TOPICAL REVIEW • OPEN ACCESS

A review of gravitational memory and BMS frame fixing in numerical relativity

To cite this article: Keefe Mitman *et al* 2024 *Class. Quantum Grav.* **41** 223001

OPEN ACCESS
IOP Publishing
Classical and Quantum Gravity
Class. Quantum Grav. **41** (2024) 223001 (40pp) https://doi.org/10.1088/1361-6382/ad83c2

Topical Review

A review of gravitational memory and BMS frame fixing in numerical relativity

Keefe Mitman[1,*], Michael Boyle[2], Leo C Stein[3], Nils Deppe[2,4], Lawrence E Kidder[2], Jordan Moxon[1], Harald P Pfeiffer[5], Mark A Scheel[1], Saul A Teukolsky[1,2], William Throwe[2] and Nils L Vu[1]

[1] Theoretical Astrophysics, Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125, United States of America
[2] Cornell Center for Astrophysics and Planetary Science, Cornell University, Ithaca, NY 14853, United States of America
[3] Department of Physics and Astronomy, University of Mississippi, University, MS 38677, United States of America
[4] Department of Physics, Cornell University, Ithaca, NY 14853, United States of America
[5] Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, D-14476 Potsdam, Germany

E-mail: kmitman@caltech.edu

Received 14 May 2024; revised 30 August 2024
Accepted for publication 1 October 2024
Published 22 October 2024

Abstract

Gravitational memory effects and the BMS freedoms exhibited at future null infinity have recently been resolved and utilized in numerical relativity simulations. With this, gravitational wave models and our understanding of the fundamental nature of general relativity have been vastly improved. In this paper, we review the history and intuition behind memory effects and BMS symmetries, how they manifest in gravitational waves, and how controlling the infinite number of BMS freedoms of numerical relativity simulations can crucially improve the waveform models that are used by gravitational wave detectors. We reiterate the fact that, with memory effects and BMS symmetries, not only can these next-generation numerical waveforms be used to observe

* Author to whom any correspondence should be addressed.

never-before-seen physics, but they can also be used to test GR and learn new astrophysical information about our Universe.

Keywords: numerical relativity, gravitational waves, gravitational memory, BMS

1. Introduction

One of the most pressing challenges for physics in the near future is performing stringent and robust tests of Einstein's theory of general relativity (GR). These tests are of the utmost importance because they will inform us about the nature of gravity within our Universe and will reveal when our long-standing theory of GR fails to explain real world phenomena. At present, the most prospective tests of GR that we can perform are those which involve analyses of the gravitational waves (GWs) that are created by binary black hole mergers (BBHs)[6]. This is because the GWs that are produced by BBHs are largely influenced by the strong-gravity regimes sourced by two coalescing black holes and should thus capture whatever deviations from GR there may be. However, to verify whether certain features in observed GWs are evidence for unknown physics, we first need to have a sound understanding of the GWs that GR predicts.

At present, the best solutions to Einstein's equations, i.e. the GW templates used to perform tests of GR, are those produced by numerical relativity (NR) simulations. Calculating the GWs sourced during the coalescence of two black holes is impossible to work out with pen and paper due to the overall complexity of the partial differential equations that need to be solved.

Furthermore, even if one uses perturbation theory to try to predict the GWs, this fails to produce reliable results during the merger phase of the binary, which is typically the loudest and most detectable part of the GW signal. Consequently, GW models that are built by using perturbation theory or GW phenomenology, like effective one-body (EOB) or Phenomenological (Phenom) GW models [1–11], always need to be calibrated against NR waveforms. Thus, NR simulations, which can achieve arbitrary accuracy with the right computational tools, are at the heart of producing accurate and robust solutions for GW-emitting phenomena in GR.

Nonetheless, NR can still fail to accurately simulate GR if the code infrastructure is not formulated correctly or if the necessary numerical resolution is not achieved. One such example of this inaccuracy was the inability of NR simulations to resolve a collection of observables in GR colloquially referred to as memory effects [12–14]. These effects are not-yet observed, nonlinear predictions of GR that physically correspond to the net displacement that two freely-falling observers will experience due to the passage of transient GWs. However, apart from being a curious prediction of GR, what makes memory effects particularly tantalizing is that they are intimately tied—through conservation laws—to the symmetry group of future null infinity—part of the asymptotic boundary of spacetime. This symmetry group is not the usual Poincaré group of special relativity, but is a larger group called the BMS group, named after Bondi, van der Burg, Metzner, and Sachs [15, 16]. Thus, there is hope that with the detection of memory we can not only conduct more stringent tests of GR, but we may even obtain a better understanding of the asymptotic structure of our encompassing Universe, which is of

[6] While the imaging of black holes, such as those performed by the EHT Collaboration, can also test Einstein's theory of GR, this typically probes much lower curvatures than are accessible by the LIGO-Virgo-KAGRA Collaboration and thus probe alternative, but also complementary, regimes of GR.

immense interest to theorists trying to formulate a theory of quantum gravity through topics such as celestial holography [17–21].

In this review, we highlight recent advancements made in the NR community to resolve memory effects as well as some work showing how fixing the BMS freedoms at future null infinity drastically improves both the accuracy and robustness of GW models and analyses. Specifically, in section 2 we begin by providing some motivation for and intuition behind the BMS group and memory effects. Next, in section 3 we provide a review of the literature on memory, the BMS group, and BMS frame fixing. Then, in section 4, we provide a more formal explanation of the origins of the BMS group and how memory effects can be understood as stemming from certain conservation laws related to the symmetries of null infinity. In section 5, we transition to a review of the code frameworks used to compute GWs at null infinity and we highlight the advancements in the NR community that has made the resolution of memory effects possible. Then, in section 6, we demonstrate the formalism from section 4 using BBH simulations. In particular, we present how the BMS conservation laws can be used to efficiently analyze gravitational waves and understand memory effects. Furthermore, in this section we also provide a review of the detectability of memory and the forecast for its detection in the coming decade. In section 7 we highlight the often-overlooked importance of fixing the BMS freedom of NR waveforms to ensure that modeling of NR waveforms is performed accurately and robustly. We show this by introducing the *superrest frame*, and demonstrating its utility by comparing NR waveforms to post-Newtonian waveforms and by fitting NR waveforms with predictions made by black hole perturbation theory. Finally, in section 8 we summarize the main points of this review and provide some outlook regarding the future of NR, memory effects, and testing the nature of gravity with GW.

2. Pedagogical approach to BMS and memory

Despite its importance and interesting characteristics, many relativists are not familiar with the BMS group or its effect on asymptotic data. In this section, we provide a pedagogical introduction to the BMS group, with the intention of making the rest of the paper more accessible. We begin by motivating the need for a coordinate system that is adapted to inertial observers, e.g. GW detectors, and discuss how such a coordinate system is provided by 'Bondi gauge'—in which the metric asymptotes to the usual Minkowski metric at large radius. This is crucial because GW waveforms are always studied in a certain coordinate system, so to provide meaningful waveforms we need a meaningful coordinate system that matches that of our inertial detectors. As we will see, it turns out that once such a coordinate system is constructed by mapping the metric to Bondi gauge, there is a residual ambiguity in the coordinates, i.e. a symmetry, which is described by the BMS group. The BMS group, however, is simply the usual Lorentz group, augmented by a generalized class of spacetime translations called *supertranslations* [15, 16]. Consequently, it is fairly straightforward to understand the BMS group once the origin of and intuition behind supertranslations is understood. Thus, once we motivate the need for a meaningful coordinate system, we will then provide some intuition for supertranslations through a few informative examples involving null rays propagating in Minkowski space. Following this, we then conclude this pedagogical overview by showing how the BMS group changes the asymptotic data that can be measured by an inertial observer. This action forms the basis for fixing the BMS frame in NR, as outlined in section 4, and also provides a unique way to study memory effects, which we utilize in section 2.6 and in the rest of the paper.

2.1. Motivation

Choosing coordinates in GR is one of the more delicate and, at times, confusing components of Einstein's theory. In fact, for decades after GR's development, researchers—including Einstein—wavered over the issue of whether or not GW were really physically observable or simply gauge artifacts [22]. Ultimately, the reason for their misgivings was that GWs are often studied in terms of components of the metric or Riemann tensors, with respect to *some basis determined by the coordinates*[7], and *expressed as functions of those coordinates*. Fortunately, this confusion regarding the observability of GWs was resolved by Pirani in 1956, who clarified their existence using tetrad methods and worldlines of particles [25]. Pirani's approach, however, was really only useful for formulating theoretical perspectives and could not be used to make statements about particular systems, like black hole mergers. For this, a more suitable framework was developed in a series of works by Bondi *et al* and Sachs [15, 16, 26–28]. Their approach to studying GWs, which we describe in section 2.2, involves constructing an explicit coordinate system and assuming a particular, but well-motivated asymptotic behavior of the spacetime metric in those coordinates.

As an alternative perspective regarding the subtleties of coordinates in GR, consider a numerical simulation. At the simplest level, numerical relativists must produce waveforms as tables of timestamps and corresponding GW strain values measured at various angular locations encompassing the source[8]. However, the meaning of those time coordinates, the angular locations at which the strain is measured, and the basis with respect to which the strain is evaluated rely on the essentially arbitrary coordinates used in the numerical simulation—coordinates imprinted by the vagaries of initial data and complicated gauges. Despite confusing declarations that may be found in the NR literature, no GW extraction method can produce 'invariant' results in the sense of being independent of the choice of coordinates [29]. Even at linear order in the size of the gauge perturbation, every waveform description is coordinate-dependent. But this issue is not unique to NR simulations. Other gravitational wave modelers, such as those working in post-Newtonian (PN) or even post-Minkowskian (PM) theory, face similar ambiguities. Ideally, we would resolve these coordinate issues in a consistent way, so that waveforms from other simulations, or other models, can be compared to each other.

Due to the diffeomorphism invariance of GR, there is a rich set of coordinate systems that, in principle, could be used to study GWs. In practice, however, working with so many possible coordinate systems is not feasible. Instead, we need some well-motivated way to limit the possible coordinate systems that we can use when studying GWs. To do this, one property that we might impose is that the coordinate systems we consider be adapted to trajectories of inertial observers. That is, curves that have constant spatial coordinates could be timelike geodesics, and the time coordinate for those curves could be the proper time measured on those geodesics. While this would certainly be possible, one issue that arises is that the coordinates we consider would then depend on their initial conditions, and would surely encounter coordinate singularities. But, it turns out that if we instead consider the region of spacetime infinitely far away

[7] See, e.g. [23, 24] for interesting efforts to find tetrads specified by the geometry rather than arbitrary coordinates.

[8] In reality, numerical relativists provide the angular dependence of the strain by representing the strain with respect to some set of angular basis functions, like spin-weighted spherical harmonics. However, to make the connection to coordinates more apparent, in this discussion we ignore this detail and instead consider the naive representation in terms of points on the two-sphere.

from the source, then it is sometimes possible[9] to find a set of coordinates that is *asymptotically* inertial. This realization is exactly what Bondi *et al* and Sachs came to in the 1960s [15, 16, 26–28]. The core idea is that one should instead model a GW source as an isolated system, with the spacetime approaching Minkowski space far from the source, so that one can then match the coordinates to the more familiar inertial trajectories of Minkowski space.

2.2. Bondi gauge and inertial observers

The Bondi–Sachs formalism[10] begins with a collection of coordinates called Bondi–Sachs coordinates that are suited to the problem of outgoing radiation. Essentially, in this formalism we have the usual spherical coordinates (θ,ϕ) as well as a null, retarded-time coordinate u, such that the u direction is orthogonal to the θ and ϕ directions, and an areal coordinate r relative to the (θ,ϕ) coordinates. Anywhere that such a set of coordinates exists, the metric can be written in Bondi–Sachs form as

$$\mathrm{d}s^2 = -Ue^{2\beta}\mathrm{d}u^2 - 2e^{2\beta}\mathrm{d}u\mathrm{d}r + r^2\gamma_{AB}\left(\mathrm{d}x^A - \mathcal{U}^A\mathrm{d}u\right)\left(\mathrm{d}x^B - \mathcal{U}^B\mathrm{d}u\right), \tag{2.1}$$

where capital Latin indices range over (θ,ϕ), and we have introduced the arbitrary functions U, β, $\mathcal{U}^A$, and γ_{AB}, each of which is a function of the coordinates (u,r,θ,ϕ).

With this set of intuitive coordinates, we then restrict the possible metrics that we allow by imposing certain boundary conditions, i.e. some asymptotic behavior in the limit of large radius. In particular, to ensure that the metric in this Bondi–Sachs coordinate system approaches the standard Minkowski metric in the large radius limit, we require that our metric functions obey

$$U \to 1, \tag{2.2a}$$

$$\beta \to 0, \tag{2.2b}$$

$$\mathcal{U}^A \to 0, \tag{2.2c}$$

$$\gamma_{AB} \to \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}. \tag{2.2d}$$

This asymptotic restriction is exactly what is meant by being in 'Bondi gauge' or some 'Bondi frame'. It provided an early notion of what is called 'asymptotic flatness'.

After this work of Bondi, various authors introduced important generalizations of this fall-off condition [36–41], most of which will be beyond the scope of this paper. But one that will be conceptually useful for this review is Penrose's notion of conformal compactification [37, 38]. This compactification, which introduces extra points to construct a boundary of

[9] This possibility rests on some fairly stringent requirements about the spacetime, including the existence of the infinite radius limit, and the fall-off behavior of the metric in that limit. In particular, these requirements rule out direct application to, for example, FLRW spacetimes. Nonetheless, recent work has sought to extend similar analyses to FLRW spacetimes [30–35].

[10] For reasons that are not immediately apparent from the literature, various aspects of this formalism are credited to various subsets of the authors of the papers in which they first appeared: Bondi *et al* and Sachs [15, 16, 26–28]. In particular, Bondi is credited for the gauge or frame; Bondi and Sachs for the coordinates, metric, and formalism generally; and Bondi, Metzner, and Sachs for the (BMS) group. For some reason, van der Burg seems to be left out of the conversation. In this work, we follow this convention without claiming to understand why.

spacetime in the $r \to \infty$ limit, is essential for formulating 'future null infinity' $\mathcal{I}^+$: the final destination of outgoing radiation[11]. This boundary is obtained by taking this $r \to \infty$ limit for fixed u, which yields a region of spacetime parameterized by (u, θ, ϕ). When studying GWs and other asymptotic data, we will be interested in the value of this data *on* this boundary. This is because, as one approaches $\mathcal{I}^+$, curves of constant (r, θ, ϕ) are *nearly* geodesics, with u *nearly* parametrizing the proper time, and errors in this geodesic approximation falling off as $1/r$. Consequently, GWs measured by a distant inertial observer can be approximated (at least over finite time spans) by GWs studied at $\mathcal{I}^+$. In fact, the Bondi frame is sometimes even referred to as the 'asymptotic inertial frame' [42].

Now, apart from constructing a coordinate system and a region of spacetime that can be used to study GWs that agree (up to $1/r$ corrections) with what a distant, inertial observer would see, the other import consequence of constructing the Bondi frame is that, by doing so, we have drastically reduced the coordinate ambiguity that would otherwise plague our GW waveforms. In particular, for the numerical simulation example, we no longer have to express the waveform in terms of the simulation's arbitrary coordinates, which are ambiguous up to the entire diffeomorphism group. Nor do we have to integrate a family of timelike geodesics and evaluate the waveform along those curves, still ambiguous up to the choice of initial conditions for each geodesic. Instead, we can simply express the waveform as a function of (u, θ, ϕ) on $\mathcal{I}^+$.

But one obvious question persists: how much ambiguity still remains in this Bondi frame description? Intuitively, we might expect that the answer would simply be the transformations that leave the usual Minkowski metric unchanged, i.e. the Poincaré group. In fact, this is indeed what Bondi, van der Burg, Metzner, and Sachs thought they would find when studying the symmetry group of future null infinity. However, this intuition turns out to be only nearly correct. Specifically, the symmetry group of future null infinity is the Poincaré group, but with the usual four spacetime translations replaced by a larger set of spacetime transformations called *supertranslations*. This is the BMS group.

2.3. The BMS group

The full BMS group is simply the set of transformations of the asymptotic coordinates—that is, (u, θ, ϕ)—that preserve the asymptotic form of the metric described in equations (2.1) and (2.2)[12]. It should be intuitively obvious that simple rotations satisfy these criteria, as do boosts. Given that the Poincaré group is the group of symmetries *in* Minkowski space, we might also expect analogs of spacetime translations to be allowed, but it turns out that our use of the *retarded* time u and the fact that we are taking the $r \to \infty$ limit complicates matters.

To see this consider the following. A time translation δt will surely affect only the retarded time via $u \mapsto u - \delta t$, but it is not obvious what to do with a space translation by some finite $\delta \vec{x}$.

[11] Note that an identical description also exists at *past* null infinity. But, because we will focus on future null infinity in this review, we ignore this subtlety for the remainder of this work.

[12] With the condition that the angular metric γ_{AB} must asymptote to the usual unit sphere metric, equation (2.2*d*), the asymptotic gauge conditions are preserved by the standard BMS group. However, by relaxing the condition on γ_{AB} so that it must asymptote to anything *conformally related to* the usual unit sphere metric (where the conformal factor can depend on both the angular and retarded-time coordinates), we instead obtain the 'extended' [43] BMS group. By relaxing the condition so that the determinant of γ_{AB} must asymptote merely to have a specified determinant (which may be a function of u, θ, and ϕ), we then obtain the 'generalized' [44–47] BMS group. While these BMS variants are certainly interesting in some contexts, like celestial holography (see, e.g. [21]), they exhibit more freedom than what is needed for the transformation of NR waveforms.

If we take (θ,ϕ) as the direction of the unit vector $\hat{n}$, then the usual space translation changes the radial coordinate via $r \mapsto r + \delta\vec{x}\cdot\hat{n}$. Clearly this has no effect on the $r \to \infty$ limit, nor does it affect the angular coordinates *in that limit*. However, it will affect the retarded time because of the mixture of time and space that the retarded time represents. Considering the prototypical retarded time $u \equiv t - r$, we can intuit the correct impact of a space translation, which is simply $u \mapsto u - \delta\vec{x}\cdot\hat{n}$. The slightly surprising feature here is that a space translation affects the retarded time in a *direction-dependent* way. While a time translation has a monopolar effect, a space translation has a dipolar effect.

Naturally, this invites the intriguing question of whether higher-order multipoles might also be permissible. In fact, if we pose

$$u \mapsto u - \alpha(\theta,\phi)\,, \tag{2.3}$$

with an arbitrary (sufficiently smooth) function $\alpha(\theta,\phi)$, we can check that the asymptotic form of the metric does not change under such a transformation. The function $\alpha(\theta,\phi)$ is exactly a supertranslation and is the ingredient that is needed for constructing the BMS group. It contains the usual spacetime translations as $\ell = 0, 1$ components, when viewed in terms of spherical harmonic modes, and proper supertranslations as $\ell \geqslant 2$ components.

Like the usual spacetime translations, we may combine two supertranslations by pointwise addition, and can thus turn the set of supertranslations into an abelian group $\mathbb{T}$. In fact, this abelian group $\mathbb{T}$ is a normal subgroup of the full BMS group, with the factor group of the BMS group by $\mathbb{T}$ being the usual restricted Lorentz group $\mathrm{SO}^+(3,1)$. The latter, however, is not a normal subgroup. Therefore, the full BMS group is formally the semidirect product $\mathbb{T} \rtimes \mathrm{SO}^+(3,1)$. Put more simply, we can express any BMS transformation as a supertranslation followed by a Lorentz transformation (see appendix for more details). For a boost with velocity $\vec{v}$, and thus a conformal factor[13]

$$k(\theta,\phi) \equiv \frac{\sqrt{1-|\vec{v}|^2}}{1-\vec{v}\cdot\hat{n}(\theta,\phi)}\,, \tag{2.4}$$

where $\hat{n}(\theta,\phi)$ is the unit normal vector to the point (θ,ϕ), the effect on the retarded time of a supertranslation $\alpha(\theta,\phi)$ followed by this Lorentz transformation is simply equation (2.3) but with a Doppler factor, i.e.

$$u \mapsto k(\theta,\phi)\,(u - \alpha(\theta,\phi))\,. \tag{2.5}$$

Note that the (θ,ϕ) angular coordinates are not affected by a supertranslation, but are altered through the more familiar rotation and relativistic aberration (see [48]).

To summarize, the BMS group of future null infinity represents the remaining spacetime coordinate freedom that must be fixed when working with asymptotic data, like GWs, in Bondi gauge. It is made up of the usual Lorentz rotations and boosts, as well as a new set of transformations called supertranslations that extend the spacetime translations of the Poincaré group. While we hope that this section has provided some motivation for why the existence of supertranslations is perhaps expected from a mathematical perspective, in the next section we provide some physical intuition for why they are expected to be symmetries of future null infinity.

[13] See appendix D of [48] for more on why boosts induce a conformal transformation of the celestial sphere.

2.4. Understanding supertranslations

For the moment, consider a proper supertranslation, i.e. a supertranslation which is not a time or space translation. In terms of spherical harmonics, this corresponds to a function $\alpha(\theta,\phi)$ whose spherical harmonic decomposition only consists of $\ell \geqslant 2$ modes. Formally, supertranslations are rather elementary: they are angle-dependent offsets in the retarded time, as illustrated through equation (2.3). Specifically, for each distant inertial observer at $\mathcal{I}^+$, a supertranslation corresponds to a simple change in the origin of the time coordinate. To understand why they are symmetries of future null infinity, we consider the following thought experiment.

Consider some asymptotically flat spacetime with an isolated astrophysical event—like a supernova explosion or a BBH—that is emitting radiation, such as photons or GWs, outward in a spherical manner. Furthermore, imagine a network of distant inertial observers that are surrounding this event, each at some finite radius from the event that need not be the same as the other observers. If these observers can communicate, then they could—in principle—use their knowledge of their locations relative to each other and to the central event to synchronize their clocks to ensure that their measurements of the central event are simultaneous in some sense. Now, consider what happens if these observers were placed at larger (but finite) radii. At these new, farther away positions, the signals that they use to synchronize their clocks will take longer to travel between them, but there is no fundamental obstacle to this synchronization *in principle*.

However, as this network of inertial observers limits to an *infinite* radius away from this central event, then they become causally disconnected from each other. Formally, what this means is that each observer approaches some generator of $\mathcal{I}^+$, and because every generator of $\mathcal{I}^+$ is causally disconnected from any other, so too are the inertial observers. In this $r \to \infty$ limit, the observers can no longer synchronize their clocks, and therefore they can no longer ensure that they receive the same radiation from the central event at the same time. That is, the invariance with respect to standard time and space translations of the Poincaré group at finite radius yields, *at infinite radius*, the invariance to the angle-dependent supertranslations. This fact, i.e. the notion that each and every point on $\mathcal{I}^+$ is causally disconnected from any other, is one way to intuit why supertranslations are indeed symmetries of future null infinity and thus elements of the BMS group.

2.5. The effects of BMS transformations

As mentioned earlier, waveforms are not invariant in any useful sense. At best, they are components of tensors defined with respect to a coordinate basis. Consequently, as we change the coordinates, the value of the waveform at each physical point will also change. Thus, when working with waveforms and their coordinate freedom we really have two main concerns: first, to transform the waveform, and second to transform the coordinates upon which the waveform is evaluated. As we have already explained, the latter can be expressed rather simply via the following.

Instead of working with Bondi coordinates, i.e. (u,θ,ϕ), it tends to be simpler (at least mathematically) to use the complex stereographic coordinate

$$\zeta \equiv e^{i\phi}\cot(\theta/2)\,. \tag{2.6}$$

With this, the action of a BMS transformation on the coordinates of future null infinity can then be written as

$$(u,\zeta) \mapsto \left(k\left(\zeta,\bar{\zeta}\right)\left(u-\alpha\left(\zeta,\bar{\zeta}\right)\right), \frac{a\zeta+b}{c\zeta+d}\right), \tag{2.7}$$

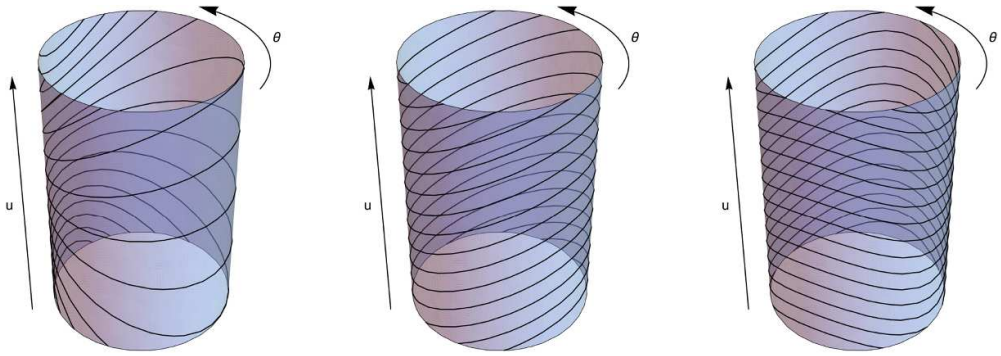

Figure 1. 'Cylinder' diagram of $\mathcal{I}^+$ to provide intuition about (i) boosts, (ii) space translations, and (iii) proper supertranslations. The retarded time coordinate u runs vertically in each plot, while the polar coordinate θ runs azimuthally. Black circles correspond to surfaces of constant u in the untransformed frame. A Lorentz boost dilates u by some factor at each θ point, while a space translation and a proper supertranslation instead shift u by some function of θ, which can be written as some combination of $\ell = 1$ spherical harmonics for space translations, and some combination of $\ell \geqslant 2$ spherical harmonics for proper supertranslations. The boost is in the $\hat{z}$ direction; the space translation is proportional to the spherical harmonic $Y_{(1,0)}(\theta,\phi)$, i.e. the $\hat{z}$ direction; and the proper supertranslation is proportional the spherical harmonic $Y_{(2,0)}(\theta,\phi)$.

where the conformal factor $k(\zeta,\bar{\zeta})$ is

$$k\left(\zeta,\bar{\zeta}\right) \equiv \frac{1+|\zeta|^2}{|a\zeta+b|^2+|c\zeta+d|^2}, \tag{2.8}$$

(a,b,c,d) are complex coefficients with $ad-bc=1$ that encode the Lorentz rotation and boost, and $\alpha(\zeta,\bar{\zeta})$ is a real, smooth function that encodes the supertranslation. See appendix for details on how the Möbius transformation (a,b,c,d) is related to usual Lorentz rotations and boosts. As an illustration of the impact of these transformations, see figure 1, which shows how an example Lorentz boost, space translation, and proper supertranslation change the retarded time u as a function of θ.

From this coordinate transformation, by examining how the coordinate tetrad on $\mathcal{I}^+$ transforms, one can then ascertain how asymptotic data on $\mathcal{I}^+$ that describes the spacetime metric transforms. But what is this *data*? Obviously, we want something on $\mathcal{I}^+$ that represents the gravitational wave strain h, since this is what our detectors measure. But is there other data that we should also be interested in, e.g. something describing the mass or the angular momentum of the spacetime that may be useful for, say, measuring the mass or the spin of an isolated black hole? While often neglected in the majority of the NR literature, the answer to this, as suggested by the isolated black hole example, is yes.

Apart from the strain, to fully reconstruct the metric on $\mathcal{I}^+$ one also needs other information about the spacetime, which is neatly encoded in the five complex Weyl scalars $\Psi_{0,1,2,3,4}$, i.e. the components of the Weyl tensor [49][14]. The Weyl tensor measures the curvature of spacetime. However, the individual Weyl scalars each have their own unique interpretation. Specifically, they can be viewed as

[14] Note that even though these are called 'scalars', the Weyl scalars are also not invariant in any sense since they are still functions of the spacetime coordinates.

- Ψ_0: ingoing radiation;
- Ψ_1: current multipole moment;
- Ψ_2: mass multipole moment;
- Ψ_3: news ($\sim\dot{h}$);
- Ψ_4: outgoing radiation ($\sim\ddot{h}$),

where dots represent time derivatives[15]. Consequently, it can be seen that Ψ_3 and Ψ_4 are actually degenerate with with the strain and thus only the strain and $\Psi_{0,1,2}$ are needed to measure physical features of the spacetime[16]. As we will see later on in section 4, Ψ_2 can be used to measure the mass of the spacetime and Ψ_1 can be used to measure the angular momentum of the spacetime.

If we now study how this asymptotic data transforms under a BMS transformation, one can ascertain through the tetrad transformation that the gravitational wave shear σ, which is related to the strain h via $h \equiv 2\bar{\sigma}$,[17] and the Weyl scalars transform as [15, 16, 36]

$$\sigma' = \frac{1}{k} e^{2i\lambda} \left[\sigma - \eth^2 \alpha\right], \tag{2.9a}$$

$$\Psi'_A = \frac{1}{k^3} e^{(2-A)i\lambda} \sum_{a=A}^{4} \binom{4-A}{a-A} \left(-\frac{1}{k}\eth u'\right)^{a-A} \Psi_a. \tag{2.9b}$$

where $A \in \{0,1,2,3,4\}$, λ is the 'spin phase'

$$e^{i\lambda} \equiv \left[\frac{\partial \bar{\zeta}'}{\partial \bar{\zeta}} \left(\frac{\partial \zeta'}{\partial \zeta}\right)^{-1}\right]^{\frac{1}{2}} = \frac{c\zeta + d}{\bar{c}\bar{\zeta} + \bar{d}}, \tag{2.10}$$

and $\eth$ and $\bar{\eth}$ are the usual GHP spin-weight operators [51]. In spherical coordinates they can be written as

$$\eth f = -\frac{1}{\sqrt{2}} (\sin\theta)^{+s} (\partial_\theta + i \csc\theta \partial_\phi) \left[(\sin\theta)^{-s} f\right], \tag{2.11a}$$

$$\bar{\eth} f = -\frac{1}{\sqrt{2}} (\sin\theta)^{-s} (\partial_\theta - i \csc\theta \partial_\phi) \left[(\sin\theta)^{+s} f\right]. \tag{2.11b}$$

When acting on spin-weighted spherical harmonics ${}_sY_{(\ell,m)}$, they yield

$$\eth \left({}_sY_{(\ell,m)}\right) = +\frac{1}{\sqrt{2}} \sqrt{(\ell - s)(\ell + s + 1)}\, {}_{s+1}Y_{(\ell,m)}, \tag{2.12a}$$

$$\bar{\eth} \left({}_sY_{(\ell,m)}\right) = -\frac{1}{\sqrt{2}} \sqrt{(\ell + s)(\ell - s + 1)}\, {}_{s-1}Y_{(\ell,m)}. \tag{2.12b}$$

Note that the conventions used here are consistent with the Moreschi–Boyle convention that is used across [48, 52–57] and the code `scri` [58–60].

[15] Note that other interpretations also exist, such as [50].

[16] Formally, because of the Bianchi identities (see equation (4.4)), one only needs the strain on $\mathcal{I}^+$ and $\Psi_{0,1,2}$ on a certain time slice of $\mathcal{I}^+$ to fully reconstruct the asymptotic spacetime metric.

[17] We introduce the shear because it makes many of the subsequent equations in this paper simpler; it is formally constructed in equation (5.1) by contracting complex dyads on the two-sphere with the angular part of the Bondi-Sachs metric.

With this information, one then has everything that is needed to transform asymptotic data on $\mathcal{I}^+$ and therefore fix the coordinate freedom of that data to match some canonical frame. This notion of mapping asymptotic data to a certain frame is called BMS frame fixing and will be reviewed in the context of NR simulations in section 7.

2.6. Pedagogical approach to memory

Without delving into the complicated mathematics of Einstein's equations, memory effects can be most easily understood as coming from conservation laws that stem from the symmetries of null infinity: the BMS group. Consequently, to provide some motivation behind why memory effects exist in GR and how they can be studied, before examining them with a more mathematical lens, as is performed in section 4, we will first provide some insight by studying them with respect to BMS transformations.

With these additional symmetries of the BMS group, Noether's theorem[18] interestingly implies that there should be a conservation law for each supertranslation. Thus, because supertranslations are effectively angle-dependent spacetime translations, one can easily imagine that such a balance law might be of the form

$$0 = \text{`change in angle-dependent mass'} + \text{`flux of angle-dependent energy'} \tag{2.13}$$

This expression, in fact, is nearly correct. The one piece of information that is missing is that these two terms on the right-hand side of equation (2.13) need not fully cancel out. One way to realize this is by considering the scattering of two particles in linearized gravity [63]. First note that, because of the linearization, there will be no energy flux. However, because the particles scatter, there will still be a change in the angle-dependent mass, i.e. a change in the mass multipole moment. Then, because a change in the mass multipole moment corresponds to a change in the strain—or, equivalently, the shear—one can intuit that there should also be a term on the left-hand side of equation (2.13) that corresponds to the shear, e.g.

$$\text{`shear'} = \text{`change in angle-dependent mass'} + \text{`flux of angle-dependent energy'} \tag{2.14}$$

In fact, if one formally works through the mathematics, as is carried out in section 4 and shown through equation (4.6*a*), one discovers that the conservation law stemming from supertranslations states exactly this.

Using equation (2.14), one can then learn a large amount of interesting physics about GWs. For one, we observe that, apart from the gravitational wave strain being sourced by the mass multipole moment, there is also a contribution from an angle-dependent energy flux. Furthermore, if we think about the net change in the strain between non-radiative regimes, i.e. the memory, we immediately find that there are two contributions: one from the angle-dependent mass and one from the angle-dependent energy. These unique contributions are the ordinary and null contributions to the memory [64][19]. They can be understood as being sourced by two types of physical processes. In particular, ordinary memory is sourced by systems that contain unbound masses, i.e. massive particles approaching timelike infinity [64, 68, 69].

[18] Note that Noether's theorem is modified for this situation since the nonzero flux of gravitational radiation implies that the charges are not conserved. We instead have *non*-conservation laws precisely quantifying how much each charge changes. The interested reader can see [61, 62] to read about this subtlety.

[19] Previously, these contributions were called linear and nonlinear (or Christodoulou). However, they have since been renamed to better reflect the way in which they are sourced [12, 13, 64–67].

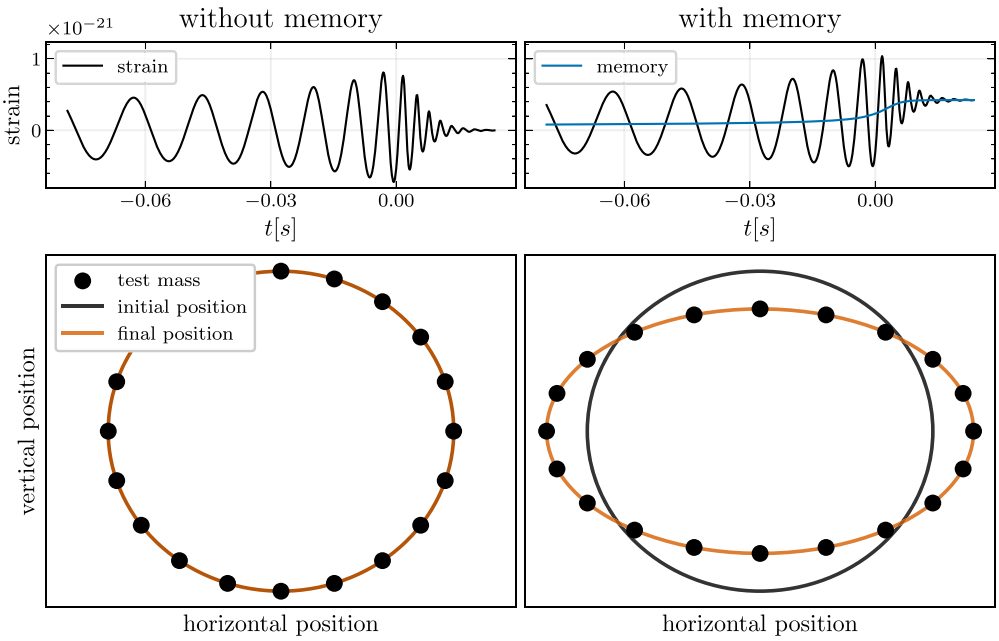

Figure 2. Top: comparing the gravitational wave computed by a simulation of a binary black hole merger when memory effects are not included (left, incorrect) and when memory effects are included (right, correct). The binary black hole simulation is an equal-mass, aligned spin system with a total mass of $60M_{\odot}$, a luminosity distance of 400 Mpc, an edge-on orientation, and equal dimensionless black hole spins of magnitude 0.6 in the direction of the orbital angular momentum. Each waveform is shown in black and in the plot on the right, we show the contribution to the gravitational wave coming from the energy flux, i.e. a proxy for the memory, in blue. Bottom: the initial (black) and final (orange) positions of a series of test particles before and after the passage of a gravitational wave without (left) and with (right) memory traveling through the plane of the figure. Because of the orientation of the binary black hole system relative to the test particles, the GW exhibits a '+' polarization.

Examples of systems that source this type of effect are hyperbolic black holes, neutron star disruptions, or even supernovae events. Meanwhile, null memory is sourced by null radiation that escapes to future null infinity, e.g. gravitational or electromagnetic waves. For an example of what the null memory looks like in the GW produced by a BBH, see figure 2.

3. Literature review

Gravitational memory effects were first realized in 1974 when Zel'dovich and Polnarev successfully calculated the gravitational radiation that is produced by two objects on flyby, i.e. hyperbolic, trajectories [12][20]. By working with Einstein's equations in linearized GR, they found that, because the stress energy tensor exhibits a net change between early and late times due to the change in the mass distribution of the flyby objects, the strain will also exhibit

[20] Even earlier, in 1966, Newman and Penrose identified that, near null infinity, the strain of surfaces of constant retarded time will change between spacelike and future timelike infinity [70].

such a net change. Consequently, because this can also be understood as stemming from a change in the Bondi mass aspect, the effect that they unearthed is what we now call ordinary displacement memory. Later, in 1985, the consequences of their result was elaborated upon by Braginski and Grishchuk who first named this effect the 'memory effect' [13]. Then, in 1987, Braginsky and Thorne found a simple equation for the memory for scattering scenarios in terms of the four-momentum of the ingoing and outgoing massive particles [65]. It says that for a system of N particles, the net change in the gravitational wave strain between early and late times is

$$\Delta h_{ij}^{\mathrm{TT}} = \frac{4}{r}\Delta\sum_{A=1}^{N}\frac{M_A}{\sqrt{1-v_A^2}}\left(\frac{v_A^i v_A^j}{1-v_A\cos(\theta_A)}\right)^{\mathrm{TT}} \tag{3.1}$$

where r is the distance from the observer to the source, M_A is the mass of particle A, $\vec{v}_A$ is the velocity with v_A^i the ith component and v_A the norm, θ_A is the angle between $\vec{v}_A$ and the observer, and the Δ in front of the sum on the right-hand side refers to the difference in this sum evaluated for the outgoing and ingoing particles.

After these early works, it was largely thought that memory effects were understood. This opinion, however, was completely overturned when in 1991 Christodoulou found that GWs themselves will also source a certain type of memory effect, through a subtle, but detectable nonlinear interaction with themselves [66][21]. Christodoulou obtained this important result by working with null hypersurface equations and asymptotic limits to obtain an equation that is similar in spirit to equation (2.14). In particular, he found that the strain is related to the flux of radiation through each point on the two-sphere. Because of this connection to the energy flux, we now call this effect the null displacement memory. A year later, in [67] Thorne identified Christodoulou's finding as equivalent to that of [65], but with the various massive particles being replaced by null radiation, i.e.

$$\Delta h_{ij}^{\mathrm{TT}} = \frac{4}{r}\int\frac{\mathrm{d}E}{\mathrm{d}\Omega'}\left(\frac{\xi^{i'}\xi^{j'}}{1-\cos(\theta')}\right)^{\mathrm{TT}}\mathrm{d}\Omega', \tag{3.2}$$

where E is the energy of the radiation, $\xi^{i'}$ is a unit vector pointing from the source toward $\mathrm{d}\Omega'$, and θ' is the angle between $\xi^{i'}$ and the observer [67]. In section 4, we will connect equations (3.1) and (3.2) to the modern interpretation of memory, which is more straightforward to understand in terms of gravitational systems and radiation.

As for the BMS group, this was realized well before memory effects in 1962 by Bondi *et al* Sachs [15, 16]. However, the connection between the BMS group and memory was not explicitly stated until [73–75] in 2014, even though this relationship has been largely understood since, e.g. [61, 70, 76–78]. What makes this history even more interesting, though, is in [75, 79] the duality between memory effects and BMS symmetries was extended to be a triangle, since it was found that BMS symmetries and memory effects could also be related to soft theorems [80]. This finding created a large stir of interest in theory communities, since soft theorems are inherently useful for studying the $\mathcal{S}$-matrix of a quantum theory, so BMS symmetries seem integral to understanding quantum gravity. The 'soft limit' means taking a particle's energy to zero, so soft particles can not be directly measured by particle colliders. However, memory effects will soon be observed through GWs with current detectors [81–83], and therefore serve

[21] This discovery was also realized by Payne (somewhat indirectly) as well as Blanchet and Damour in [71] and [72].

as a natural probe of this exciting realm of physics. This excitement was only further enhanced once more memory effects were unearthed through this connection. In particular, in [84–86], by studying extensions of the BMS group to include 'superrotations' [87–90], i.e. extensions of the Lorentz transformations, two new memory effects were found: the spin memory and the center-of-mass memory. These two effects correspond to the net displacement that two observers with an initial relative velocity will experience due to the passage of transient radiation[22]. Ever since, the field of celestial holography, which aims to establish a kind of holographic dictionary between gravitational scattering in asymptotically flat spacetimes and a conformal field theory on the celestial sphere, has reached unprecedented levels of interest [17–21].

At the same time as these theoretical developments regarding memory were occurring, significant progress in resolving memory effects in numerical simulations was also being achieved. In particular, in early 2010 [92] was able to successfully simulate the memory sourced by a BBH using a more robust method for extracting the NR waveforms at future null infinity: Cauchy-characteristic evolution (CCE)[23]. Later, in 2020, [93] performed a similar series of simulations using a more efficient version of the code and calculated the individual contributions to the gravitational wave strain in terms of the charges and fluxes of equations (2.14) and (4.6). This showed that, as expected, the memory in BBH is indeed sourced by the null memory.

One complexity that arose, however, was, with these new NR waveforms that contained memory, it was not exactly obvious how to compare these finite waveforms to PN waveforms that had information about the entire past history of the binary inspiral and thus had a larger prediction for the instantaneous value of the strain [94]. This was important because if one wanted to construct a hybridization between these NR waveforms and PN waveforms to build a waveform model from, the hybridization would fail because the two predictions would not agree (see, e.g. figure 6). This was resolved in a series of works which established a program now called the BMS frame fixing program [52, 53, 95]. Ever since, numerous findings and advancements in gravitational wave physics using these developments have been made [96–100].

4. Mathematical overview

Recall that equation (2.1) provides the general form of the metric in Bondi gauge, in terms of the functions $U, \beta, \mathcal{U}^A$, and γ_{AB}, where A and B range over (θ, ϕ). In general, these functions can each depend on all four coordinates. However, Bondi gauge demands specific behavior in the limit of $r \to \infty$. To proceed, we expand this behavior in powers of $1/r$:

$$U = 1 - \frac{2m}{r} - \frac{2\mathcal{M}}{r^2} + \mathcal{O}\left(r^{-3}\right), \tag{4.1a}$$

$$\beta = \frac{\beta_0}{r} + \frac{\beta_1}{r^2} + \frac{\beta_2}{r^3} + \mathcal{O}\left(r^{-4}\right), \tag{4.1b}$$

$$\mathcal{U}^A = \frac{U^A}{r^2} + \frac{1}{r^3}\left[-\frac{2}{3}N^A + \frac{1}{16}D^A\left(C_{BC}C^{BC}\right) + \frac{1}{2}C^{AB}D^C C_{BC}\right] + \mathcal{O}\left(r^{-4}\right), \tag{4.1c}$$

$$\gamma_{AB} = q_{AB} + \frac{C_{AB}}{r} + \frac{D_{AB}}{r^2} + \frac{E_{AB}}{r^3} + \mathcal{O}\left(r^{-4}\right), \tag{4.1d}$$

[22] They can also be viewed as the magnetic and electric components of the more general drift memory [91].

[23] We will elaborate more on this waveform extraction procedure, as well as others, in section 5.

where the various coefficients on the right-hand sides are functions of (u,θ^A) only, and $q_{AB}(\theta^A)$ is the metric on the two-sphere, i.e. in the usual spherical coordinates $q_{AB}(\theta,\phi)\mathrm{d}x^A\mathrm{d}x^B = \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2$. Of these functions, the most important ones are the Bondi mass aspect m, the Bondi angular momentum aspect N^A, and finally the shear tensor C_{AB}, whose retarded time derivative is the Bondi news tensor $N_{AB} \equiv \partial_u C_{AB}$, which characterizes the presence of radiation in a spacetime.

At this point one can then impose Einstein's equations to obtain evolution equations for the various functions. Specifically, from the $\mathcal{O}(uu, r^{-2})$ and $\mathcal{O}(uA, r^{-2})$ parts[24] of the Einstein tensor one obtains [101]

$$\partial_u m = \frac{1}{4}D_A D_B N^{AB} - \frac{1}{8}N_{AB}N^{AB}, \tag{4.2a}$$

$$\partial_u \hat{N}_A = \frac{1}{4}\left(D_B D_A D_C C^{BC} - D^2 D^B C_{AB}\right) - \left[\frac{3}{8}\left[\left(N_{AB}D_C C^{BC} - C_{AB}D_C N^{BC}\right)\right.\right.$$
$$\left.\left. - \frac{1}{8}\left(N^{BC}D_B C_{AC} - C^{BC}D_B N_{AC}\right)\right] - uD_A\partial_u m, \tag{4.2b}$$

where $\hat{N}^A$ is the angular momentum aspect with the Wald–Zoupas correction [62]:

$$\hat{N}_A \equiv N_A - uD_A m - \frac{1}{4}C_{AB}D_C C^{BC} - \frac{1}{16}D_A\left(C_{BC}C^{BC}\right). \tag{4.3}$$

By contracting these evolution equations with dyads on the two-sphere[25] and making use of the Bianchi identities for the Weyl scalars, i.e.

$$\dot{\Psi}_0 = \eth\Psi_1 + 3\sigma\Psi_2, \tag{4.4a}$$
$$\dot{\Psi}_1 = \eth\Psi_2 + 2\sigma\Psi_3, \tag{4.4b}$$
$$\dot{\Psi}_2 = \eth\Psi_3 + \sigma\Psi_4, \tag{4.4c}$$
$$\mathrm{Im}\left[\Psi_2\right] = -\mathrm{Im}\left[\eth^2\bar{\sigma} + \sigma\dot{\bar{\sigma}}\right], \tag{4.4d}$$
$$\Psi_3 = -\eth\dot{\bar{\sigma}}, \tag{4.4e}$$
$$\Psi_4 = -\ddot{\bar{\sigma}}, \tag{4.4f}$$

one can then rewrite equations (4.2) rather simply in terms of spin-weighted functions as

$$\partial_u m = -\partial_u \mathrm{Re}\left[\Psi_2 + \sigma\dot{\bar{\sigma}}\right], \tag{4.5a}$$
$$\partial_u \hat{N} = -\partial_u\left[\Psi_1 + \sigma\eth\bar{\sigma} + u\eth m + \frac{1}{2}\eth\left(\sigma\bar{\sigma}\right)\right]. \tag{4.5b}$$

Then, by reorganizing terms (for equation (4.6*a*)) or following the derivation in [93] (for equation (4.6*b*)), one can rewrite equations (4.2) with the shear on the left-hand side as

$$\mathrm{Re}\left[\eth^2\bar{\sigma}\right] = m + \mathcal{E}, \tag{4.6a}$$
$$\mathrm{Im}\left[\eth^2\bar{\sigma}\right] = \partial_u \mathrm{Im}\left[\eth^{-1}\left(\hat{N} + \mathcal{J}\right)\right], \tag{4.6b}$$

[24] Here the first argument corresponds to the component of the metric tensor that is being examined, while the second argument corresponds to the relevant term in the $1/r$ expansion.

[25] See, e.g. [93] for more details.

where

$$\mathcal{E} \equiv \int_{-\infty}^{u} |\dot{\sigma}|^2 \, \mathrm{d}u, \tag{4.7a}$$

$$\mathcal{J} \equiv \int_{-\infty}^{u} \frac{1}{2} \left(3\dot{\sigma}\eth\bar{\sigma} - 3\sigma\eth\dot{\bar{\sigma}} + \bar{\sigma}\eth\dot{\sigma} - \dot{\bar{\sigma}}\eth\sigma\right) \mathrm{d}u. \tag{4.7b}$$

In equations (4.7), $\mathcal{E}$ can be thought of as the total energy flux measured at $\mathcal{I}^+$ and $\mathcal{J}$ as the angular momentum flux. Equation (4.6*a*) is called the supertranslation conservation law and can be thought of as the conservation law that stems from the supertranslation symmetry of $\mathcal{I}^+$; equation (4.6*b*), however, is not often presented in the literature, in part because it does not have as clear of an interpretation as equation (4.6*a*) as being the conservation law corresponding to a symmetry contained in the BMS group. Nonetheless, other symmetry groups of future null infinity have also been proposed that contain more symmetries than the BMS group for which equation (4.6*b*) can be understood as a conservation law [87–90]. Specifically, equation (4.6*b*) has been referred to as the superrotation conservation law. These other groups are extensions of the BMS group. They are obtained by relaxing the fall off conditions in equations (4.1) to be less restrictive, which ends up enabling the existence of other symmetries that extend the usual Lorentz transformations. But, because in this review we restrict our attention to focus on the BMS group, for the remainder of this paper we will not consider these other symmetry groups and instead refer the interested reader to explore these extensions further in [87–90].

To provide more motivation as to why, at least to some degree, equations (4.6) should be viewed as conservation laws, let us first consider what the Bondi mass aspect and the angular momentum aspect represent[26]. In particular, from these aspects one can construct Poincaré charges, i.e. the translation, rotation, and boost charges [61, 102–104]

$$P^a(u) = \frac{1}{4\pi} \int_{S^2} n^a m \, \mathrm{d}\Omega, \tag{4.8a}$$

$$J^a(u) = \frac{1}{4\pi} \int_{S^2} \mathrm{Im}\left[\left(\bar{\eth} n^a\right)\hat{N}\right] \mathrm{d}\Omega, \tag{4.8b}$$

$$K^a(u) = \frac{1}{4\pi} \int_{S^2} \mathrm{Re}\left[\left(\bar{\eth} n^a\right)\hat{N}\right] \mathrm{d}\Omega, \tag{4.8c}$$

where n^a with $a \in t,x,y,z$ is the four vector

$$n^t = 1 = \sqrt{4\pi} Y_{(0,0)}, \tag{4.9a}$$

$$n^x = \sin\theta\cos\phi = \sqrt{\frac{4\pi}{3}}\left[\frac{1}{\sqrt{2}}\left(Y_{(1,-1)} - Y_{(1,+1)}\right)\right], \tag{4.9b}$$

$$n^y = \sin\theta\sin\phi \tag{4.9c}$$

$$= \sqrt{\frac{4\pi}{3}}\left[\frac{i}{\sqrt{2}}\left(Y_{(1,-1)} + Y_{(1,+1)}\right)\right], \tag{4.9d}$$

$$n^z = \cos\theta \tag{4.9e}$$

$$= \sqrt{\frac{4\pi}{3}} Y_{(1,0)}, \tag{4.9f}$$

[26] Note that the word aspect here corresponds to the fact that the mass and angular momentum aspects can be thought of as angle-dependent generalizations of mass and angular momentum.

with each component being a spin-weight 0 function. Meanwhile, the quantities $\mathcal{E}$ and $\mathcal{J}$ defined in equations (4.7) are the usual energy and angular momentum fluxes. So, if one takes equations (4.6) and instead writes them in terms of spherical harmonics, it can readily be seen that, since $\eth^2\bar{\sigma}$ has no $\ell = 0$ or $\ell = 1$ components, these equations are the four and angular momentum conservation laws that correspond to the translation and rotation symmetries. Obtaining such charges can be achieved by following the prescription of Wald and Zoupas [62] or the derivation of these exact charges by Dray and Streubel [61].

As for the $\ell \geqslant 2$ components of equations (4.6), these parts correspond to the supertranslation (and superrotation) conservation laws mentioned earlier. This can be seen, for example, from the fact that here the energy flux that appears in equation (4.6*a*) is a function of angular coordinates and thus corresponds to the energy radiated at each point on the celestial two-sphere, which is reminiscent of the pedagogical example presented in section 2.6. However, to view equation (4.6*a*) as a supertranslation conservation law, we should also identify the charge that corresponds to supertranslations. Naively, one might expect that the supertranslation charge is simply equation (4.8*a*), but with n^a replaced by some arbitrary function on the two-sphere, e.g.

$$\mathcal{P}_{(\ell,m)}(u) = \frac{1}{4\pi}\int_{S^2}\left[\sum_{\ell\geqslant 0,|m|\leqslant\ell}\alpha_{(\ell,m)}Y_{(\ell,m)}\right] m\,\mathrm{d}\Omega, \tag{4.10}$$

where $\alpha_{(\ell,m)}$ are spherical harmonic coefficients for some real-valued, smooth function $\alpha(\theta,\phi)$. This, in fact, is a reasonable hypothesis for this charge. In particular, the only possible expressions for the supertranslation charge, or what some call the 'supermomentum', are

$$\mathcal{P}_{p,q}(u,\theta,\phi) = \Psi_2 + \sigma\dot{\bar{\sigma}} + p\left(\eth^2\bar{\sigma}\right) - q\left(\bar{\eth}^2\sigma\right), \tag{4.11}$$

where p and q are real numbers [61]. This was pointed out by Dray and Streubel in [61] and was also later independently realized by Wald and Zoupas in [62]. From this supermomentum expression, it can be easily shown that if $p+q=1$ then the supermomentum is real and if $p=q$ there is no supermomentum flux in Minkowski space, both of which are nice properties [61]. Consequently, the natural choice of supermomentum is the Geroch (G) supermomentum with $p=q=\frac{1}{2}$ [40, 61]

$$\mathcal{P}_G(u,\theta,\phi) \equiv \Psi_2 + \sigma\dot{\bar{\sigma}} + i\mathrm{Im}\left[\eth^2\bar{\sigma}\right] = -m, \tag{4.12}$$

which, up to a sign, is exactly the Bondi mass aspect. Thus, the naive guess for the supermomentum being that which appears in equation (4.10) is indeed correct.

At this point, having established the BMS charges, i.e. equations (4.8) as well as equation (4.10), we can now return to the examination of equations (4.6) as the BMS conservation laws. In particular, since $\eth^2\bar{\sigma}$ has no $\ell = 0$ or $\ell = 1$ components when written in terms of spherical harmonics, we readily find that the $\ell = 0$ component of equation (4.6*a*) is stating energy conservation, while the $\ell = 1$ component is stating linear momentum conservation. As for equation (4.6*b*), the $\ell = 0$ component of this equation is trivially zero. But, the $\ell = 1$ is stating angular momentum conservation. And, finally, if we now examine the $\ell \geqslant 2$ components of equation (4.6*a*), we readily find a statement regarding the conservation of supermomentum. For equation (4.6*b*), this has sometimes been viewed as a statement regarding the conservation of super angular momentum, but we stress that to do so requires extending the symmetry group of future null infinity to be larger than the BMS group [105].

Now, besides providing a clear and straightforward connection between the various BMS symmetries and the conservation of charges and fluxes at null infinity, equations (4.6) also provide a unique and useful means for studying GWs. In particular, because the right-hand side of equations (4.6) contain the shear, which is related to the gravitational wave strain h via $h = 2\bar{\sigma}$, one can readily use equations (4.6) to study contributions to the strain in terms of BMS charges and fluxes. In particular, as explained in section 2.6, the contribution from the charges in equations (4.6) correspond to the ordinary memory while the flux contributions corresponds to the null memory.

5. Numerical code frameworks

In this section we briefly outline the code frameworks that are required to simulate a BBH and extract the waveform data to future null infinity. Readers who are primarily interested in examining the waveforms output by the numerical simulations and how they can be studied using equations (4.6), rather than the details of the numerical simulation, may skip to section 6.

When simulating a BBH, there are two types of numerical evolutions that need to be run[27]. The first, and the bulk of the computation, is what is typically called the Cauchy evolution. This part of the NR simulation involves solving Einstein's equations on a finite region of spacetime near the binary black holes to obtain the metric and its derivatives at a finite radius.

Once the Cauchy evolution is complete and the metric and its derivatives are obtained at a finite radius, it is then possible to run a waveform extraction to obtain the waveform data at future null infinity. This is what we use as a proxy for the data that GW detectors should observe on Earth, because the extra information from being at a finite distance from the GW-sourcing event are higher order in the $1/r$ expansion of the angular part of the metric and should therefore be highly subdominant. Ultimately, there are two ways this can be performed:

- extrapolation, which consists of fitting metric data (γ_{AB} in equation (4.1*d*)) to polynomials in $1/r$ to extract waveform data at future null infinity [107]; or
- a characteristic evolution, which consists of solving Einstein's equations on hypersurfaces that connect a finite radius worldtube to future null infinity.

The primary output of these extraction methods is the gravitational wave strain h, which, with respect to the shear tensor C_{AB} in equation (4.1*d*), is defined via

$$h(u,\theta,\phi) \equiv \bar{q}_A \bar{q}_B C_{AB} = \sum_{\ell \geqslant 2, |m| \leqslant \ell} h_{(\ell,m)}(u)\, {}_{-2}Y_{(\ell,m)}(\theta,\phi) \tag{5.1}$$

and is a spin-weight -2 function often decomposed into spin-weight -2 spherical harmonics ${}_{-2}Y_{(\ell,m)}$ with the complex coefficients $h_{(\ell,m)}(u)$. In equation (5.1), q_A is the dyad $q_A \equiv -(1, i\sin(\theta))/\sqrt{2}$ (see, e.g. [93]).

As one may imagine, the characteristic evolution, albeit more challenging to perform, is much more accurate than extrapolation: both in terms of the numerical precision that can be

[27] In practice constructing initial data for the binary that one aims to simulate is also an important part of the simulation. However, because here we aim to review the parts of the simulation that produce meaningful waveform data, we will restrict ourselves to a discussion of the different types of evolutions. The interested reader can look to [106] for more on initial data construction.

achieved as well as ensuring that the expected physics is accurately captured[28]. Effectively, the way that the characteristic evolution works, which is formally called a CCE, is through the following. First, treat the worldtube at some finite radius and the initial null hypersurface that connects the worldtube to future null infinity as two sets of initial data. Then, reduce Einstein's equations to a series of ordinary differential equations (ODEs). Finally, solve this series of ODEs by integrating in retarded time to obtain data for the waveforms on subsequent null hypersurfaces and eventually the whole of future null infinity. By performing this sequence of tasks, one can then accurately compute the strain as well as the Weyl scalars at future null infinity. One subtlety, however, is that choosing the initial data on the first null hypersurface is a highly nontrivial problem. An incorrect choice of this initial data effectively amounts to putting the output waveform data in some arbitrary BMS frame that needs to be manipulated in order to perform robust waveform model comparisons or analyses. This issue is covered in more detail in section 7.1.

The idea of CCE was first theorized in 1996 [108, 109]. However, it was not until 2009 when a CCE code was first used on a binary black hole simulation [92, 110–112]. This version of CCE was run using the finite-difference Pitt Null code [92, 110–112]. Later, in 2014, an improved version of CCE using spectral methods was incorporated into the SpEC code [113–115]. And, finally, throughout 2020 and 2021 an even more-improved version of CCE that could extract the Weyl scalars was developed by [116] and incorporated into the SpECTRE code [117] by [118]. This version of CCE is the most advanced version and is what will be used throughout this review.

6. Numerical waveforms and memory effects

6.1. Examining conservation laws in equations (4.6)

With the pedagogical and mathematical theory behind the BMS group, its corresponding conservation laws, and memory effects presented in earlier sections, we now turn to a numerical study of GWs in the context of asymptotics. In particular, to highlight the usefulness of equations (4.6), we first present figure 3 to illustrate how the gravitational wave strain can be easily interpreted in terms of BMS charges and fluxes. In this figure we show four spin weight -2 spherical harmonic modes of the gravitational wave strain as well as the contributions to the strain coming from the BMS charges and fluxes in the right-hand sides of equations (4.6). The system considered is a mass ratio $q = 1.22$ binary black hole, whose black hole spins are parallel to the system's orbital angular momentum and have the dimensionless magnitudes $\chi_z^{(1)} = 0.33$ and $\chi_z^{(2)} = -0.44$. These parameters resemble the most likely parameters of the first gravitational wave detection, GW150914 [119].

Let us first consider what is shown in the top left panel, i.e. the $(\ell, m) = (2, 2)$ mode. As can be seen, the strain is nearly entirely sourced by the Bondi mass aspect. This, however, is expected because the Bondi mass aspect can effectively be thought of as the mass multipole moment. So all that this plot is showing is an illustration of the fact that GWs are predominantly sourced by the mass quadrupole moment. If we now examine the top right panel, we see a similar phenomenon, but now illustrating the fact that gravitational waves are also sourced by the current multipole moment. In this panel we show the $(3, 2)$ mode of the strain as well as the two contributions from the angular momentum aspect and the angular momentum flux. Like

[28] In particular, waveform data produced using extrapolation is known to not capture memory effects. We will discuss this failure, and how it can be mitigated, more in section 6.3.

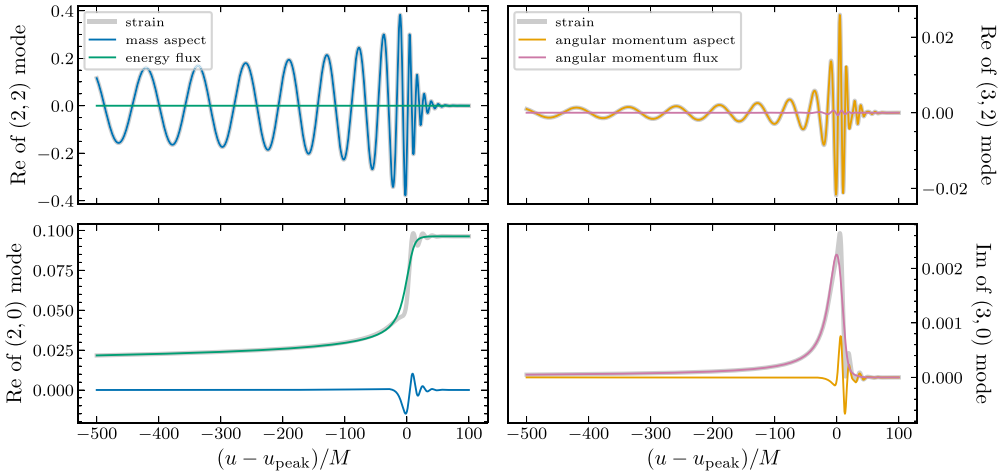

Figure 3. Four different spin-weight -2 spherical harmonic modes of the gravitational wave strain and the contributions from the BMS charges and fluxes appearing in the right-hand side of equations (4.6). The system is a mass ratio $q = 1.22$ binary black hole, whose black hole spins are parallel to the system's orbital angular momentum and have the values $\chi_z^{(1)} = 0.33$ and $\chi_z^{(2)} = -0.44$. Note that this system (SXS:BBH:0305) has parameters which are consistent with those of the first GW detection GW150914. Top left: the real part of the $(\ell, m) = (2, 2)$ mode of the strain. The Bondi mass aspect's and energy flux's contributions are shown in blue and green. Bottom left: identical to that shown in the top left panel, but for the real part of the $(2, 0)$ mode. Top right: the real part of the $(3, 2)$ mode of the strain. The angular momentum aspect's and flux's contributions are shown in orange and purple. Bottom right: identical to that shown in the top right panel, but for the imaginary part of the $(3, 0)$ mode. The horizontal axis for each plot is the retarded time u, with $u_{\rm peak}$ the peak of the L^2 norm of the news over the two sphere.

the top left panel, we find that the strain is primarily sourced by the charge in equations (4.6). And, since the angular momentum aspect can be related to the current multipole moment, this plot also highlights the fact that the current multipole moment also sources the strain, albeit subdominantly.

The perhaps more interesting panel in figure 3, however, is the bottom left panel. In this plot we now show the $(2, 0)$ mode of the strain as well as the contributions from both the Bondi mass aspect and the energy flux. As can be seen, instead of the strain being sourced by the Bondi mass aspect, we instead find that there is a large and dominating contribution from the energy flux. Furthermore, there is the unique phenomenon that the strain no longer decays to zero; this is the memory effect[29]! More specifically, this is the term that we associate with the memory. What we mean by this is the following. Formally, gravitational wave memory is a phenomenon that can only be measured between two non-radiative regions of spacetime, e.g. before and after the passage of a burst of gravitational radiation. In practice, however, gravitational wave

[29] One also may observe that there is a nonnegligible contribution from the Bondi mass aspect near the peak of the strain. This is because when the black holes are merging, they form an excited remnant black hole that emits GWs in a process called the ringdown. The contribution from the mass aspect is exactly this ringing that occurs as the remnant black hole settles to be in a state of equilibrium, i.e. a Kerr black hole.

detectors are not really freely-falling[30] and analyze the strain in the frequency domain. Thus, to measure memory we need to associate the memory with some nonzero frequency, which is most naturally whatever the frequency of the energy flux is, because it is the source of the memory for these binary systems, as illustrated through the bottom left panel of figure 3. So, for the remainder of the paper when we refer to memory what we will really be referring to is the evolution of the memory, as measured through the BMS fluxes.

Finally, in the bottom right panel of figure 3 we again highlight a similar result to that of the bottom left panel, but now illustrated through the imaginary part of the $(3,0)$ mode of the strain. As is shown, the strain is again predominantly sourced by the flux contribution, but now the contribution looks like a delta function, rather than a step function; this is the spin memory effect. As outlined in section 3, unlike the displacement memory, which affects initially co-moving observers, the spin memory instead affects observers with a non-zero initial relative velocity. This can be understood in part by thinking about the time integral of the strain, in which case this contribution would instead manifest as a step function.

6.2. *Detectability of memory effects*

While figure 3 is useful in that it clearly illustrates the fact that the energy flux is the contribution responsible for the memory exhibited by the gravitational wave strain, it does not provide the overall magnitude of this effect. This is because when the strain is evaluated at a point on the celestial sphere, e.g. if the strain were observed by a gravitational wave detector, one needs to consider the sum of the strain's modes weighted by the spin weighted spherical harmonics. As a result, while the memory looks fairly prominent in the bottom left panel of figure 3, if one evaluates the strain at a point on the sky the memory can be noticeably suppressed. This is in part because the $(2,0)$ spin weight -2 spherical harmonic is

$$ {}_{-2}Y_{(2,0)}(\theta,\phi) = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\sin(\theta)^2, \tag{6.1}$$

so the memory is maximized for systems that, from the binary's viewpoint, are viewed edge-on ($\theta = \pi/2$) and is minimized for systems that are viewed face-on ($\theta = 0$). This fact is clearly highlighted in figure 2, which shows what the whole gravitational wave strain looks like for a binary black hole system when evaluated at the point on the sky with $(\theta,\phi) = (\pi/2, 0)$. As can be seen, even for an ideal orientation (as well as an ideal binary, i.e. equal mass and large black hole spins in the direction of the orbital angular momentum), the net memory that is induced on the gravitational wave detector is $\lesssim 50\%$ of the magnitude of the full signal.

Furthermore, a unique challenge for observing memory in a real-world detector is the fact that detectors do not analyze the GW strain in the time domain, but rather the frequency domain. This is an issue because the part of the strain that sources the memory, even though it can have a large amplitude, is a low-frequency effect since it resembles a step function and is not as oscillatory as the other modes of the strain. We highlight this in figure 4, which shows the Amplitude Spectral Density (ASD), i.e. the root of the Power Spectral Density (PSD), of the gravitational wave strain evaluated at $(\theta,\phi) = (\pi/2, 0)$. We also show the contributions to the strain from the $(2,\pm 2)$ modes and the energy flux. The NR system is the same as that in figure 2. Before Fourier transforming the waveform we first pre-process it following [120], i.e. we subtract a linear function from the waveform to remove the offset due to the memory

[30] This is because the test masses in a gravitational wave detector are influenced by actuators that are always working to restore the initial configuration of the detector.

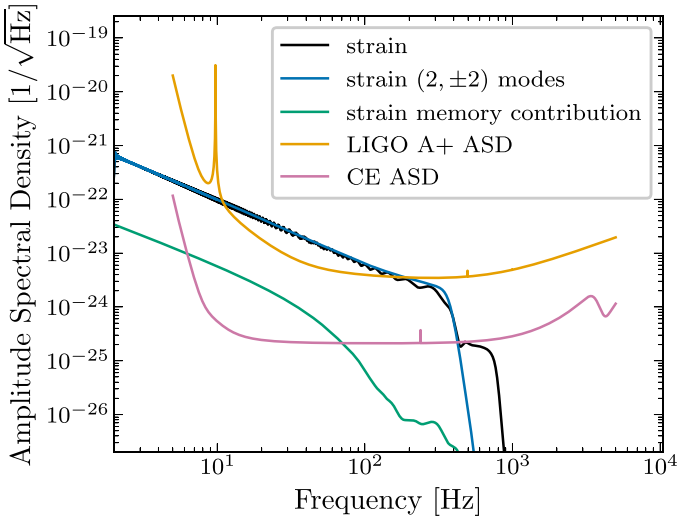

Figure 4. Amplitude spectral density of the GW strain (black) evaluated at $(\theta,\phi) = (\pi/2,0)$ and the contributions to the strain from the $(2,\pm2)$ modes (blue) and energy flux (green). The NR system is an equal-mass, aligned spin binary with a total mass of $60\,M_\odot$, a luminosity distance of 400 Mpc, and equal dimensionless black hole spins of magnitude 0.6 in the direction of the orbital angular momentum. This simulation is used to show a more optimistic observation of the memory. As a reference, we include the LIGO A+ noise curve in orange and the CE noise curve in magenta.

effect and we apply a Planck window with an ϵ of 10^{-4} [121]. As is shown, the ASD of the strain is primarily represented by the $(2,\pm2)$ modes, while the energy flux, i.e. the memory, only contributes to the ASD at frequencies below 10 Hz. This is what makes memory challenging to observe in current ground-based detectors. Because of seismic noise, LVK detectors are not sensitive to signals below $\sim$10 Hz. In particular, the signal-to-noise ratio (SNR) ρ of the strain and the memory contribution to the strain, i.e.

$$\rho = \sqrt{4\int_{f_{\rm min}}^{f_{\rm max}} \frac{|\tilde{h}(f)|^2}{S_n(f)}\mathrm{d}f}, \tag{6.2}$$

where $\tilde{h}(f)$ is the Fourier transform of the strain, $S_n(f)$ is the noise PSD, are $\approx$65 and $\approx$2. Because of this low SNR, current efforts to observe memory rely on a procedure called 'stacking' which combine the SNR estimates of the memory from various events to instead compute a type of population measurement [81–83]. Put differently, these stacking efforts aim to show that memory has been detected in a population of events. To detect memory in a single event, we will most-certainly need to rely on future ground-based detectors like the Einstein Telescope or Cosmic Explorer (CE) or perhaps even one of the planned space-based detectors like LISA, which are more susceptible to the lower frequency regimes.

6.3. Memory correction

Although we have so far primarily presented memory as being easily resolvable in NR simulations, this is not the case. In particular, prior to the creation of a CCE code framework by [92] in 2010 and a more efficient framework by [93] in 2020, there was no memory in the

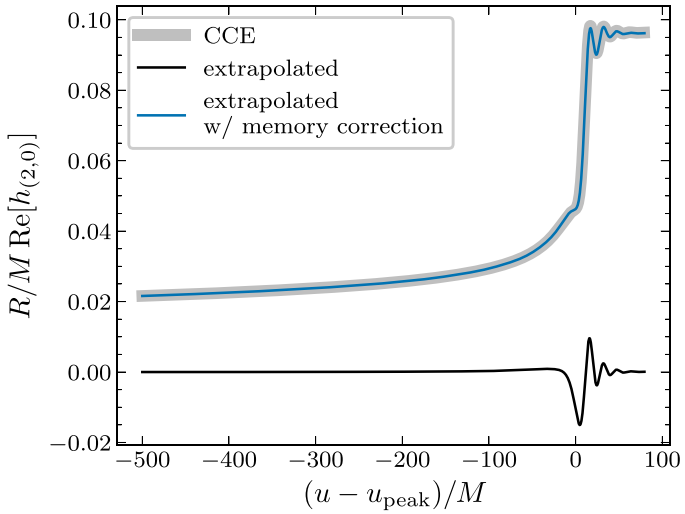

Figure 5. Comparison of the real part of the $(2,0)$ mode of the strain extracted using CCE (gray) to that of the strain extracted using the extrapolation procedure (black) as well as the its memory-corrected version (blue). We apply the memory correction by computing the energy flux, equation (4.7*a*), using the extrapolated strain waveform and then adding it to this waveform. The system used is the same as in figure 3.

waveforms that were produced by NR simulations. This was because, when not using CCE, NR simulations relied on extrapolation to produce predictions for what the gravitational wave strain should be at null infinity. However, because extrapolation is not an exact solution to Einstein's equations, it is unable to accurately resolve memory effects. This important fact is shown in figure 5, which shows the $(2,0)$ mode of the strain for waveforms extracted using CCE and extrapolation. As can be seen, the CCE waveform captures both the memory and the oscillations induced in the merger and ringdown phases, but the extrapolated waveform only captures the later.

Even so, [122] showed that extrapolated waveforms can be corrected, through a post-processing technique, to include the memory contribution and exhibit much better agreement with CCE waveforms[31]. In [122], it was found that for a wide range of binary simulations, the extrapolated strain waveforms simply seem to fail to capture the energy flux contribution in equation (4.6*a*). Thus, the authors argued that since the energy flux is only a function of the strain, the extrapolated strain waveforms can be self-consistently corrected to include the missing memory contribution. In doing so, they found that with such a correction the extrapolated waveforms then satisfy the supertranslation conservation law to a higher degree and better match the CCE waveforms. For completeness, we also show this phenomenon in figure 5. By eye, one can easily see that once the extrapolated strain waveform is corrected to include the missing energy flux contribution, it agrees much better with the CCE waveform.

Unfortunately, even with this useful memory correction, [122] found that extrapolated waveforms still do not outperform CCE waveforms in terms of their accuracy and violation of the Bianchi identities. Consequently, future waveform models should use CCE waveforms.

[31] Talbot *et al* [123] also performed a correction to extrapolated waveforms to include memory, but did not compare the corrected waveforms to CCE waveforms, which naturally exhibit this effect.

7. BMS frame fixing

Up until this point, we have primarily been focused on utilizing the charge/flux perspective enlightened through the BMS group to understand memory effects and the means by which the gravitational wave strain is sourced. For the rest of this review, we will turn our attention to how the transformations of the BMS group are crucial for performing robust analyses with NR waveforms and building waveform models to test Einstein's theory of GR with gravitational wave detector data.

7.1. Fixing the frame with BMS charges

As is the case in every field of physics, fixing the frame of the system that one is studying is vitally important to ensure that the observed phenomena are physical and not just gauge artifacts. For BBH, the situation is no different. In particular, whenever we study the gravitational radiation from a binary merger, we are often implicitly fixing part of the frame without knowing it. For example, by specifying a $\hat{z}$-axis, e.g. the direction of the binary's orbital angular momentum or the direction of the remnant black hole's spin axis, with respect to which we construct spherical harmonics to decompose the gravitational wave strain in, we are inherently fixing the rotation freedom by constructing a canonical direction. However, while this aspect of the frame fixing may seem trivial, there are other freedoms that need to be fixed that are more subtle.

When fixing the frame of a system, one needs to fix the transformations that are contained within the system's symmetry group. For gravitational radiation, in which the symmetry group is the BMS group, this means that to fix the frame of, say, the waveform radiated to future null infinity in a BBH, one must not only fix the rotation freedom, but also the boost and the supertranslation freedom. But, for an arbitrary system, how should these freedoms be fixed? In principle, there is no canonical frame, since GR has no preferred frame. In practice, however, whatever analysis one would like to conduct on a system often has its own canonical frame, in the sense that there is a certain frame which makes the analysis much simpler. For example, when using the quadrupole formula to compute the gravitational wave emitted by a binary system, it tends to be easier to do the calculation in the frame in which the binary system's orbital angular momentum is aligned with the $\hat{z}$-axis. If a different orientation is used, then certain simplifications that would otherwise occur do not and one is instead met with many more terms than are necessary to explain the underlying physics. For studying numerical simulations, the same holds true. In particular, because working with NR results often consists of comparing NR waveforms to certain perturbative solutions to Einstein's equations, like PN in the early inspiral phase or black hole perturbation theory (BHPT) in the ringdown, the ideal frame is typically set by the perturbative result. Then, to map the NR system to this canonical frame one must work with the charges corresponding to the frame freedoms, i.e. the BMS charges in equations (4.8) and (4.10), and find the transformation which maps these charges in the NR system to the values expected by the canonical frame. For example, to map to the system's center-of-mass frame, one can simply take the translation and boost charges in equations (4.8) and find the translation and boost which maps these charges to zero.

While at first this business of fixing the frame may seem trivial, we stress that performing this procedure is of the utmost importance to extract meaningful physics. To help highlight this, we present figure 6, which shows the consequence of not accounting for the difference in frames between a NR and a PN system. One important aspect of waveform modeling is constructing waveforms which span the entire frequency band of GW detectors. This means that one must build waveforms that contain thousands of orbits. NR simulations, however,

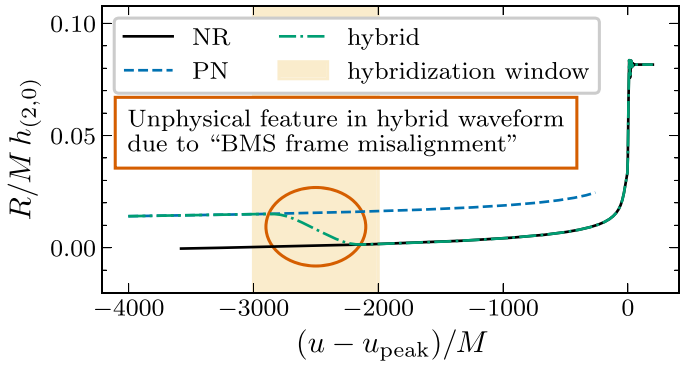

Figure 6. The $(2,0)$ mode of the strain waveform computed in a NR system (solid/black), a PN system (dashed/blue), and the corresponding NR/PN hybrid waveform (dash-dotted/green). The hybridization window is the highlighted region in yellow. An unphysical feature in the NR/PN hybrid waveform as a result of an improper BMS frame alignment is circled in red. The system is the same as that in figures 3 and 5.

typically only contain tens or perhaps hundreds of orbits, but not nearly as many as required to span the frequency bands of current and future observatories. Therefore, there is a constant need to 'hybridize' finite NR waveforms with perturbative solutions that are accurate during the early inspiral regimes and can be more efficiently calculated. Often these hybrid waveforms are built by hybridizing finite NR waveforms with PN waveforms. This results in waveforms which are PN during the early inspiral phase, a smooth blend of PN and NR over a window called the hybridization window, and NR for the remainder of the binary coalescence. As is shown by figure 6, however, if one does not account for the freedoms resulting from the BMS group, i.e. if one does not map the NR system to be in the same BMS frame as the PN system, then this hybridization procedure will fail in the sense that it will introduce unphysical features in the hybrid waveform. Fortunately, this issue can be mitigated rather easily by simply mapping the NR system to be in the right frame. But, in order to do so, one needs to know how the various BMS symmetries transform the asymptotic coordinates as well as the asymptotic data of interest.

Fortunately, this was already presented in section 2.5. With both equations (2.7) and (2.9), one has all of the information that is needed to transform asymptotic data. The only ingredient that remains for fixing the frame of the asymptotic data is which charges one should use to constrain the BMS freedoms[32]. Normally, one could just use the charges presented in equations (4.8) and (4.10). However, because it is often the case that one wishes to compare NR systems to PN predictions for which there is only the strain and not the Weyl scalars, it turns out that there are more convenient BMS charges that can be used. In particular, rather than working with the translation and boost charges, i.e. equation (4.8*a*) and equation (4.8*c*), it is useful to construct the center-of-mass charge [105, 128]

$$G^a(u) = \left(K^a + uP^a\right)/P^t = \frac{1}{4\pi}\int_{S^2} \mathrm{Re}\left[\left(\bar{\eth} n^a\right)\left(\hat{N} + u\eth m\right)\right] \mathrm{d}\Omega/P^t. \tag{7.1}$$

[32] Note that throughout this review we ignore the subtle fact that the BMS charges presented in equations (4.8) and in the remainder of this section are not supertranslation-invariant [66, 124–127].

This measures the spacetime's center-of-mass motion[33]. It is useful since we will always want to examine systems in their center-of-mass frame, which can be mapped to by finding the translation and boost which minimize equation (7.1). As for the rotation charge, again because we typically do not have access to either of the Ψ_1 or Ψ_2 Weyl scalars in PN theory, one typically utilizes a rotation-like charge that can be computed from the strain. In particular, in [59] such a rotation charge was built by finding the angular velocity which keeps the radiative fields, e.g. the strain, as constant as possible in the corotating frame of the binary system. This 'charge' is

$$\vec{\omega}(u) = -\langle \vec{L}\vec{L}\rangle^{-1} \cdot \langle \vec{L}\partial_t \rangle, \tag{7.2}$$

where

$$\langle \vec{L}\partial_t \rangle^a \equiv \sum_{\ell,m,m'} \mathrm{Im}\left[\bar{f}_{(\ell,m')}\langle \ell, m'|L^a|\ell,m\rangle \dot{f}_{(\ell,m)}\right], \tag{7.3a}$$

$$\langle \vec{L}\vec{L}\rangle^{ab} \equiv \sum_{\ell,m,m'} \bar{f}_{(\ell,m')}\langle \ell, m'|L^{(a}L^{b)}|\ell,m\rangle f_{(\ell,m)}, \tag{7.3b}$$

and $f(u,\theta,\phi)$ is some function corresponding to the asymptotic radiation, such as the GW strain or the news. In equation (7.2), $\vec{L}$ is the infinitesimal generator of rotations.

When the system only consists of one black hole, as is the case, for example, when studying the ringdown phase, this prescription can break down. In this case, it is more useful to fix the rotation freedom using the intrinsic spin of the black hole, to ensure that one is in the frame of the individual black hole. To compute this, one can either map to the center-of-mass frame of the black hole and compute the angular momentum charge, or, for simplicity, they can instead compute [54]

$$\vec{\chi}(u) = \frac{\gamma}{M_B^2}\left(\vec{J} + \vec{v}\times\vec{K}\right) - \frac{\gamma - 1}{M_B^2}\left(\hat{v}\cdot\vec{J}\right)\hat{v}, \tag{7.4}$$

which achieves the same result. Here

$$\gamma(u) \equiv \left(1 - |\vec{v}|^2\right)^{-1/2} \tag{7.5}$$

is the Lorentz factor,

$$M_B(u) \equiv \sqrt{-\eta_{\mu\nu}P^\mu P^\nu} \tag{7.6}$$

is the Bondi mass,

$$\vec{v}(u) \equiv \vec{P}/P^t \tag{7.7}$$

is the velocity vector, and the vectors $\vec{J}$ and $\vec{K}$ are the angular momentum and boost charges found in equations (4.8*b*) and (4.8*c*). With this 'charge', one can then map to the frame of the black hole by, say, finding the rotation which aligns this charge with the positive $\hat{z}$-axis. Again, we stress that equations (7.2) and (7.4) have been introduced to help with fixing the rotation

[33] One can think of the boost charge K^a as measuring the translation away from the origin and the uP^a term as measuring the change in center-of-mass due to a non-zero velocity.

freedom by constructing a preferred axis, either through the angular velocity vector of a PN waveform or the spin of an isolated black hole.

For fixing the supertranslation, since for fixing the Poincaré transformations we have been able to rely on the charges in equations (4.8), one may naively think that we can simply use the supermomentum charge in equation (4.10). This charge, however, turns out to be supertranslation-invariant in nonradiative regimes of future null infinity, i.e. regimes where the news is zero. As a result, it cannot be used to fix the supertranslation freedom[34]. Instead, one must construct a different supertranslation 'charge' which transforms in a meaningful way. In [55–57, 129], such a charge was presented. It is simply equation (4.11) with $p=1$ and $q=0$ and is the Moreschi supermomentum

$$\mathcal{P}_{\rm M}(u,\theta,\phi) \equiv \Psi_2 + \sigma\dot{\bar{\sigma}} + \eth^2\bar{\sigma} = -m + \mathrm{Re}\left[\eth^2\bar{\sigma}\right] = \int_{-\infty}^{u} |\dot{\sigma}|^2\,\mathrm{d}u - M_{\rm ADM}, \tag{7.8}$$

where $M_{\rm ADM}$ is the ADM mass [130]. An important property of the Moreschi supermomentum is that, because it transforms as

$$\mathcal{P}'_{\rm M} = \frac{1}{k^3}\left(\mathcal{P}_{\rm M} - \eth^2\bar{\eth}^2\alpha\right) = \frac{1}{k^3}\left(-m + \mathrm{Re}\left[\bar{\eth}^2\left(\sigma - \eth^2\alpha\right)\right]\right), \tag{7.9}$$

in Minkowski space where $m=0$, minimizing the Moreschi supermomentum corresponds to finding the supertranslation that ensures the spacetime is really Minkowski, not supertranslated Minkowski. Furthermore, this property is true so long as the $\ell \geqslant 2$ components of the mass aspect are zero, which is the same as there being no Geroch supermomentum. This is useful because isolated black holes cannot have Geroch supermomentum, so if one instead uses the Moreschi supermomentum to fix the BMS freedom, they ensure that they are always working in the frame that is the most natural frame from the perspective of the individual black hole. We will call the BMS frame in which the Moreschi supermomentum has zero-valued $\ell \geqslant 1$ modes the *superrest frame*, as it is in some sense an extension of the notion of a rest frame. To map to such a frame, one can simply solve equation (7.9) for $\mathcal{P}'_M = -M_B$, which yields

$$\eth^2\bar{\eth}^2\alpha = \mathcal{P}_M(u=\alpha,\theta,\phi) + k_{\rm rest}(\alpha,\theta,\phi)^3 M_B(\alpha) \tag{7.10}$$

where $k_{\rm rest}$ is a special case of the conformal factor in equation (2.4) in the sense that it is the conformal factor for a boost whose velocity matches the momentum charge:

$$k_{\rm rest}(u,\theta,\phi) \equiv \frac{1}{\gamma\left(1-\vec{v}\cdot\vec{r}\right)} = \frac{M_B}{P_a n^a}. \tag{7.11}$$

Furthermore, in [57], the authors proved that, for a condition on the energy flux, which is always obeyed in nonradiative regimes of future null infinity, equation (7.10) always has a regular solution. It was also shown that equation (7.10) can be solved iteratively. That is, if one wishes to find the supertranslation that maps a system to the superrest frame at some time u_0, they can evaluate the right-hand side of equation (7.10) at time $u=u_0$, solve for α, evaluate the right-hand side at time $u=\alpha$, solve for a new α, etc until α converges to a solution. This is the underpinning of the frame fixing program. Specifically, by utilizing this fact, one can always solve for the infinite degrees of freedom in the BMS transformation and map to a target

[34] This phenomenon should not necessarily come as a surprise, as this also happens for the translation-invariant momentum charge.

BMS frame in an iterative fashion. For example, to map a NR system to the PN BMS frame, one can

(i) Construct a window during the early inspiral phase over which to align a NR system to a PN system;
(ii) Iteratively solve equation (7.10) for the supertranslation which maps the $\ell \geqslant 2$ components of the difference of the NR/PN Moreschi supermomenta to zero; this is equivalent to mapping to the superrest frame at $u \to -\infty$, i.e. in the infinite past of the binary;
(iii) Find the frame rotation that maps the NR system's angular velocity vector (see equation (7.2)) to match that of the PN system;
(iv) Find the translation and boost which minimize the center-of-mass charge (see equation (7.1));
(v) Perform a time and phase translation optimization to align the NR and PN waveforms in the window;
(vi) Repeat steps (i)–(v), until the error between the NR and PN waveforms converges.

For mapping to the frame of an individual black hole, like a remnant black hole, the process is identical, but in step (ii) one can minimize the Moreschi supermomentum, since the target Moreschi supermomentum is $-M_B$ and in step (iii) the rotation charge should instead be the spin of the black hole (see equation (7.4)). Also, step (v) is no longer necessary since there is no preferred way to fix the time or phase freedom from the perspective of the remnant black hole. In the subsequent sections, we explain in more detail how to map a system to either the PN BMS frame or the superrest frame of an individual black hole and provide some numerical results showing why frame fixing is important and the success of this procedure at mapping NR systems to some reasonable BMS frame.

7.2. PN BMS frame

At the end of section 7.1 the steps for mapping to the PN BMS frame through an iterative process using the BMS charges was outlined. Ultimately, to map to this frame one must simply find the BMS transformation which maps the charges of the NR system to match those of the PN system. This, however, requires that one has knowledge of the charges in PN theory. Unfortunately, most PN calculations focus on computing the strain and not the Weyl scalars, so the majority of the canonical BMS charges can not be computed in a PN formulation. Therefore, one can either compute these charges in PN or try to use alternative charges that are functions of only the strain. Because the later is easier, this is what is typically used in BMS frame fixing. But, even if one uses charges that are only functions of the strain, these charges must still be computed in PN. Fortunately, the center-of-mass charge, which is zero for PN systems, and the angular velocity vector, which can easily be computed numerically from the PN strain, are painless to obtain. The PN Moreschi supermomentum, however, because it involves a time integral, cannot be computed numerically and must instead be worked out analytically. This was first performed in [53], which computed it from the PN strain using equation (7.8) to 3PN order without spins and 2PN order with spins. They then implemented this iterative procedure for fixing the frame of NR systems in the publicly available python module `scri` [58].

In figure 7, we demonstrate the overall success of the BMS frame fixing procedure by mapping a NR system to the PN BMS frame and computing the error between the NR waveform and the PN waveform in this frame. This figure is identical to figure 6, but correctly utilizes the BMS freedoms at future null infinity to perform the waveform alignment. As can be seen, by

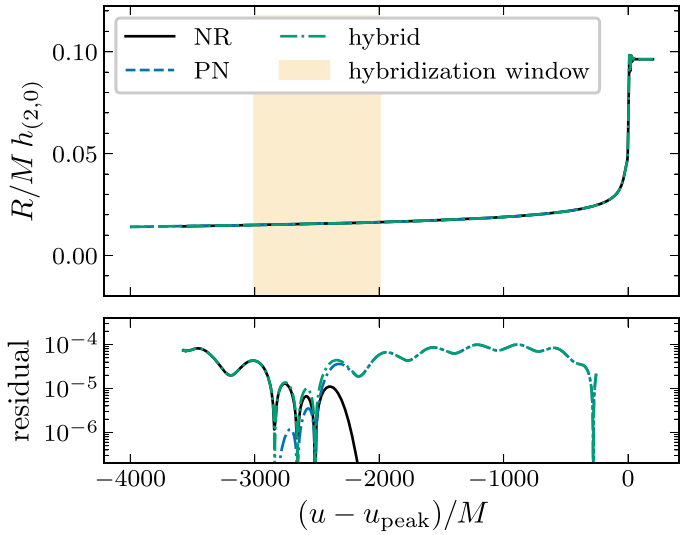

Figure 7. The $(2, 0)$ mode of the strain waveform computed in a NR system (solid/-black), a PN system (dashed/blue), and the corresponding NR/PN hybrid waveform (dash-dotted/green). The hybridization window is the highlighted region in yellow. In the bottom panel, we show he absolute error between the NR and PN waveforms (green), the hybrid and NR (black), and the hybrid and the PN (blue). The NR simulation is the same as that used in figures 3, 5 and 6.

mapping the NR waveform to the PN BMS frame, the absolute error between the two waveforms is decreased by three orders of magnitude. Also, as seen through the hybrid waveform built from these two inputs, by fixing the frame properly the hybrid waveform no longer has an unphysical feature that could bias data analyses that used this waveform. Furthermore, we stress that although here the effects of fixing the frame are a bit pronounced due to us studying the $(2, 0)$ mode, if one instead studies other modes they will still find that the absolute error is improved.

7.3. Superrest frame

Apart from comparing NR systems to PN predictions to construct more effective hybrid waveforms and models, NR simulations of black holes are also particularly useful for extracting the quasi-normal mode (QNM) amplitudes expected by perturbation theory [131, 132]. A subtle issue, however, is that when a perturbed black hole is formed, either through a black hole merger or stellar collapse, it is not in the frame that BHPT is typically performed in. Put differently, the remnant black hole formed in some astrophysical event may not be described by the usual Kerr metric, but instead a boosted or a supertranslated version. Consequently, if one tries to study the perturbations of the remnant black hole formed in a NR simulation without accounting for the difference in BMS frames, then the analysis will fail in the same way that the hybridization shown in figure 6 fails.

In particular, in figure 8, we show exactly this. In the top row we show two fits of QNMs to the $(2, 0)$ mode of the NR strain waveform from the simulation shown in figures 3, 5 and 6. Meanwhile, in the bottom row we show the residuals. The QNM model that we consider has

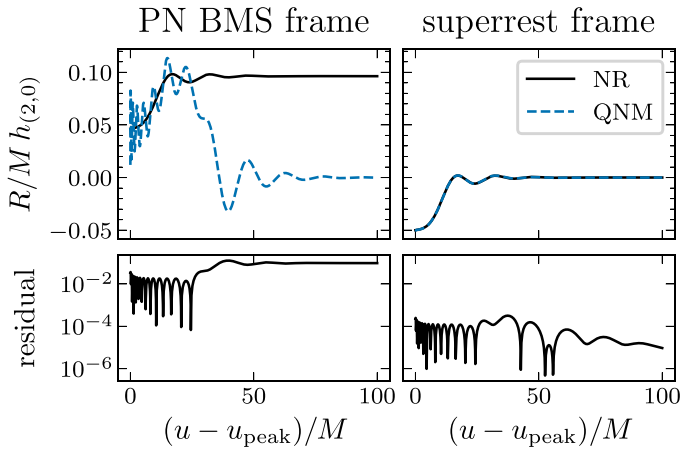

Figure 8. Top row: the $(2,0)$ mode of the strain waveform for a NR system (black) and the best fit QNM model (blue). The QNM model contains seven $(2,0)$ overtones. On the left, the NR waveform is in the PN BMS frame, while on the right, it is instead in the superrest frame of the remnant black hole. Bottom row: residual between the NR waveform and the fit.

seven overtones and both mirror modes[35]. As seen through the left-most column, which corresponds to the NR waveform in the PN BMS frame, i.e. the wrong frame for BHPT analyses, when the NR system is not in the superrest frame of the remnant black hole the strain can not be fit with QNMs, in part because of the nonzero asymptotic value resulting from the memory effect. But, if one maps the NR system to the superrest frame of the remnant black hole, then the QNM fit can be performed successfully and the residual is reduced by three orders of magnitude. This illustrates that fixing the frame is essential for extracting physical QNM amplitudes.

Another important result that can be seen through figure 8 and the bottom-left panel of figure 3 is an example of the nonlinear nature of GR and a highlight of the need for predictions from second-order perturbation theory to study the ringdown of NR waveforms. More specifically, since a large portion of the $(2,0)$ mode of the strain is sourced by the energy flux, i.e. equation (4.7*a*), this means that first-order perturbation theory is not sufficient to explain the ringdown excitations of the strain near the point of peak luminosity $u_{\rm peak}$. In particular, because the energy flux is inherently nonlinear, i.e. it goes as the strain squared, this means it can only be modeled by second-order BHPT, e.g. by the quadratic QNMs studied in [96, 133].

8. Discussion

In this review, we have presented a vast amount of information regarding memory effects and the BMS group. In particular, we began in section 2.1 by studying where in a NR simulation one should extract gravitational wave information for gravitational wave detectors. While it may be tempting to extract GWs on some timelike wordline, like that of an inertial detector,

[35] We stress that here we use this model only for simplicity and to show the importance of fixing the frame. In reality, when trying to correctly fit a NR waveform with QNMs, one should take precaution to ensure that the QNMs are meaningfully resolved.

we showed how this can quickly lead to coordinate ambiguities and thus motivated the need to extract GWs at future null infinity $\mathcal{I}^+$—the final destination of outgoing radiation. By working with data on $\mathcal{I}^+$, instead of in the bulk, one is free from the full diffeomorphism invariance of GR and obtains a more suitable set of coordinates to work with. However, as we outlined in section 2.2, even on $\mathcal{I}^+$ that are still troubling coordinate ambiguities that are present because of the symmetry group of null infinity: the BMS group. Consequently, to work with data on $\mathcal{I}^+$, one must have a robust understanding of the BMS group and the way in which it transforms coordinates and asymptotic data.

In section 2.3 we reviewed the infamous BMS group; specifically how it can be viewed as the Poincaré group, but with the usual spacetime translations replaced by an infinite group of transformations called *supertranslations*. These can be viewed as direction-dependent translations and arise from the fact that each point on $\mathcal{I}^+$, which can each be thought of as a single observer, ends up being casually disconnected from any other. Then, in section 2.5, we showed how an arbitrary BMS transformation changes the spacetime coordinates that the asymptotic data will be functions of. We also explained what this data on $\mathcal{I}^+$ should be, i.e. not only the gravitational wave strain, but also the five complex Weyl scalars, which encode information about the spacetime curvature.

Finally, in section 2.6, we used our understanding of the symmetries of $\mathcal{I}^+$ to construct an intuition for why memory effects should exist in GR. In particular, because of Neother's theorem [62], one might expect that the supertranslation symmetries of $\mathcal{I}^+$ yield some kind of conservation law between an angle-dependent mass and an angle-dependent energy flux, like that of equation (2.13). However, because changes to the $\ell \geqslant 2$ harmonics of system's mass multipole moment source GWs, one needs an extra ingredient on the left-hand side of equation (2.13), i.e. that of equation (2.14), to balance this fact. Therefore, the conservation law that stems from supertranslations, equation (2.14) or equations (4.6), is really a statement about the memory that is induced on the spacetime by any system that emits gravitational radiation.

After our pedagogical overview, in section 3 we reviewed the literature behind memory effects and the BMS group. Focusing on the memory first, we highlighted how this effect has been known since 1974, but did not receive much attention until its connection to asymptotic symmetries and soft theorems was found in 2013 and 2014 [73–75, 79]. Ever since this realization, there has been a surge of interest and discovery in tangentially related fields, such as celestial holography [17–21].

In section 4, we presented the more formal mathematics behind our earlier discussions. In particular, we showed how the Bondi-Sachs metric and Einstein's equations yield a series of evolution equations for both the mass and angular momentum aspects (see equations (4.2)), which, when integrated with respect to time, yield the formal version of the pedagogical conservation law of equation (2.14), i.e. equations (4.6). These equations show that the real and imaginary parts of the gravitational wave strain (or shear) can be written in terms of mass and angular momentum charges and fluxes, which help illustrate the different ways in which the strain or memory can be sourced. That is, equations (4.6) present the modern interpretation of memory: it can be sourced by a change in the system's charges (*linear* or *ordinary* memory) or a change in the fluxes (*nonlinear*, *Christodoulou*, or *null* memory).

In section 6, we implemented our mathematical results numerically for simulations of BBH. In section 6.1 we took equations (4.6) and evaluated them for a GW150914-like NR simulation in figure 3. This showed how useful equations (4.6) can be for studying NR simulations

and extracting the memory contribution to the strain, as highlighted in the lower left panel of figure 3. In section 6.2, in figure 4 we showed what the frequency spectrum of a typical memory signal looks like relative to the LIGO and CE amplitude spectral densities. As shown, even though memory is fairly pronounced in the time domain (see, e.g. figure 2), in the frequency domain memory occupies the lower part of the band which makes it challenging to detect in single events with current detectors. However, by stacking events and trying to detect memory in a population [81–83] or by relying on future detectors like CE, ET, or LISA [83, 134], the chances for observing memory for the first time are much higher.

Last, in section 7 we highlighted the importance of controlling the BMS freedoms for studying NR waveforms. First, in figure 6, we showed how NR and PN waveforms naturally disagree with one another, due to the fact that PN waveforms have information about the past history of the compact binary's inspiral, while numerical simulations do not. Furthermore, figure 6 shows that, without resolving this issue, if one tries to construct a hybrid waveform which is needed for building waveform models for the LVK collaboration, then the hybrid waveform exhibits an unphysical feature due to this disagreement. Fortunately, this disagreement simply turns out to be an issue of the two waveforms being in different BMS frames, i.e. they differ by some arbitrary BMS transformation. In the text surrounding figure 6, we outlined how one can use BMS charges to determine this transformation and fix the BMS frame of the NR waveform to match that of the PN waveform. After doing so, one obtains figure 7 which shows much better agreement between NR and PN strain waveforms. Finally, in figure 8, we also highlighted why fixing the BMS frame is important for studying the ringdown phase of NR simulations with BHPT. In particular, because of the memory, the remnant black hole formed in a black hole merger is not exactly Kerr, but *supertranslated Kerr*. Therefore, if one wishes to fit the ringdown phase with QNMs, they need to find the BMS transformation that maps the remnant black hole to the same frame of the usual Kerr metric: the *superrest* frame. After doing so, one can then perform meaningful ringdown fits and analyses, as is illustrated through the right-hand panels of figure 8.

It is our hope that this review will serve as a resource for those interested in learning more about memory effects and the BMS group in the context of NR simulations and GW modeling more generally. As new and more accurate GW detectors come online, understanding the complexities of the GWs that GR predicts, such as those we have covered here, will be of the utmost importance, as they will likely be key to observing never-before-seen physics that will futher illuminate how our Universe works.

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

The authors thank Eanna Flanagan for informative conversations about the history of memory effects and Alexander Grant for intuition behind BMS balance laws. Computations for this work were performed with the Wheeler cluster at Caltech. This work was supported by the Sherman Fairchild Foundation and NSF Grants Nos. PHY-2011968, PHY-2011961, PHY-2309211, PHY-2309231, OAC-2209656 at Caltech. The work of LCS was partially supported by NSF CAREER Award PHY-2047382 and a Sloan Foundation Research Fellowship.

Appendix. Expressing a BMS transformation in terms of rotations, boosts, and supertranslations

In this appendix, we briefly outline how one can write an arbitrary BMS transformation as a supertranslation followed by a Lorentz transformation. This is useful for mapping NR waveforms to a particular BMS frame because it enables one to compose BMS transformations. We begin by studying Lorentz transformations.

A.1. $\mathrm{SL}(2,\mathbb{C})$ *representation of a Lorentz transformation*

Let us begin with rotations. For a rotation R by an angle θ about the axis $\hat{r} = (r_x, r_y, r_z)$, one may write this as the quaternion[36]

$$\mathbf{q} = \exp\left(\frac{1}{2}\theta\hat{r}\right) = \cos(\theta/2)\,\mathbf{I} + (r_x\mathbf{i} + r_y\mathbf{j} + r_z\mathbf{k})\sin(\theta/2)\,, \tag{A1}$$

where $\mathbf{I}$, $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are the elementary quaternions obeying the usual multiplication rules. Using spin matrices, i.e. elements of $\mathrm{SL}(2,\mathbb{C})$, one may write these quaternions as

$$\boldsymbol{\sigma} \equiv (\mathbf{I},\mathbf{i},\mathbf{j},\mathbf{k}) = \left(\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & i\\ i & 0\end{pmatrix}, \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix}\right). \tag{A2}$$

Therefore, the rotation R has the $\mathrm{SL}(2,\mathbb{C})$ representation

$$\tilde{R} = (\cos(\theta/2)\,, \hat{r}\sin(\theta/2))\cdot\boldsymbol{\sigma}. \tag{A3}$$

Note that in equation (A3) $\tilde{R}$ is a unitary matrix. Meanwhile, for boosts, since a boost B by rapidity w along the axis $\hat{v} = (v_x, v_y, v_z)$ is nothing more than a rotation by the angle iw about the axis $\hat{v}$, a boost has the representation

$$\tilde{B} = (\cosh(w/2)\,, i\hat{v}\sinh(w/2))\cdot\boldsymbol{\sigma}. \tag{A4}$$

Note that in equation (A4) $\tilde{B}$ is a Hermitian matrix. Because $\tilde{L}$—the $\mathrm{SL}(2,\mathbb{C})$ representation of some Lorentz transformation—is invertible, it has a unique polar decomposition. Correspondingly, a general Lorentz transformation can be written as a composition of a boost followed by a rotation, which in the $\mathrm{SL}(2,\mathbb{C})$ representation is

$$\tilde{L} = \tilde{R}\cdot\tilde{B} = [(\cos(\theta/2)\,, \hat{r}\sin(\theta/2))\cdot\boldsymbol{\sigma}]\cdot[(\cosh(w/2)\,, i\hat{v}\sinh(w/2))\cdot\boldsymbol{\sigma}]\,. \tag{A5}$$

Note though that what is shown above is for a *passive* Lorentz transformation. Put simply, the transformation is acting on the coordinates themselves. Under L, our original coordinates X are transformed to an intermediate coordinate system X' by a boost, and then to a final coordinate system X' by a rotation. The parameters of the rotation should be interpreted as *a rotation in the X' coordinate system*, not in the X coordinate system.

[36] For a review of quaternions, see [135].

A.2. Decomposing an $\mathrm{SL}(2,\mathbb{C})$ *matrix into rotation and boost*

Because $\mathrm{SL}(2,\mathbb{C})$ is a double cover of the restricted Lorentz group $\mathrm{SO}^+(3,1)$, we can always write the $\mathrm{SL}(2,\mathbb{C})$ representation of a Lorentz transformation L as

$$\tilde{L} \equiv \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{A6}$$

for some $a,b,c,d \in \mathbb{C}$ with $ad-bc=1$. We would like to perform a polar decomposition as either $\tilde{L}=\tilde{R}\cdot\tilde{B}$ or $\tilde{L}=\tilde{B}'\cdot\tilde{R}'$, where $\tilde{R}$ or $\tilde{R}'$ is unitary and thus a rotation, and where $\tilde{B}$ or $\tilde{B}'$ is hermitian and thus a boost, as seen in equations (A3) and (A4). This is easily accomplished with a singular value decomposition,

$$\tilde{L} = U\cdot\Sigma\cdot V^\dagger, \tag{A7}$$

where U and V are complex unitary matrices and Σ is diagonal, and in this case Hermitian. Using the inverse properties of U and V, we can write $\tilde{L}$ in the more suggestive form

$$\tilde{L} = \left(U\cdot V^\dagger\right)\cdot\left(V\cdot\Sigma\cdot V^\dagger\right) \tag{A8}$$

or

$$\tilde{L} = \left(U\cdot\Sigma\cdot U^\dagger\right)\cdot\left(U\cdot V^\dagger\right), \tag{A9}$$

which give our desired decompositions,

$$\tilde{L}=\tilde{R}\cdot\tilde{B} \text{ with } \tilde{R}=U\cdot V^\dagger \text{ and } \tilde{B}=V\cdot\Sigma\cdot V^\dagger \tag{A10}$$

or

$$\tilde{L}=\tilde{B}'\cdot\tilde{R}' \text{ with } \tilde{B}'=U\cdot\Sigma\cdot U^\dagger \text{ and } \tilde{R}'=U\cdot V^\dagger. \tag{A11}$$

A.3. Composition of BMS elements

Equipped with the $\mathrm{SL}(2,\mathbb{C})$ representation of arbitrary Lorentz transformations, we also want a decomposition of BMS group elements as compositions of supertranslations followed by Lorentz transformations, or vice versa. This will make it computationally convenient to compose BMS transformations. Such a decomposition is possible since supertranslations $\mathbb{T}$ are a normal subgroup of the BMS group $\mathfrak{B}$. Namely, we can construct the homomorphism to the quotient group $\varphi:\mathfrak{B}\to\mathfrak{B}/\mathbb{T}\simeq\mathrm{SO}(3,1)$. We can say that a BMS group element s is a 'pure supertranslation' if $s\in\ker\varphi$.

Now let a BMS transformation be

$$g = l\circ s, \tag{A12}$$

for l a Lorentz transformation and s a supertranslation. The Lorentz transformation component of a BMS transformation $l=\varphi(g)$ is clearly well-defined. Alternatively we can write a g as

$$g = \hat{s}\circ l \tag{A13}$$

for $\hat{s}=l\circ s\circ l^{-1}$ some other pure supertranslation.

We can now consider composing $g_1, g_2 \in \mathfrak{B}$,

$$g = g_2 \circ g_1 = l_2 \circ s_2 \circ l_1 \circ s_1. \tag{A14}$$

Because s_2 is in the normal subgroup, we can write it as conjugate to another pure supertranslation, namely

$$s_2 = l_1 \circ s' \circ l_1^{-1} \tag{A15}$$

for some $s' \in \mathbb{T}$. Thus

$$g = (l_2 \circ l_1) \circ (s' \circ s_1) \tag{A16}$$

with $s' = l_1^{-1} \circ s_2 \circ l_1$. Note that here the composition of the Lorentz transformations can be carried out via the $\mathrm{SL}(2,\mathbb{C})$ representation, whereas the composition of supertranslations is simply additive.

ORCID iDs

Keefe Mitman https://orcid.org/0000-0003-0276-3856
Michael Boyle https://orcid.org/0000-0002-5075-5116
Leo C Stein https://orcid.org/0000-0001-7559-9597
Nils Deppe https://orcid.org/0000-0003-4557-4115
Lawrence E Kidder https://orcid.org/0000-0001-5392-7342
Jordan Moxon https://orcid.org/0000-0001-9009-6955
Harald P Pfeiffer https://orcid.org/0000-0001-9288-519X
Mark A Scheel https://orcid.org/0000-0001-6656-9134
Saul A Teukolsky https://orcid.org/0000-0001-9765-4526
William Throwe https://orcid.org/0000-0001-5059-4378
Nils L Vu https://orcid.org/0000-0002-5767-3949

References

[1] Buonanno A and Damour T 1999 Effective one-body approach to general relativistic two-body dynamics *Phys. Rev.* D **59** 084006
[2] Husa S, Khan S, Hannam M, Pürrer M, Ohme F, Forteza X J and Bohé A 2016 Frequency-domain gravitational waves from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal *Phys. Rev.* D **93** 044006
[3] Khan S, Husa S, Hannam M, Ohme F, Pürrer M, Forteza X J and Bohé A 2016 Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era *Phys. Rev.* D **93** 044007
[4] Pratten G *et al* 2021 Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes *Phys. Rev.* D **103** 104056
[5] Estellés H, Colleoni M, García-Quirós C, Husa S, Keitel D, Mateu-Lucena M, Planas M de L and Ramos-Buades A 2022 New twists in compact binary waveform modeling: a fast time-domain model for precession *Phys. Rev.* D **105** 084040
[6] Taracchini A, Pan Y, Buonanno A, Barausse E, Boyle M, Chu T, Lovelace G, Pfeiffer H P and Scheel M A 2012 Prototype effective-one-body model for nonprecessing spinning inspiral-merger-ringdown waveforms *Phys. Rev.* D **86** 024011
[7] Lackey B D, Kyutoku K, Shibata M, Brady P R and Friedman J L 2014 Extracting equation of state parameters from black hole-neutron star mergers: aligned-spin black holes and a preliminary waveform model *Phys. Rev.* D **89** 043009

[8] Pan Y, Buonanno A, Taracchini A, Kidder L E, Mroué A H, Pfeiffer H P, Scheel M A and Szilágyi B 2014 Inspiral-merger-ringdown waveforms of spinning, precessing black-hole binaries in the effective-one-body formalism *Phys. Rev.* D **89** 084006

[9] Bohé A *et al* 2017 Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors *Phys. Rev.* D **95** 044028

[10] Ossokine S *et al* 2020 Multipolar effective-one-body waveforms for precessing binary black holes: construction and validation *Phys. Rev.* D **102** 044055

[11] Ramos-Buades A, Buonanno A, Estellés H, Khalil M, Mihaylov D P, Ossokine S, Pompili L and Shiferaw M 2023 Next generation of accurate and efficient multipolar precessing-spin effective-one-body waveforms for binary black holes *Phys. Rev.* D **108** 124037

[12] Zel'dovich Y B and Polnarev A G 1974 Radiation of gravitational waves by a cluster of superdense stars *Sov. Astron.* **18** 17

[13] Braginsky V B and Grishchuk L P 1985 Kinematic resonance and memory effect in free mass gravitational antennas *Sov. Phys. - JETP* **62** 427–30

[14] Favata M 2010 The gravitational-wave memory effect *Class. Quantum Grav.* **27** 084036

[15] Bondi H, Van der Burg M G J and Metzner A W K 1962 Gravitational waves in general relativity, VII. Waves from axi-symmetric isolated system *Proc. R. Soc.* A **269** 21–52

[16] Sachs R 1962 Gravitational waves in general relativity. VIII. Waves in asymptotically flat space-time *Proc. R. Soc.* A **270** 103–26

[17] Strominger A 2017 Lectures on the infrared structure of gravity and gauge theory (arXiv: 1703.05448 [hep-th])

[18] Pasterski S 2019 Implications of superrotations *Phys. Rep.* **829** 1–35

[19] Raclariu A-M 2021 Lectures on celestial holography (arXiv:2107.02075 [hep-th])

[20] Pasterski S 2021 Lectures on celestial amplitudes *Eur. Phys. J.* C **81** 1062

[21] Pasterski S, Pate M and Raclariu A-M 2021 Celestial holography (arXiv:2111.11392 [hep-th])

[22] Cervantes-Cota J L, Galindo-Uribarri S and Smoot G-F 2016 A brief history of gravitational waves *Universe* **2** 22

[23] Nerozzi A, Beetle C, Bruni M, Burko L M and Pollney D 2005 Towards wave extraction in numerical relativity: the quasi-Kinnersley frame *Phys. Rev.* D **72** 024014

[24] Nerozzi A 2017 Spin coefficients and gauge fixing in the Newman-Penrose formalism *Phys. Rev.* D **95** 064012

[25] Pirani F A E 1956 On the physical significance of the Riemann tensor *Acta Phys. Polon.* **15** 389–405

[26] Bondi H 1960 Gravitational waves in general relativity *Nature* **186** 535–535

[27] Sachs R 1961 Gravitational waves in general relativity. VI. The outgoing radiation condition *Proc. R. Soc.* A **264** 309–38

[28] Sachs R 1962 Asymptotic symmetries in gravitational theory *Phys. Rev.* **128** 2851–64

[29] Lehner L and Moreschi O M 2007 Dealing with delicate issues in waveforms calculations *Phys. Rev.* D **76** 124040

[30] Bonga B and Prabhu K 2020 Bms-like symmetries in cosmology *Phys. Rev.* D **102** 104043

[31] Rojo M E and Schröder T 2023 Asymptotic symmetries and memories of gauge theories in FLRW spacetimes *J. High Energy Phys.* JHEP01(2023)011

[32] Jokela N, Kajantie K and Sarkkinen M 2022 Gravitational wave memory and its tail in cosmology *Phys. Rev.* D **106** 064022

[33] Rojo M E and Heckelbacher T 2021 Asymptotic symmetries in spatially flat FRW spacetimes *Phys. Rev.* D **103** 064009

[34] Fernández-Álvarez F and Senovilla J M M 2022 Asymptotic structure with vanishing cosmological constant *Class. Quantum Grav.* **39** 165011

[35] Rojo M E, Heckelbacher T and Oliveri R 2023 Asymptotic dynamics and charges for FLRW spacetimes *Phys. Rev.* D **107** 024039

[36] Newman E T and Unti T W J 1962 Behavior of asymptotically flat empty spaces *J. Math. Phys.* **3** 891–901

[37] Penrose R 1963 Asymptotic properties of fields and space-times *Phys. Rev. Lett.* **10** 66–68

[38] Penrose R 1965 Zero rest-mass fields including gravitation: asymptotic behaviour *Proc. R. Soc.* A **284** 159–203

[39] Hawking S W and Ellis G F R 1976 *The Large Scale Structure of Space-Time* (Cambridge University Press)

[40] Geroch R 1977 Asymptotic structure of space-time *Asymptotic Structure of Space-Time* ed F P Esposito and L Witten (Springer) pp 1–105
[41] Geroch R and Horowtiz G T 1978 Asymptotically simple does not imply asymptotically Minkowskian *Phys. Rev. Lett.* **40** 203–6
[42] Mädler T and Winicour J 2016 Bondi-Sachs formalism *Scholarpedia* **11** 33528
[43] Barnich G and Troessaert C 2010 Aspects of the BMS/CFT correspondence *J. High Energy Phys.* JHEP05(2010)062
[44] Campiglia M and Laddha A 2014 Asymptotic symmetries and subleading soft graviton theorem *Phys. Rev.* D **90** 124028
[45] Campiglia M and Laddha A 2015 New symmetries for the gravitational S-matrix *J. High Energy Phys.* JHEP04(2015)076
[46] Compère G, Fiorucci A and Ruzziconi R 2018 Superboost transitions, refraction memory and super-Lorentz charge algebra *J. High Energy Phys.* JHEP11(2018)200
[47] Flanagan E É, Prabhu K and Shehzad I 2020 Extensions of the asymptotic symmetry algebra of general relativity *J. High Energy Phys.* JHEP01(2020)002
[48] Boyle M 2016 Transformations of asymptotic gravitational-wave data *Phys. Rev.* D **93** 084031
[49] Newman E and Penrose R 1962 An Approach to gravitational radiation by a method of spin coefficients *J. Math. Phys.* **3** 566–78
[50] Szekeres P 1965 The gravitational compass *J. Math. Phys.* **6** 1387–91
[51] Geroch R P, Held A and Penrose R 1973 A space-time calculus based on pairs of null directions *J. Math. Phys.* **14** 874–81
[52] Mitman K *et al* 2021 Fixing the BMS frame of numerical relativity waveforms *Phys. Rev.* D **104** 024051
[53] Mitman K *et al* 2022 Fixing the BMS frame of numerical relativity waveforms with BMS charges *Phys. Rev.* D **106** 084029
[54] Iozzo D A B *et al* 2021 Comparing remnant properties from horizon data and asymptotic data in numerical relativity *Phys. Rev.* D **103** 124029
[55] Moreschi O M 1988 Supercenter of mass system at future null infinity *Class. Quantum Grav.* **5** 423–35
[56] Moreschi O M and Dain S 1998 Rest frame system for asymptotically flat space-times *J. Math. Phys.* **39** 6631–50
[57] Dain S and Moreschi O M 2000 General existence proof for rest frame systems in asymptotically flat space-time *Class. Quantum Grav.* **17** 3663–72
[58] Boyle M, Iozzo D, Stein L, Khairnar A, Rüter H, Scheel M, Varma V and Mitman K 2024 scri (Zenodo) (https://doi.org/10.5281/zenodo.11001966)
[59] Boyle M 2013 Angular velocity of gravitational radiation from precessing binaries and the corotating frame *Phys. Rev.* D **87** 104006
[60] Boyle M, Kidder L E, Ossokine S and Pfeiffer H P 2014 Gravitational-wave modes from precessing black-hole binaries (arXiv:1409.4431 [gr-qc])
[61] Dray T and Streubel M 1984 Angular momentum at null infinity *Class. Quantum Grav.* **1** 15–26
[62] Wald R M and Zoupas A 2000 A general definition of 'conserved quantities' in general relativity and other theories of gravity *Phys. Rev.* D **61** 084027
[63] Kovacs S J and Thorne K S 1978 The generation of gravitational waves. 4. Bremsstrahlung *Astrophys. J.* **224** 62–85
[64] Bieri L and Garfinkle D 2014 Perturbative and gauge invariant treatment of gravitational wave memory *Phys. Rev.* D **89** 084039
[65] Braginsky V B and Thorne K S 1987 Gravitational-wave bursts with memory and experimental prospects *Nature* **327** 123–5
[66] Christodoulou D 1991 Nonlinear nature of gravitation and gravitational wave experiments *Phys. Rev. Lett.* **67** 1486–9
[67] Thorne K S 1992 Gravitational-wave bursts with memory: the Christodoulou effect *Phys. Rev.* D **45** 520–4
[68] Mädler T and Winicour J 2016 The sky pattern of the linearized gravitational memory effect *Class. Quantum Grav.* **33** 175006
[69] Zhang P M, Zhao Q L and Horvathy P A 2024 Displacement memory for flyby (arXiv:2407.10787 [gr-qc])
[70] Newman E T and Penrose R 1966 Note on the Bondi-Metzner-Sachs group *J. Math. Phys.* **7** 863–70

[71] Payne P N 1983 Smarr's zero frequency limit calculation *Phys. Rev.* D **28** 1894–7
[72] Blanchet L and Damour T 1992 Hereditary effects in gravitational radiation *Phys. Rev.* D **46** 4304–19
[73] Winicour J 2014 Global aspects of radiation memory *Class. Quantum Grav.* **31** 205003
[74] Flanagan E É and Nichols D A 2015 Observer dependence of angular momentum in general relativity and its relationship to the gravitational-wave memory effect *Phys. Rev.* D **92** 084057
Flanagan E É and Nichols D A 2016 *Phys. Rev.* D **93** 049905 (erratum)
[75] Strominger A and Zhiboedov A 2016 Gravitational memory, BMS supertranslations and soft theorems *J. High Energy Phys.* JHEP01(2016)086
[76] Geroch R P and Winicour J 1981 Linkages in general relativity *J. Math. Phys.* **22** 803–12
[77] Ashtekar A and Streubel M 1981 Symplectic geometry of radiative modes and conserved quantities at null infinity *Proc. R. Soc.* A **376** 585–607
[78] Ludvigsen M 1989 Geodesic deviation at null infinity and the physical effects of very long wave gravitational radiation *Gen. Relativ. Gravit.* **21** 1205–12
[79] Strominger A 2014 On BMS invariance of gravitational scattering *J. High Energy Phys.* JHEP07(2014)152
[80] Weinberg S 1965 Infrared photons and gravitons *Phys. Rev.* **140** B516–24
[81] Hübner M, Talbot C, Lasky P D and Thrane E 2020 Measuring gravitational-wave memory in the first LIGO/Virgo gravitational-wave transient catalog *Phys. Rev.* D **101** 023011
[82] Hübner M, Lasky P and Thrane E 2021 Memory remains undetected: updates from the second LIGO/Virgo gravitational-wave transient catalog *Phys. Rev.* D **104** 023004
[83] Grant A M and Nichols D A 2023 Outlook for detecting the gravitational-wave displacement and spin memory effects with current and future gravitational-wave detectors *Phys. Rev.* D **107** 064056
Grant A M and Nichols D A 2023 *Phys. Rev.* D **108** 029901 (erratum)
[84] Cachazo F and Strominger A 2014 Evidence for a new soft graviton theorem (arXiv:1404.4091 [hep-th])
[85] Pasterski S, Strominger A and Zhiboedov A 2016 New gravitational memories *J. High Energy Phys.* JHEP12(2016)053
[86] Nichols D A 2018 Center-of-mass angular momentum and memory effect in asymptotically flat spacetimes *Phys. Rev.* D **98** 064032
[87] de Boer J and Solodukhin S N 2003 A holographic reduction of Minkowski space-time *Nucl. Phys.* B **665** 545–93
[88] Banks T 2003 A critique of pure string theory: heterodox opinions of diverse dimensions (arXiv:hep-th/0306074)
[89] Barnich G and Troessaert C 2010 Symmetries of asymptotically flat 4 dimensional spacetimes at null infinity revisited *Phys. Rev. Lett.* **105** 111103
[90] Barnich G and Troessaert C 2010 Supertranslations call for superrotations *Proc. Sci.* **CNCFG2010** 010
[91] Grant A M 2023 Persistent gravitational wave observables: nonlinearities in (non-)geodesic deviation (arXiv:2401.00047 [gr-qc])
[92] Pollney D and Reisswig C 2011 Gravitational memory in binary black hole mergers *Astrophys. J. Lett.* **732** L13
[93] Mitman K, Moxon J, Scheel M A, Teukolsky S A, Boyle M, Deppe N, Kidder L E and Throwe W 2020 Computation of displacement and spin gravitational memory in numerical relativity *Phys. Rev.* D **102** 104007
[94] Wiseman A G and Will C M 1991 Christodoulou's nonlinear gravitational wave memory: evaluation in the quadrupole approximation *Phys. Rev.* D **44** R2945–9
[95] Magaña Zertuche L *et al* 2022 High precision ringdown modeling: multimode fits and BMS frames *Phys. Rev.* D **105** 104015
[96] Mitman K *et al* 2023 Nonlinearities in black hole ringdowns *Phys. Rev. Lett.* **130** 081402
[97] Yoo J *et al* 2023 Numerical relativity surrogate model with memory effects and post-Newtonian hybridization *Phys. Rev.* D **108** 064027
[98] Grant A M and Mitman K 2023 Higher memory effects in numerical simulations of binary black hole mergers (arXiv:2312.02295 [gr-qc])
[99] Zhu H *et al* 2023 Black hole spectroscopy for precessing binary black hole coalescences (arXiv:2312.08588 [gr-qc])

[100] Zhu H *et al* 2024 Nonlinear effects in black hole ringdown from scattering experiments I: spin and initial data dependence of quadratic mode coupling (arXiv:2401.00805 [gr-qc])
[101] Flanagan E É and Nichols D A 2017 Conserved charges of the extended Bondi-Metzner-Sachs algebra *Phys. Rev.* D **95** 044002
Flanagan E É and Nichols D A 2023 *Phys. Rev.* D **108** 069902 (erratum)
[102] Gómez López L A and Quiroga G D 2017 Asymptotic structure of spacetime and the Newman-Penrose formalism: a brief review *Rev. Mex. Fis.* **63** 275 (arXiv:1711.11381 [gr-qc])
[103] Dray T 1985 Momentum flux at null infinity *Class. Quantum Grav.* **2** L7–L10
[104] Streubel M 1978 "Conserved" quantities for isolated gravitational systems *Gen. Relativ. Gravit.* **9** 551–61
[105] Compère G, Oliveri R and Seraj A 2020 The Poincaré and BMS flux-balance laws with application to binary systems *J. High Energy Phys.* JHEP10(2020)116
Compère G, Oliveri R and Seraj A 2024 *J. High Energy Phys.* JHEP06(2024)045 (erratum)
[106] Cook G B 2000 Initial data for numerical relativity *Living Rev. Relativ.* **3** 5
[107] Boyle M *et al* 2019 The SXS Collaboration catalog of binary black hole simulations *Class. Quantum Grav.* **36** 195006
[108] Bishop N T, Gomez R, Lehner L and Winicour J 1996 Cauchy-characteristic extraction in numerical relativity *Phys. Rev.* D **54** 6153–65
[109] Bishop N T, Gomez R, Lehner L, Maharaj M and Winicour J 1997 High powered gravitational news *Phys. Rev.* D **56** 6298–309
[110] Reisswig C, Bishop N T, Pollney D and Szilagyi B 2010 Characteristic extraction in numerical relativity: binary black hole merger waveforms at null infinity *Class. Quantum Grav.* **27** 075014
[111] Reisswig C, Bishop N T, Pollney D and Szilagyi B 2009 Unambiguous determination of gravitational waveforms from binary black hole mergers *Phys. Rev. Lett.* **103** 221101
[112] Babiuc M C, Szilagyi B, Winicour J and Zlochower Y 2011 A characteristic extraction tool for gravitational waveforms *Phys. Rev.* D **84** 044057
[113] Handmer C J and Szilagyi B 2015 Spectral characteristic evolution: a new algorithm for gravitational wave propagation *Class. Quantum Grav.* **32** 025008
[114] Handmer C J, Szilágyi B and Winicour J 2015 Gauge invariant spectral Cauchy characteristic extraction *Class. Quantum Grav.* **32** 235018
[115] Handmer C J, Szilágyi B and Winicour J 2016 Spectral Cauchy characteristic extraction of strain, news and gravitational radiation flux *Class. Quantum Grav.* **33** 225007
[116] Moxon J, Scheel M A and Teukolsky S A 2020 Improved Cauchy-characteristic evolution system for high-precision numerical relativity waveforms *Phys. Rev.* D **102** 044052
[117] Deppe N *et al* 2024 `SpECTRE v2024.04.12` (Zenodo) (https://doi.org/10.5281/zenodo.10967177)
[118] Moxon J, Scheel M A, Teukolsky S A, Deppe N, Fischer N, Hébert F, Kidder L E and Throwe W 2023 SpECTRE Cauchy-characteristic evolution system for rapid, precise waveform extraction *Phys. Rev.* D **107** 064013
[119] Abbott B P *et al* LIGO Scientific, Virgo 2016 Observation of gravitational waves from a binary black hole merger *Phys. Rev. Lett.* **116** 061102
[120] Chen Y *et al* 2024 Improved frequency spectra of gravitational waves with memory in a binary-black-hole simulation (arXiv:2405.06197 [gr-qc])
[121] McKechan D J A, Robinson C and Sathyaprakash B S 2010 A tapering window for time-domain templates and simulated signals in the detection of gravitational waves from coalescing compact binaries *Class. Quantum Grav.* **27** 084020
[122] Mitman K *et al* 2021 Adding gravitational memory to waveform catalogs using BMS balance laws *Phys. Rev.* D **103** 024031
[123] Talbot C, Thrane E, Lasky P D and Lin F 2018 Gravitational-wave memory: waveforms and phenomenology *Phys. Rev.* D **98** 064031
[124] Chen P-N, Wang M-T, Wang Y-K and Yau S-T 2022 BMS charges without supertranslation ambiguity *Commun. Math. Phys.* **393** 1411–49
[125] Javadinezhad R, Kol U and Porrati M 2022 Supertranslation-invariant dressed Lorentz charges *J. High Energy Phys.* JHEP04(2022)069
[126] Javadinezhad R and Porrati M 2023 Supertranslation-invariant formula for the angular momentum flux in gravitational scattering *Phys. Rev. Lett.* **130** 011401
[127] Javadinezhad R and Porrati M 2024 Three puzzles with covariance and supertranslation invariance of angular momentum flux and their solutions *Phys. Rev. Lett.* **132** 151604

[128] Kozameh C N and Quiroga G D 2016 Center of mass and spin for isolated sources of gravitational radiation *Phys. Rev.* D **93** 064050
[129] Moreschi O M 1986 On angular momentum at future null infinity *Class. Quantum Grav.* **3** 503
[130] Arnowitt R L, Deser S and Misner C W 1959 Dynamical structure and definition of energy in general relativity *Phys. Rev.* **116** 1322–30
[131] Teukolsky S A 1973 Perturbations of a rotating black hole. 1. Fundamental equations for gravitational electromagnetic and neutrino field perturbations *Astrophys. J.* **185** 635–47
[132] Berti E, Cardoso V and Starinets A O 2009 Quasinormal modes of black holes and black branes *Class. Quantum Grav.* **26** 163001
[133] Cheung M H-Y *et al* 2023 Nonlinear effects in black hole ringdown *Phys. Rev. Lett.* **130** 081401
[134] Islo K, Simon J, Burke-Spolaor S and Siemens X 2019 Prospects for memory detection with low-frequency gravitational wave detectors (arXiv:1906.11936 [astro-ph.HE])
[135] Doran C and Lasenby A 2007 *Geometric Algebra for Physicists* (Cambridge University Press)