejecta, is potentially important for r-process nucleosynthesis and kilonovae, although there are two important differences. Both come from the fact that dynamical ejecta comes not only from the tidal tail, but also from the collision interface. Much of this material is polar, leading to the first major difference from black hole–neutron star ejecta: the distribution of ejecta is much more isotropic [580, 582]. It has even been suggested that the ejecta forms a cocoon around the central object that can collimate the GRB outflow [583]. Additional ejecta is released in the following milliseconds as the hot remnant neutron starsettles. These extra sources of early-time ejecta lead to a second difference from the black hole–neutron star binary case, namely that ejecta can be hotter and more neutrino-processed. Due to n + e⁻ + νe → p + e⁻ reactions, a portion of ejecta with Y_e ∼ 0.3–0.4 can be created. The wide range of ejecta Y_e can produce all three r-process peaks without the need for a subsequent disk wind [138]. Numerical relativity has also been used to study systems with pre-merger neutron star spin [573, 574, 584–586] and orbital eccentricity [579, 585, 587, 588].

MHD simulations of neutron star–neutron star binary mergers have struggled with the difficulty of resolving the MRI in high-density regions and, even more challenging, the growth of the magnetic field in Kelvin–Helmholtz vortices. Early studies [589–591] demonstrated the ability of numerical relativity MHD codes to follow neutron star–neutron star binary mergers but could not resolve these effects, although they could resolve the MRI in a post-collapse torus [592] and showed the possibility of jet formation from the black hole–torus system (leading perhaps to a short duration gamma ray burst). A series of unprecedentedly high-resolution (as low as δx = 17.5 m) by Kiuchi et al [593, 594] succeeded in resolving these effects well enough to demonstrate amplification of the average field by a factor of 10³, but even this is taken as a lower limit.

Although this small-scale amplification cannot be adequately resolved, numerical relativity MHD studies continue to study large-scale processes during merger. For example, a few studies investigate the difference in merger scenarios between ideal and resistive MHD [153, 155, 645]. Most recently, a series of papers beginning with Endrizzi et al [595–597] has surveyed binary properties, EoS, and seed field effects on magnetized neutron star–neutron star binaries. Among their findings is a characteristic large-scale field structure appearing in many cases. For cases with black hole formation, these simulations did not observe polar magnetic jets, but Ruiz et al [598] does see them for stronger seed fields (allowing better MRI resolution) and longer integration times.

6.3.3. General relativistic effects: prompt collapse threshold and post-merger gravitational waves. Having outlined the development of neutron star–neutron star binary merger simulations, we turn to their findings concerning distinctly relativistic effects. The threshold mass M_th for prompt (i.e. on a dynamical timescale) collapse after merger to a black hole is found to be 30%–70% above M_TOV_max, depending on the EoS. Bauswein et al find that M_th = (2.43 − 3.38C_TOV-max)M_TOV_max to reasonable accuracy for all EoS studied, where C_TOV_max is the compaction of the TOV maximum mass configuration [572]. For neutron star–neutron star binary systems with mass below M_th, the massive (perhaps hypermassive) differentially rotating neutron star remnant is dynamically stable.

The remnant is formed with a strong quadrupolar distortion, emitting strong—and, what’s even better, EoS-sensitive—gravitational waves [567, 580, 599, 600]. Recall that the remnant is differentially rotating, so the overall rotating quadrupole in the matter profile is an m = 2 density mode, not solid body rotation. In fact, the mode angular frequency is close to, but slightly higher than, the maximum angular frequency in the star, for reasons that remain unclear [574, 576, 585, 577, 589].
...polytropes in 2D (axisymmetry) starting from a neutron star of the mass $M > \frac{3}{4} M_{\odot}$. On longer timescales of total angular momentum [580], or to the radius at some other fiducial mass [580, 581]. Simulations also consistently find a weaker signal at a lower frequency $f_1$, usually in the range 1.2–2.5 kHz, which has been explained in terms of oscillations in the distance between the two cores still distinct for a short time after merger [561], as a spiral density wave [600]. For nonlinear perturbations of this sort, these interpretations may not be mutually exclusive. For a wide range of EoS, $f_1$ seems to obey a universal relation to the total angular compaction of the premerger neutron stars [567]. Measuring these frequencies would significantly constrain the neutron star EoS. Unfortunately, this would only be possible for close mergers ($< 40$ Mpc) with Advanced LIGO or with a next generation gravitational wave observatory.

One worry about the above studies is that they ignore magnetic field-related stresses. Simulations by Palenzuela et al [604] with hot EoS, neutrino cooling, and MHD included find that the large-scale magnetic fields (those resolved in global merger simulations) are too weak to affect the post-merger waveform during the first 10 ms, even when Kelvin–Helmholtz amplification is included in an approximate way via subgrid modeling. As we shall shortly see, the effects of subgrid-scale MHD turbulence may be a different story.

6.3.4. Longer term evolution of remnants: subgrid scale modeling. The subsequent evolution of the remnant depends on secular processes that drive the star from one equilibrium to another. During the first tens of milliseconds, hydrodynamic torques redistribute angular momentum outward while gravitational radiation drains the star’s total angular momentum [580]. On longer timescales of $\sim 10^3 – 10^5$ ms, magnetic processes, namely magnetic winding and turbulent motions triggered by the MRI, also redistribute angular momentum outward [367]. If the core depends on rotational support, any of these might trigger collapse to a black hole–torus system. However, simulations with finite-temperature nuclear EoS tend to find thermally supported hypermassive remnants, so collapse of hypermassive remnants may be delayed until the neutrino cooling timescale, which would be of order seconds [131, 605]. Winds driven by magnetic fields [606] or neutrinos [555] may carry off a small amount of mass ($10^{-3} – 10^{-2} M_{\odot}$) and some angular momentum during that time. Remnants that are merely supramassive can survive beyond this time and collapse much later from angular momentum loss due to pulsar spindown.

Direct modeling of this evolution is, for the time being, out of reach, due to the multi-scale nature of the problem. We have seen how small-scale growth of the MRI and momentum loss due to pulsar spindown. During that time. Remnants that are merely supramassive can carry off a small amount of mass ($10^{-3} – 10^{-2} M_{\odot}$) and some angular momentum during that time. Remnants that are merely supramassive can survive beyond this time and collapse much later from angular momentum loss due to pulsar spindown.

6.3.4. Longer term evolution of remnants: subgrid scale modeling. The subsequent evolution of the remnant depends on secular processes that drive the star from one equilibrium to another. During the first tens of milliseconds, hydrodynamic torques redistribute angular momentum outward while gravitational radiation drains the star’s total angular momentum [580]. On longer timescales of $\sim 10^3 – 10^5$ ms, magnetic processes, namely magnetic winding and turbulent motions triggered by the MRI, also redistribute angular momentum outward [367]. If the core depends on rotational support, any of these might trigger collapse to a black hole–torus system. However, simulations with finite-temperature nuclear EoS tend to find thermally supported hypermassive remnants, so collapse of hypermassive remnants may be delayed until the neutrino cooling timescale, which would be of order seconds [131, 605]. Winds driven by magnetic fields [606] or neutrinos [555] may carry off a small amount of mass ($10^{-3} – 10^{-2} M_{\odot}$) and some angular momentum during that time. Remnants that are merely supramassive can survive beyond this time and collapse much later from angular momentum loss due to pulsar spindown.

Direct modeling of this evolution is, for the time being, out of reach, due to the multi-scale nature of the problem. We have seen how small-scale growth of the MRI and momentum loss due to pulsar spindown.
A disadvantage of the Navier–Stokes equations is that it results in a parabolic system that violates causality. A decade after Duez et al.’s work, Shibata et al. [608] returned to this problem, using a version of the Israel-Stewart formalism for relativistic viscosity [609], which introduces an evolution equation for $T^{\text{visc}}_{\mu \nu}$ and respects causality. Like in the earlier study, they evolve hypermassive stars in 2D. A notable finding is that, for high viscosities (roughly $\alpha \sim \gamma \Omega/P > 10^{-3}$, with $\Omega$ the angular frequency and $P$ the pressure), outflows driven by viscous heating may be a major source of expelled matter, with outflow masses comparable to that of the dynamical ejecta.

Both of the above simulations use artificial initial data. Recently, Radice [610] has performed neutron star–neutron star binary merger simulations using a subgrid turbulence model very similar to a viscosity. Because the cores of neutron star–neutron star binary remnants are slowly rotating, transport effects spin up the core but spin down the inner envelope, so that collapse can be delayed or accelerated, depending on the strength of the effective viscosity. Meanwhile, Shibata and Kiuchi [611] find that the effect of viscosity (with a reasonable $\alpha \sim 10^{-3}$) on nonaxisymmetric deformations is dramatic, with these and their corresponding gravitational wave signals damping on a viscous timescale of around 5 ms.

Subscale effects can also affect the large-scale magnetic field (e.g. the ‘alpha effect’ in dynamo theory). Giacomazzo et al. [612] have taken a step to incorporate these effects in numerical relativity, adding a subgrid EMF to the induction equation. The added term is designed to grow the magnetic field to equipartition with the turbulent kinetic energy (mean-field term) on a viscous timescale of around 5 ms.

Simulations with subgrid terms depend on the reliability of the subgrid model, an assumption that cannot easily be relaxed, but they do allow long-term evolutions including effects that would otherwise be inaccessible.

### 7. Comparison to observations

#### 7.1. Gravitational wave astronomy

To date, the LIGO and Virgo collaborations have announced five confirmed neutron star binary merger observations, GW150914 [4], GW151226 [6], GW170104 [7], GW170608 [8], and GW170817 [9]. The data itself does not, at the time of writing, exclude the possibility that one component was a black hole, but known black hole formation scenarios would not produce black holes of such low mass.

The exact allowed range of masses depends on whether one allows the possibility of large binary component spins. If spins are assumed to be not larger than those observed Galactic binary neutron stars, the component masses are inferred to lie in the range $1.16\pm1.62 \, M_\odot$. Allowing for large spins, the range expands to $1.00\pm1.89 \, M_\odot$ [618]. The data itself does not, at the time of writing, exclude the possibility that one component was a black hole, but known black hole formation scenarios would not produce black holes of such low mass.

On August 17, 2017, LIGO-Virgo made the first gravitational wave detection of a late-inspiral neutron star–neutron star binary system, labeled GW170817 (see figure 21) [10]. The identification as a neutron star–neutron star binary system could be made from the masses of the binary components, which from the waveform were estimated to be in the range $1.17\sim1.60 \, M_\odot$, with a total mass of $M_{\text{total}} = 2.73\sim2.78 \, M_\odot$ and mass ratio in the range $0.7\sim1.9$.

The GRB was unusually dim (isotropic) luminosity $L \sim 10^{57} \text{ erg s}^{-1}$, perhaps due to some combination of being seen off-axis and various forms of interaction between the jet and enveloping matter ejected during merger. Comparisons to numerical relativity results have mostly focused on the optical and infrared AT2017gfo signal, which strongly resembles an anticipated kilonova signal. Recall that, for the $v \sim 0.1c$ outflows expected from neutron star–neutron star binary mergers, neutron-rich $Y_e < 0.25$ outflow will synthesize lanthanide elements, have high opacity, and is expected to peak after a week in the near infrared. Less neutron-rich outflow (perhaps made so by neutrino processing) will have lower opacity and is expected to peak after about a day in optical wavelengths. AT2017gfo showed signs of both signals, an early UV-blue signal with a near-IR tail [619], which can naturally be explained if outflows of both kinds are present, and the high emissions consistent with the predictions of general relativity.

GW170817: The age of multimessenger astronomy begins

19The exact allowed range of masses depends on whether one allows the possibility of large binary component spins. If spins are assumed to be not larger than those observed Galactic binary neutron stars, the component masses are inferred to lie in the range $1.16\sim1.62 \, M_\odot$. Allowing for large spins, the range expands to $1.00\sim1.89 \, M_\odot$ [618]. The data itself does not, at the time of writing, exclude the possibility that one component was a black hole, but known black hole formation scenarios would not produce black holes of such low mass.
opacity material does not completely occult the low opacity material. Kilonova models can accommodate the observed model with two components to the ejecta: a $M = 0.01 M_\odot$, $v = 0.3c$ lanthanide-poor ejecta for the ‘blue’ component and a $M = 0.04 M_\odot$, $v = 0.1c$ lanthanide-rich ejecta for the ‘red’ component [620–622].

These numbers can be rather directly compared to numerical predictions. A first conclusion is that dynamical ejecta from tidal forces and the collision shock are inadequate. For the range of realistic EoS, the total dynamical ejecta mass does not exceed about $0.02 M_\odot$ [623, 624]. More ejecta can be produced by outflows from the stellar remnant or surrounding accretion disk. Numerical relativity has already told us something interesting: the merged object must be of a kind to give such outflows.

Next, numerical relativity provides another crucial piece of information. If the remnant collapses promptly, it will leave a black hole and very low-mass disk, lower than the remaining mass that needs to be ejected. (A warning is in order here. If the binary was very asymmetric, of the order $q \approx 0.7$, it may be possible to get massive disks even in a high-mass, prompt-collapse scenario. The inferences below mostly assume that this was not the case.) Therefore, it is surmised that the remnant did not promptly collapse. This means the binary’s mass, which we know, is below the threshold mass for prompt collapse, which is loosely connected to the neutron star maximum mass. Alternatively, the need to avoid prompt collapse and small disk led Bauswein et al [646] to set a lower limit on the radius of a $1.6 M_\odot$ neutron star of 10.6 km and Radice et al [624] to place a lower limit on the tidal deformability of $\Lambda > 400$, both based on numerical relativity neutron star–neutron star binary simulations with a variety of realistic EoS.

Next, we may ask whether the remnant was above or below the supramassive–hypermassive cutoff mass. (See section 1.4 above.) Arguments have been made to the effect that the remnant must have suffered delayed collapse, meaning that it was above this mass. First, a long-lived magnetar would have released energy via dipole radiation at $L \sim 10^{50} \text{ erg s}^{-1} (B/10^{15} \text{ G})^2$, which would accelerate the ejecta to $v \approx c$ and produce bright emissions not matching observations [625]. Second, the collapse to a black hole with disk may be needed to explain the GRB. The supramassive mass limit is related to the TOV mass limit; the former is about 20% larger than the latter. (Note, though, that thermal effects effectively alter the EoS and hence the supramassive limit, an additional source of uncertainty.) Reasoning of this sort has been used in several papers [625] to set an upper limit to the neutron star maximum mass in the range $2.16–2.28 M_\odot$ [625, 626].

Finally, there may be clues in the presence of the blue kilonova. Shibata et al [623] compare merger simulations with the stiff DD2 EoS with those of the soft SFHo EoS, combining 3D merger simulations with 2D viscous simulations of the subsequent $\sim$ s of secular evolution, with neutrino transport included in both. SFHo has a lower maximum mass and so predicts prompt collapse, while DD2 results in a long-lived stellar remnant. Viscosity drives ejecta from both the hypermassive remnant (for DD2) and the disk. However, a long-lived remnant seems to be necessary to provide the neutrino irradiation needed for a high-$Y_e$ outflow component that produces the blue kilonova. Once again, the kilonova combined with numerical relativity constrains the EoS generally in the direction of excluding soft EoS. However, the EoS-related parameter to which the neutron star–neutron star binary outcome is most sensitive is the maximum mass.

8. Conclusion

The birth of gravitational wave astronomy presents a remarkable story of long-term planning and investment, with numerical relativity being only one of the fields built up largely in
anticipation of discoveries known to be decades away. The investment of time, money, and careers has now been vindicated. Although it is more broadly useful, the majority of effort in numerical relativity has been devoted where it was needed by LIGO-Virgo, to the three types of compact object binaries. For each type, numerical relativity has established some definite results that are beyond the reach of Newtonian physics or perturbation theory: the possibility of super-kicks in black hole–black hole binary mergers and the threshold mass for prompt collapse in neutron star–neutron star binary merger, to name just two. The study of these systems is not yet finished. Of the three system types, the simulation of black hole–black hole binaries is the most mature, but here also the binary parameter space is most intimidating, and the application of numerical relativity to devising templates for the full 7D space is ongoing work. Neutron star–neutron star and black hole–neutron star binary simulations are less accurate, and their ability to capture the multi-scale magnetic and neutrino effects of the post-merger evolution might not even be qualitatively adequate.

However, even where numerical relativity simulations of black hole–neutron star binary and neutron star–neutron star binary mergers are inadequate, there is no longer anything distinctively relativistic about the problems. Newtonian simulations of neutron star–neutron star binary mergers are no more advanced; they face all the same difficulties. In fact, Newtonian simulations of these systems are becoming less common. Since it is not much more difficult, why not just work in general relativity? With the advent of open-source, publicly available numerical relativity codes such as the Einstein Toolkit [41, 145], the barrier to an interested astrophysicist doing numerical relativity work has never been lower. This is for the best. One could say that the goal of numerical relativity all along has been to abolish itself as a distinct subfield, to make solving the Einsteins equations as routine as solving Poisson’s equation, and thus to dissolve into computational astrophysics.

And yet, numerical relativity will also remain as a tool for addressing questions in gravitational physics. Numerical experiments are used to investigate features of black hole physics such as cosmic censorship and the generic structure of spacetime singularities [627]. Perhaps more important, as gravitational wave detections grow in number and accuracy, numerical relativity will aid in testing alternative theories of gravity. Ultimately, to test general relativity (and—dare we hope?—supersede it), we will need to compare general relativity predictions, say of black hole–black hole binary mergers, with more general possibilities, to identify the signatures of new physics that cannot be reproduced in general relativity by tinkering with parameters in the vast 7D black hole–black hole binary parameter space. A systematic generalization of general relativity (analogous to the parameterized post-Newtonian formalism) does not exist, and any generalization will expand an already prohibitive parameter space. Nevertheless, exploratory simulations using scalar-tensor [628] and Chern–Simons [629] theories of gravity have already been carried out. Neutron star systems are ‘messier’, with EoS uncertainties mixing with gravitational uncertainties, but at least in scalar-tensor theories, they also allow distinct signatures if the neutron star undergoes spontaneous scalarization [630], so these systems may be useful gravity test probes as well.

When all else is done, numerical relativity results leave us with a problem of their own, the problem of their interpretation. The asymptotic gravitational wave output is gauge invariant, but the dynamics of the interior is encoded in tensor functions on grids in an evolved coordinate system. It is actually remarkable that the movies numerical relativists produce of our mergers look so qualitatively reasonable, an artifact of gauges designed to minimize unnecessary coordinate dynamics. When one wants to know something concrete about the strong-field region, the difficulty of disentangling coordinate effects becomes acute. Even to prove that the result of a black hole–black hole binary merger settles to a Kerr spacetime is a surprisingly intricate affair [631]. Apparent horizons are foliation-dependent. Event horizons are not, but even they will be visualized on some arbitrary coordinate system. Ray tracing can be used to reconstruct what a nearby observer would actually see from a black hole–black hole binary merger [632]. To try to give some intuition for the physics of the merger, Nichols et al [633] have proposed plotting (gauge-dependent) field lines to illustrate the local tidal stretches and twists. The challenge of making sense of a background-independent field theory extends more widely in theoretical physics, but it confronts us in a particularly concrete form in numerical relativity.

Acknowledgments

We thank Andreas Bauswein, John Baker, Manuela Campanelli, Koutarou Kyutoku, Carlos Lousto, Luciano Rezzolla, Richard O’Shaughnessy, Frans Pretorius, and Masaru Shibata for careful reading of this manuscript. We thank the referees and the editorial board member for their many helpful suggestions. MD gratefully acknowledge the NSF for financial support from Grant PHY-1806207. YZ gratefully acknowledge the NSF for financial support from Grants No. PHY-1607520, No. PHY-1707946, No. PHY-1726215, No. OAC-1811228, No. ACI-1550436, No. AST-1516150, and No. ACI-1516125.

ORCID iDs

Matthew D Duez https://orcid.org/0000-0002-0050-1783
Yosef Zlochower https://orcid.org/0000-0002-7541-6612

References

[1] Pretorius F 2005 Phys. Rev. Lett. 95 121101
[2] Campanelli M, Lousto C O, Marronetti P and Zlochower Y 2006 Phys. Rev. Lett. 96 111101
[3] Baker J G, Centrella J, Choi D I, Koppitz M and van Meter J 2006 Phys. Rev. Lett. 96 111102
[4] Abbott B P et al (Virgo, LIGO Scientific) 2016 Phys. Rev. Lett. 116 061102
[5] Abbott B P et al (Virgo, LIGO Scientific) 2016 Phys. Rev. Lett. 116 241102
[6] Abbott B P et al (Virgo, LIGO Scientific) 2016 Phys. Rev. Lett. 116 241103
[7] Abbott B P et al (Virgo, LIGO Scientific) 2017 Phys. Rev. Lett. 118 221101
[8] Abbott B P et al (Virgo, LIGO Scientific) 2017 Astrophys. J. 851 L35
[9] Abbott B P et al (Virgo, LIGO Scientific) 2017 Phys. Rev. Lett. 119 141101
[10] Abbott B P et al (Virgo, LIGO Scientific) 2017 Phys. Rev. Lett. 119 161101
[11] Abbott B P et al (GROND, SALT Group, OzGrav, DFN, INTEGRAL, Virgo, Insight-Hxmt, MAXI Team, Fermi-LAT, J-GEM, RATIR, IceCube, CAASTRO, LWA, ePESSTO, GRAWITA, RIMAS, SKA South Africa/ MeerKAT, H.E.S.S., 1M2H Team, IKIGW Follow-up, Fermi GBM, Pi of Sky, DWF (Deeper Wider Fainter Future Program), Dark Energy Survey, MASTER, Astrosat Cadmium Zinc Telluride Imager Team, Swift, Pierre Auger, ASKAP, VINROUGE, JAGWARK, Chandra Team at McGill University, TTU-NRAO, GROWTH, AGILE Team, MWA, ATCA, AST3, TOROS, Pan-STARRS, NuSTAR, ATLAS Telescopes, BOOTES, CaltechNRAO, LIGO Scientific, High Time Resolution Universe Survey, Nordic Optical Telescope, Las Cumbres Observatory Group, TZAC Consortium, LOFAR, IPN, DLT40, Texas Tech University, HAWC, ANTARES, KU, Dark Energy Camera GWEM, CALET, Euro VLBI Team, ALMA) 2017 Astrophys. J. 848 L12
[12] Abbott B P et al (Virgo, Fermi-GBM, INTEGRAL, LIGO Scientific) 2017 Astrophys. J. 848 L13
[13] Hartle J B (ed) 2003 Gravity (San Francisco, CA: Addison-Wesley) XXII+582p
[14] Schutz B 2009 A First Course in General Relativity by Bernard Schutz (Cambridge: Cambridge University Press)
[15] Carroll S M (ed) 2004 Spacetime and Geometry (San Francisco, CA: Addison-Wesley)
[16] Wald R M 1984 (Chicago, IL: University of Chicago Press) 504
[17] Schwarzschild K 1916 Sitzungsber. Dtsch. Akad. Wiss. Berlin, Kl. Math. Phys. Tech.
[18] Schwarzschild K 1916 Sitzungsber. Dtsch. Akad. Wiss. Berlin, Kl. Math. Phys. Tech.
[19] Kerr R P 1963 Phys. Rev. Lett. 11 237
[20] Thorne J 2007 Living Rev. Relativ. 10 3
[21] Tolman R C 1939 Phys. Rev. 55 364
[22] Oppenheimer J R and Volkoff G M 1939 Phys. Rev. 55 374
[23] Font J A, Goodale T, Iyer S, Miller M A, Rezzolla L, Seidel E, Stergioulas N, Suen W M and Tobias M 2002 Phys. Rev. D 65 084024
[24] Cook G B, Shapiro S L and Teukolsky S A 1994 Astrophys. J. 422 227
[25] Stergioulas N and Friedman J 1995 Astrophys. J. 444 306
[26] Cook G B, Shapiro S L and Teukolsky S A 1994 Astrophys. J. 424 823
[27] Friedman J L, Ipser J R and Sorkin R D 1988 Astrophys. J. 325 722
[28] Takami K, Rezzolla L, Yoshida S-I 2011 Mon. Not. R. Astron. Soc. 416 L1–5
[29] Baumgarte T W, Shapiro S L and Shibata M 2000 Astrophys. J. 528 L29
[30] Gourgout D J O, Haensel P and Zdunik J L 1997 Astron. Astrophys. 321 822
[31] Kaplan J D, Ott C D, O'Connor E P, Kiuchi K, Roberts L and Duez M 2014 Astrophys. J. 790 19
[32] Weih L R, Most E R and Rezzolla L 2018 Mon. Not. R. Astron. Soc. 473 L126
[33] Bauswein A and Stergioulas N 2017 Mon. Not. R. Astron. Soc. 471 4956
[34] Arnowitt R, Deser S and Misner C W 1962 Gravitation: an Introduction to Current Research ed L Witten (New York: Wiley) pp 227–65
[35] Baumgarte T W and Shapiro S L 2010 Numerical Relativity: Solving Einstein’s Equations on the Computer (Cambridge: Cambridge university Press)
[36] Alcubierre M (ed) 2008 Introduction to 3 + 1 Numerical Relativity (Oxford: Oxford University Press)
[37] Shibata M (ed) 2016 Numerical Relativity (Singapore: World Scientific Publishing Co)
[38] York J W Jr 1979 Sources of Gravitational Radiation ed L L Smarr pp 83–126
[39] Husa S, Gonzalez J A, Hannam M, Bruegmann B and Sperhake U 2008 Class. Quantum Grav. 25 105006
[40] Lousto C O and Zlochower Y 2008 Phys. Rev. D 77 024034
[41] Löffler F et al 2012 Class. Quantum Grav. 29 115001
[42] Gustafsson B, Kreiss H O and Oliger J (ed) 1995 Time-Dependent Problems and Difference Methods 2nd edn (New York: Wiley-Interscience)
[43] Baker J G, Campanelli M and Lousto C O 2002 Phys. Rev. D 65 044001
[44] Zlochower Y, Baker J G, Campanelli M and Lousto C O 2005 Phys. Rev. D 72 024021
[45] Zilhão M and Noble S C 2014 Class. Quantum Grav. 31 065013
[46] Berger M J and Oliger J 1984 J. Comput. Phys. 53 484
[47] Schnetter E, Hawley S H and Hawke I 2004 Class. Quantum Grav. 21 1465
[48] Imbiriba B, Baker J, Choi D I, Centrella J, Fiske D R, Brown J D, van Meter J R and Olson K 2004 Phys. Rev. D 70 124025
[49] Bruegmann B, Gonzalez J A, Hannam M, Husa S, Sperhake U and Tichy W 2008 Phys. Rev. D 77 024027
[50] Anderson M, Hirschmann E W, Lehner L, Liebling S L, Motl P M, Neilsen D, Palenzuela C and Tohline J E 2008 Phys. Rev. D 77 024006
[51] Courant R, Friedrichs K and Lewy H 1928 Math. Ann. 100 32
[52] Kidder L E, Scheel M A, Teukolsky S A, Carlson E D and Cook G B 2000 Phys. Rev. D 62 084032
[53] Boyle M, Lindblom L, Pfeiffer H, Scheel M and Kidder L E 2007 Phys. Rev. D 75 024006
[54] Scheel M A, Pfeiffer H P, Lindblom L, Kidder L E, Rinne O and Teukolsky S A 2006 Phys. Rev. D 74 104006
[55] Ansorg M, Bruegmann B and Tichy W 2004 Phys. Rev. D 70 064011
[56] Tichy W 2006 Phys. Rev. D 74 084005
[57] Boyd J P 1999 Chebyshev and Fourier Spectral Methods 2nd edn (New York: Dover Publications)
[58] Fornberg B 1996 A Practical Guide to Pseudospectral Methods (New York: Cambridge University Press)
[59] Wald R M 1984 General Relativity (Chicago, IL: Chicago University Press)
[60] Brodbeck O, Fritelli S, Hubner P and Reula O A 1999 J. Math. Phys. 40 909
[61] Tiglio M, Lehner L and Neilsen D 2004 Phys. Rev. D 70 104018
[62] Calabrese G, Pullin J, Sarbach O, Tiglio M and Reula O 2003 Commun. Math. Phys. 240 377
[63] Sarbach O and Tiglio M 2002 Phys. Rev. D 66 064023
[64] Calabrese G, Lehner L and Tiglio M 2002 Phys. Rev. D 65 104031
[65] Calabrese G, Pullin J, Sarbach O and Tiglio M 2002 Phys. Rev. D 66 064011
[66] Kelly B J, Laguna P, Lockitch K, Pullin J, Schnetter E, Shoemaker D and Tiglio M 2001 Phys. Rev. D 64 084013
[67] Frittelli S 2006 Phys. Rev. D 73 124001
[68] Frittelli S 2004 Phys. Rev. D 70 044029
[69] Frittelli S and Gomez R 2000 J. Math. Phys. 41 5535
[70] Neilsen D, Lehner L, Sarbach O and Tiglio M 2006 Lect. Notes Phys. 692 223–49 (Proc. 319th WE-Heraeus Seminar: Mathematical Relativity: New Ideas and Developments (Bad Honnef, Germany, 1–5 March 2004))
[495] Kawaguchi K, Kyutoku K, Nakano H, Owaka H, Shibata M and Taniguchi K 2015 Phys. Rev. D 92 024014

[496] Belczynski K, Taam R E, Rantsiou E and van der Sluys M 2008 Astroph. J. 682 474

[497] Kyutoku K, Shibata M and Taniguchi K 2010 Phys. Rev. D 82 044049

[498] Kyutoku K, Shibata M and Taniguchi K 2011 Phys. Rev. D 84 049902 (Erratum)

[499] Lackey B D, Kyutoku K, Shibata M, Brady P R and Friedman J L 2014 Phys. Rev. D 89 043009

[500] Panzarale F, Bert E, Kyutoku K and Shibata M 2013 Phys. Rev. D 88 084011

[501] Farrar R E, Baker J, Cabezue J, Rosswog S, Davies M B, Thielemann F K, Piran T 2000 Phys. Rev. D 61 044049

[502] Etienne Z B, Paschalidis V and Shapiro S L 2012 Mon. Not. R. Astron. Soc. 435 502

[503] Ferrand D, Fernández M and Metzger B D 2013 Mon. Not. R. Astron. Soc. 449 390

[504] Berti E, Kyutoku K, Shibata M and Taniguchi K 2015 Phys. Rev. D 92 084050

[505] Wilson J R and Mathews G J 1995 Phys. Rev. Lett. 75 4161

[506] Wilson J R, Mathews G J and Marrone G P 1996 Phys. Rev. D 54 1317

[507] Flanagan E E 1999 Phys. Rev. Lett. 82 1354

[508] Oechslin R, Rosswog S and Thielemann F K 2002 Phys. Rev. D 65 103005

[509] Ruffert M, Janka H T and Schaefer G 1997 Astron. Astrophys. 311 532

[510] Ruffert M, Janka H T, Takahashi K and Schaefer G 1997 Astron. Astrophys. 319 122

[511] Ruffert M and Janka H T 2001 Astron. Astrophys. 380 544

[512] Rosswog S and Davies M B 2003 Mon. Not. R. Astron. Soc. 345 1077

[513] Rosswog S and Liebendörfer M 2003 Mon. Not. R. Astron. Soc. 342 673

[514] Rosswog S, Ramirez-Ruiz E and Davies M B 2003 Mon. Not. R. Astron. Soc. 345 1077

[515] Dessart L, Ott C, Burrows A, Rosswog S and Livne E 2009 Astrophys. J. 690 1681

[516] Perego A, Rosswog S, Cabezon R M, Korobkin O, Keppler R, Arcones A and Liebendörfer M 2014 Mon. Not. R. Astron. Soc. 443 3134

[517] Rosswog S and Liebendörfer M, Thielemann F K, Davies M B, Benz W and Piran T 1999 Astron. Astrophys. 341 499

[518] Freiburghaus C, Rosswog S and Thielemann F K 1999 Astrophys. J. Lett. 525 L121

[519] Rosswog S, Davies M B, Thielemann F K and Piran T 2000 Astron. Astrophys. 360 171

[520] Shibata M and Uryu K 2000 Phys. Rev. D 61 064001

[521] Shibata M and Uryu K 2003 Phys. Rev. D 68 084020

[522] Baiotti L, Giacomazzo B and Rezzolla L 2008 Phys. Rev. D 78 084033

[523] Bernuzzi S, Thierfelder M and Bruegmann B 2012 Phys. Rev. D 85 104030

[524] Shibata M, Taniguchi K and Uryu K 2005 Phys. Rev. D 71 084021

[525] Hotokezaka K, Kyutoku K, Okawa H, Shibata M and Taniguchi K 2011 Phys. Rev. D 83 124008

[526] Hotokezaka K, Kyutoku K, Okawa H, Shibata M and Taniguchi K 2013 Phys. Rev. D 87 024001
[566] Read J S, Baiotti L, Creighton J D E, Friedman J L, Giacomazzo B, Kyutoku K, Markakis C, Rezzolla L, Shibata M and Taniguchi K 2013 Phys. Rev. D 88 044042

[567] Takami K, Rezzolla L and Baiotti L 2014 Phys. Rev. Lett. 113 091104

[568] Kiuchi K, Sekiguchi Y, Shibata M and Taniguchi K 2009 Phys. Rev. D 80 064037

[569] Rezzolla L, Baiotti L, Giacomazzo B, Link D and Font J A 2010 Class. Quantum Grav. 27 114105 (Microphysics in Computational Relativistic Astrophysics. Proc., Workshop (Copenhagen, Denmark, 24–28 August 2009))

[570] Giacomazzo B, Perna R, Rezzolla L, Troja E and Lazzati D

[571] Sekiguchi Y, Kiuchi K, Kyutoku K and Shibata M 2011 Phys. Rev. Lett. 107 211101

[572] Bauswein A, Baumgarte T W and Janka H T 2013 Phys. Rev. Lett. 111 131101

[573] Kastaun W and Stergioulas N 2015 Mon. Not. R. Astron. Soc. 457 4150

[574] Kawamura T, Giacomazzo B, Kastaun W, Ciolfi R, Endrizzi A, Ciolfi R, Giacomazzo B, Kastaun W and Kiuchi K, Kyutoku K, Sekiguchi Y, Shibata M and Wada T 2013 Astrophys. J. 762 L18

[575] Bauswein A, Janka H T, Hebeler K and Schwenk A 2012 Phys. Rev. D 85 044012

[576] Stergioulas N, Bauswein A, Zagkouris K and Janka H T 2011 Mon. Not. R. Astron. Soc. 418 L6

[577] Ruiz M, Lang R N, Paschalidis V and Shapiro S L 2016 Astrophys. J. 824 L6

[578] Bauswein A and Janka H T 2012 Phys. Rev. Lett. 108 011101

[579] Ruiz M, Lang R N, Paschalidis V and Shapiro S L 2016 Astrophys. J. 824 L6

[580] Stergioulas N, Bauswein A, Zagkouris K and Janka H T 2011 Mon. Not. R. Astron. Soc. 418 427

[581] Radice D, Bernuzzi S and Ott C D 2016 Phys. Rev. D 94 044011

[582] Lehner L, Liebling S L, Palenzuela C and Motl P M 2016 Phys. Rev. D 94 044003

[583] Bauswein A and Stergioulas N 2015 Phys. Rev. D 91 124056

[584] Palenzuela C, Liebling S L, Neilsen D, Lehner L, Caballero O L, O’Connor E and Anderson M 2015 Phys. Rev. D 92 044045

[585] Paschalidis V, Etienne Z B and Shapiro S L 2012 Phys. Rev. D 86 064032

[586] Siegel D M, Ciolfi R and Rezzolla L 2014 Astrophys. J. 785 L6

[587] Duez M D, Liu Y T, Shapiro S L and Stephens B C 2004 Phys. Rev. D 69 104030

[588] Shibata M, Kiuchi K and Sekiguchi Y I 2017 Phys. Rev. D 95 083005

[589] Israel W and Stewart J M 1979 Ann. Phys. 118 341

[590] Radice D 2017 Astrophys. J. 838 L2

[591] Shibata M and Kiuchi K 2017 Phys. Rev. D 95 123003

[592] Giacomazzo B, Zrake J, Duffell P, MacFadyen A I and Perna R 2015 Astrophys. J. 809 39

[593] Zrake J and MacFadyen A I 2015 Astrophys. J. 769 L29

[594] Abbott B P et al (Virgo, LIGO Scientific) 2016 Phys. Rev. X 6 041015

[595] Lange J et al 2017 Phys. Rev. D 96 104041

[596] Healy J et al 2018 Phys. Rev. D 97 064027

[597] Abbott B P et al (Virgo, LIGO Scientific) 2016 Phys. Rev. Lett. 116 221101

[598] Abbott B P et al (Virgo, LIGO Scientific) 2018 arXiv:1805.11579 [gr-qc]

[599] Soares-Santos M et al (Dark Energy Camera GW-EM, DES) 2017 Astrophys. J. 848 L16

[600] Cowperthwaite P S et al 2017 Astrophys. J. 848 L17

[601] Nicholl M et al 2017 Astrophys. J. 848 L18

[602] Chornock R et al 2017 Astrophys. J. 848 L19

[603] Shibata M, Fujibayashi S, Hotokezaka K, Kiuchi K, Yukutoku K, Sekiguchi Y and Tanaka M 2017 Phys. Rev. D 96 123012

[604] Radice D, Perego A, Zappa F and Bernuzzi S 2018 Astrophys. J. 852 L29

[605] Margalit B and Metzger B D 2017 Astrophys. J. 850 L19

[606] Ruiz M, Shapiro S L and Tsoharos A 2018 Phys. Rev. D 97 021501

[607] Berger B K 2002 Living Rev. Relativ. 5 1

[608] Healy J, Bode T, Haas R, Pagos E, Laguna P, Shoemaker D M and Yunus N 2012 Class. Quantum Grav. 29 232002

[609] Okounkova M, Stein L C, Scheel M A and Hemberger D A 2017 Phys. Rev. 96 044020

[610] Damour T and Esposito-Farese G 1996 Phys. Rev. D 54 1474

[611] Campanelli M, Lousto C O and Zlochower Y 2009 Phys. Rev. D 79 084012

[612] Bonn A, Throwe W, Hébert F, Henriksson K, Bunandar D, Scheel M A and Taylor N W 2015 Class. Quantum Grav. 32 065002

[613] Nichols D A et al 2011 Phys. Rev. D 84 124014

[614] Rezzolla L and Zanotti O (ed) 2013 Relativistic Hydrodynamics (Oxford: Oxford University Press)
Matthew D Duez, is an associate professor at Washington State University. He received his PhD from the University of Illinois at Urbana-Champaign in 2005. He then took a postdoctoral position at Cornell University, where he joined the SXS collaboration. His current position at WSU began in 2010. His research focuses on the application of numerical relativity to non-vacuum (hydrodynamic) systems such as compact binaries in which at least one of the components is a neutron star.

Yosef Zlochower, is an associate professor at the Rochester Institute of Technology. He received his PhD from the University of Pittsburgh in 2002 and held postdoctoral positions at the University of Pittsburgh, the University of Texas at Brownsville, and the Rochester Institute of Technology. He has spent the past decade studying numerical simulations of black-hole spacetimes and was one of the principle developers of the breakthrough “moving punctures” approach for evolving multiple black holes in full numerical relativity. His research focuses on binary dynamics and generating waveforms for LIGO data analysis.