Princeton University
Department of Physics

Junior Paper

Black hole ringdown and gravitational waves: a review

Connor Hainje

Advised by
Prof. Herman Verlinde

This paper represents my work in accordance with University regulations.
/s/ Connor Hainje

Contents

Abstract

In recent years, methods to detect gravitational waves have been developed, making the physics of binary black hole mergers increasingly relevant. In particular, it has become important to better understand the gravitational waves produced by such events and how to analyze them. This study begins by reviewing the physics relevant to black holes and gravitational waves from relatively foundational principles. It then studies the waveforms detected from gravitational waves and how they may be analyzed. The final goal of this review is to understand the waveforms and analysis of the ringdown of black hole merger events, where the predictions of general relativity can be directly tested.

1 Introduction

In 1915, Einstein released his theory of general relativity. A landmark theory, it was able to explain many of the known—and then-unknown—inconsistencies with Newtonian physics. Another result, however, was the prediction of many new physical phenomena, which people sought as "tests" of the theory. Two such predictions that are important to us today are black holes and gravitational waves.

Black holes were first predicted by the work of Karl Schwarzschild, very shortly after general relativity was published, although it was not recognized as such at the time. Roughly twenty years later, Chandrasekhar, Landau, and others derived new results which more directly suggested that the final stage of the evolution of large stars was a black hole, although they continued to be ignored and shrugged off as largely nonphysical for many years. Finally, in the late 1950s and 1960s, black holes became the subject of serious investigation with the works of Wheeler, Kerr, and others [1]. Today, we finally have somewhat direct methods of probing the existence of black holes, with breakthroughs in gravitational wave detection and the Event Horizon Telescope project.

Similarly, the presence of gravitational waves was for a long time rather contested. Until 2015, the strongest evidence for the existence of gravitational waves came from the 1974 study of the binary pulsar PSR 1913+16 [1]. Researchers found that anomalies in the pulsar frequency could be explained by a Doppler shift and gravitational redshift and the decay of the orbit due to the emission of gravitational waves. As these measurements have become increasingly accurate over the years, they have only served to strengthen the evidence for the existence of gravitational waves.

In 2015, however, gravitational waves were directly detected and measured at LIGO [2]. These direct detections give us a new way to probe the universe for the existence of many exotic astrophysical bodies, such as black holes. These bodies are predicted to have very unique gravitational wave signatures, and so we can directly detect their existence by measuring gravitational waves and verifying that they have these signatures.

The gravitational waves that we measure here on Earth are typically the product of very highly energetic merger events. In fact, the most common merger events seen to date have been those of black holes. Thus, much work has been done in understanding how the physics of a black hole merger event relates to the waveform we detect here. Recently, studies have been performed to understand how we can better model the *ringdown*, the last stage of a merger event, to discover more about the mergers from the waveform. Ringdown has thus far not been very well understood, but studies like [3] show the importance of this stage to gravitational waveform analyses.

This study intends to review the above ideas, from the general relativistic description of black holes and gravitational waves, to the measurements of gravitational waves on Earth, to the analysis of these measured ringdown waveforms, and the resulting tests of general relativity that can be performed.

2 Black holes

We begin our review by examining black holes themselves. To do so, we go back to the conception of the idea, and survey the relevant topics from general relativity.

2.1 General relativity

In Newtonian physics, one thinks of gravity as a somewhat mysterious force which acts at a distance on objects with mass. In relativity, however, gravity is not a force acting on massive objects, but rather the product of the deformation of spacetime by massive objects themselves.[1] Particles move, then, not under the action of a force, but rather by following the shortest path through the curved spacetime.

As such, we desire in relativity to describe the geometry of spacetime in order to explain the movements and interactions of objects. To do so, we define a **metric**: a differential element representing "distance" in spacetime which is invariant under frame-of-reference transformations. The simplest of these is that of special relativity:[2]

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2. \tag{1}$$

Often, we let $x^\mu = (t, x, y, z)$ and write the metric above in tensor notation like so:[3]

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu, \tag{2}$$

where we are letting $\eta_{\mu\nu}$ be the **metric tensor**. $\eta_{\mu\nu}$ is a rather special metric, known as the **Minkowski metric**. The Minkowski metric describes "flat" spacetime, where there are no sources of gravity and everything remains undisturbed. Thus, if the universe were empty, one could use the Minkowski metric to describe spacetime everywhere. More generally, however,

[1] I say *massive* objects distort spacetime, but this is not entirely true. From the principle of mass-energy equivalence, it is really *energetic* objects which distort spacetime. This is why, in relativity, the effects of gravity on light are unambiguous. Of course, Einstein's principle of mass-energy equivalence makes this point moot, but I just wanted to emphasize this for clarity.

[2] Throughout this paper, we typically use a system of geometrized units in which G and c are set to 1. This is in keeping with many authors on the subject.

[3] In tensor notation, we use Einstein summation notation, where an index which is written both "down" and "up" is implicitly summed over. For more details, see Appendix A.

the Minkowski metric is only used under the assumption that one is considering a point in spacetime far away from any sources.

In more general terms, we can write *any* metric with the following form:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu, \tag{3}$$

where x^μ is a vector describing the coordinates and $g_{\mu\nu}$ is the metric tensor.

With the metric defined as such, how do we deal with the presence of gravity? As stated before, massive objects distort spacetime, so the problem in general relativity is to determine how to describe the spacetime resulting from a system of massive objects. Thus, given a distribution of mass and energy, we need to find the metric. Doing so requires solving Einstein's Field Equations, which are nonlinear and famously difficult to solve. Solutions have been found for some systems, however, and we shall consider two of these.

The **Schwarzschild metric** describes the spacetime surrounding a sphere of mass M which is chargeless, nonrotating, and unmoving.[4] Here we use a coordinate system of (t, r, θ, ϕ), known as Boyer-Lindquist coordinates. Again, note that we use units with $G = c = 1$. The metric is given by

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2, \tag{4}$$

where M is the mass of the spherical body. Notice that this metric breaks down for $r = 2M$; this value for r, denoted r_S, is known as the **Schwarzschild radius**.

When the radius of the spherical body, R, is larger than r_S, then there are no issues, as the metric adequately describes all of spacetime outside of the sphere. What happens inside the sphere is a little more complicated, as this depends on the mass distribution and other properties of the body itself. But this does raise the question: what happens for a spherical body with radius *less than* r_S? It turns out that this describes a **black hole**, or, more precisely, a Schwarzschild black hole. For $R < r_S$, the metric gives that *all* objects inside of r_S will be pulled to a singularity at the point $r = 0$. As such, r_S is often known as the **event horizon**. Even light, after crossing the horizon, cannot escape the gravitational pull of the singularity; it is from this phenomenon that black holes derived their name.

A generalized case of the Schwarzschild metric describes spacetime containing an axisymmeteric body rotating with angular momentum J. This metric is called the **Kerr metric** and is given by

$$\begin{aligned} ds^2 = &-\left(1 - \frac{2Mr}{\Sigma}\right) dt^2 - \frac{4aMr\sin^2\theta}{\Sigma} dt\, d\phi + \frac{\Sigma}{\Delta} dr^2 \\ &+ \Sigma \, d\theta^2 + \left(r^2 + a^2 + \frac{2Mra^2\sin^2\theta}{\Sigma}\right) \sin^2\theta \, d\phi^2, \end{aligned} \tag{5}$$

where we have introduced the constants

$$a \equiv \frac{J}{M}, \quad \Delta \equiv r^2 - 2Mr + a^2, \quad \Sigma \equiv r^2 + a^2\cos^2\theta. \tag{6}$$

[4] In the language of general relativity, one would call it *stationary* and *static*.

As the Schwarzschild metric described stationary, nonrotating black holes, the Kerr metric describes stationary, rotating black holes with an axis of rotation about the ϕ-axis. Notice that it reduces to the Schwarzschild metric for $J = a = 0$.

Similar to before, we seek the condition for which the metric breaks down; this defines the horizon. For the Kerr metric, this occurs when $\Delta = 0$. We can solve this equation for r to find two solutions; the larger of the two turns out to be the event horizon for Kerr black holes, occuring at

$$r_+ = M + \sqrt{M^2 - a^2}. \tag{7}$$

It is interesting—and nontrivial—to note that the properties of a black hole may be specified with just two quantities: its mass, M, and angular momentum, J.[5] Much of the dynamics of more complicated systems, in fact, depend on these two quantities, as we shall see later.

2.2 Binary black holes

We have worked out the appearance of black holes in general relativity, but we now turn our focus to systems of two black holes. These systems are known as **binary black holes**, abbreviated BBHs. All BBHs are unstable, meaning that all BBHs will collide and merge to form a single remaining black hole. Thus, we may categorize the stages of the life of a BBH in three stages: inspiral, merger, and ringdown [4].

The **inspiral** is the stage where the two black holes are orbiting each other, getting progressively closer. The main phenomenon by which their orbit decays is the emission of gravitational waves. In this regime, the energy of the emitted gravitational waves is initially rather low, but increases as the black holes become closer and gain velocity.

The **merger** is the stage where the black holes finally merge. In this regime, the gravitational waves are being emitted with such high energy that the system is no longer adiabatic. The length of the phase is on the order of one or two gravitational wave cycles, so it is very short. The result of the merger is a single rotating black hole which is not axisymmetric, and so may be modeled as a perturbed Kerr black hole.

The **ringdown** is the stage in which the perturbed Kerr black hole "settles down" to an unperturbed Kerr black hole. This regime may also be defined as the phase in which the gravitational waves emitted by the remnant black hole can be written with good precision as a superposition of **quasinormal modes** (QNMs). Recent research has shown that much information may be gleaned from the gravitational wave shape in the ringdown by modeling the wave using a superposition of many quasinormal modes; the purpose of this study is to review this procedure. A more in-depth description of QNMs may be found below, in the section on gravitational waves.

[5] Black holes may also be charged, but the charge of these systems is quickly neutralized as the black holes attract charged particles of the opposite charge and so we do not consider this case [1].

3 Gravitational waves

Gravitational waves are fundamentally a result of general relativity; in the Newtonian picture, spacetime is not a sort of "fabric" which may be distorted. Simply put, a gravitational wave is a ripple in the curvature of spacetime, propagating at the speed of light.

3.1 Description of gravitational waves

Gravitational waves are quite similar to water waves. Water waves can be thought of as ripple propagating along the surface of some body of water. Similarly, gravitational waves are a "ripple" propagating through spacetime, with one key difference: in the case of water, waves are simply motions in a material, but gravitational waves are ripples in *spacetime itself.*

The description of gravitational waves also has many parallels with that of electromagnetic waves. One of the key differences, however, is that electromagnetic fields are vector fields, where the gravitational fields we consider are tensor. A key result of this difference is that electromagnetic radiation is primarily dipolar, meaning that it is polarized in two perpendicular directions, but gravitational radiation is quadrupolar, so it is polarized in two ways, called plus ($+$) and cross ($\times$). Plus-polarized gravitational waves distort matter inward along the y-axis as they distort matter outword along the x-axis, and vice versa. Cross-polarized gravitational waves do much the same, but oriented at a $45°$ angle from the plus-oriented wave.

It is illustrative to consider how gravitational waves impact the matter they interact with and to compare this to an electromagnetic wave interacting with a charge particle. For a charged particle, we know that a test charge will, when hit by a plane wave polarized in the x-direction, oscillate in the x-direction, and similarly for y. To consider the effects of a general gravitational wave, it is best to consider how a ring of test particles is deformed. See Figure 1 for a graphical depiction of this.

For the following discussion, we follow closely the work of Shapiro and Teukolsky [1]. To describe a gravitational wave more mathematically, we first consider the problem of a gravitational wave propagating through free space, so we make the assumption that gravity is weak and the metric is well-described by $\eta_{\mu\nu}$. The gravitational wave is a weak perturbation of spacetime in this region, and so we describe it with the metric tensor

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{8}$$

where $h_{\mu\nu}$ is small and describes the wave. We know that gravitational waves are polarized in two ways, $+$ and $\times$, and so we may describe $h_{\mu\nu}$ with two amplitudes h_+ and $h_\times$ by introducing the **polarization tensors**. In (t, x, y, z) coordinates, if we assume that the wave is propagating in the z direction, then we define the polarization tensors e^+ and $e^\times$ by

$$e^+ = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad e^\times = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{9}$$

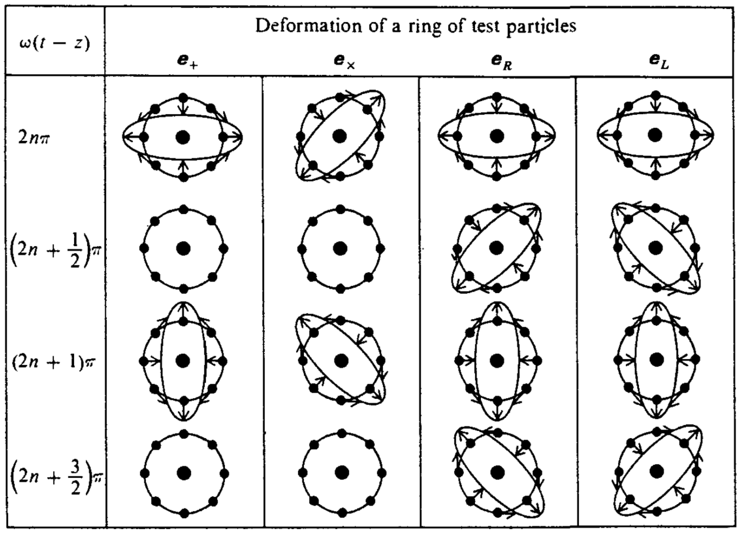

Figure 1: The deformation of a ring of test particles by gravitational waves with different polarizations. e_+ and $e_\times$ are the plus and cross polarizations discussed in the text. e_R and e_L are the *unit circular polarization tensors*, defined as $e_R = (e_+ + ie_\times)/\sqrt{2}$ and $e_L = (e_+ - ie_\times)/\sqrt{2}$. Figure reproduced from [5] (where it is Figure 35.2).

and we can write the gravitational wave as

$$h_{jk} = h_+ e^+_{jk} + h_\times e^\times_{jk}. \tag{10}$$

Note the use of latin letters j and k here, which denote that we are only considering the spacelike indices 1, 2, and 3. We may then solve the following system to find the form of $h_{\mu\nu}$ describing the wave:

$$h_{0\mu} = 0, \quad h_{jk,k} = 0, \quad \Box h_{jk} = 0, \tag{11}$$

as well as the condition that $h_{\mu\nu}$ be trace-free. These equations are rather complicated, despite their appearance. The first of these equations states that all elements of the first column of the tensor $h_{\mu\nu}$ must be 0. The second equation may be expanded as

$$h_{jk,k} = 0 \iff \frac{\partial h_{j1}}{\partial x^1} + \frac{\partial h_{j2}}{\partial x^2} + \frac{\partial h_{j3}}{\partial x^3} = 0, \tag{12}$$

where this last one is really three equations: one for each value of j. The third equation uses the $\Box$ operator, which is described in Appendix A. It is actually Einstein's field equation for this system.

For the production of gravitational waves by black hole mergers, however, it is not sufficient to find only the solution for gravitational waves in free space. We also need the solution for gravitational waves in a black hole spacetime. We still treat the wave as a perturbation $h_{\mu\nu}$ to the metric, but now we use one of the black hole metrics discussed above. To find this, we expand the wave to a combination of *quasinormal modes*, as was first mentioned before. The following section is dedicated to their discussion.

3.2 Quasinormal modes

As the ringdown regime of the black hole merger is described by a Kerr spacetime, we concern ourselves with the Kerr metric. The quasinormal modes of the Schwarzschild metric exist and are rather well-studied, but are not terribly relevant for our following discussions, so we omit them.

As before, we consider the gravitational waves to be perturbations, small relative to the spacetime. Thus, let $g^0_{\mu\nu}$ be the Kerr metric tensor and let $h_{\mu\nu}$ describe the gravitational waves. The full metric, then, is $g^0_{\mu\nu} + h_{\mu\nu}$. We require that this metric satisfies Einstein's field equations. Following [6], we assume that we can decompose the scalar quantities from the tensor $h_{\mu\nu}$ into a form like

$$\chi(t, r, \theta, \phi) = \sum_{\ell,m} \frac{\chi_{\ell m}(r,t)}{r} Y_{\ell m}(\theta, \phi). \tag{13}$$

Then, χ may be found by solving the following scalar wave equation [7]:

$$\begin{aligned}\left[\frac{(r^2+a^2)^2}{\Delta} - a^2\sin^2\theta\right]\frac{\partial^2\chi}{\partial t^2} + \frac{4Mar}{\Delta}\frac{\partial^2\chi}{\partial t\partial\varphi} + \left[\frac{a^2}{\Delta} - \frac{1}{\sin^2\theta}\right]\frac{\partial^2\chi}{\partial\varphi^2} \\ - \Delta^{-\sigma}\frac{\partial}{\partial r}\left(\Delta^{\sigma+1}\frac{\partial\chi}{\partial r}\right) - \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\chi}{\partial\theta}\right) - 2\sigma\left[\frac{a(r-M)}{\Delta} + \frac{i\cos\theta}{\sin^2\theta}\right]\frac{\partial\chi}{\partial\varphi} \\ - 2\sigma\left[\frac{M(r^2-a^2)}{\Delta} - r - ia\cos\theta\right]\frac{\partial\chi}{\partial t} + \left(\sigma^2\cot^2\theta - \sigma\right)\chi = 4\pi\Sigma T,\end{aligned} \tag{14}$$

with $\sigma = 0, 1,$ or 2 for scalar, electromagnetic, or gravitational perturbations respectively, and T describing the sources. To help unpack this, consider the case where $T = 0$ (i.e. the wave is now propagating far from the Kerr black hole). Then, we find that the wave equation is separable in the form

$$\chi = e^{-i\omega t}e^{im\phi}S(\theta)R(r), \tag{15}$$

and the separated differential equations for S and R are

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dS}{d\theta}\right) + \left(a^2\omega^2\cos^2\theta - 2a\omega\sigma\cos\theta - \frac{(m+\sigma\cos\theta)^2}{\sin^2\theta} + \sigma + A\right)S = 0, \tag{16}$$

$$\Delta^{-\sigma}\frac{d}{dr}\left(\Delta^{s+1}\frac{dR}{dr}\right) + \left(\frac{K^2 - 2is(r-M)K}{\Delta} + 4is\omega r - \lambda\right)R = 0, \tag{17}$$

where $K \equiv (r^2+a^2)\omega - am$ and $\lambda \equiv A + a^2\omega^2 - 2am\omega$. Here, A is the separation constant of the differential equations, and the angular equation (16) is an eigenvalue problem for A with eigenfunctions ${}_sS_{\ell m}$, which are the spin-weighted spheroidal harmonics. Solving these equations for S and R is rather involved, but results in a set of complex frequencies, $\omega_{\ell mn}$.

When $T \neq 0$, it is found that the radial equation will be identical on the left-hand side, but with a source term on the right-hand side. Further, it is found that for outgoing waves at infinity, the strain of the wave (described in greater detail below) goes like $h_+ + ih_\times$, with h_+ and $h_\times$ the amplitudes of the two polarizations of the wave [7]; a useful result that shall be used in later sections.

3.3 Detecting gravitational waves

With all this talk of gravitational waves and how they may be analyzed, it is worth taking some time to discuss how these waves are actually detected. In fact, it was only recently that gravitational waves were directly detected for the first time. The first direct detection came from the Laser Interferometer Gravitational-Wave Observatory, or LIGO [8], experiment in 2015 [2]. To this day, the LIGO experiment is still the source of the majority of all known gravitational wave observations.

LIGO, as is implied in the name, uses laser interferometry to measure incident gravitational waves from high energy merger events. They do this at multiple observatory sites across Earth—in particular, there are two primary observatories: one in Hanford, WA and one in Livingston, LA. Each observatory consists mostly of two 4-kilometer "arms," which are long pipes housing the laser-beam paths, situated at a right angle to one another. A single-frequency laser is then directed into a 50/50 beamsplitter, such that one-half of the power of the beam is directed down one the beam path of one arm and one-half down the beam path of the other. At the ends of each arm are very reflective and highly stabilized mirrors. The beams are then reflected back toward the beam splitter, where they are recombined and measured.

The arms are set in such a way that under normal conditions, one of the beams will pick up a phase shift relative to the other one and the two will destructively interfere when recombined. If a gravitational wave is incident on the detectors, however, then the lengths of the arms are changed by a very small amount—on the order of 10^{-18} meters. Surprisingly, this is detectable, as it changes the relative phase shift of the two beam components by enough to prevent total destructive interference. In fact, the amplitude of the recombined beam is related directly to the amplitude of the gravitational wave: a strong gravitational wave will change the relative phase difference by a larger amount, resulting in less destructive interference.

The setup at LIGO includes other measures to stabilize the mirrors, to stabilize the power of the laser, to keep the beam paths at ultra-high vacuum, and to recycle the lasers themselves, among other things. A detailed description of this is outside of the scope of this review, but may be found in [8] and [9].

In Figure 2, one can see an example of a gravitational wave signal as detected by observatories here on Earth. What is plotted is known as the *relative strain* of the wave, often denoted h. This quantity describes the amount by which particles are displaced by the gravitational wave, and is thus related directly to the amplitude of the wave projected onto the detector. One can see in the waveform a slow build-up of frequency and amplitude: this regime corresponds to the inspiral. Then, the period of peak frequency is the merger. This regime is rather short, lasting only one or two wave cycles [4]. Third, the amplitude dies back down. This is the ringdown of the resultant black hole.

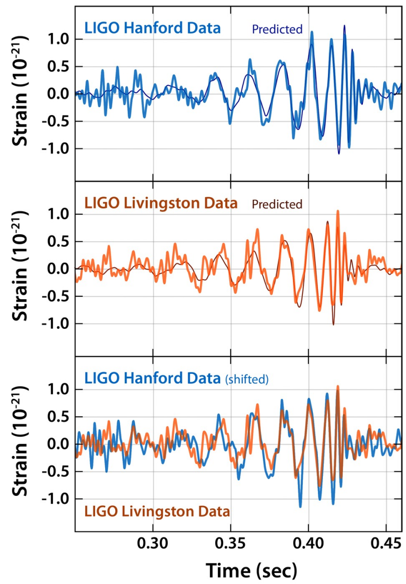

Figure 2: The signal of the first gravitational wave detected by LIGO, GW150914, as observed by both the Hanford and Livingston observatories. Image taken from ligo.org/detections.php, with colors inverted.

4 Ringdown

As we discussed previously, ringdown is the last stage in the lifecycle of a BBH merger. By this point, the black holes have merged but not equilibriated, so the region is well-described by a perturbed Kerr spacetime. More precisely, we have from [4] the definition of the ringdown phase as "the time after which the gravitational waves [may] be written entirely as a superposition of quasinormal modes of the final black hole." Unfortunately, though, this is not as simple as it sounds.

4.1 Description of waveform

As we found in Section 3, a gravitational wave consists of a plus, h_+, and a cross, $h_\times$, component. From [10], we find that these components of the *ringdown* waveform may be written in terms of the radial Teukolsky function $R_{\ell m\omega}$ as

$$h_+ + ih_\times = -\frac{2}{r^4}\int_{-\infty}^{\infty}\frac{1}{\omega^2}e^{i\omega t}\sum_{\ell,m}\big(S_{\ell m}R_{\ell m\omega}(r)\big)\,d\omega. \tag{18}$$

Here, $S_{\ell m}$ are spin-weighted spheroidal harmonics and $\omega = \omega_{\ell mn} + i/\tau_{\ell mn}$ is the complex frequency, whose real part is the gravitational wave frequency and imaginary part is the reciprocal of the damping time of the oscillation.

It is noted by [10] that we may now assume that this expands as a linear superposition of damped exponentials, replacing Eq. 18 by

$$h_+ + ih_\times = \frac{M}{r}\sum_{\ell,m,n}\mathcal{A}_{\ell mn}e^{i(\omega_{\ell mn}t+\phi_{\ell mn})}e^{-t/\tau_{\ell mn}}S_{\ell mn}, \tag{19}$$

where M is the mass of the black hole, $\mathcal{A}_{\ell mn}$ and $\phi_{\ell mn}$ are the real amplitude and phase of the oscillation.

Here, [10] notes that, while it is common to assume the $\ell = m = 2$ mode (corresponding to purely quadrupolar waveforms) is the dominant mode, this assumption is flawed in general. However, by the analysis of a detected gravitational waveform in [3], it is experimentally found that $\ell = m = 2$ *is* the dominant mode, at least for the specific waveform analyzed, and other modes may be ignored. They too note that this assumption may not be generally valid, but holds at least epirically for the waveforms analyzed to date. As such, we continue our analysis assuming this mode alone.

Assuming the dominance of this mode, we let $\ell = m = 2$ and find

$$h_+ + ih_\times = \frac{M\mathcal{A}_{\ell mn}}{r}e^{i(\omega_{\ell mn}t+\phi_{\ell mn})}e^{-t/\tau_{\ell mn}}S_{\ell mn}. \tag{20}$$

Breaking this into the individual plus and cross components, we have found

$$h_+ = \frac{M}{r}\Re\big[\mathcal{A}_{\ell mn}e^{i(\omega_{\ell mn}t+\phi_{\ell mn})}e^{-t/\tau_{\ell mn}}S_{\ell mn}\big], \tag{21}$$

$$h_\times = \frac{M}{r}\Im\big[\mathcal{A}_{\ell mn}e^{i(\omega_{\ell mn}t+\phi_{\ell mn})}e^{-t/\tau_{\ell mn}}S_{\ell mn}\big]. \tag{22}$$

4.2 Detection

With the plus and cross components written as an expansion of the n-indexed overtones of the $\ell = m = 2$ quasinormal mode, we can describe the relative strain measured by a detector as

$$h = h_+F_+(\theta,\phi,\psi) + ih_\times F_\times(\theta,\phi,\psi), \tag{23}$$

where F_+ and $F_\times$ are "pattern functions," depending on the orientation of the detector relative to the source of the wave [11]. They are written explicitly as

$$F_+(\theta,\phi,\psi) = \frac{1}{2}(1+\cos^2\theta)\cos(2\phi)\cos(2\psi) - \cos(\theta)\sin(2\phi)\sin(2\psi), \tag{24}$$

$$F_\times(\theta,\phi,\psi) = \frac{1}{2}(1+\cos^2\theta)\cos(2\phi)\sin(2\psi) + \cos(\theta)\sin(2\phi)\cos(2\psi), \tag{25}$$

where we have introduced the variables θ, ϕ, and ψ. These variables describe the orientation of the gravitational wave relative to the x- and y-arms of the detector. Their definitions are best described by Figure 3.

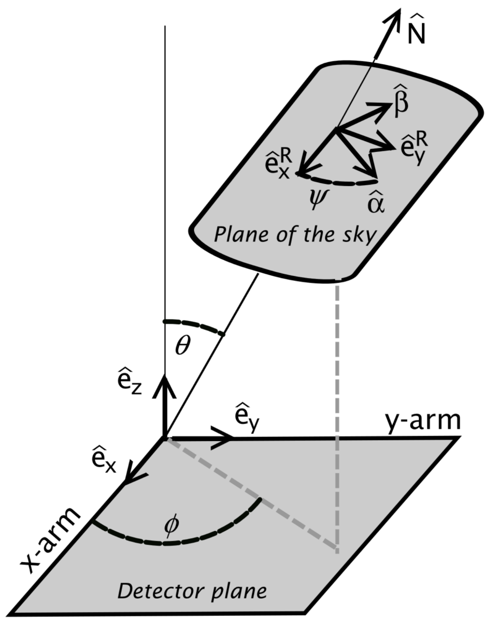

Figure 3: Definition of the variables θ, ϕ, and ψ describing the relationship between the orientation of the detector and a gravitational wave oriented along the $\hat{\alpha}$ and $\hat{\beta}$ axes. Image taken from [11].

4.3 Remnant parameter recovery

Now we consider how the above description can be used to perform an analysis of a detected ringdown waveform, following in large part the work in [3]. We assume that we know the F_+ and $F_\times$ parameters well, and so we can write the $\ell = m = 2$ mode of the detected strain of the gravitational wave as

$$h_{22}(t) = F_+ h_{+,22}(t) + iF_\times h_{\times,22}(t) = \sum_{n=0}^{N} A_{22n} e^{i(\omega_{22n}(M,J)t+\phi_{22n})} e^{-t/\tau_{22n}(M,J)}, \tag{26}$$

where the non-script $A_{\ell mn}$ is a real coefficient describing the amplitude, which has absorbed the other coefficients from before. The $h_{+,22}$ and $h_{\times,22}$ are the $\ell = m = 2$ modes of the h_+ and $h_\times$ components of the wave we have seen before. The amplitudes A_{22n} and phases ϕ_{22n} are not known beforehand, but are determined by using a least-squares fit from the detected waveform. Also note the explicit dependence of $\omega_{\ell mn}$ and $\tau_{\ell mn}$ on the mass, M, and spin, J, of the remnant black hole. In general, these can be fully determined by perturbation theory methods like those described in Section 3.2.

Lastly, note that the sum includes not only the fundamental mode ($n = 0$), but also N overtones ($n > 0$). This N may be arbitrarily chosen; in [3], Giesler et al. analyze the contributions of overtones $n = 1$ through 7 for a specific waveform similar to the GW150914 event and discuss the improvements to remnant parameter recovery using $N = 3$. The inclusion of overtones turns out to be important, [3] finds, despite the oft-used assumption that the fundamental mode is dominant. In fact, at early times in the ringdown, it was found that overtones dominate the waveform with larger amplitudes than the fundamental mode. Making use of this increases the amount of the ringdown waveform that can be accurately analyzed and further increases the signal-to-noise ratio.

Using this model, we thus have the free variables M and J which need to be determined by minimizing the mismatch between the h_{22} waveform and the detected waveform, which we will let be h. Following [3], we define the mismatch to be

$$\mathcal{M} = 1 - \frac{\langle h, h_{22} \rangle}{\sqrt{\langle h, h \rangle \langle h_{22}, h_{22} \rangle}}, \tag{27}$$

where the inner product is defined as

$$\langle x, y \rangle = \int_0^T x(t)\overline{y(t)}\, dt, \tag{28}$$

with T a time late in the waveform, but before it has decayed to noise, and the overline indicating complex conjugation. Thus, the procedure to recover remnant parameters is as follows: choose M and J, determine the A_n and ϕ_n values from a least-squares fit, and compute the mismatch with the detected waveform. One can then use Markov chain Monte Carlo methods to minimize the mismatch over values of M and J. The values which minimize $\mathcal{M}$ are then the recovered mass and angular momentum of the resultant black hole.

and recover M and J from the waveform.

4.4 Testing general relativity

The methods described above also have the potential to directly test predictions of general relativity. The primary of these tests is known as the **two-mode test**, as it relies on the ability to determine the contributions of two QNMs. The test is rather simple. Each mode corresponds to a range of possible pairs of black hole masses and angular momenta. If we recover a set of several modes (at least two), then the two modes' possible remnant parameters should intersect at minimally one point. If they do not, then either the object observed is not an isolated black hole, or there is a contradiction in the predictions of general relativity. Due to noise and other experimental problems, there will likely never be exact agreement, but much work has been done to determine how badly the two modes must disagree before we conclude that general relativity has failed [12].

Until recently, most work has focused on performing the two-mode test with the fundamental modes of two different angular harmonics; i.e. two modes whose ℓ or m values differ. This is only feasible for specific systems which present a strong secondary angular mode. Further, the fundamental modes are only dominant late in the waveform [3], where the signal is weaker. With the revelations of [3] regarding the importance of $\ell = m = 2$ overtones, though, we see that multi-overtone tests will likely be more accessible and powerful for a test of general relativity.

When using overtones in conjunction with the fundamental mode, we can also obtain more precise predictions for the remnant parameters. (Note the difference between determining the range of possible remnant parameters independently for each mode, and using the overtones in conjunction with the fundamental mode for a more precise prediction of remnant parameters.) This is important to compare against the results of analysis of the inspiral. If there are significant, unpredicted differences between the results of the two analyses, this could also hint at a possible contradiction in the theory.

5 Summary

With the recent advances made in gravitational wave detection, binary black hole mergers have become one of the most important phenomena to understand to test our understanding of gravity near some of the strongest gravitational sources: black holes. As such, this paper has, from relatively foundational principles, reviewed some of the relevant physics involved in black hole studies and gravitational waves. From this foundation, we have tried to develop a reasonably good understanding of black hole binary mergers and the relevant analysis that may be performed on the ringdown stage. This resulted in an understanding of the importance of quasinormal mode overtones in ringdown analysis, as well as an overview of the uses of quasinormal modes for testing general relativity.

Acknowledgments

I would like to thank Professor Herman Verlinde for so generously agreeing to advise me on this paper and for all his help to understand the math and physics that make up these studies. I'd also like to thank Professor Frans Pretorius for his advice and suggestions during the work.

References

[1] S. L. Shapiro, *Black holes, white dwarfs, and neutron stars: the physics of compact objects*. New York: Wiley, 1983.

[2] LIGO Scientific Collaboration and Virgo Collaboration, "Observation of Gravitational Waves from a Binary Black Hole Merger," *Physical Review Letters*, vol. 116, p. 061102, Feb. 2016.

[3] M. Giesler, M. Isi, M. Scheel, and S. Teukolsky, "Black hole ringdown: the importance of overtones," *arXiv:1903.08284 [gr-qc]*, Mar. 2019. arXiv: 1903.08284.

[4] F. Pretorius, "Binary Black Hole Coalescence," *arXiv:0710.1338 [astro-ph, physics:gr-qc, physics:hep-th]*, Oct. 2007. arXiv: 0710.1338.

[5] C. W. Misner, *Gravitation*. Princeton, N.J.: Princeton University Press, 2017.

[6] K. D. Kokkotas and B. G. Schmidt, "Quasi-Normal Modes of Stars and Black Holes," *Living Reviews in Relativity*, vol. 2, p. 2, Dec. 1999. arXiv: gr-qc/9909058.

[7] S. A. Teukolsky, "The Kerr Metric," *Classical and Quantum Gravity*, vol. 32, p. 124006, June 2015. arXiv: 1410.2130.

[8] B. P. Abbott et al. (LIGO Collaboration), "LIGO: the Laser Interferometer Gravitational-Wave Observatory," *Reports on Progress in Physics*, vol. 72, p. 076901, June 2009.

[9] J. Aasi et al. (LIGO Collaboration), "Advanced LIGO," *Classical and Quantum Gravity*, vol. 32, p. 074001, Mar. 2015.

[10] E. Berti, V. Cardoso, and C. M. Will, "Gravitational-wave spectroscopy of massive black holes with the space interferometer LISA," *Physical Review D*, vol. 73, p. 064030, Mar. 2006.

[11] B. S. Sathyaprakash and B. F. Schutz, "Physics, Astrophysics and Cosmology with Gravitational Waves," *Living Reviews in Relativity*, vol. 12, p. 2, Dec. 2009. arXiv: 0903.0338.

[12] O. Dreyer, B. Kelly, B. Krishnan, L. S. Finn, D. Garrison, and R. Lopez-Aleman, "Black-hole spectroscopy: testing general relativity through gravitational-wave observations," *Class. Quantum Grav.*, vol. 21, Jan. 2004.

[13] D. J. Raine, *Black holes: a student text.* Hackensack, NJ: Imperial College Press, 3rd ed., 2015.

A Notation

One of the most difficult facets of general relativity to become acquainted with is often the notation, so I have decided to do my best to collect the notation here.

A.1 Index notation

In relativity, one describes an *event* as a four-dimensional point in spacetime. Similar to how the coordinates of a point in space are specified by a vector in $\mathbb{R}^3$ of the form (x, y, z), the coordinates of an event in spacetime are specified by a vector in $\mathbb{R}^4$ of the form (t, x, y, z). Note that the timelike coordinate, t, occupies the first index, often numbered 0, and the spacelike coordinates take indices 1, 2, and 3. We are used to representing spatial vectors by $\vec{x}$, where the arrow above the x indicates that it is a vector. In relativity, however, it is common to denote an event vector by x^μ, where the superscript μ is an indication that the quantity is a vector. (Note that the superscript is *not* an exponent.)

Typically, one can use any Greek letter to specify the four-dimensional spacetime vector; i.e. $x^\mu = (t, x, y, z)$ could equally well be called x^ν. However, Latin letters are reserved for talking about the spatial coordinates only (x, y, and z).

This index-based notation is intimately related to **Einstein summation notation**. The idea behind Einstein summation notation is that any indices which are repeated are implicitly summed over. As an example, letting x^μ be defined as above, we get

$$x^\mu x_\mu = x_\mu x^\mu \equiv \sum_{\mu=0}^{3} x^\mu x_\mu = t^2 + x^2 + y^2 + z^2. \tag{29}$$

This is where the distinction between Greek and Latin letters becomes most relevant. Conventionally, an Einstein sum over Latin indices would mean that the sum goes over the indices 1, 2, and 3, where a sum over Greek indices goes over 0, 1, 2, and 3. In symbols,

$$x^k x_k = x_k x^k \equiv \sum_{k=1}^{3} x^k x_k = x^2 + y^2 + z^2. \tag{30}$$

Like we have just seen, one can use this notation to write the dot product of two vectors rather succinctly, as

$$x^i y_i \equiv \sum_{i=1}^{3} x^i y_i = x_1 y_1 + x_2 y_2 + x_3 y_3. \tag{31}$$

Similar to the notation used for vectors, one defines a notation to describe tensors by using multiple indices. For instance, the familiar Minkowski metric tensor is a four-by-four matrix, denoted $\eta_{\mu\nu}$, and we can write the Minkowski metric as

$$\eta_{\mu\nu} dx^\mu dx^\nu \equiv \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} \eta_{\mu\nu} dx^\mu dx^\nu = -dt^2 + dx^2 + dy^2 + dz^2. \tag{32}$$

Suppose we have two tensors A_{ij} and B_{ij}. One could write their tensor product as $A_{ij} B_{jk} = C_{ik}$, where C_{ik} is the resulting tensor. One could also write the product of a vector, x_μ, and a tensor, $A_{\mu\nu}$, as $x_\mu A_{\mu\nu} = y_\nu$, where y_ν is the resultant vector.

For a matrix A_{ij}, it is conventional to write A^{ij} as the matrix inverse of A_{ij}. One can also easily write the trace of a matrix $A_{\mu\nu}$ as $A_{\mu\mu}$, where, since the i is repeated, a sum over the diagonal elements is implied.

A.2 Derivatives and differentials

In this index notation, we often also use shorthands for derivatives. One of the most common is the use of ∂_μ, which is $\partial/\partial x^\mu$. Note that the derivative with a *down* index corresponds to differentiation with respect to an *up* index. Sometimes *this* is shortened even further; consider the derivative of a metric tensor $g_{\mu\nu}$ with respect to some x^γ. Then one may write

$$g_{\alpha\beta,\gamma} \equiv \frac{\partial g_{\alpha\beta}}{\partial x^\gamma}. \tag{33}$$

Another bit of notation worth discussing are differential operators. A familiar one from vector calculus is the **gradient** operator, ∇. This is defined, in Cartesian coordinates, by

$$\nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right). \tag{34}$$

A closely related operator is the **Laplacian**, which is written ∇^2 or Δ. This is defined by

$$\Delta = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{35}$$

Lastly, we have the **d'Alembert** or **box** operator, written as $\Box$ and defined in the usual coordinates as

$$\Box = \frac{\partial^2}{\partial t^2} - \Delta, \tag{36}$$

or, in traditional units,

$$\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta. \tag{37}$$

The box operator is also sometimes called the *wave operator*, because for the electromagnetic four-potential, $A^\mu = (\phi, \vec{A})$, $\Box A^\mu = 0$ is exactly the wave equation for electromagnetic radiation in vacuum.