University of Copenhagen

Niels Bohr Institute

Black Holes and Gravitational Waves

Jose María Ezquiaga

(jose.ezquiaga@nbi.ku.dk)

and

Maarten van de Meent

(maarten.meent@nbi.ku.dk)

Last Modification: March 3, 2026

Contents

Prelude

These lecture notes are very much under construction. They will be updated regularly. Please bear with us!

A few points before we get started:

Throughout the lectures notes we will use different **conventions** for units. The general principle is that whenever some dimensionfull number needs to be computed, we will include all the Gs, cs and $\hbar$s. When this is not the case, typically during analytical derivations, We will set $c = 1$. Sometimes we also set $G = \hbar = 1$. We will also try to avoid using too many **acronyms** in different chapters, but if you get lost in the next page you can find a compilation of all acronyms.

Similarly, We will try to add relevant **references** to both seminal works and recent results. Given the amount of literature we will certainly miss many works though. Citations are not meant to be exhaustive but rather indicative. You can find the full bibliography at the end of the lecture notes. At the start of many sections you will find a box pointing to additional literature. For example:

Literature:

Additional literature

These point to other textbooks and/or lecture notes treating the same subject, which you can refer to for additional explanations. Reading these texts is not mandatory.

Exercises will be marked with boxes within the text of each chapter. For instance:

Exercise 0.1: An example exercise

The content of the example exercise to do

The list of all exercises can be found at the end of the document. The exercise sheets for evaluation will be sent separately.

This course is short and the field of black holes and gravitational waves very rich and active nowadays. Therefore we will inevitably only cover some basic points and highlights of current research lines. In the lectures notes we invite you to **explore** further some other topics. Those will be marked in boxes, both to excite you but mostly to identify them as separate from the standard curriculum. An example:

Explore: An example topic

An example text for further exploration

Since these parts are outside of the main curriculum, we will add further details as we move along. This means, keep checking previous chapters for new edits.

Enjoy the course!

Acronyms

AGN active galactic nucleus. 65

BBH binary black hole. 79

BH black hole. 48

BNS binary neutron star. 79

CBC compact binary coalescence. 79

CL confidence level. 127

EHT Event Horizon Telescope. 66

EM electromagnetic. 64

EoS equation of state. 52

FAP false alarm probability. 126

FRW Friedmann-Robertson-Walker. 53, 96

GR general relativity. 48, 68

GW gravitational wave. 43, 66, 68, 92, 113

IBCO inner bound stable circular orbit. 25

ISCO inner most stable circular orbit. 25, 89

NS neutron star. 56

NSBH neutorn star black hole. 79

PBH primordial black hole. 55

PISN pair instability supernova. 60

PM post-Minkowskian. 75

PN post-Newtonian. 75

PSD power spectral density. 115, 121, 134

QNM Quasinormal mode. 46

SMBH super massive black hole. 62

SNR signal-to-noise ratio. 121, 134

SPA stationary phase approximation. 86, 89, 107, 109

TOV Tolman–Oppenheimer–Volkoff. 52, 59

TT transverse-traceless. 70

WD white dwarf. 56

WKB Wentzel–Kramers–Brillouin. 94

CHAPTER

1

Penrose Diagrams

Literature:

- Harvey Reall, "Part 3 Black Holes" [1] , Chapter 5
- Sean M. Carroll, "Spacetime and Geometry" [2], Sec. 5.6, 5.7 and App. H

In this chapter we will develop the tools to describe the causal structure of a spacetime, allowing us to define what it means for a spacetime to contain a black hole.

1.1 Conformal transformations

Given a metric $g_{\mu\nu}$ on a spacetime M, we can define a new metric $\bar{g}_{\mu\nu}$ as $\bar{g}_{\mu\nu} = \omega(x)^2 g_{\mu\nu}$, where $\omega : M \to \mathbb{R}$ is a positive smooth function. This new metric $\bar{g}_{\mu\nu}$ in general defines a new spacetime geometry. Nonetheless, this new geometry will agree with the old geometry on the notions of curves being "timelike", "spacelike", or "null", i.e. they will agree on the possible causal relations between different spacetime events; the **causal structure** of the spacetime. The transformation from $g_{\mu\nu}$ to $\bar{g}_{\mu\nu}$ encoded by the function $\omega(x)$ is known as a **conformal transformation**, since it preserves the angles (but not lengths) of the geometry.

The idea is to use these conformal transformations to take a (generally infinite) spacetime and map it to a compact spacetime (which we can draw on a page) with the same causal structure.

1.2 Minkowski space

We start with the Minkowski space. In spherical coordinates the Minkowski metric reads

$$\eta_{\mu\nu} = -dt^2 + dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2), \tag{1.2.1}$$

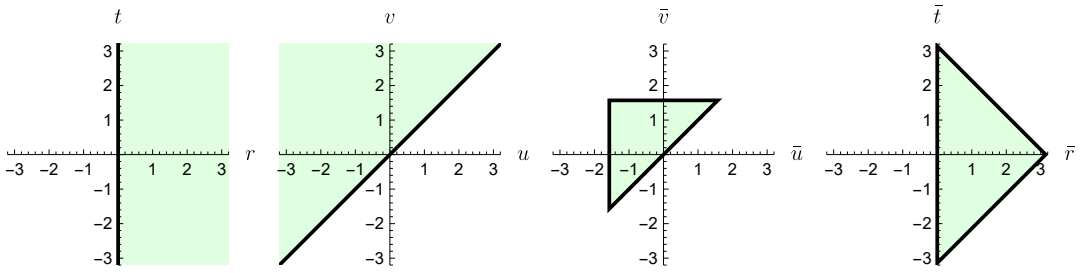


Figure 1.1. Illustration of the ranges of the different coordinates used in this section.

with the coordinate ranges (which will become important later on) $t \in (-\infty, \infty)$, $r \in [0, \infty)$, $\theta \in [0, \pi]$, and $\phi \in [-\pi, \pi]$. Our first step is to define lightcone coordinates

$$u = t - r, \qquad \text{and} \qquad v = t + r, \tag{1.2.2}$$

which gives the metric

$$\eta_{\mu\nu} = -dudv + \frac{1}{4}(v-u)^2(d\theta^2 + \sin^2\theta d\phi^2), \tag{1.2.3}$$

with ranges $-\infty < u \leqslant v < \infty$. We can compactify the range of our coordinates by choosing new coordinates $(\bar{u}, \bar{v})$ defined by,

$$u = \tan\bar{u}, \qquad \text{and} \qquad v = \tan\bar{v}, \tag{1.2.4}$$

which reduces the ranges to $-\pi/2 < \bar{u} \leqslant \bar{v} < \pi/2$, and gives the metric

$$\eta_{\mu\nu} = (2\cos\bar{u}\cos\bar{v})^{-2}\left(-4d\bar{u}d\bar{v} + \sin^2(\bar{v}-\bar{u})(d\theta^2 + \sin^2\theta d\phi^2)\right). \tag{1.2.5}$$

To find the conformal compatification of Minkowski space, we finally define the function

$$\omega = 2\cos\bar{u}\cos\bar{v}, \tag{1.2.6}$$

and the coordinates

$$\bar{t} = \bar{u} + \bar{v}, \qquad \text{and} \qquad \bar{r} = \bar{v} - \bar{u}, \tag{1.2.7}$$

to find the metric

$$\bar{\eta}_{\mu\nu} = -d\bar{t}^2 + d\bar{r}^2 + \sin^2\bar{r}(d\theta^2 + \sin^2\theta d\phi^2). \tag{1.2.8}$$

The astute reader may recognize this as the line element of the Einstein static universe, $\mathbb{R} \times S^3$ (see Harmark Ex. 4.3). However, the crucial difference lies in the range of the coordinates, in particular $0 \leqslant \bar{r} < \pi$ and $-\pi + \bar{r} < \bar{t} < \pi - \bar{r}$ (see Fig. 1.1). By projecting onto the $(\bar{r}, \bar{t})$-plane, we obtain the **Penrose diagram** for Minkowski space, shown in Fig. 1.2. Each interior point of the shown region represents a 2-sphere in the full spacetime. The dashed boundary at $\bar{r} = 0$ represents the symmetry axis at $r = 0$

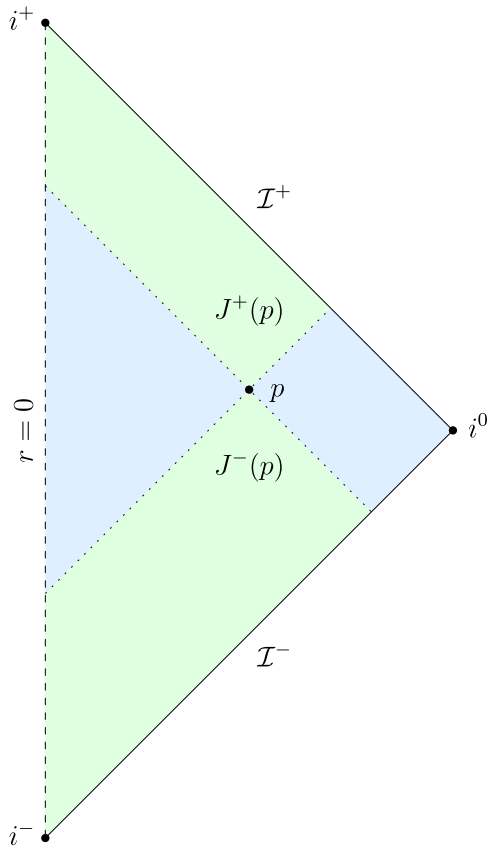


Figure 1.2. The Penrose diagram for Minkowski space.

of the original coordinates, and the outer boundary denotes the points originally "at infinity".

To understand the structure of "infinity" we can consider the behavior of radial geodesics in the diagram. By construction of the diagram all radial null geodesics travel along 45 degree lines. Any null geodesic will originate at $\mathcal{I}^-$, travel along an inwards 45 degree straight line until it reaches $r = 0$, and reflect back to an outwards 45 degree line until it reaches $\mathcal{I}^+$. We therefore refer to $\mathcal{I}^-$ as **past null infinity**, and $\mathcal{I}^+$ as **future null infinity**. In a similar vein, all radial spacelike geodesics start and end at i^0, which we call **spatial infinity**, and timelike geodesics start at i^- and end at i^+, which are called **past** and **future timelike infinity**, respectively.

A spacetime whose Penrose diagram has a structure near infinity resembling[1] that of Minkowski space is called **asymptotically flat**.

Finally, we use the Penrose diagram of Minkowski space to visualize two more important concepts related to the causal structure of spacetime. Given an event $p \in M$, the set of all events that can be connected to p with a future-directed causal (i.e. timelike or null) curve ending in p, is called the **causal past of** p, and denoted J^-. In the Penrose diagram the intersection of $J^-(p)$ with a constant (θ, ϕ)-slice is simply the 45-degree wedge below the point p as shown in Fig. 1.2. Similarly, the set of all events that can be connected to p with a future-directed causal (i.e. timelike or null) curve starting at p, is called the **causal future of** p, and denoted $J^+(p)$.

Exercise 1.1: Conformal transformations

a.The angle α between two spacelike vectors x^μ and y^ν is defined by

$$\cos\alpha = \frac{x^\mu y^\nu g_{\mu\nu}}{\sqrt{x^\mu x_\mu y^\nu y_\nu}}.$$

Show that α remains invariant when change the metric by a conformal transformation.

b. Above we showed that

$$\bar{\eta}_{\mu\nu} = -d\bar{t}^2 + d\bar{r}^2 + \sin^2\bar{r}(d\theta^2 + \sin^2\theta d\phi^2)$$

is related to the Minkowski metric by a conformal transformation. Calculate the Ricci tensor $R_{\mu\nu}$ for $\bar{\eta}_{\mu\nu}$. Does $\bar{\eta}_{\mu\nu}$ satisfy the vacuum Einstein equation?

1.3 Schwarzschild

In the course by Harmark you were introduced to the Schwarzschild metric,

[1]For a more precise technical discussion of what "resembling" means in this context, see Chapter 5 of Reall's lecture notes.

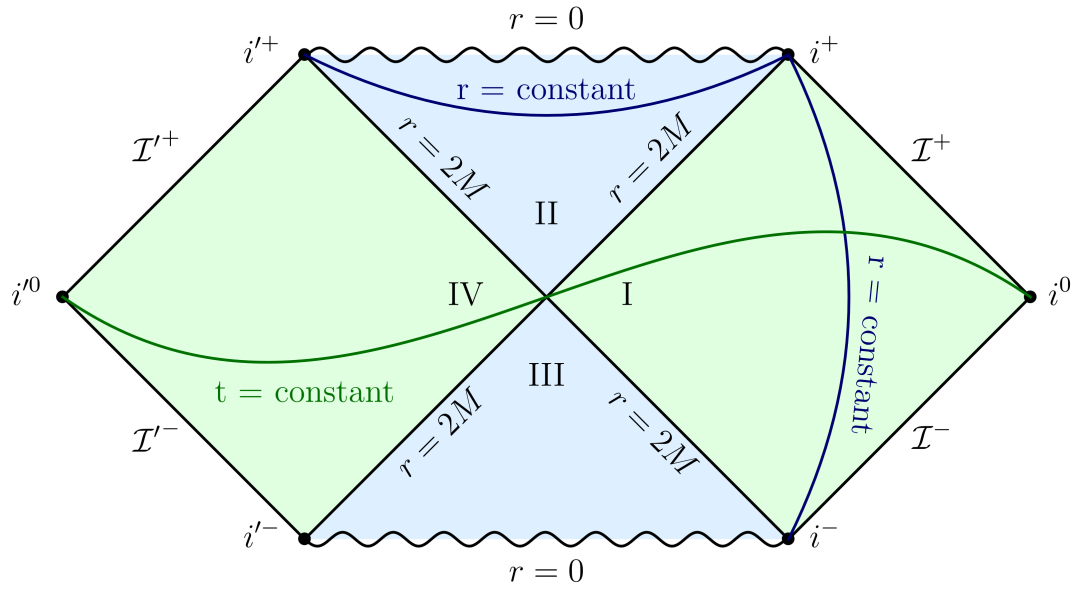


Figure 1.3. The Penrose diagram for the maximally extended Schwarzschild spacetime.

$$g_{\mu\nu} = -(1 - \frac{2M}{r})dt^2 + \frac{dr^2}{1 - \frac{2M}{r}} + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{1.3.1}$$

This metric has a (coordinate) singularity at $r = 2M$. Consequently, the a priori domain of the coordinates is $-\infty < t < \infty$ and $2M < r < \infty$. By changing the time coordinate to (advanced) Eddington-Finkelstein time

$$v = t + r^* \quad \text{with} \quad r^* = r + 2M \log\left(\frac{r - 2M}{2M}\right), \tag{1.3.2}$$

you were shown that the validity of the Schwarzschild metric can be extended to $0 < r < \infty$ for infalling observers. The same reasoning can be applied when using the retarded Eddington-Finkelstein time coordinate $u = t - r^*$ to show that there exists an (inequivalent) extension of Schwarzschild for the past of outgoing observers. When using both u and v to form double null coordinates we obtain

$$g_{\mu\nu} = -\left(1 - \frac{2M}{r(u,v)}\right) dudv + r(u,v)^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{1.3.3}$$

Even with the maximal coordinate range $-\infty < u, v < \infty$ these coordinates correspond only to the $2M < r < \infty$ Schwarzschild patch. The $r = 2M$ boundary corresponds to the $u \to \infty$ and $v \to -\infty$ limits. To find coordinates that cover all of the extended spacetime we introduce rescaled null coordinates

$$u = -4M \log(-\underline{u}) \quad \text{and} \quad v = 4M \log(\underline{v}). \tag{1.3.4}$$

In these coordinates the original Schwarzschild patch is covered by $-\infty < \underline{u} < 0$ and $0 < \underline{v} < \infty$. With these coordinates the Schwarzschild metric takes its **Kruskal-Szekeres** form,

$$g_{\mu\nu} = -\frac{32M}{r(\underline{u},\underline{v})} e^{-\frac{r}{2M}} d\underline{u} d\underline{v} + r(\underline{u},\underline{v})^2 (d\theta^2 + \sin^2\theta d\phi^2), \tag{1.3.5}$$

where r is defined implicitly through the relation $\underline{u}\,\underline{v} = (1 - r/(2M))e^{r/(2M)}$. This form is manifestly smooth at $\underline{u} = 0$ and $\underline{v} = 0$ (i.e. $r = 2M$), and can be extended beyond this point.

To produce the Penrose diagram we can again compactify the range of $\underline{u}$ and $\underline{v}$ by introducing

$$\underline{u} = \tan\bar{u}, \qquad \text{and} \qquad \underline{v} = \tan\bar{v}, \tag{1.3.6}$$

and an appropriate conformal factor ω such that we obtain a compactified metric of the form

$$\bar{g}_{\mu\nu} = -2d\bar{u}d\bar{v} + R(\bar{u},\bar{v})^2 (d\theta^2 + \sin^2\theta d\phi^2). \tag{1.3.7}$$

(the explicit form of ω and R is messy and not particularly illuminating).

The resulting Penrose diagram is shown in Fig. 1.3. The original Schwarzschild patch (marked "I") is asymptotically flat. The original (advanced) Eddington-Finkelstein coordinates extend the metric to region "II" allowing objects to fall into the black hole and reach the singularity at $r = 0$. The corresponding construction with the retarded time extends the metric to region "III", allowing the worldline to extend to the singularity in their past.

Finally, we are presented with a second asymptotically flat region ("IV"), complete with its own copies of future and past null infinity ($\mathcal{I}'^+$ and $\mathcal{I}'^-$), future and past timelike infinity (i'^+ and i'^-), and spacelike infinity i'^0. There are no causal curves that connect regions I and IV. So, for an observer in region I, region IV might as well not exist, and vice versa. There are however spacelike curves that connect regions I and IV. This "worm hole" between "parallel universes" is known as an **Einstein-Rosen bridge**. Because it consists solely of spacelike curves, there is no way to transverse it.

Exercise 1.2: Kruskal-Szekeres coordinates

The Schwarschild metric in Kruskal-Szekeres coordinates is given by

$$ds^2 = \frac{32M^3}{r} e^{-\frac{r}{2M}} (-d\underline{t}^2 + d\underline{r}^2) + r(\underline{t},\underline{r})^2 (d\theta^2 + \sin^2\theta d\phi^2)$$

with $r(\underline{t},\underline{r})$ defined implicitly by

$$\underline{\mathrm{t}}^2 - \underline{\mathrm{r}}^2 = \left(1 - \frac{r}{2M}\right) e^{\frac{r}{2M}}.$$

In region I of the Schwarzschild spacetime, the Kruskal coordinates $\underline{\mathrm{t}}$ and $\underline{\mathrm{r}}$ are related to the Schwarzschild coordinates t and r trough

$$\underline{\mathrm{t}} = \left(\frac{r}{2M} - 1\right)^{1/2} e^{\frac{r}{4M}} \sinh\left(\frac{t}{4M}\right) \tag{1.3.8}$$

$$\underline{\mathrm{r}} = \left(\frac{r}{2M} - 1\right)^{1/2} e^{\frac{r}{4M}} \cosh\left(\frac{t}{4M}\right) \tag{1.3.9}$$

a. Show by explicit substitution that (1.3.8) and (1.3.9) relate the Kruskal metric in region I ($0 < \underline{\mathrm{t}} + \underline{\mathrm{r}} < \infty$ and $0 > \underline{\mathrm{t}} - \underline{\mathrm{r}} > -\infty$) is isometric to the exterior ($r > 2M$) patch of the Schwarzschild metric. Pay explicit attention to the ranges of the coordinates.
b. The coordinate transformation (1.3.8) and (1.3.9) does not make sense when $r < 2M$. Why? Find the coordinate transformation that relates the Schwarzschild interior patch ($r < 2M$) to region II ($0 < \underline{\mathrm{t}} + \underline{\mathrm{r}} < \infty$ and $0 < \underline{\mathrm{t}} - \underline{\mathrm{r}} < \infty$).
c. Find the coordinate transformations that relate regions III ($0 > \underline{\mathrm{t}} + \underline{\mathrm{r}} > -\infty$ and $0 > \underline{\mathrm{t}} - \underline{\mathrm{r}} > -\infty$) and IV ($0 > \underline{\mathrm{t}} + \underline{\mathrm{r}} > -\infty$ and $0 < \underline{\mathrm{t}} - \underline{\mathrm{r}} < \infty$) of the Kruskal metric to parts of the Schwarzschild metric. Are these regions isometric to the interior or exterior patch of the Schwarzschild metric?

1.3.1 Definition of a Black Hole

We are now ready to give a precise definition of a black hole:

Definition 1. Let $(M, g_{\mu\nu})$ be a spacetime that is asymptotically flat at null infinity. The **black hole region** is $\mathcal{B} = M \setminus J^-(\mathcal{I}^+)$. The boundary of $\mathcal{B}$, $\partial\mathcal{B}$ is called the **future event horizon**, $\mathcal{H}^+$. Similarly, $\mathcal{W} = M \setminus J^+(\mathcal{I}^-)$ is the **white hole region**, and its boundary $\partial\mathcal{W}$, the **past event horizon**, $\mathcal{H}^-$.

In more plain English a black hole consists of those events in spacetime from which no signal can reach future null infinity, a white hole consists of those events which cannot be reached by any signal starting at past null infinity. A direct consequence of this definition is that an event horizon is always a null-surface.

Concretely, when we look at the Penrose diagram for Schwarzschild in Fig. 1.3, we see that the black hole region $\mathcal{B}$ corresponding to $\mathcal{I}^+$ consists of the union of regions II and IV, and the future event horizon $\mathcal{H}^+$ is the boundary $\bar{u} = 0$, where $r = 2M$. The white hole region $\mathcal{W}$ is the union of regions III and IV, and the past event horizon is $\mathcal{H}^-$ is the boundary $\bar{v} = 0$, where $r = 2M$ as well.

Note that since the maximally extended Schwarzschild spacetime has two asymptotically flat regions, each region has its own black (white) hole regions. Relative to $\mathcal{I}'^+$ the

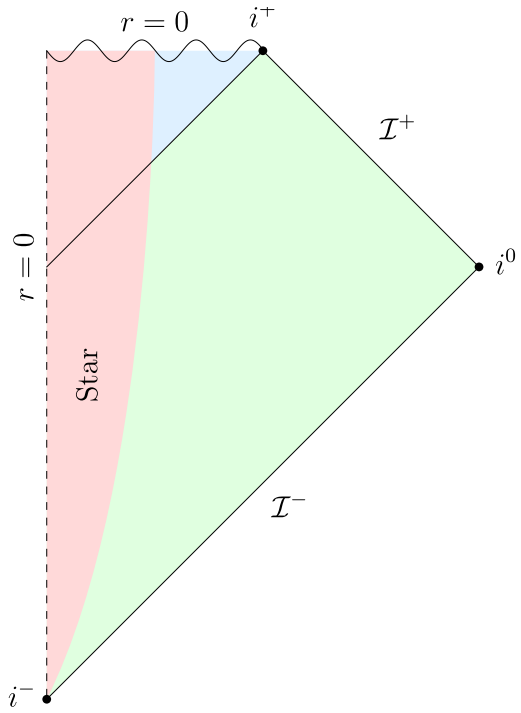


Figure 1.4. The Penrose diagram for a spherically symmetric star collapsing to a black hole.

black hole consists of the unions of I and II, and relative to $\mathcal{I}'^-$ the white hole consists of regions I and III.

1.3.2 Spherical collapse

The maximally extended Schwarzschild solution contains some fantastical features (worm-holes, parallel universes, etc.). A good question to ask ourselves is which of these features arise from the strict symmetry constraints (vacuum; spherical symmetry; static), and which features survive in a more realistic scenario. To this end we display the Penrose diagram for a spherically symmetric star collapsing to a black hole in Fig. 1.4. Outside of the star, the spacetime is spherically symmetric and vacuum, and therefore — by Birkhoff's theorem — given by the Schwarzschild solution. The interior of the star, which encapsulates $r = 0$ at early times, is given by a more regular solution.

We see that the only fantastical feature that remains is the formation of an event horizon, black hole region, and singularity. All the more science fiction sounding features like white holes, worm holes, and parallel universes have disappeared due to the inclusion of a matter distribution.

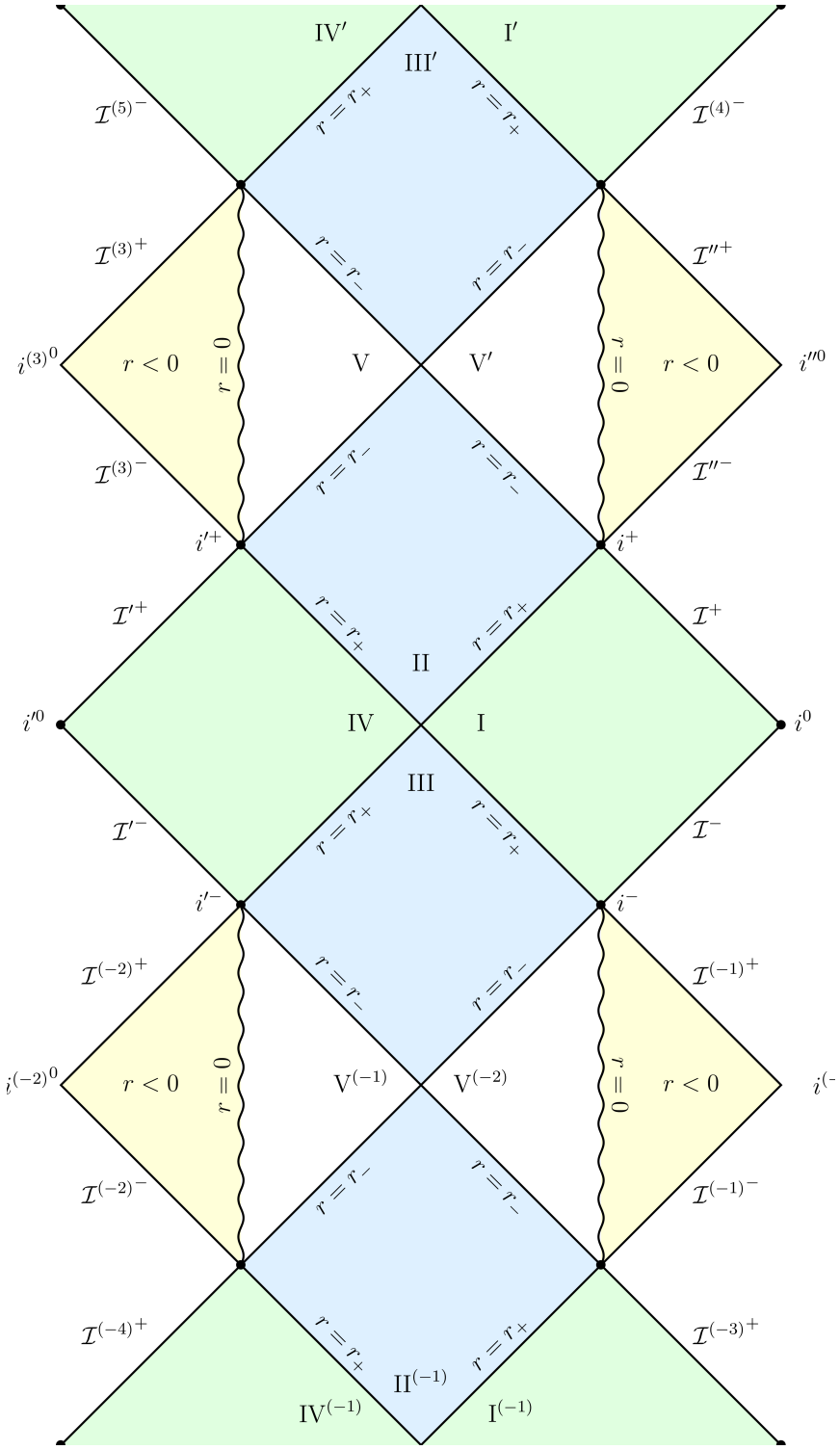


Figure 1.5. The Penrose diagram for a maximally extended Kerr black hole.

1.4 The Kerr Black Hole

The Kerr solution for a rotating black hole was introduced in Sec. 3.3 of Harmark's notes. In Boyer-Lindquist coordinates it is given by

$$g_{\mu\nu} = -\left(1 - \frac{2Mr}{\Sigma}\right) dt^2 - \frac{4Mar}{\Sigma} \sin^2\theta dt d\phi \\ + \frac{(r^2+a^2)^2 - a^2\Delta\sin^2\theta}{\Sigma} \sin^2\theta d\phi^2 + \frac{\Sigma}{\Delta} dr^2 + \Sigma d\theta^2, \tag{1.4.1}$$

with

$$\Sigma(r,\theta) = r^2 + a^2\cos^2\theta, \quad \text{and} \tag{1.4.2}$$

$$\Delta(r) = r^2 - 2Mr + a^2 = (r - r_+)(r - r_-). \tag{1.4.3}$$

This metric has a trivial coordinate singularity at $\theta = 0$ and $\theta = \pi$, coordinate singularities at $r = r_\pm = M \pm \sqrt{M^2 - a^2}$ corresponding to horizons, and a true curvature ring singularity at $r = 0$ and $\theta = \pi/2$.

The coordinate singularities at $r = r_\pm$ divide the metric in three disjoint patches ($r_+ < r < \infty$, $r_- < r < r_+$, and $-\infty < r < r_-$) that, a priori, are unrelated. The different patches can be related to each other using a similar procedure as in Schwarzschild using advanced and retarded time coordinates. We will not discuss the details here[2], instead we simply present the resulting Penrose diagram in Fig. 1.5.

From region I we can again traverse into a black hole by crossing $r = r_+$ into region II. In region II the r-direction is timelike and we can only proceed to smaller values of r until we cross $r = r_-$ into region V (or V'). In this region the r-direction is again spacelike and worldlines can both proceed to smaller r or (as will be the case for almost all geodesics) back outward to larger r until we cross $r = r_-$ again and enter region III'. Here r is again timelike and we can only proceed further outward until we cross $r = r_+$ into region I', a new distinct copy of region I (or its parallel universe region IV'). From here we can continue ad infinitum to find an infinite stack of new asymptotically flat regions.

In region V we could have also proceeded to $r = 0$. Only for $\theta = \pi/2$ is this a singularity. For all other values of θ we are free to cross into the $r < 0$ region. This region is again asymptotically flat but now in the $r \to -\infty$ limit. However, in this region there is no horizon to shield future null infinity from the singularity at $r = 0$, the singularity is **naked**. To make matters worse, this region also has closed timelike curves, breaking causality.

Exercise 1.3: Negative Kerr

We discussed that in the maximally extended Kerr solution we can access the $r < 0$ region by passing through the ring singularity at $r = 0$.

[2]For the full details see [3]

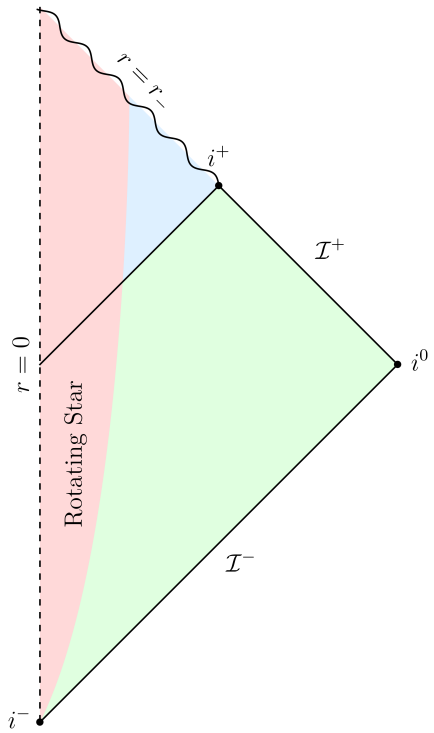


Figure 1.6. A sketch of the Penrose diagram for a rotating star collapsing to a Kerr black hole.

a. Show that the $r < 0$ patch of the Kerr metric is isometric to the $r > 0$ patch of the Kerr metric with negative mass.
b. Consider the closed loop described by taking t, r, and θ constant. Calculate the norm of tangent vector of this loop.
c. Show that for small negative values of r this norm can become negative. What does this imply?

1.4.1 Collapse of a rotating black hole

Maximally extended Kerr is even more fantastical than its non-rotating counterpart. It is therefore a good idea to regain our grounding and look at a more realistic collapse scenario. Unfortunately, there is no equivalent of Birkhoff's theorem for rotating geometries. Consequently, we are not guaranteed that the spacetime outside of the star is exactly Kerr. Instead we have to make do with the much weaker result from the no hair theorem that after the collapse the region outside the horizon has to settle down to a member of the Kerr(-Newman) family.

Most of the really fantastical features of maximally extended Kerr manifest themselves after crossing the $r = r_-$ boundary. How much should we really trust the solution

after this point. The first thing to note is that this boundary is a **Cauchy horizon**, meaning that if we try to construct the solution for initial data on some timelike slice this is the absolute limit we can evolve or solutions before singularities start to appear somewhere in our domain of dependence. We should be very wary of the analytically extended version of the solution beyond this point, because it cannot be obtained through a natural evolution. Second, we note that if we take an event on the Cauchy horizon between region II and V' its causal history $J^-(p)$ will actually contain all of region I (for some value of θ). An observer crossing this line is therefore confronted with an infinite amount of history in a finite amount of time. Consequently, even the tiniest perturbation of the spacetime is likely to be blue shifted to something singular. It is therefore conjectured that this Cauchy surface will become singular for generic perturbations away from Kerr.

Based on these conjectures we sketch what the Penrose diagram for a rotating collapsing star likely looks like in Fig. 1.6. The final result looks remarkably like the Schwarzschild case, avoiding the most fantastical behaviour of the maximally extended Kerr.

CHAPTER

2

Geodesics in Kerr Spacetime

2.1 The geodesic equation

Literature:

- Troels Harmark, "General Relativity and Cosmology" [4] , Sec. 1.3.3 and 1.4.6
- Sean M. Carroll, "Spacetime and Geometry" [2], Sec. 3.3 and 3.4

In the first GR course you saw the geodesic equation

$$\frac{D}{ds}\frac{\mathrm{d}x^\mu}{\mathrm{d}s} = \frac{\mathrm{d}x^\alpha}{\mathrm{d}s}\nabla_\alpha\frac{\mathrm{d}x^\mu}{\mathrm{d}s} = \frac{\mathrm{d}^2x^\mu}{\mathrm{d}s^2} + \Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}s}\frac{\mathrm{d}x^\beta}{\mathrm{d}s} = 0. \tag{2.1.1}$$

Solving this equation can be made easier by first identifying constants of motion, the existence of which is intertwined with the notion of spacetime symmetries.

2.2 Symmetries and Killing vectors

Literature:

- Sean M. Carroll, "Spacetime and Geometry" [2], Sec. 3.8 and App. B,
- Robert Wald, "General Relativity" [5], Appendix C

A symmetry of a spacetime M is a automorphism $\phi : M \to M$ that leaves the spacetime "invariant". Infinitesimal automorphisms are given by vector fields in the following way. Suppose we have a (smooth) vector field V^μ, we can construct a family of automorphisms $\phi^t_{V^\mu}$ indexed by a variable $t \in \mathbb{R}$ by mapping each event $p \in M$ to

a new event p' by following the integral curves of V^μ for a time t. (I.e. we solve the differential equation $\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = V^\mu$ with initial condition $x(0) = p$, and set $p' = x(t)$.)

We can try to ask ourselves the question how does a $T_{\mu_1\dots\mu_n}$ change along the integral lines of V^μ. Naively one may try to write the down the derivative

$$\lim_{t\to 0} \frac{T_{\mu_1\dots\mu_n}(p') - T_{\mu_1\dots\mu_n}(p)}{t}. \tag{2.2.1}$$

However, such an expression does not make any mathematical sense. The tensors $T_{\mu_1\dots\mu_n}(p')$ and $T_{\mu_1\dots\mu_n}(p)$ belong to (the tensor product of) the (co)tangent space at different points in M. Consequently, we cannot add (or subtract) them. To get another object that lives at p, we can consider the map induced by $\phi^t_{V^\mu}$ on the tensor bundles (the push forward), or more specifically its inverse (the pull back) $(\phi^t_{V^\nu})^*$, which in terms of components is given by

$$\left(\left(\phi^t_{V^\nu}\right)^* T\right)_{\mu_1\dots\mu_n}(p) = \frac{\partial p'^{\alpha_1}}{\partial p^{\mu_1}} \cdots \frac{\partial p'^{\alpha_n}}{\partial p^{\mu_n}} T_{\alpha_1\dots\alpha_n}(p'). \tag{2.2.2}$$

This allows us to define a derivative that is defined covariantly and expresses how much a tensor field changes in the direction of a vector field V^μ.

Definition 2. Given a vector field V^μ and a tensor field $T_{\mu_1\dots\mu_n}$ we can define the **Lie**[1] **derivative** of $T_{\mu_1\dots\mu_n}$ w.r.t. V^μ at an event p as follows,

$$\mathcal{L}_{V^\nu} T_{\mu_1\dots\mu_n}(p) = \lim_{t\to 0} \frac{\left((\phi^t_{V^\nu})^* T_{\mu_1\dots\mu_n}\right)(p) - T_{\mu_1\dots\mu_n}(p)}{t} \tag{2.2.3}$$

Note that this notion of derivative does not require the existence of a metric. This makes it suitable to explore the symmetries of the metric itself. In particular, if we calculate the Lie derivative of a metric tensor $g_{\mu\nu}$ we find

$$\begin{aligned}
\mathcal{L}_{V^\lambda} g_{\mu\nu} &= \lim_{t\to 0} \frac{\left((\phi^t_{V^\lambda})^* g_{\mu\nu}\right)(p) - g_{\mu\nu}(p)}{t} && (2.2.4)\\
&= \lim_{t\to 0} \frac{\frac{\partial p'^\alpha}{\partial p^\mu}\frac{\partial p'^\beta}{\partial p^\nu} g_{\alpha\beta}(p') - g_{\mu\nu}(p)}{t} && (2.2.5)\\
&= \lim_{t\to 0} \frac{\left(\delta^\alpha_\mu + t\partial_\mu V^\alpha\right)\left(\delta^\beta_\nu + t\partial_\nu V^\beta\right)\left(g_{\alpha\beta}(p) + tV^\gamma \partial_\gamma g_{\alpha\beta}(p)\right) - g_{\mu\nu}(p)}{t} && (2.2.6)\\
&= g_{\alpha\nu}\partial_\mu V^\alpha + g_{\mu\beta}\partial_\nu V^\beta + V^\gamma \partial_\gamma g_{\mu\nu} && (2.2.7)\\
&= \partial_\mu V_\nu - V^\alpha \partial_\mu g_{\alpha\nu} + \partial_\nu V_\mu - V^\beta \partial_\nu g_{\mu\beta} + V^\gamma \partial_\gamma g_{\mu\nu} && (2.2.8)\\
&= \partial_\mu V_\nu + \partial_\nu V_\mu - V^\alpha \left(\partial_\mu g_{\alpha\nu} + \partial_\nu g_{\mu\alpha} - \partial_\alpha g_{\mu\nu}\right) && (2.2.9)\\
&= \partial_\mu V_\nu + \partial_\nu V_\mu - 2V_\alpha \Gamma^\alpha_{\mu\nu} && (2.2.10)\\
&= \nabla_\mu V_\nu + \nabla_\nu V_\mu = 2\nabla_{(\mu} V_{\nu)}. && (2.2.11)
\end{aligned}$$

We are now ready to introduce the notion of a symmetry of spacetime $(M, g_{\mu\nu})$.

[1] Pronounced "Lee" after 19th century Norwegian mathematician Sophus Lie.

Definition 3. Let K^μ be a vector field on a (pseudo)-Riemannian manifold $(M, g_{\mu\nu})$. K^μ is called a **Killing**[2] vector (field) if equivalently:

1. $\mathcal{L}_{K^\lambda} g_{\mu\nu} = 0$
2. $\nabla_{(\mu} K_{\nu)} \equiv \frac{1}{2}(\nabla_\mu K_\nu + \nabla_\nu K_\mu) = 0$

Killing vectors encode the symmetries of a spacetime geometry. Sometimes coordinates make it easy to find Killing vectors, as described by the following lemma.

Lemma 1. If the components of a metric $g_{\mu\nu}$ in some particular coordinates do not depend on the coordinate k, then $\left(\frac{\partial}{\partial k}\right)^\mu$ is a Killing vector field.

Proof. Left as an exercise to the reader. □

This leads to the main result that will help us solve the geodesic equation in Kerr.

Lemma 2. Let $x^\mu(s)$ be a geodesic on a (pseudo)-Riemannian manifold $(M, g_{\mu\nu})$, and let K^μ be a Killing vector field, then the quantity $\mathcal{K} = K_\alpha \frac{\mathrm{d}x^\alpha}{\mathrm{d}s}$ is conserved along the geodesic $x^\mu(s)$.

Proof.

$$\frac{\mathrm{d}\mathcal{K}}{\mathrm{d}s} = \frac{\mathrm{d}x^\alpha}{\mathrm{d}s}\nabla_\alpha \mathcal{K} \tag{2.2.12}$$

$$= \frac{\mathrm{d}x^\alpha}{\mathrm{d}s}\nabla_\alpha \left(K_\beta \frac{\mathrm{d}x^\beta}{\mathrm{d}s} \right) \tag{2.2.13}$$

$$= \frac{\mathrm{d}x^\alpha}{\mathrm{d}s}\frac{\mathrm{d}x^\beta}{\mathrm{d}s}\nabla_\alpha K_\beta + K_\beta \frac{\mathrm{d}x^\alpha}{\mathrm{d}s}\nabla_\alpha \frac{\mathrm{d}x^\beta}{\mathrm{d}s} \tag{2.2.14}$$

$$= \frac{\mathrm{d}x^\alpha}{\mathrm{d}s}\frac{\mathrm{d}x^\beta}{\mathrm{d}s}\nabla_{(\alpha} K_{\beta)} + 0 \tag{2.2.15}$$

$$= 0. \tag{2.2.16}$$

□

The notion of a Killing vector field has a generalization to tensors for higher rank that satisfy similar properties.

Definition 4. Let $K_{\mu\nu}$ be a symmetric tensor field on a (pseudo)-Riemannian manifold $(M, g_{\mu\nu})$. $K_{\mu\nu}$ is called a **Killing tensor** (field) if

$$\nabla_{(\lambda} K_{\mu\nu)} = 0. \tag{2.2.17}$$

There are a number of trivial examples of Killing tensors.

Example 1. For any (pseudo)-Riemannian manifold $(M, g_{\mu\nu})$, the metric tensor itself $g_{\mu\nu}$ is a Killing tensor field.

[2] After the 19th century German mathematician, Wilhelm Killing.

Example 2. Suppose V^μ and W^μ are Killing vector fields, then $K_{\mu\nu} = V_{(\mu}W_{\nu)}$ is a Killing tensor.

Unlike Killing vectors, Killing tensors do not have an interpretation in terms of spacetime symmetries. Nonetheless, they still lead to constants of motion.

Lemma 3. Let $x^\mu(s)$ be a geodesic on a (pseudo)-Riemannian manifold $(M, g_{\mu\nu})$, and let $K^{\mu\nu}$ be a Killing tensor field, then the quantity $\mathcal{K} = K_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}$ is conserved along the geodesic $x^\mu(s)$.

Proof. Left as an exercise to the reader. □

Example 3. When applied to the metric $g_{\mu\nu}$ this lemma reproduces the familiar result that the norm of the tangent vector to a geodesic $g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}$ is preserved along a geodesic.

Exercise 2.1: Killing-Yano tensors

An anti-symmetric rank-2 tensor $\omega_{\mu\nu} = -\omega_{\nu\mu}$ is called a **Killing-Yano tensor** if it satisfies

$$\nabla_\alpha\omega_{\beta\gamma} + \nabla_\beta\omega_{\alpha\gamma} = 0.$$

a. Show that for any Killing-Yano tensor $\omega_{\alpha\beta}$, the symmetric tensor $\omega_{\alpha\gamma}\omega_\beta{}^\gamma$ is a Killing tensor. (In some sense Killing-Yano tensors can be thought of as the "square root" of a Killing tensor.))
b. Show that for any Killing-Yano tensor $\omega_{\mu\nu}$ and geodesic $x^\mu(s)$, the (co-)vector $V_\mu = \omega_{\mu\alpha}\frac{dx^\alpha}{ds}$ is parallel transported along the geodesic, i.e.

$$\frac{dx^\alpha}{ds}\nabla_\alpha V^\mu = 0. \tag{2.2.18}$$

Note: Kerr also has a Killing-Yano tensor.

2.3 Constants of Motion of Kerr geodesics

From the explicit expression for the Kerr metric in Boyer-Lindquist coordinates (1.4.1), we can immediately infer that the Kerr metric has two Killing vector fields $\left(\frac{\partial}{\partial t}\right)^\mu$ and $\left(\frac{\partial}{\partial\phi}\right)^\mu$, related to the time translation and axial symmetries of the spacetime, respectively. These are all the independent Killing vector fields that the Kerr metric has. (More precisely, all Killing vectors of Kerr can be written as a linear combination of $\left(\frac{\partial}{\partial t}\right)^\mu$ and $\left(\frac{\partial}{\partial\phi}\right)^\mu$).

These Killing symmetries lead to two constants of motion

$$\mathcal{E} = -\left(\frac{\partial}{\partial t}\right)_\alpha\frac{\mathrm{d}x^\alpha}{\mathrm{d}s} = -g_{t\alpha}\frac{\mathrm{d}x^\alpha}{\mathrm{d}s}, \tag{2.3.1}$$

$$\mathcal{L} = \left(\frac{\partial}{\partial\phi}\right)_\alpha\frac{\mathrm{d}x^\alpha}{\mathrm{d}s} = g_{\phi\alpha}\frac{\mathrm{d}x^\alpha}{\mathrm{d}s}. \tag{2.3.2}$$

When the affine parameter s is chosen such that $\frac{\mathrm{d}x^\mu}{\mathrm{d}s}$ equals the four momentum p^μ of the particle following the geodesic, $\mathcal{E}$ is equal to the energy of the particle and $\mathcal{L}$ is equal to the component of the orbital angular momentum along the symmetry axis of the Kerr geometry. By analogy, we will refer to $\mathcal{E}$ and $\mathcal{L}$ as the **energy** and **angular momentum** regardless of the chosen affine parameter.

As for any spacetime the metric $g_{\mu\nu}$ is a Killing tensor leading to our third constant of motion, the invariant mass squared

$$\mu = -g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}. \tag{2.3.3}$$

It turns out that Kerr spacetime has an additional "hidden" symmetry in the form of a Killing tensor. This Killing tensor is given by

$$K_{\mu\nu} = \Sigma\left(\ell_\mu n_\nu + \ell_\nu n_\mu\right) + r^2 g_{\mu\nu}, \tag{2.3.4}$$

where ℓ^μ and n^ν are principal null vectors[3] of the Kerr geometry. In Boyer-Lindquist coordinates their components can be written,

$$\ell^\mu = \left(\frac{r^2+a^2}{\Delta}, 1, 0, \frac{a}{\Delta}\right), \quad \text{and} \tag{2.3.5}$$

$$n^\mu = \left(\frac{r^2+a^2}{2\Sigma}, -\frac{\Delta}{2\Sigma}, 0, \frac{a}{2\Sigma}\right). \tag{2.3.6}$$

This gives rise to a fourth constant of motion, the **Carter constant**

$$\mathcal{K} = K_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}. \tag{2.3.7}$$

This constant of motion has the rough interpretation as the "total angular momentum squared" of the particle.

Any functional combination of constants of motion is also a constant of motion. It is conventional to use this freedom to replace $\mathcal{K}$ with the constant

$$\mathcal{Q} = \mathcal{K} - \left(\mathcal{L} - a\mathcal{E}\right)^2, \tag{2.3.8}$$

which (somewhat confusingly) is also referred to as "the Carter constant". (We will be following this tradition.)

[3]This means that they are null vectors k^μ, which satisfy $k^\alpha k^\beta k_{[\mu}C_{\nu]\alpha\beta[\lambda}k_{\rho]} = 0$, which expresses that k^μ is an eigenvector of the Weyl curvature tensor $C_{\mu\nu\rho\sigma}$ in some suitable sense.

2.4 Separating the geodesic equations

So with constants of motion $(\mu, \mathcal{E}, \mathcal{L}, \mathcal{Q})$ in hand, we have four equations that a geodesic must satisfy

$$\mu = -g_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}s}\frac{\mathrm{d}x^\beta}{\mathrm{d}s}, \tag{2.4.1}$$

$$\mathcal{Q} + (\mathcal{L} - a\mathcal{E})^2 = K_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}s}\frac{\mathrm{d}x^\beta}{\mathrm{d}s}, \tag{2.4.2}$$

$$\mathcal{E} = -g_{t\alpha}\frac{\mathrm{d}x^\alpha}{\mathrm{d}s}, \tag{2.4.3}$$

$$\mathcal{L} = g_{\phi\alpha}\frac{\mathrm{d}x^\alpha}{\mathrm{d}s}. \tag{2.4.4}$$

So, if we fix values of $(\mu, \mathcal{E}, \mathcal{L}, \mathcal{Q})$, we have four equations for four unknowns (the components of $\frac{\mathrm{d}x^\mu}{\mathrm{d}s}$), and we can solve these equations. After some straightforward (but tedious) algebra we find

$$\left(\frac{\mathrm{d}r}{\mathrm{d}s}\right)^2 = \frac{\left(\mathcal{E}(r^2+a^2) - a\mathcal{L}\right)^2 - \Delta\left(\mathcal{Q} + (\mathcal{L}-a\mathcal{E})^2 + \mu r^2\right)}{\Sigma^2} \tag{2.4.5}$$

$$\left(\frac{\mathrm{d}\cos\theta}{\mathrm{d}s}\right)^2 = \frac{\mathcal{Q} - \cos^2\theta\Big(a^2(\mu - \mathcal{E}^2)\sin^2\theta + \mathcal{L}^2 + \mathcal{Q}\Big)}{\Sigma^2} \tag{2.4.6}$$

$$\frac{\mathrm{d}t}{\mathrm{d}s} = \frac{\frac{r^2+a^2}{\Delta}\big(\mathcal{E}(r^2+a^2) - a\mathcal{L}\big) - a^2\mathcal{E}\sin^2\theta + a\mathcal{L}}{\Sigma}, \text{ and} \tag{2.4.7}$$

$$\frac{\mathrm{d}\phi}{\mathrm{d}s} = \frac{\frac{a}{\Delta}\left(\mathcal{E}(r^2+a^2) - a\mathcal{L}\right) + \frac{\mathcal{L}}{\sin^2\theta} - a\mathcal{E}}{\Sigma}. \tag{2.4.8}$$

This is already a huge improvement over the general geodesic equations in that only first order derivatives appear. Moreover, t and ϕ do not appear at all on the right hand side of the equations. If we can manage to solve the equations for r and θ the solutions for t and ϕ can be found by direct integration. Finally, the equations for r and θ appear almost entirely decoupled, apart from the Σ in the denominators the right hand side of the radial (r) equation depends only on r and the polar (θ) equation depends only on θ. Having decoupled equations would be nice since we could then solve each equation independently. To fully decouple the equations we introduce a new time parameter, the **Mino-Carter time** λ defined by

$$\frac{\mathrm{d}\lambda}{\mathrm{d}s} = \frac{1}{\Sigma}. \tag{2.4.9}$$

With this choice our equations become

$$\begin{aligned}
\left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2 &= \Big(\mathcal{E}(r^2+a^2) - a\mathcal{L}\Big)^2 - \Delta\left(\mathcal{Q} + (\mathcal{L}-a\mathcal{E})^2 + \mu r^2\right) &&= P_r(r) \\
\left(\frac{\mathrm{d}\cos\theta}{\mathrm{d}\lambda}\right)^2 &= \mathcal{Q} - \cos^2\theta\Big(a^2(\mu-\mathcal{E}^2)\sin^2\theta + \mathcal{L}^2 + \mathcal{Q}\Big) &&= P_\theta(\cos\theta) \\
\frac{\mathrm{d}t}{\mathrm{d}\lambda} &= \frac{r^2+a^2}{\Delta}(\mathcal{E}(r^2+a^2) - a\mathcal{L}) - a^2\mathcal{E}\sin^2\theta + a\mathcal{L} &&= T_r(r) + T_\theta(\cos\theta) \\
\frac{\mathrm{d}\phi}{\mathrm{d}\lambda} &= \frac{a}{\Delta}\Big(\mathcal{E}(r^2+a^2) - a\mathcal{L}\Big) + \frac{\mathcal{L}}{\sin^2\theta} - a\mathcal{E} &&= \Phi_r(r) + \Phi_\theta(\cos\theta).
\end{aligned}$$

These equations are completely decoupled. Moreover, the right hand side of the t and ϕ equations separate as the sum of something that depends on r with something that depends on θ. Consequently, these can be integrated separately for different radial and polar solutions.

The equations for r and $\cos\theta$ each individually take the form of a particle moving in a 1-dimensional potential. The functions P_r and P_θ are therefore known as the **radial** and **polar potentials**. Note that instead of the equations involving the square of the first derivatives (for r and θ) we can easily return to second order differential equations (which now will also be decoupled) by taking the derivative with respect to λ to yield

$$\frac{\mathrm{d}^2 r}{\mathrm{d}\lambda^2} = \frac{1}{2}P_r'(r), \quad \text{and} \tag{2.4.10}$$

$$\frac{\mathrm{d}^2\cos\theta}{\mathrm{d}\lambda^2} = \frac{1}{2}P_\theta'(\cos\theta). \tag{2.4.11}$$

2.5 The polar equation

$$\left(\frac{\mathrm{d}\cos\theta}{\mathrm{d}\lambda}\right)^2 = \mathcal{Q} - \cos^2\theta\Big(a^2(\mu-\mathcal{E}^2)\sin^2\theta + \mathcal{L}^2 + \mathcal{Q}\Big) = P_\theta(\cos\theta) \tag{2.5.1}$$

Because the left-hand side of the equation is a square, we can only find solutions if the right-hand side is not negative, i.e. when $P_\theta(\cos\theta) \geqslant 0$. We can therefore learn about the possible solutions by studying the roots of the polynomial $P_\theta(z)$. Since $P_\theta(z)$ is a fourth order polynomial in $z = \cos\theta$, there at most 4 real roots. Moreover, since $P_\theta(z)$ is an even function of z these roots come in pairs $z = \pm z_i$. Consequently,

$$P_\theta(z) = a^2(\mu-\mathcal{E}^2)(z^2-z_1^2)(z^2-z_2^2). \tag{2.5.2}$$

If we evaluate $P_\theta(z)$ at the poles $z = \pm 1$, we find that $P_\theta(\pm 1) = -\mathcal{L}^2$. We thus immediately learn geodesics can only reach the poles when $\mathcal{L} = 0$. When we evaluate $P_\theta(z)$ on the equator $\cos\theta = 0$ we find that $P_\theta(0) = \mathcal{Q}$. Therefore, we only find solutions that visit the equator when $\mathcal{Q} \geqslant 0$. If $\mathcal{Q} > 0$ and $\mathcal{L} \neq 0$, there must be an odd number of zeroes between 0 and 1, and the same number between -1 and 0. Since there are at most 4 zeroes, there is exactly one such zero z_1, and the only solutions oscillate around

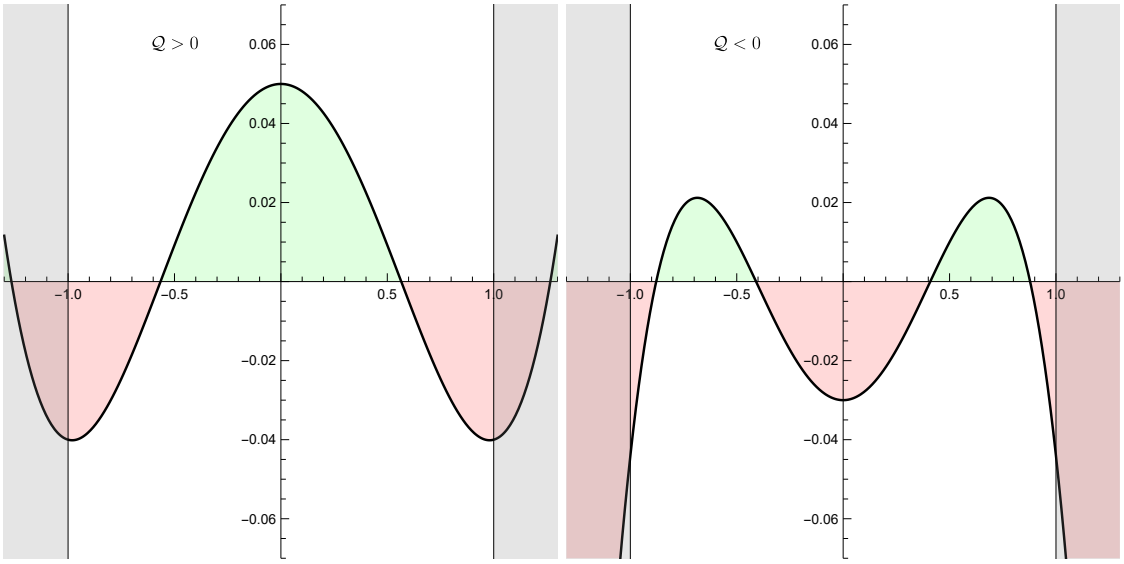


Figure 2.1. The polar potential P_θ when $\mathcal{Q}$ is positive (on the left) or negative (on the right). In the positive $\mathcal{Q}$ case, we find solutions where $\cos\theta$ oscillates around the equator ($\cos\theta = 0$) between $\pm z_1$. In the negative $\mathcal{Q}$ case, we can get vortical solutions where $\cos\theta$ oscillates between $z_1 < z_2$ (or $-z_2 < -z_1$), but never crosses the equator.

the equator between $\pm z_1$. In the Schwarzschild ($a \to 0$) limit, these solutions describe trajectories whose orbital plane is inclined relative to the equator of the coordinate system. As such these solutions are generally known as **inclined** trajectories. In Kerr this orbital plane precesses, leading to another common name **precessing** trajectories.

If $\mathcal{Q} = 0$, then $z = 0$ is a double root of P_θ. This implies that $x = \cos\theta = 0$ is constant is a solution to the differential equation. These solutions stay in the Kerr equatorial plane and are therefore known as **equatorial** trajectories.

If $\mathcal{Q} < 0$ then no solution is possible near the equator. In this case, it is only possible for P_θ to be positive somewhere in the range $-1 < \cos\theta < 1$, if P_θ has exactly two zeroes $0 < z_1 < z_2 \leqslant 1$. We can easily see that this is only possible if $\mu < \mathcal{E}^2$ by considering the behavior of P_θ at large z. At large z, $P_\theta = a^2(\mu - \mathcal{E}^2)z^4 + \mathcal{O}(z^2)$. So, if $\mu > \mathcal{E}^2$ the polar potential becomes positive at large z. Consequently, there must be at least one zero between $z = 1$ (where P_θ is negative) and $z = \infty$, and there cannot be four roots in the range $-1 < z < 1$. The condition $\mu < \mathcal{E}^2$ is by itself not sufficient for the existence of two zeroes $0 < z_1 < z_2 \leqslant 1$. One can show (which is left as an exercise) that such roots exist if and only if the following conditions are met

$$\mathcal{L}^2 \leqslant a^2(\mathcal{E}^2 - \mu) \text{ and } -\left(|\mathcal{L}| - |a|\sqrt{\mathcal{E}^2 - \mu}\right)^2 \leqslant \mathcal{Q} \leqslant 0. \tag{2.5.3}$$

Solutions of this type are referred to as **vortical** trajectories.

Type	$\mu - \mathcal{E}^2$	$\frac{\mathcal{L}^2}{a^2(\mathcal{E}^2-\mu)}$	$\mathcal{Q}$	$z = \cos\theta$
Equatorial	any	any	0	0
Inclined	any	any	> 0	$-z_1 < z < z_1$
Vortical	< 0	$\leqslant 1$	$-1 \leqslant \frac{\mathcal{Q}}{\left(\lvert\mathcal{L}\rvert - \lvert a\rvert\sqrt{\mathcal{E}^2-\mu}\right)^2} \leqslant 0$	$z_1 < \lvert z\rvert < z_2$

Table 2.1. Classification of solutions of the polar equation.

Exercise 2.2: Vortical solutions

Let's examine the conditions for vortical solutions.
a. Define

$$\hat{\mathcal{Q}} = \frac{\mathcal{Q}}{a^2(\mathcal{E}^2-\mu)}$$
$$\hat{\mathcal{L}}^2 = \frac{\mathcal{L}^2}{a^2(\mathcal{E}^2-\mu)}$$
$$u = \cos^2\theta$$
$$\hat{P}_\theta = -\frac{P_\theta}{a^2(\mathcal{E}^2-\mu)}$$

Show that with this notation we get

$$\hat{P}_\theta(u) = u^2 + \left(\hat{\mathcal{Q}} + \hat{\mathcal{L}}^2 - 1\right)u - \hat{\mathcal{Q}}.$$

b. Find the zeroes of $\hat{P}_\theta(u)$.
c. Show that in order for both roots u_1 and u_2 real and lie in the range $0 < u_1 < u_2 < 1$ we need that $\hat{\mathcal{L}}^2 \leqslant 1$ and $-\hat{\mathcal{Q}} \geqslant \left(\left|\hat{\mathcal{L}}\right| - 1\right)^2$? Can you recover the conditions for vortical solutions listed in Table 2.1?

2.6 The radial equation

The decoupled equation for radial motion is given by

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2 = \left(\mathcal{E}(r^2+a^2) - a\mathcal{L}\right)^2 - \Delta\left(\mathcal{Q} + (\mathcal{L}-a\mathcal{E})^2 + \mu r^2\right) = P_r(r). \tag{2.6.1}$$

It has the same, overall structure as the polar equation. In particular, we can only have solutions when the radial potential P_r is non-negative, and the radial potential is a fourth order polynomial in r, which therefore has four zeroes r_i (0,2, or 4 of which may be real-valued) and can be written

$$P_r = (\mathcal{E}^2-\mu)(r-r_1)(r-r_2)(r-r_3)(r-r_4). \tag{2.6.2}$$

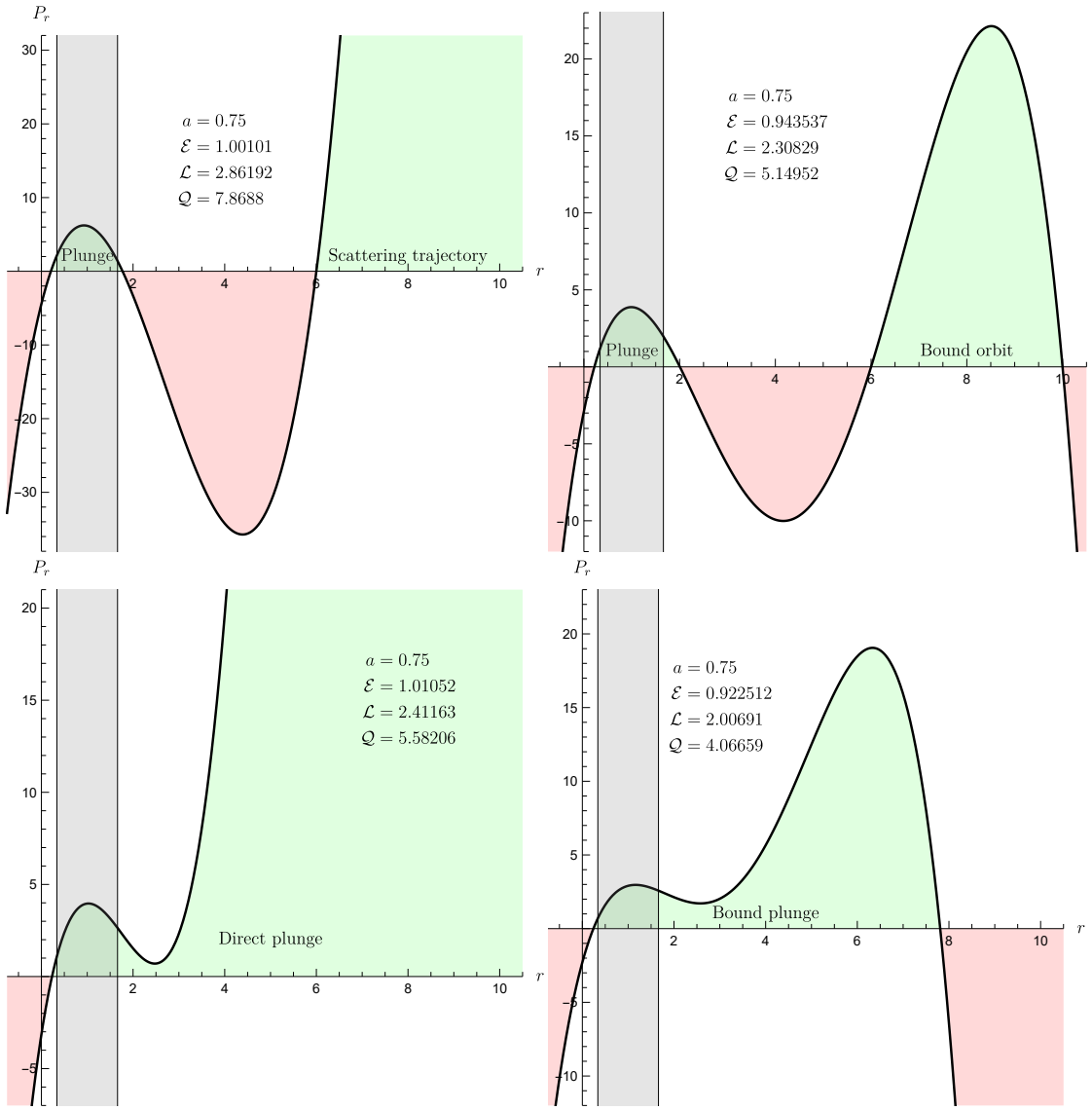


Figure 2.2. The radial potential P_r for a scattering trajectory (top left), bound orbit (top right), direct plunge (bottom left), and bound plunge (bottom right). Note the existence of secondary "deeply bound" plunge solutions in the top plots.

As with the polar equation it will be useful to examine the value of P_r at specific values of r. We start with $r = 0$, and observe that $P_r(0) = -a^2\mathcal{Q}$. Consequently, solutions to the radial equation can only reach $r = 0$ if $\mathcal{Q} \leqslant 0$, i.e. only equatorial and vortical trajectories can ever reach $r = 0$. Of these, only the equatorial ($\mathcal{Q} = 0$) solutions can ever reach the curvature singularity at $\theta = \pi/2$. We thus find that of all geodesics in Kerr only a measure zero subset hit the singularity. All others miss it entirely, some of them passing through the ring singularity to the $r < 0$ region (all of them vortical) and some of them scattering back due to centrifugal barrier.

Next we turn our attention to the horizon $r_\pm$, since Δ vanishes here we find

$$P_r = \left(\mathcal{E}(r_\pm^2 + a^2) - a\mathcal{L}\right)^2 \geqslant 0. \tag{2.6.3}$$

Consequently, for all values of $(\mu, \mathcal{E}, \mathcal{L}, \mathcal{Q})$ that allow the existence of solutions to the polar equation, we have solutions that cross the horizons. Moreover, since r is timelike between the horizons, we cannot have any zeroes between $r_\pm$ for timelike $\mu > 0$ and null $\mu = 0$ solutions. This also implies that if $\mathcal{Q} > 0$ then there must exist at least one (and at most three) zeroes between the inner horizon r_- and $r = 0$.

Last we turn our attention to the behavior of P_r as r approaches $\pm\infty$. Expand P_r we find $P_r = (\mathcal{E}^2 - \mu)r^4 + \mathcal{O}(r^3)$. Consequently, in order for a geodesic to reach infinity $(\mathcal{I}^\pm, i^\pm, i^0)$ it must have $\mathcal{E}^2 > \mu$. For this reason geodesics satisfying this condition are referred to as **unbound**. Note that for null geodesics $\mu = 0$; they are necessarily unbound.

For unbound $\mathcal{E}^2 > \mu$ geodesics the radial potential must have an even number of zeroes between the outer horizon r_+ and $r = \infty$. When there are no zeroes in this region the solution corresponds to a trajectory that starts at infinity and dives directly into the black hole. These solutions are known as **direct plunges**.

If there are two zeroes $r_1 > r_2$ in the region outside of the black hole, solutions correspond to trajectories starting from infinity, and scattering of the black hole potential back to infinity. This trajectories are known as **scattering** or **hyperbolic** trajectories.

Geodesics with $\mathcal{E}^2 < \mu$ are only possible for timelike trajectories with $\mu > 0$, and known as **bound** trajectories because they do not possess enough energy to reach infinity. The existence of a polar solution implies that for bound orbits we necessarily have that $\mathcal{Q} \geqslant 0$, and therefore that there exists at least one zero inside the inner horizon. Since the radial potential changes sign between the outer horizon r_+ and infinity there must be either one or three zeroes outside the outer horizon.

If there exists just one zero $r_+ < r_1 < \infty$ the solution corresponds to a trajectory starting at the past horizon moving outward to r_1 and then diving back into the black hole. These trajectories are known as **(bound) plunges**. Of course, these geodesics do not stop at the horizon. If followed into the black hole, they will also cross the inner horizon before encountering a turning point in the inner region, and proceeding back outward through the inner horizon and exiting the outer horizon into a new parallel universe in the maximally extended Kerr solution, after which the above repeats ad infinitum with the geodesic eventually (according to its own proper time) visiting infinitely many copies of the external universe.

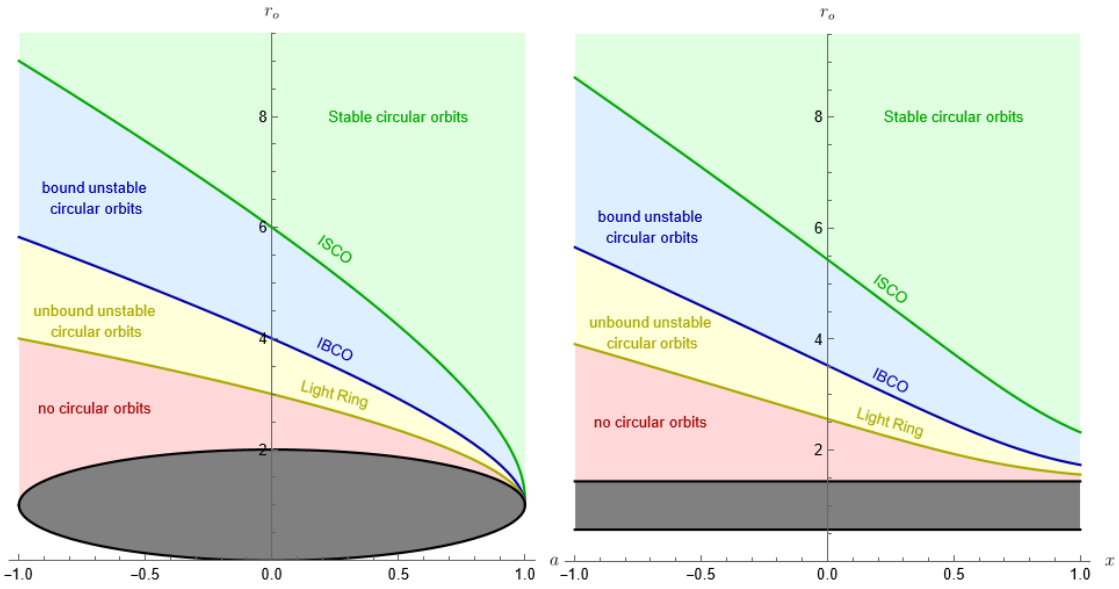


Figure 2.3. The location of the ISCO, IBCO, and light ring in Kerr. On the left, equatorial orbits, on the right at fixed spin $a = 0.9M$.

If there are three zeroes $r_+ < r_3 < r_2 < r_1 < \infty$ then the solution oscillates between r_1 and r_2 to form a **bound (eccentric) orbit**.

Note that both in the case of bound orbits and scattering orbits, there exists a secondary solution with the same constants of motion which oscillates between the remaining turning point outside the horizon and the zero inside the inner horizon (if $\mathcal{Q} > 0$). These solutions behave similarly to the bound plunges and are therefore known as **deeply bound plunges**.

Exercise 2.3: Vortical solutions revisited

We now continue our examination of vortical geodesics from Exercise 2.2 and turn the radial part of vortical solution.

d. Show that for vortical solutions

$$P_r \geqslant 2Mr(\mathcal{Q} + (\mathcal{L} - a\mathcal{E})^2 + \mu r^2).$$

e. Show that for vortical null geodesics ($\mu = 0$) that $P_r > 0$ when $r \geqslant 0$. What does this mean for their radial solutions? Sketch a vortical null trajectory in the maximally extended Penrose diagram for Kerr.

f. Show that for vortical timelike geodesics ($\mu > 0$) we always have that $P_r > 0$ outside r_-. What does this tells us about their radial solutions?

2.6.1 Circular orbits

Circular orbits are geodesics for which the radial potential allows a constant solution r_o. This requires not only that $P_r(r_o) = 0$ (which ensures that $\frac{\mathrm{d}r}{\mathrm{d}\lambda} = 0 = \frac{\mathrm{d}r}{\mathrm{d}s}$), but also that $P_r'(r_o) = 0$ (which ensures that $\frac{\mathrm{d}^2r}{\mathrm{d}\lambda^2} = 0 = \frac{\mathrm{d}^2r}{\mathrm{d}s^2}$). This therefore requires that r_o is a double zero of the radial potential. A double zero can be either a maximum or minimum of the radial potential. In the former case the circular orbit is **stable**; any small perturbation will produce an eccentric orbit with small eccentricity. In the later case, the circular orbit is **unstable**; any perturbation leads to a wildly different orbit.

Stable circular orbits occur as the limit of eccentric orbits as the zeroes r_1 and r_2 merge. Unstable circular orbits are formed by merging the inner turning point of a scattering or bound orbit with the outer turning point of a deeply bound plunge.

Stable circular orbits can only be found outside a certain radius, the **inner most stable circular orbit (ISCO)**. Inside this radius one can still find circular orbits, but the become increasingly unstable with larger and larger energies. Initially, these unstable circular orbits have low enough energies, that even after a perturbation the particle will stay bound to the black hole. However, below another radius, the **inner bound stable circular orbit (IBCO)**, the energy becomes large enough that perturbations can send the particle scattering to infinity. Finally, the energy of the circular orbits diverges at the **light ring**, this radius can only be approached if one simultaneously allows the mass μ to go to zero, and the geodesic to become null. Inside the light ring no circular orbits are possible (outside the event horizon).

2.7 Characterizing bound orbits

In the previous sections we have found the bound orbits ($\mu > \mathcal{E}^2$) are necessarily either inclined ($\mathcal{Q} > 0$) or equatorial ($\mathcal{Q} = 0$), and corresponds to a solution oscillating in the radial direction between $r_1 \geqslant r_2 > r_+$, and in the polar direction between $\pm z_1$. There are a number of characterizations of these orbits that are useful in identifying them

- The **constants of motion** $(\mathcal{E}, \mathcal{L}, \mathcal{Q})$. For any sets of values of the constants of motion, if a bound orbit exists then that solution is unique. (There will always also exist a deeply bound plunge with the same constants of motion.)

- The **turning points** (r_1, r_2, z_1) uniquely characterize a bound orbit up to the direction of motion: with the rotation of the black hole (**prograde**) or against the rotation of the black hole (**retrograde**). This degeneracy is conventionally broken by the convention that the angular momentum $\mathcal{L}$ is always positive, and letting the sign of the spin determine the sense of the motion. The turning point r_1 is known as the **apocenter** and r_2 as the **pericenter**.

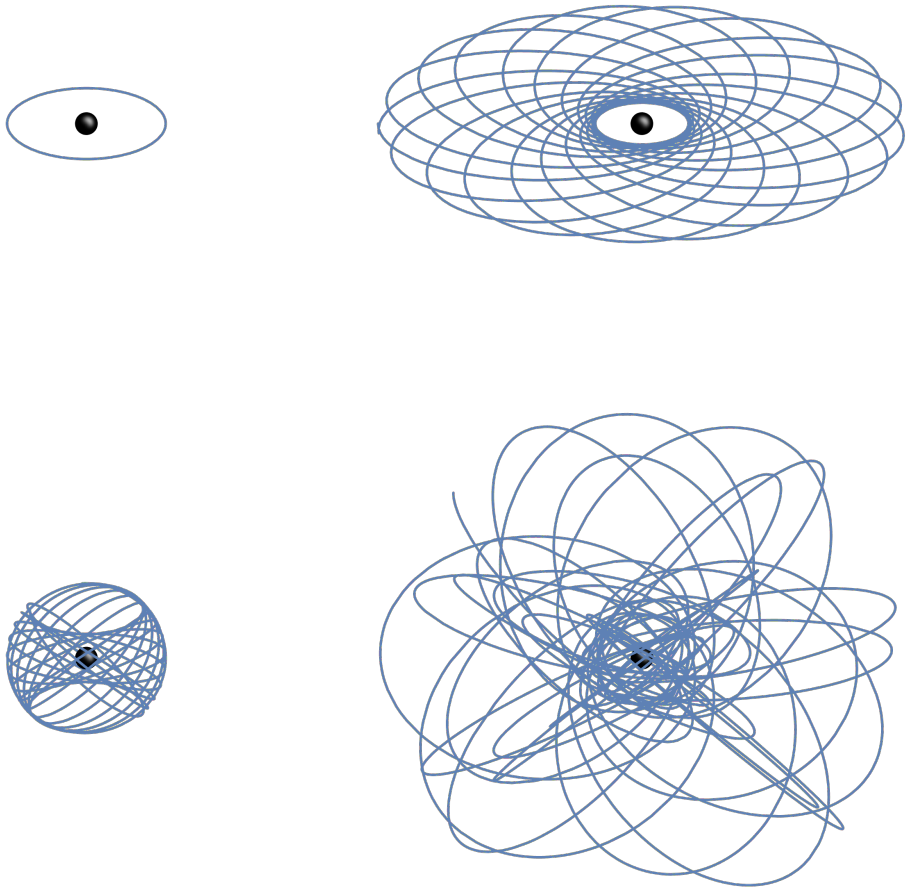


Figure 2.4. Four renderings of bound orbits around a Kerr black hole with spin $a = 0.9M$. **Top left:** A circular equatorial orbit with radius $r = 10M$. **Top right:** An eccentric equatorial orbit with semilatus rectum $p = 10M$ and eccentricity $e = 0.7$. **Bottom left:** A circular inclined orbit with radius $r = 10M$ and polar turning point $z_1 = 0.7$. **Right left:** A eccentric inclined orbit with semilatus rectum $p = 10M$, eccentricity $e = 0.7$, and polar turning point $z_1 = 0.7$.

- A Keplerian-like parametrization consting of the

$$p = \frac{2r_1 r_2}{r_1 + r_2} \qquad \textbf{semi-latus rectum}$$
$$e = \frac{r_1 - r_2}{r_1 + r_2} \qquad \textbf{eccentricity}$$
$$x = \begin{cases} \sqrt{1 - z_1^2} & \text{prograde} \\ -\sqrt{1 - z_1^2} & \text{retrograde} \end{cases} \qquad \textbf{inclination}$$

  In this parametrization $e = 0$ corresponds to circular motion, and $x = \pm 1$ corresponds to equatorial motion.

- A more phenomenological characterization is in terms of **periods** of orbit
  - Λ_r The (Mino time) period of the radial motion.
  - Λ_θ The (Mino time) period of the polar motion.
  - Λ_ϕ The average Mino time needed to complete an orbital cycle.

  The triple $(\Lambda_r, \Lambda_\theta, \Lambda_\phi)$, uniquely identifies a (relativistic) bound orbit. This may come as a surprise to those familiar with Newtonian orbital dynamics where we always have $\Lambda_r = \Lambda_\theta = \Lambda_\phi$.

- Similarly, the **Mino time frequencies**

$$\Upsilon_r = \frac{2\pi}{\Lambda_r}$$
$$\Upsilon_\theta = \frac{2\pi}{\Lambda_\theta}$$
$$\Upsilon_\phi = \frac{2\pi}{\Lambda_\phi} = \lim_{\Lambda\to\infty} \frac{1}{2\Lambda} \int_{-\Lambda}^{\Lambda} \frac{\mathrm{d}\phi}{\mathrm{d}\lambda}\, \mathrm{d}\lambda.$$

In the same spirit as the azimuthal frequency we can define the average advance of the Boyer-Lindquist time t,

$$\Upsilon_t = \lim_{\Lambda\to\infty} \frac{1}{2\Lambda} \int_{-\Lambda}^{\Lambda} \frac{\mathrm{d}t}{\mathrm{d}\lambda}\, \mathrm{d}\lambda.$$

This allows us to define the **Boyer-Lindquist frequencies** of the orbit.

$$\Omega_r = \frac{\Upsilon_r}{\Upsilon_t}$$
$$\Omega_\theta = \frac{\Upsilon_\theta}{\Upsilon_t}$$
$$\Omega_\phi = \frac{\Upsilon_\phi}{\Upsilon_t}.$$

The Boyer-Lindquist frequencies are frequencies that a distant observer at rest would infer when observing the system. Unlike the Mino-time frequencies the triple $(\Omega_r, \Omega_\theta, \Omega_\phi)$ does not always uniquely identify a bound orbit. In the strong field regime, we typically encounter pairs of distinct **isofrequency** orbits with the same Boyer-Lindquist frequencies.

Two other commonly observed quantities are :

- The **pericenter advance** $\Upsilon_\phi \Lambda_r - 2\pi$
- The **nodal advance** $\Upsilon_\phi \Lambda_\theta - 2\pi$

Although we are not going to derive them here, explicit expressions for all these quantities (and the corresponding geodesic solutions) can be obtained analytically by solving the ODEs. These solutions can be obtained e.g. from the Black Hole Perturbation Toolkit [6].

2.8 Null geodesics and black hole shadows

Null geodesics describe the trajectories of light rays through the spacetime. From the analysis in the previous sections we note that setting $\mu = 0$ implies that null geodesics are always unbound. As observers far away from a typical black hole we are interested in light rays that can reach infinity. We found that there are two main possibilities if we follow such a ray back in time either:

A The ray follows a scattering trajectory with $r \geqslant r_2 \geqslant r_3 > r_+$. In this case the polar motion must be either equatorial or inclined.

or

B The ray follows a (time reversed) direct plunge originating from the past horizon (with no zeroes outside the horizon). In this case the polar motion can be any of equatorial, inclined, or vortical.

Let us first consider the equatorial case. In this case the radial potential specializes to

$$P(r) = \Big(\mathcal{E}(r^2 + a^2) - a\mathcal{L}\Big)^2 - \Delta(\mathcal{L} - a\mathcal{E})^2 \tag{2.8.1}$$

By rescaling everything by a factor $\mathcal{E}^2$ and introducing $b = \mathcal{L}/\mathcal{E}$ we obtain

$$\tilde{P}(r) \equiv \frac{P(r)}{\mathcal{E}^2} = \Big((r^2 + a^2) - ab\Big)^2 - \Delta(b - a)^2. \tag{2.8.2}$$

We find that where the radial equation for equatorial null geodesics has only one independent degree of freedom, b, whereas the timelike case had two, $\mathcal{E}$ and $\mathcal{L}$. That is, in Kerr spacetime there is a one parameter family of equatorial null geodesics, parameterized by b. The parameter b, known as the **impact parameter**, has a simple geometric

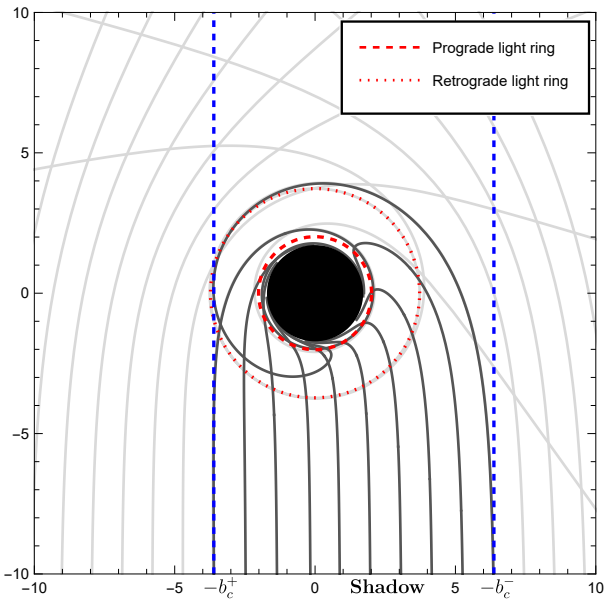


Figure 2.5. Light rays traced-back in the equatorial plane of a Kerr black hole rotating in the positive direction with spin $a = 0.70M$). Light rays with an impact parameter b between b_c^+ and b_c^- emerged from the white hole in the past, while rays outside this range trace their origin back to the celestial sphere at infinity. Consequently, when viewed from a distance the region between b_c^+ and b_c^- will appear as a shadow.

interpretation: Near infinity light rays will travel along a straight line. If we imagine extending this line towards the black hole, ignoring the curvature of the Kerr spacetime, the impact parameter b is the distance of this imaginary line to the center of the black hole. Note that this is different from closest approach of the actual null geodesic to the black hole!

In a Newtonian sense this can be understood as follows. In terms of Cartesian 3-vectors the angular momentum $\vec{L}$ is given by

$$\vec{L} = \vec{r} \times \vec{p}, \tag{2.8.3}$$

where $\vec{r}$ and $\vec{p}$ are the position and linear momentum 3-vectors. If $\vec{r}$ and $\vec{p}$ are perpendicular we get

$$\left\|\vec{L}\right\| = \|\vec{r}\|\|\vec{p}\|. \tag{2.8.4}$$

Noting that $\mathcal{L} = \left\|\vec{L}\right\|$ and for null rays $\mathcal{E} = \|\vec{p}\|$, we find that $b = \mathcal{L}/\mathcal{E} = \|\vec{r}\|$. If we van extend our convention that negative values of $\mathcal{L}$ correspond to retrograde geodesics, to b by having retrograde orbits have $b < 0$.

In figure 2.5 we plot light rays travelling in the equatorial plane of a Kerr black hole, viewed top down such that the rays are parallel travelling down near the bottom of the diagram. In this setup, the light rays cross the bottom of the diagram (approximately)

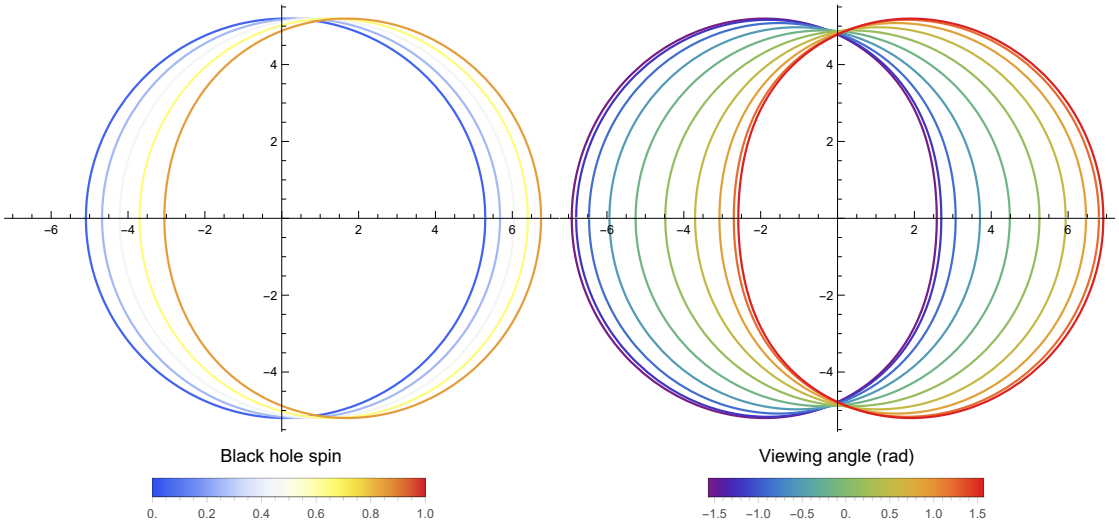


Figure 2.6. The shadow of a Kerr black hole for varying values of the black hole spin (on the left, viewed edge on) and viewing angle (on the right, at fixed $a = 0.95M$).

at $x = -b$. We see that there two regimes. If b is larger than a critical value b_c^+ (or smaller than the a critical value b_c^-) the light ray is deflected by the black hole, but originates somewhere from infinity, i.e. the null geodesic follows a scattering trajectory. On the other hand, light rays with $b_c^- < b < b_c^+$ originate from the (white) black hole. If we imagine the black hole being illuminated from all directions, no light would reach this region, it is the **shadow** of the black hole.

Near the critical value b_c^+, scattering null rays can be made to make an arbitrary large number of orbits close to the (prograde) light ring by letting b approach b_c^+ (from above). Similarly,scattering null rays approaching the b_c^- critical impact parameter (from below) can make an arbitrary number of loops around the retrograde light ring.

This picture extends to generic null geodesics as follows. Rather than having a single impact parameter, null geodesics are parameterized by a 2-dimensional **impact vector** in an emphimpact plane, with a similar geometric interpretation regarding describing how the null rays are separated from a center line through the black hole. In this plane, there is a critical loop describing the edge of the black hole shadow. Impact vectors outside the shadow correspond to scattering geodesics, while those inside correspond to plunges (or their time reverse).

In figure 2.6, we show the shape of the black hole shadow for various values of the black hole spin a and the viewing angle of the observer relative to the spin direction of the black hole. Since the shape of the black hole shadow depends on the spin, we can, in principle, measure the spin of a black hole observing its shadow, as was done by the Event Horizon Telescope. However, this is very difficult in practice since the dependence of the shape on the spin is very mild, and the main effect of the spin is to shift the location of the shadow.

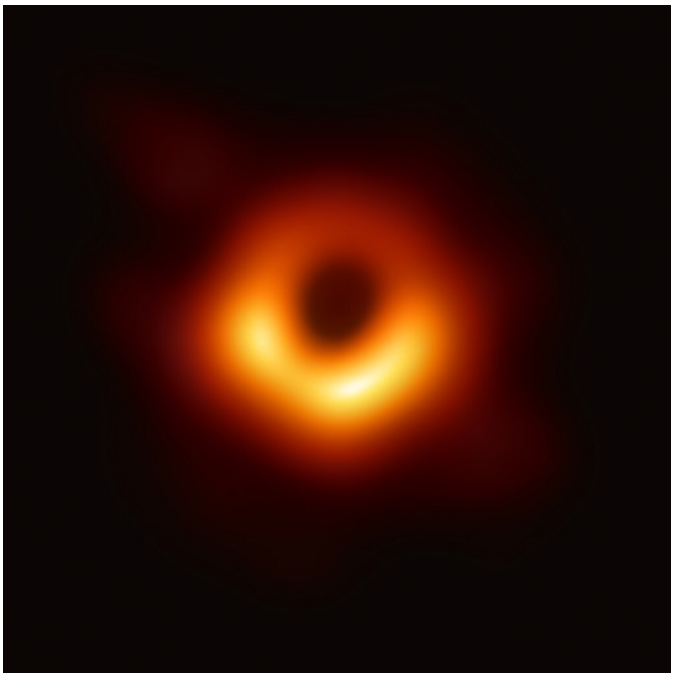


Figure 2.7. A picture of the black hole shadow of the supermassive black hole M87* made by the Event Horizon Telescope.

CHAPTER

3

Black Hole Perturbation Theory

3.1 Perturbation theory in General Relativity

Globally speaking, when applying perturbation theory to general relativity we want to make the statement that some given spacetime $(M, g_{\mu\nu})$ is equal to some background spacetime $(\bar{M}, \bar{g}_{\mu\nu})$ plus something "small". However, the equation

$$(M, g_{\mu\nu}) = (\bar{M}, \bar{g}_{\mu\nu}) + \text{"something small"}$$

doesn't really make much mathematical sense. To make sense of such a comparison we will need a map $\phi : \bar{M} \to M$ that identifies the points of $\bar{M}$ with those of M. This makes it possible to compare any tensors (such as the metric) that "live" on M with tensors that live in $\bar{M}$ by considering the pull back of the object. E.g. we can consider the difference between the pull-back of $g_{\mu\nu}$ and compare it with $\bar{g}_{\mu\nu}$

$$\delta_\phi g_{\mu\nu} = \phi^* g_{\mu\nu} - \bar{g}_{\mu\nu}, \tag{3.1.1}$$

and ask ourselves whether this is small. However, we should now wonder how this comparison depends on the choice of ϕ. What would change if we had chosen a different map $\varphi : \bar{M} \to M$ to identify the points of $\bar{M}$ and M? To answer this question note that if ϕ and φ are both diffeomorphisms then there exists a unique automorphism $\psi : \bar{M} \to \bar{M}$ such that $\varphi = \phi \circ \psi$, which is generated by some vector field ξ^μ.

To compare $\delta_\phi g_{\mu\nu}$ and $\delta_\varphi g_{\mu\nu}$ we first note that starting from the same image point $p \in M$, the differences $\delta_\varphi g_{\mu\nu}$ and $\delta_\phi g_{\mu\nu}$ "live" at different points $\bar{p}, \bar{p}' \in \bar{M}$ satisfying $\bar{p}' = \psi(\bar{p})$. So to compare the two we need to pullback $\delta_\phi g_{\mu\nu}$ to $\bar{p}$ using ψ^*. In the limit

that ξ^μ is small we get

$$\begin{aligned}
\delta_\varphi g_{\mu\nu} - \psi^* \delta_\phi g_{\mu\nu} &= \varphi^* g_{\mu\nu} - \bar{g}_{\mu\nu} - \psi^*(\phi^* g_{\mu\nu} - \bar{g}_{\mu\nu}) && (3.1.2)\\
&= (\phi \circ \psi)^* g_{\mu\nu} - \bar{g}_{\mu\nu} - \psi^*(\phi^* g_{\mu\nu} - \bar{g}_{\mu\nu}) && (3.1.3)\\
&= \psi^*\phi^* g_{\mu\nu} - \bar{g}_{\mu\nu} - \psi^*\phi^* g_{\mu\nu} + \psi^* \bar{g}_{\mu\nu} && (3.1.4)\\
&= \psi^* \bar{g}_{\mu\nu} - \bar{g}_{\mu\nu} && (3.1.5)\\
&= \mathcal{L}_{\xi^\mu} \bar{g}_{\mu\nu} + \mathcal{O}(\xi^2) && (3.1.6)\\
&= \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu + \mathcal{O}(\xi^2) && (3.1.7)
\end{aligned}$$

So, fundamentally the comparison of $g_{\mu\nu}$ with $\bar{g}_{\mu\nu}$ is ambiguous up to a Lie derivative of $\bar{g}_{\mu\nu}$. In the context of perturbation theory this is known as a **gauge** ambiguity. More generally, when comparing any tensor $T_{\mu_1 \cdots \mu_n}$ with a background tensor $\bar{T}_{\mu_1 \cdots \mu_n}$ that comparison has a gauge ambiguity $\mathcal{L}_{\xi^\mu} \bar{T}_{\mu_1 \cdots \mu_n}$.

Lemma 4. If a background tensor $\bar{T}_{\mu_1 \cdots \mu_n}$ is zero, then its perturbations are free from gauge ambiguities. A quantity free from gauge ambiguities is called **gauge invariant**.

Note that this gauge freedom in perturbation theory is logically distinct from the general coordinate freedom of general relativity. However, in practice the two are closely related, since we tend to construct the identification map ϕ by choosing similar coordinates on $\bar{M}$ and M and identifying the points with the same coordinate values. Consequently, the ambiguity ψ in this identification map simply becomes the ambiguity of choosing the coordinates.

Moving forward we will drop the fancy notation involving identification maps, and work with the understanding that a suitably identification map has been chosen and used to pull back all tensors to the background manifold $\bar{M}$. Consequently, all tensors will be tensors on the manifold $\bar{M}$ and all raising and lowering operations are understood to use the background metric $\bar{g}_{\mu\nu}$. Moreover, any perturbations are understood as being ambiguous up to gauge transformations defined by gauge vectors ξ^μ

Using these conventions we are now ready to write that

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \epsilon h_{\mu\nu} + \mathcal{O}(\epsilon^2) \tag{3.1.8}$$

where $h_{\mu\nu}$ is a symmetric rank-2 tensor living on $\bar{M}$ known as the (first-order) **metric perturbation**, and $\epsilon > 0$ is a small number that we will use to keep track of the orders in our perturbation theory. The metric perturbations are ambiguous up to transformations

$$h_{\mu\nu} \mapsto h_{\mu\nu} + \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu. \tag{3.1.9}$$

We can try to fix this gauge freedom by imposing additional conditions on $h_{\mu\nu}$.

So, suppose now we want to know what the inverse perturbed metric $g^{\mu\nu}$ looks like. Note $g^{\mu\nu}$ is now considered a tensor of $\bar{M}$, but cannot be obtained by simply raising the indices on $g_{\mu\nu}$. We should therefore be using a different symbol. Let us write

$$g^{\mu\nu} = A^{\mu\nu} + \epsilon B^{\mu\nu} + \mathcal{O}(\epsilon^2). \tag{3.1.10}$$

The inverse metric will still have to satisfy

$$\delta_\mu{}^\nu = g_{\mu\alpha}g^{\alpha\nu} \tag{3.1.11}$$
$$= (\bar{g}_{\mu\alpha} + \epsilon h_{\mu\alpha})(A^{\alpha\nu} + \epsilon B^{\alpha\nu}) \tag{3.1.12}$$
$$= \bar{g}_{\mu\alpha}A^{\alpha\nu} + \epsilon(h_{\mu\alpha}A^{\alpha\nu} + \bar{g}_{\mu\alpha}B^{\alpha\nu}) + \mathcal{O}(\epsilon^2). \tag{3.1.13}$$

This equation will have to be satisfied order-by-order in ϵ, giving us

$$\delta_\mu{}^\nu = \bar{g}_{\mu\alpha}A^{\alpha\nu}, \quad \text{and} \tag{3.1.14}$$
$$0 = h_{\mu\alpha}A^{\alpha\nu} + \bar{g}_{\mu\alpha}B^{\alpha\nu}, \tag{3.1.15}$$

which we can solve to find

$$g^{\mu\nu} = \bar{g}^{\mu\nu} - \epsilon h^{\mu\nu} + \mathcal{O}(\epsilon^2). \tag{3.1.16}$$

3.2 Linearized Einstein Equation

Literature:

- Robert Wald, "General Relativity" [5], Sec. 7.5

We now want to make our way to the perturbed form of the Einstein equation. For this we first observe that both ∇_μ and $\bar{\nabla}_\mu$ are covariant derivatives, but compatible with different metrics. Consequently, we have that their difference

$$\left(\nabla_\mu - \bar{\nabla}_\mu\right) V^\nu = (\Gamma^\nu_{\mu\alpha} - \bar{\Gamma}^\nu_{\mu\alpha})V^\alpha = \epsilon C^\nu{}_{\mu\alpha}V^\alpha \tag{3.2.1}$$

is represented by a rank-3 tensor $C^\nu{}_{\mu\alpha}$.

Equivalently, we could express this as

$$\nabla_\mu V^\nu = \bar{\nabla}_\mu V^\nu + \epsilon C^\nu{}_{\mu\alpha}V^\alpha. \tag{3.2.2}$$

This is similar as the usual expression for the covariant derivative, but with the partial derivative replaced by the background covariant derivative and the Christoffel symbols $\Gamma^\nu_{\mu\alpha}$ replaced by the tensor $C^\nu{}_{\mu\alpha}$.

Compatibility, of ∇_μ with the metric $g_{\mu\nu}$, yields an expression for $C^\nu{}_{\mu\alpha}$ in the same way we obtained our expression for the Christoffel symbols of the Levi-Civita connection (see e.g. Theorem 3.1.1 in Wald)

$$\epsilon C^\lambda{}_{\mu\nu} = \frac{1}{2}g^{\lambda\alpha}\left(\bar{\nabla}_\mu g_{\nu\alpha} + \bar{\nabla}_\nu g_{\alpha\mu} - \bar{\nabla}_\alpha g_{\mu\nu}\right) \tag{3.2.3}$$
$$= \epsilon\frac{1}{2}\bar{g}^{\lambda\alpha}\left(\bar{\nabla}_\mu h_{\nu\alpha} + \bar{\nabla}_\nu h_{\alpha\mu} - \bar{\nabla}_\alpha h_{\mu\nu}\right) + \mathcal{O}(\epsilon^2). \tag{3.2.4}$$

The Riemann tensor will now be given

$$R^{\lambda}{}_{\sigma\mu\nu} = \bar{R}^{\lambda}{}_{\sigma\mu\nu} + \epsilon\left(\bar{\nabla}_{\mu}C^{\lambda}{}_{\nu\sigma} - \bar{\nabla}_{\nu}C^{\lambda}{}_{\mu\sigma}\right) + \epsilon^{2}\left(C^{\lambda}{}_{\mu\alpha}C^{\alpha}{}_{\nu\sigma} - C^{\lambda}{}_{\nu\alpha}C^{\alpha}{}_{\mu\sigma}\right). \tag{3.2.5}$$

The Einstein equation can be written

$$R_{\mu\nu} = 8\pi\left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T\right), \tag{3.2.6}$$

where $T = g^{\alpha\beta}T_{\alpha\beta}$ the trace of the energy-momentum tensor. Let us assume we are expanding around a background metric $\bar{g}_{\mu\nu}$ that is itself a solution to the vacuum Einstein equation (e.g. Kerr), such that $\bar{R}_{\mu\nu} = 0$. We then find

$$\begin{aligned}
R_{\mu\nu} &= R^{\alpha}{}_{\mu\alpha\nu} && (3.2.7)\\
&= \epsilon\left(\bar{\nabla}_{\alpha}C^{\alpha}{}_{\mu\nu} - \bar{\nabla}_{\nu}C^{\alpha}{}_{\alpha\mu}\right) + \mathcal{O}(\epsilon^{2}) && (3.2.8)\\
&= \frac{\epsilon}{2}\bar{g}^{\alpha\beta}(\bar{\nabla}_{\alpha}\left(\bar{\nabla}_{\mu}h_{\nu\beta} + \bar{\nabla}_{\nu}h_{\beta\mu} - \bar{\nabla}_{\beta}h_{\mu\nu}\right) && (3.2.9)\\
&\qquad - \bar{\nabla}_{\nu}\left(\bar{\nabla}_{\alpha}h_{\mu\beta} + \bar{\nabla}_{\mu}h_{\beta\alpha} - \bar{\nabla}_{\beta}h_{\alpha\mu}\right)) + \mathcal{O}(\epsilon^{2}) && (3.2.10)\\
&= \frac{\epsilon}{2}\left(-\bar{\Box}h_{\mu\nu} - \bar{\nabla}_{\nu}\bar{\nabla}_{\mu}h + 2\bar{\nabla}^{\alpha}\bar{\nabla}_{(\mu}h_{\nu)\alpha}\right) + \mathcal{O}(\epsilon^{2}) && (3.2.11)\\
&= \frac{\epsilon}{2}\left(-\bar{\Box}h_{\mu\nu} - \bar{\nabla}_{\nu}\bar{\nabla}_{\mu}h + 2\bar{\nabla}_{(\mu}\bar{\nabla}^{\alpha}h_{\nu)\alpha} + 2\bar{R}^{\alpha}{}_{\mu\nu}{}^{\beta}h_{\alpha\beta}\right) + \mathcal{O}(\epsilon^{2}), && (3.2.12)
\end{aligned}$$

where $\bar{\Box} = \bar{\nabla}^{\alpha}\bar{\nabla}_{\alpha}$, and $h = h_{\alpha\beta}\bar{g}^{\alpha\beta}$ is the trace of the metric perturbation, and in the last line we used that $\bar{\nabla}_{\mu}\bar{\nabla}_{\nu}h_{\alpha\beta} = \bar{\nabla}_{\nu}\bar{\nabla}_{\mu}h_{\alpha\beta} + \bar{R}_{\mu\nu\beta}{}^{\gamma}h_{\alpha\gamma} + \bar{R}_{\mu\nu\alpha}{}^{\gamma}h_{\gamma\beta}$. Moreover, since the background was vacuum the energy momentum tensor $T_{\mu\nu}$ must also be order ϵ. Consequently, the order ϵ of the Einstein equation becomes,

$$\frac{1}{2}\bar{\Box}h_{\mu\nu} + \frac{1}{2}\bar{\nabla}_{(\mu}\bar{\nabla}_{\nu)}h - \bar{\nabla}_{(\mu}\bar{\nabla}^{\alpha}h_{\nu)\alpha} - \bar{R}^{\alpha}{}_{\mu\nu}{}^{\beta}h_{\alpha\beta} = -8\pi\left(T_{\mu\nu} - \frac{1}{2}\bar{g}_{\mu\nu}T\right). \tag{3.2.13}$$

We can further simplify this by using the gauge freedom in the metric perturbation. By choosing the Lorenz[1] gauge condition

$$\bar{\nabla}^{\alpha}h_{\mu\alpha} = \frac{1}{2}\bar{\nabla}_{\mu}h, \tag{3.2.14}$$

the linearized Einstein equation becomes

$$\frac{1}{2}\bar{\Box}h_{\mu\nu} - \bar{R}^{\alpha}{}_{\mu\nu}{}^{\beta}h_{\alpha\beta} = -8\pi\left(T_{\mu\nu} - \frac{1}{2}\bar{g}_{\mu\nu}T\right). \tag{3.2.15}$$

In this gauge, the linearized Einstein equation explicitly takes the form of a wave equation. However, we should not be deceived by the apparent simplicity of this equation, it still consists of 10 coupled second order partial differential equations, and solving them

[1] After the 19th century Danish physicist Ludvig Lorenz, not the 19th century Dutch physicist Hendrik Lorentz.

is generally hard. Sometimes, the symmetries of the background help. For example, for a Schwarzschild background the equations can be solved through separation of variables, allowing the solution to be found mode-by-mode. For each mode, it then reduces to a set of 10 coupled ordinary differential equations. But even in Kerr, there is no known way the separate the variables.

It is therefore worth to consider a different approach.

Exercise 3.1: Lorenz gauge

a. Let $h_{\mu\nu}$ be a metric perturbation. What condition should the a gauge vector ξ_μ satisfy in order for $h_{\mu\nu} + \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu$ tot satisfy the Lorenz gauge condition.
b. Show that if a gauge vector ξ_μ satisfies $\Box \xi_\mu = 0$, it preserves the Lorenz gauge condition for perturbations of vacuum spacetimes.

3.3 The Penrose Wave Equation

Instead of a wave equation for the metric we will pursue a wave equation for the curvature. The Riemann curvature tensor $R_{\mu\nu\alpha\beta}$ in 4 dimensions has 20 degrees of freedom. Ten of these are encoded by the Ricci tensor $R_{\mu\nu} = R^{\alpha}{}_{\mu\alpha\nu}$, which through the Einstein equation are algebraically determined by the energy-momentum content. The remaining 10 degrees of freedom are captured by the trace-free part of the Riemann tensor, the so-called Weyl tensor

$$C_{\mu\nu\alpha\beta} = R_{\mu\nu\alpha\beta} - g_{\alpha[\mu} R_{\nu]\beta} + g_{\beta[\mu} R_{\nu]\alpha} + \frac{R}{3} g_{\alpha[\mu} g_{\nu]\beta}. \tag{3.3.1}$$

It is defined in such a way that any contraction of its indices produces zero, while retaining the symmetries of $R_{\mu\nu\alpha\beta}$,

$$C_{\mu\nu\alpha\beta} = C_{[\mu\nu][\alpha\beta]}, \tag{3.3.2}$$
$$C_{\mu\nu\alpha\beta} = C_{\alpha\beta\mu\nu}. \tag{3.3.3}$$
$$C_{\mu\nu\alpha\beta} = C_{\mu[\nu\alpha\beta]}. \tag{3.3.4}$$

Moreover, when $R_{\mu\nu} = 0$, we get that $C_{\mu\nu\alpha\beta} = R_{\mu\nu\alpha\beta}$.

The Weyl tensor therefore must capture all the propagating curvature degrees of freedom, and it would be great if we could write a wave equation for it. To achieve this we first consider the Bianchi identity

$$\nabla_\gamma R_{\alpha\beta\mu\nu} + \nabla_\alpha R_{\beta\gamma\mu\nu} + \nabla_\beta R_{\gamma\alpha\mu\nu} = 0. \tag{3.3.5}$$

Taking the divergence of this identity yields

$$\Box R_{\alpha\beta\mu\nu} + \nabla^\gamma \nabla_\alpha R_{\beta\gamma\mu\nu} + \nabla^\gamma \nabla_\beta R_{\gamma\alpha\mu\nu} = 0. \tag{3.3.6}$$

The box operator already gives this the appearance of a wave equation. We can commute the two covariant derivatives using the Ricci identity

$$\begin{aligned}\nabla_\mu\nabla_\nu R_{\alpha\beta\gamma\delta} - \nabla_\nu\nabla_\mu R_{\alpha\beta\gamma\delta} = \\ R_{\mu\nu\alpha}{}^\lambda R_{\lambda\beta\gamma\delta} + R_{\mu\nu\beta}{}^\lambda R_{\alpha\lambda\gamma\delta} + R_{\mu\nu\gamma}{}^\lambda R_{\alpha\beta\lambda\delta} + R_{\mu\nu\delta}{}^\lambda R_{\alpha\beta\gamma\lambda}\end{aligned} \tag{3.3.7}$$

This yields

$$\begin{aligned}\Box R_{\alpha\beta\mu\nu} + 2R^\gamma{}_{\alpha\mu}{}^\lambda R_{\gamma\beta\nu\lambda} - 2R^\gamma{}_{\alpha\nu}{}^\lambda R_{\gamma\beta\mu\lambda} - 2R^\gamma{}_{\alpha\beta}{}^\lambda R_{\gamma\lambda\mu\nu} \\ - R_\alpha{}^\lambda R_{\lambda\beta\mu\nu} + R_\beta{}^\lambda R_{\lambda\alpha\mu\nu} - \nabla_\alpha\nabla^\gamma R_{\gamma\beta\mu\nu} + \nabla_\beta\nabla^\gamma R_{\gamma\alpha\mu\nu} = 0.\end{aligned} \tag{3.3.8}$$

Next by contracting the Bianchi identity itself we find the following identity for the divergence of the Riemann tensor

$$\nabla^\gamma R_{\gamma\nu\alpha\beta} = \nabla_\alpha R_{\beta\nu} - \nabla_\beta R_{\alpha\nu} = 2\nabla_{[\alpha} R_{\beta]\nu}. \tag{3.3.9}$$

Using this last identity, we can eliminate the divergences of the Riemann curvature in favour of the Ricci tensor to obtain the **Penrose wave equation**

$$\begin{aligned}\Box R_{\alpha\beta\mu\nu} + 2R^\gamma{}_{\alpha\mu}{}^\lambda R_{\gamma\beta\nu\lambda} - 2R^\gamma{}_{\alpha\nu}{}^\lambda R_{\gamma\beta\mu\lambda} + 2R^\gamma{}_{\alpha\beta}{}^\lambda R_{\gamma\lambda\mu\nu} \\ - R_\alpha{}^\lambda R_{\lambda\beta\mu\nu} + R_\beta{}^\lambda R_{\lambda\alpha\mu\nu} - 2\nabla_\alpha\nabla_{[\mu} R_{\nu]\beta} + 2\nabla_\beta\nabla_{[\mu} R_{\nu]\alpha} = 0.\end{aligned} \tag{3.3.10}$$

This equation is built from identities that hold for any pseudo-Riemannian manifold, and therefore this equation is satisfied for the curvature tensor obtained from any metric. If in addition we impose that the metric satisfies the vacuum Einstein equation $R_{\mu\nu} = 0$, all the terms in the second line vanish. So, for vacuum solutions this turns into a wave equation for the Weyl tensor.

$$\Box C_{\alpha\beta\mu\nu} + 2C^\gamma{}_{\alpha\mu}{}^\lambda C_{\gamma\beta\nu\lambda} - 2C^\gamma{}_{\alpha\nu}{}^\lambda C_{\gamma\beta\mu\lambda} + 2C^\gamma{}_{\alpha\beta}{}^\lambda C_{\gamma\lambda\mu\nu} = 0. \tag{3.3.11}$$

More generally, we could replace all occurrences of the Ricci tensor with the trace reversed energy-momentum tensor using the Einstein equation to find the wave equation for the curvature coupled to matter.

Let us now consider perturbations around a vacuum solution of the Einstein equation with an energy-momentum source that is order ϵ. Writing $C_{\alpha\beta\mu\nu} = \bar{C}_{\alpha\beta\mu\nu} + \epsilon\, \delta C_{\alpha\beta\mu\nu} + \mathcal{O}(\epsilon^2)$, the background Weyl curvature will satisfy Eq. (3.3.11).

To obtain the linear in ϵ part of the equation (3.3.10) we make the following observations

- The linear in ϵ part of the second line in Eq. (3.3.10) is constructed entirely from $\bar{C}_{\alpha\beta\mu\nu}$, $T_{\mu\nu}$, the background metric $\bar{g}_{\mu\nu}$, and the background covariant derivative $\bar{\nabla}_\mu$.

- The linear in ϵ part of the Riemann curvature tensor $\delta R_{\alpha\beta\mu\nu}$, consists of $\delta C_{\alpha\beta\mu\nu}$ plus terms constructed from $T_{\mu\nu}$ and the background metric $\bar{g}_{\mu\nu}$.
- We can write the action of the box operator on a generic 4-tensor $T_{\alpha\beta\mu\nu}$ as $\Box T_{\alpha\beta\mu\nu} = \bar{\Box} T_{\alpha\beta\mu\nu} + \epsilon B[h]_{\alpha\beta\mu\nu}{}^{\alpha'\beta'\mu'\nu'} T_{\alpha'\beta'\mu'\nu'}$ where $B[h]_{\alpha\beta\mu\nu}{}^{\alpha'\beta'\mu'\nu'}$ is formed as a linear operator on $h_{\mu\nu}$ constructed entirely from $\bar{g}_{\mu\nu}$ and $\bar{\nabla}_\mu$.
- Combining the last two observations we note that we can write the linear in ϵ part of $\Box R_{\alpha\beta\mu\nu}$ as $\bar{\Box}\delta C_{\alpha\beta\mu\nu} + B[h]_{\alpha\beta\mu\nu}{}^{\alpha'\beta'\mu'\nu'} \bar{C}_{\alpha'\beta'\mu'\nu'}$ plus terms depending only on $\bar{C}_{\alpha\beta\mu\nu}$, $T_{\mu\nu}$, the background metric $\bar{g}_{\mu\nu}$, and the background covariant derivative $\bar{\nabla}_\mu$.
- The linear in ϵ part of the Riemann curvature tensor $\delta R_{\alpha\beta\mu\nu}$, consists of $\delta C_{\alpha\beta\mu\nu}$ plus terms constructed from $T_{\mu\nu}$ and the background metric $\bar{g}_{\mu\nu}$.

So, if in the linear in ϵ part of the equation (3.3.10), we take all the terms that depend only on $T_{\mu\nu}$, the background metric $\bar{g}_{\mu\nu}$, and the background covariant derivative $\bar{\nabla}_\mu$, move them to the right hand side and collectively call them $S[T]_{\alpha\beta\mu\nu}$, we get

$$\begin{aligned}\bar{\Box}\delta C_{\alpha\beta\mu\nu} + B[h]_{\alpha\beta\mu\nu}{}^{\alpha'\beta'\mu'\nu'} \bar{C}_{\alpha'\beta'\mu'\nu'} + 2\bar{C}^{\gamma}{}_{\beta\nu}{}^{\lambda}\delta C_{\gamma\alpha\mu\lambda} + 2\bar{C}^{\gamma}{}_{\alpha\mu}{}^{\lambda}\delta C_{\gamma\beta\nu\lambda} \\ -2\bar{C}^{\gamma}{}_{\beta\mu}{}^{\lambda}\delta C_{\gamma\alpha\nu\lambda} - 2\bar{C}^{\gamma}{}_{\alpha\nu}{}^{\lambda}\delta C_{\gamma\beta\mu\lambda} + 2\bar{C}^{\gamma}{}_{\mu\nu}{}^{\lambda}\delta C_{\gamma\lambda\alpha\beta} + 2\bar{C}^{\gamma}{}_{\alpha\beta}{}^{\lambda}\delta C_{\gamma\lambda\mu\nu} \\ = S[T]_{\alpha\beta\mu\nu}.\end{aligned} \tag{3.3.12}$$

This does not seem like much progress. While it looks like a wave equation for $\delta C_{\alpha\beta\mu\nu}$ it mixes with the equally unknown metric perturbation $h_{\mu\nu}$ in a non-trivial way. To make this useful in someway, we are going to need a miracle.

3.4 The Weyl scalars

The Weyl curvature tensor with its many components seems a highly inefficient way of capturing 10 degrees of freedom. The Newman-Penrose formalism provides a more compact way of describing these degrees of freedom. It starts by choosing a tetrad of null vectors $(l^\mu, n^\mu, m^\mu, \bar{m}^\mu)$ with properties that:

- Each vector is null, i.e. $l^\alpha l_\alpha = n^\alpha n_\alpha = m^\alpha m_\alpha = \bar{m}^\alpha \bar{m}_\alpha = 0$
- $l^\alpha n_\alpha = -1$ and $m^\alpha \bar{m}_\alpha = 1$
- All other legs are orthogonal.
- m^μ and $\bar{m}^\mu$ are complex valued and $\bar{m}^\mu$ is the complex conjugate of m^μ.

These conditions imply that

$$g_{\mu\nu} = -2l_{(\mu}n_{\nu)} + 2m_{(\mu}\bar{m}_{\nu)}.$$

This tetrad can used to encode the 10 degrees of freedom of the Weyl curvature in 5 complex scalar fields

$$\psi_0 = C_{\alpha\beta\gamma\delta}\, \ell^\alpha m^\beta \ell^\gamma m^\delta, \tag{3.4.1a}$$
$$\psi_1 = C_{\alpha\beta\gamma\delta}\, \ell^\alpha m^\beta \ell^\gamma n^\delta, \tag{3.4.1b}$$
$$\psi_2 = C_{\alpha\beta\gamma\delta}\, \ell^\alpha m^\beta \bar{m}^\gamma n^\delta, \tag{3.4.1c}$$
$$\psi_3 = C_{\alpha\beta\gamma\delta}\, \ell^\alpha n^\beta \bar{m}^\gamma n^\delta, \tag{3.4.1d}$$
$$\psi_4 = C_{\alpha\beta\gamma\delta}\, n^\alpha \bar{m}^\beta n^\gamma \bar{m}^\delta, \tag{3.4.1e}$$

It is always possible to choose l^μ such that $\psi_0 = 0$. Such a null vector is called a **principal null vector** of the spacetime. Generically, a 4-dimensional spacetime is expected to have 4-linearly independent principal null vectors. If two principal null vectors happen to coincide then ψ_1 will vanish in addition to ψ_0, and l^μ is called a double principal null vector.

Kerr spacetime has two double principal null vectors[2]. If we choose both l^μ and n^μ to be a double principal null vector then we get that $\psi_0 = \psi_1 = \psi_3 = \psi_4 = 0$.

A popular choice for the tetrad in Kerr that achieves this is the **Kinnersley tetrad**

$$\ell^\alpha = \left(\frac{r^2+a^2}{\Delta}, 1, 0, \frac{a}{\Delta}\right), \tag{3.4.2}$$
$$n^\alpha = \left(\frac{r^2+a^2}{2\Sigma}, -\frac{\Delta}{2\Sigma}, 0, \frac{a}{2\Sigma}\right), \tag{3.4.3}$$
$$m^\alpha = \frac{1}{\sqrt{2}(r+ia\cos\theta)}\left(ia\sin\theta, 0, 1, \frac{i}{\sin\theta}\right) \qquad \text{and} \tag{3.4.4}$$
$$\bar{m}^\alpha = \frac{-1}{\sqrt{2}(r-ia\cos\theta)}\left(ia\sin\theta, 0, -1, \frac{i}{\sin\theta}\right). \tag{3.4.5}$$

When considering perturbed spacetimes, we need to choose a tetrad both in the background spacetime and in the perturbed spacetime. The requirement that the tetrad remains null orthonormal fixes both tetrads up to a local Lorentz transformation. This remaining freedom introduces an additional gauge freedom.

3.5 Teukolsky Equation

A miracle happens when we try to project equation (3.3.12) on to a tetrad basis adapted to the double principal null vectors of Kerr spacetime. When considering the ψ_0 component, first it happens that commuting the tetrad legs with the background box operator produces terms which exactly cancel the $h_{\mu\nu}$ dependent terms in the B tensor. Second, the dependence on all other components of $\delta C_{\alpha\beta\mu\nu}$ drops out, leaving a decoupled wave

[2] In the Petrov classification of spacetimes this means that Kerr is type D.

equation for ψ_0. A similar thing happens when trying to project on $(r - ia\cos\theta)^4\psi_4$. Introducing, the notation $\Phi_2 = \psi_0$ and $\Phi_{-2} = (r - ia\cos\theta)^4\psi_4$, using the Kinnersley tetrad the decoupled equations are given by[3],

$$\begin{aligned} -\left[\frac{(r^2+a^2)^2}{\Delta} - a^2\sin^2\theta\right]\frac{\partial^2\Phi_s}{\partial t^2} - \frac{4Mar}{\Delta}\frac{\partial^2\Phi_s}{\partial t\partial\varphi} - \left(\frac{a^2}{\Delta} - \frac{1}{\sin^2\theta}\right)\frac{\partial^2\Phi_s}{\partial\varphi^2} \\ +\Delta^{-s}\frac{\partial}{\partial r}\left(\Delta^{s+1}\frac{\partial\Phi_s}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Phi_s}{\partial\theta}\right) \\ +2s\left[\frac{a(r-M)}{\Delta} + \frac{i\cos\theta}{\sin^2\theta}\right]\frac{\partial\Phi_s}{\partial\varphi} + 2s\left[\frac{M(r^2-a^2)}{\Delta} - r - ia\cos\theta\right]\frac{\partial\Phi_s}{\partial t} \\ -(s^2\cot^2\theta - s)\Phi_s = S_s[T_{\mu\nu}], \end{aligned} \tag{3.5.1}$$

where $S_s[T_{\mu\nu}]$ is a second order differential operator that produces the projected source term for $T_{\mu\nu}$. This is the **Teukolsky equation**. (It turns out that this same equation with $s = \pm 1$ describes solutions to the Maxwell-equations on a curved background. With $s = 0$ it describes a massless scalar field, and with $s = 1/2$ a massless Dirac field.)

The perturbed Weyl scalars ψ_0 and ψ_4 have some useful properties in Kerr spacetimes.

First of all, they are gauge invariant. Since they are scalar quantities whose background values are zero, they are insensitive to normal gauge transformations. Moreover, infinitesimal rotations of the tetrad only mix the background values of ψ_0 and ψ_1 into the perturbed value of ψ_0 (or ψ_4 and ψ_3 into ψ_4). Since both background values vanish, infinitesimal rotations of the tetrad cannot change ψ_0 (or ψ_4).

Second, it turns out that ψ_0 (or ψ_4) contain almost all there is to know about a metric perturbation in Kerr spacetime, as expressed by the following theorem.

Theorem (Wald [8])**.** Suppose $h_{\mu\nu}$ and $h'_{\mu\nu}$ are (suitably well-behaved) solutions to the Einstein equation, such that the perturbation to the Weyl scalar ψ_0 (or ψ_4) is the same for both perturbations. Then

$$h_{\mu\nu} - h'_{\mu\nu} = \nabla_{(\mu}\xi_{\nu)} + c_M\frac{\partial g^{\text{Kerr}}}{\partial M} + c_a\frac{\partial g^{\text{Kerr}}}{\partial a}, \tag{3.5.2}$$

for some gauge vector ξ_μ and constants c_M and c_a.

So, ψ_0 "knows" everything about the vacuum metric perturbation except the gauge (since it is itself gauge invariant) and perturbations of the mass and spin of the Kerr background (since $\psi_0 = 0$ on any Kerr background.) Furthermore, we get,

Corollary.

$$\psi_0[h] = \psi_0[h'] \Leftrightarrow \psi_4[h] = \psi_4[h']. \tag{3.5.3}$$

[3]For the gory details of how to get here see [7].

3.5.1 Separation of Variables

Note that the left-hand side of the Teukolsky equation does not contain any explicit dependence on t and ϕ. We can thus do a Fourier transform with respect to t and ϕ to get

$$\Phi_s(t,r,\theta,\phi) = \int \mathrm{d}\omega \sum_m \hat{\Phi}_{sm\omega}(r,\theta) e^{i(m\phi-\omega t)}, \tag{3.5.4}$$

and

$$\begin{aligned} &\Delta^{-s}\frac{\partial}{\partial r}\left(\Delta^{s+1}\frac{\partial \hat{\Phi}_{sm\omega}}{\partial r}\right) \\ &+\Big(\frac{\big(\omega(r^2+a^2)-am-is(r-M)\big)^2+s^2(r-M)^2}{\Delta}+2am\omega+4is\omega r\Big)\hat{\Phi}_{sm\omega} \\ &+\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\hat{\Phi}_{sm\omega}}{\partial\theta}\right)-\Big(\frac{(m+s\cos\theta)^2}{\sin^2\theta}+a^2\omega^2\sin^2\theta+2sa\omega\cos\theta-s\Big)\hat{\Phi}_{sm\omega} \\ &= \hat{S}_{sm\omega}[T_{\mu\nu}]. \end{aligned} \tag{3.5.5}$$

The left-hand side of the equation separates into a part with only explicit dependence on r and a second part that depends on θ.

The polar operator appearing on the left-hand side,

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)-\frac{(m+s\cos\theta)^2}{\sin^2\theta}-a^2\omega^2\sin^2\theta-2sa\omega\cos\theta+s, \tag{3.5.6}$$

is a Sturm-Liouville operator. Consequently, it has real eigenvalues $(-\lambda_{slm\omega}-2ma\omega)$ with $l \geqslant l_0\max(|s|,|m|)$ such that $0 \leqslant \lambda_{sl_0m\omega} < \lambda_{s(l_0+1)m\omega} < \ldots$, and corresponding real-valued eigenfunctions $S_{slm\omega}(\theta)$ that satisfy

$$\int_{-1}^{1} S_{slm\omega}(\theta)S_{sl'm\omega}(\theta)\,\mathrm{d}\cos\theta = \delta_{ll'}, \tag{3.5.7}$$

and span the space of continuous functions of the interval $[0,\pi]$.

These functions are known as **spin-weighted spheroidal harmonics** with spin-weight s and spheroidicity $a\omega$. In the limit $a\omega \to 0$, these reduce to $\lambda_{slm\omega} = l(l+1)-s(s+1)$, and the more familiar spin-weighted spherical harmonics $Y_{slm}(\theta,\phi)$ (when paired with the $e^{im\phi}$ from the Fourier expansion).

We can therefore further expand $\hat{\Phi}_{sm\omega}(r,\theta)$ and source $\hat{S}_{sm\omega}[T_{\mu\nu}]$ in the basis of spin-weighted spheroidal harmonics

$$\hat{\Phi}_{sm\omega}(r,\theta) = \sum_{l\geqslant|s|} R_{slm\omega}(r)S_{slm\omega}(\theta), \tag{3.5.8}$$

and

$$\hat{S}_{sm\omega}[T_{\mu\nu}](r,\theta) = \sum_{l\geqslant|s|} \hat{S}_{slm\omega}[T_{\mu\nu}](r)S_{slm\omega}(\theta). \tag{3.5.9}$$

Plugging this into (3.5.5), we find that the radial coefficients satisfy

$$\Delta^{-s}\frac{\partial}{\partial r}\left(\Delta^{s+1}\frac{\partial R_{slm\omega}}{\partial r}\right) - V_{slm\omega}R_{slm\omega} = \hat{S}_{slm\omega}[T_{\mu\nu}], \tag{3.5.10}$$

with

$$V_{slm\omega} = -\frac{\left(\omega(r^2+a^2) - am - is(r-M)\right)^2 + s^2(r-M)^2}{\Delta} - 4is\omega r + \lambda_{slm\omega}. \tag{3.5.11}$$

This is the **radial Teukolsky equation**.

3.5.2 Homogeneous solutions, Asymptotic Behaviour, and Boundary conditions

Consider the homogeneous (i.e.$\hat{S}_{slm\omega} = 0$) radial Teukolsky equation

$$\Delta^{-s}\frac{\partial}{\partial r}\left(\Delta^{s+1}\frac{\partial R_{slm\omega}}{\partial r}\right) - V_{slm\omega}R_{slm\omega} = 0. \tag{3.5.12}$$

To study the asymptotic behaviour of solutions to the homogeneous radial Teukolsky equation near the horizon and near infinity, it is useful to make the following transformation

$$\tilde{R} = \Delta^{s/2}\sqrt{r^2+a^2}R \tag{3.5.13}$$

$$\frac{\mathrm{d}r^*}{\mathrm{d}r} = \frac{r^2+a^2}{\Delta}, \tag{3.5.14}$$

$$r^* = r + \frac{Mr_+}{\sqrt{M^2-a^2}}\log\frac{r-r_+}{2M} - \frac{Mr_-}{\sqrt{M^2-a^2}}\log\frac{r-r_-}{2M}, \tag{3.5.15}$$

where r^* is the Kerr tortoise coordinate. This turns the homogeneous radial equation into (suppressing the subscripts $slm\omega$ for brevity),

$$\frac{\mathrm{d}^2\tilde{R}}{\mathrm{d}r^{*2}} + U\tilde{R} = 0, \tag{3.5.16}$$

with

$$U = \frac{\left(\omega(r^2+a^2) - am - is(r-M)\right)^2 + s^2(r-M)^2 + \Delta(4is\omega r - \lambda)}{(r^2+a^2)^2} - G^2 - \frac{\mathrm{d}G}{\mathrm{d}r^*}, \tag{3.5.17}$$

and

$$G = \frac{s(r-M)}{r^2+a^2} + \frac{r\Delta}{(r^2+a^2)^2}. \tag{3.5.18}$$

For large r (and thus $r^* \to \infty$), we get that

$$U = \omega^2 + \frac{2is\omega}{r} + \mathcal{O}(r^{-2}) \tag{3.5.19}$$

Consequently, we get the asymptotic solutions near infinity

$$\tilde{R} \underset{r\to\infty}{\propto} \frac{1}{r^{\pm s}} \exp\{(\pm i\omega r^*)\}, \tag{3.5.20}$$

or translating back to our original radial field R

$$R \underset{r\to\infty}{\propto} \frac{1}{r^{2s+1}} \exp\{(i\omega r^*)\} \quad \text{or} \quad R \underset{r\to\infty}{\propto} \frac{1}{r} \exp\{(-i\omega r^*)\}. \tag{3.5.21}$$

Near the horizon as $r \to r_+$ (and $r^* \to -\infty$), we get

$$U = \left(k - is\frac{\sqrt{M^2-a^2}}{2Mr_+}\right)^2 + \mathcal{O}\,\Delta, \tag{3.5.22}$$

where

$$k = \omega - m\Omega_+ \tag{3.5.23}$$

is the frequency of the mode shifted by the rotation frequency of the outer horizon $\Omega_+ = \frac{a}{2Mr_+}$. Consequently, the asymptotic behaviour of the solutions near the horizon is

$$\tilde{R} \underset{r\to r_+}{\propto} \exp\left(\pm i\left(k - is\frac{\sqrt{M^2-a^2}}{2Mr_+}\right)r^*\right) \tag{3.5.24}$$

$$= \exp\left(\pm s\frac{\sqrt{M^2-a^2}}{2Mr_+}r^*\right)\exp\left(\pm ikr^*\right) \tag{3.5.25}$$

$$\approx \Delta^{\pm s/2}\exp\left(\pm ikr^*\right), \tag{3.5.26}$$

where we used that near the horizon

$$r^* \approx \frac{Mr_+}{\sqrt{M^2-a^2}}\log\frac{r-r_+}{2M} \approx \frac{Mr_+}{\sqrt{M^2-a^2}}\log\Delta. \tag{3.5.27}$$

For our original field R this translates to

$$R \underset{r\to r_+}{\propto} \exp\left(ikr^*\right) \quad \text{or} \quad R \underset{r\to r_+}{\propto} \Delta^{-s}\exp\left(-ikr^*\right). \tag{3.5.28}$$

So, we find that a general solution of the homogeneous radial Teukolsky equation $R_{slm\omega}$ has the following asymptotic behaviour (for $\omega \neq 0$)

$$R_{slm\omega} = \begin{cases} \dfrac{A^{\mathcal{I}^+}_{slm\omega}}{r^{2s+1}}\exp\{(i\omega r^*)\} + \dfrac{A^{\mathcal{I}^-}_{slm\omega}}{r}\exp\{(-i\omega r^*)\} & \text{as } r\to\infty \\ A^{\mathcal{H}^-}_{slm\omega}\exp\{(ikr^*)\} + A^{\mathcal{H}^+}_{slm\omega}\Delta^{-s}\exp\{(-ikr^*)\} & \text{as } r\to r_+ \end{cases} \tag{3.5.29}$$

The A coefficients are related to gravitational waves (GWs) entering or leaving the system:

- $A^{\mathcal{I}^+}_{slm\omega}$ is (related to) the amplitude of GWs in the solution $R_{slm\omega}$ traveling **out** to future null infinity $\mathcal{I}^+$.

- $A^{\mathcal{I}^-}_{slm\omega}$ is (related to) the amplitude of GWs in the solution $R_{slm\omega}$ coming **in** from past null infinity $\mathcal{I}^-$.

- $A^{\mathcal{H}^+}_{slm\omega}$ is (related to) the amplitude of GWs in the solution $R_{slm\omega}$ falling **down** the future event horizon $\mathcal{H}^+$.

- $A^{\mathcal{H}^-}_{slm\omega}$ is (related to) the amplitude of GWs in the solution $R_{slm\omega}$ coming **up** from the past event horizon $\mathcal{H}^-$.

The homogeneous radial Teukolsky equation is a second-order ordinary differential equation. Hence it has a two dimensional space of solutions. We can use our knowledge about the asymptotic behaviour of a general solution at infinity and the horizon to impose boundary conditions to identify the elements of a basis for this solution space. The most popular choice involves putting boundary conditions on past null infinity $\mathcal{I}^-$ and the past horizon $\mathcal{H}^-$. The two basis elements in this basis are:

- The **ingoing** solution $R^{\text{in}}_{slm\omega}$ is the homogeneous solution that only has waves coming in from past null infinity $\mathcal{I}^-$, but no waves coming up from the past horizon $\mathcal{H}^-$. It is defined by setting $A^{\mathcal{H}^-}_{slm\omega} = 0$ and normalized by $A^{\mathcal{H}^+}_{slm\omega} = 1$.

- The **upgoing** solution $R^{\text{up}}_{slm\omega}$ is the homogeneous solution that has waves coming up from the past event horizon $\mathcal{H}^-$, but no waves coming in from past null infinity $\mathcal{I}^-$. It is defined by setting $A^{\mathcal{I}^-}_{slm\omega} = 0$ and normalized by $A^{\mathcal{H}^+}_{slm\omega} = 1$.

Together these two span the solutions for the homogeneous radial Teukolsky equation for any real non-zero frequency ω.

In addition people sometimes use the "time-reversed" solutions for their basis:

- The **outgoing** solution $R^{\text{out}}_{slm\omega}$ defined by setting $A^{\mathcal{H}^+}_{slm\omega} = 0$ and normalized by $A^{\mathcal{H}^-}_{slm\omega} = 1$.

- The **downgoing** solution $R^{\text{down}}_{slm\omega}$ defined by setting $A^{\mathcal{I}^+}_{slm\omega} = 0$ and normalized by $A^{\mathcal{H}^-}_{slm\omega} = 1$.

Generically, any pair of $\left(R^{\text{in}}_{slm\omega}, R^{\text{out}}_{slm\omega}, R^{\text{up}}_{slm\omega}, R^{\text{down}}_{slm\omega}\right)$ can be used as a basis, and the other two can be written in terms of them, e.g.

$$R^{\text{down}}_{slm\omega} \propto R^{\text{in}}_{slm\omega} - R^{\text{up}}_{slm\omega}, \quad \text{and} \tag{3.5.30}$$

$$R^{\text{up}}_{slm\omega} \propto R^{\text{out}}_{slm\omega} - R^{\text{down}}_{slm\omega}. \tag{3.5.31}$$

3.5.3 Particular solutions

$$\Delta^{-s}\frac{\partial}{\partial r}\left(\Delta^{s+1}\frac{\partial R_{slm\omega}}{\partial r}\right) - V_{slm\omega}R_{slm\omega} = \hat{S}_{slm\omega}[T_{\mu\nu}], \tag{3.5.32}$$

To find particular solutions to the inhomogeneous Teukolsky equation we can follow a Green's function approach. If we first solve

$$\Delta^{-s}\frac{\partial}{\partial r}\left(\Delta^{s+1}\frac{\partial G_{slm\omega}(r,r_0)}{\partial r}\right) - V_{slm\omega}G_{slm\omega}(r,r_0) = \delta(r-r_0), \tag{3.5.33}$$

the particular solution for a general source is then given by

$$R_{slm\omega} = \int_{r_+}^{\infty} G(r,r_0)\hat{S}_{slm\omega}[T_{\mu\nu}]\,\mathrm{d}r_0. \tag{3.5.34}$$

To solve for the Green's function we first need to choose boundary conditions. Usually the physically relevant choice are the "retarded" boundary conditions, where for large r, the Green's function $G(r,r_0)$ satisfies upgoing boundary conditions (i.e. no waves from $\mathcal{I}^-$), and ingoing boundary conditions (no waves from $\mathcal{H}^-$) near the horizon. With these choices the Green's function is given by

$$G_{slm\omega}(r,r_0) = \begin{cases} \dfrac{R^{\mathrm{up}}_{slm\omega}(r_0)R^{\mathrm{in}}_{slm\omega}(r)}{\Delta(r_0)W[R^{\mathrm{in}}_{slm\omega},R^{\mathrm{up}}_{slm\omega}](r_0)} & r<r_0 \\ \dfrac{R^{\mathrm{in}}_{slm\omega}(r_0)R^{\mathrm{up}}_{slm\omega}(r)}{\Delta(r_0)W[R^{\mathrm{in}}_{slm\omega},R^{\mathrm{up}}_{slm\omega}](r_0)} & r>r_0, \end{cases} \tag{3.5.35}$$

where $W[R^{\mathrm{in}}_{slm\omega},R^{\mathrm{up}}_{slm\omega}](r_0)$ is the Wronskian

$$W[R^{\mathrm{in}}_{slm\omega},R^{\mathrm{up}}_{slm\omega}](r) = R^{\mathrm{in}}_{slm\omega}\frac{\mathrm{d}R^{\mathrm{up}}_{slm\omega}}{\mathrm{d}r} - \frac{\mathrm{d}R^{\mathrm{in}}_{slm\omega}}{\mathrm{d}r}R^{\mathrm{up}}_{slm\omega}. \tag{3.5.36}$$

Exercise 3.2: Wronskian

a. Show that

$$\Delta^{s+1}W[R^{\mathrm{in}}_{slm\omega},R^{\mathrm{up}}_{slm\omega}](r)$$

is constant.

b. Use this to show that

$$W[R^{\mathrm{in}}_{slm\omega},R^{\mathrm{up}}_{slm\omega}](r) = \frac{2i\omega A^{\mathrm{in},\mathcal{I}^-}_{slm\omega}A^{\mathrm{up},\mathcal{I}^+}_{slm\omega}}{\Delta^{s+1}} = \frac{2(2ikMr_+ + s\sqrt{M^2-a^2})\omega A^{\mathrm{up},\mathcal{H}^-}_{slm\omega}}{\Delta^{s+1}}.$$

3.6 Quasinormal Modes

In the last section, we discussed using $R^{\mathrm{in}}_{slm\omega}$ and $R^{\mathrm{up}}_{slm\omega}$ as a basis for all vacuum, and how they can be used to construct general particular solutions. However, what happens when

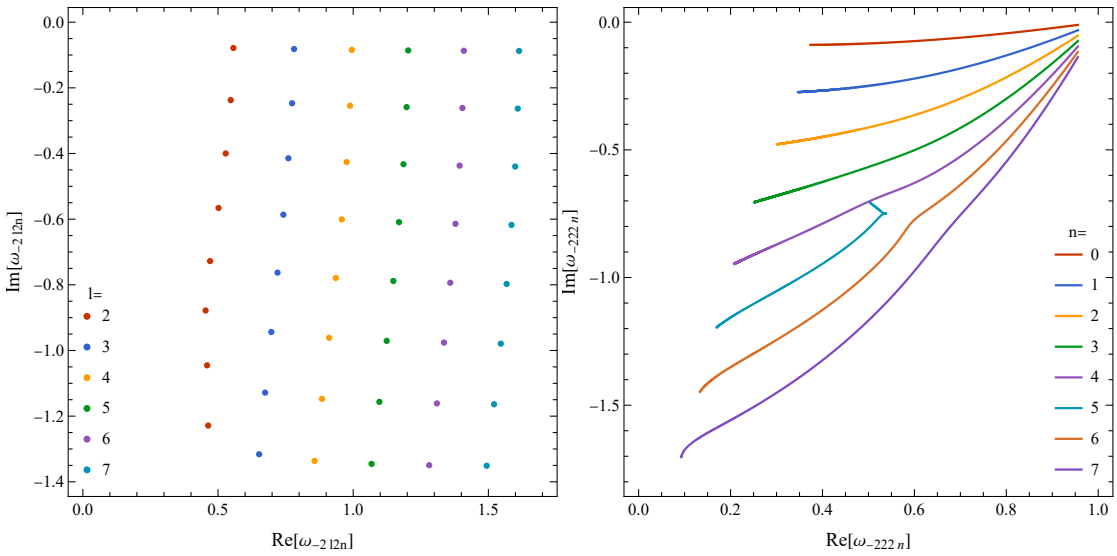


Figure 3.1. On the left the QNMs for a Kerr black hole with spin $a = 0.75M$ and $m = 2$. On the right the QNMs with $l = m = 2$ for Kerr black holes with spin between $a = 0$ and $a = 0.999M$. (Data from `https://pages.jh.edu/eberti2/ringdown/`.)

$R^{\text{in}}_{slm\omega} = R^{\text{up}}_{slm\omega}$? Can there be vacuum solutions that simultaneously satisfy upgoing and ingoing boundary conditions, i.e. where there are no incoming waves from either past null infinity or the past horizon? For generic ω the answer is no, but for special values of ω the answer is yes. These frequencies are known as the **Quasinormal modes (QNMs)** of a black hole. These QNMs act like a "finger print" of the black hole.

We start by citing an important result regarding a Kerr black hole's QNMs

Theorem (Kerr mode stability [9])**.** All quasinormal modes of a Kerr black hole have $\operatorname{im}\omega < 0$.

This theorem guarantees that the time dependence of any Kerr QNM ($\exp(-i\omega t)$) is decaying in nature, and there are no vacuum modes that grow out of control.

This also means that for each l and m there will be a QNM with the largest imaginary frequency. This will be the longest lived mode with those l and m, and is called the **fundamental** mode. The QNMs with a certain l and m are numbered by natural numbers n in order of descending imaginary part (i.e. $n = 0$ is the fundamental mode.) The modes with $n > 0$ are referred to as **overtones** (despite generally having a lower real part of the frequency.)

Lemma 5. If ω_{lmn} is a quasinormal mode, then so is $-\omega^*_{l(-m)n}$.

The intuitive picture of quasinormal modes is that they consist of waves travelling along the lightring of the Kerr black hole, decaying because of the unstable nature of this orbit. For small enough wave packets, i.e. large l, we indeed see that $\operatorname{re}\omega \approx l\Omega_{LR}$ and $\operatorname{im}\omega$ roughly corresponds to the Lyapunov exponent characterizing how unstable it is.

The QNM frequencies are determined by the parameters of the background black hole (i.e. the mass M and the spin a). The spectrum of QNMs thus forms a unique "fingerprint" for the black hole.

If we perturb a black hole (e.g. by dropping something into it) it will settle down exponentially to a member of the Kerr family, with the final stage being a linear combination of the quasi-normal modes. This process is called **ringdown**.

CHAPTER

4

Black holes in our Universe: how they form, how we find them

Black holes (BHs) are incredible predictions of general relativity (GR). From a theoretical perspective, they are unique objects to explore the limits of GR. What it is more impressive though is that over the past decades we have accumulated overwhelming evidence of their existence. Therefore, black holes are no longer only a theoretical playground to push the limit of our theories, but rather an actual laboratory where we can test different predictions. In this chapter we are going to have a **brief overview** of how black holes form in our Universe and the different evidence that we have about their existence. By its own nature, this will be a chapter that will cover very different physical phenomena. Therefore, instead of providing detailed derivations we will focus mostly on highlighting the key processes and their order of magnitude implications.

4.1 Gravitational Collapse

In simple terms, in order to form a black hole we need to accumulate enough energy-density in a small enough region so that the gravitational pull cannot be stopped by any other force. In the astrophysical context such dense regions essentially only occur in stars. There, the battle is between the pressure and the gravity of its interior. We are going to take a glimpse of how this looks like in realistic astrophysical set up in Sec. 4.2. For the moment we will focus on understanding some general principles.

We aim to study the gravitational collapse of a spherically symmetric star. Irrespectively of the material properties of the interior, Birkhoff's theorem guarantees us that outside of the star (i.e. in vacuum) the metric is described by the Schwarzschild metric:

$$ds^2 = -\left(1 - \frac{r_{\rm Sch}}{r}\right) dt^2 + \left(1 - \frac{r_{\rm Sch}}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \,, \tag{4.1.1}$$

where

$$r_{\rm Sch} = \frac{2GM}{c^2} \simeq 3 \left(\frac{M}{M_\odot}\right) {\rm km} \,, \tag{4.1.2}$$

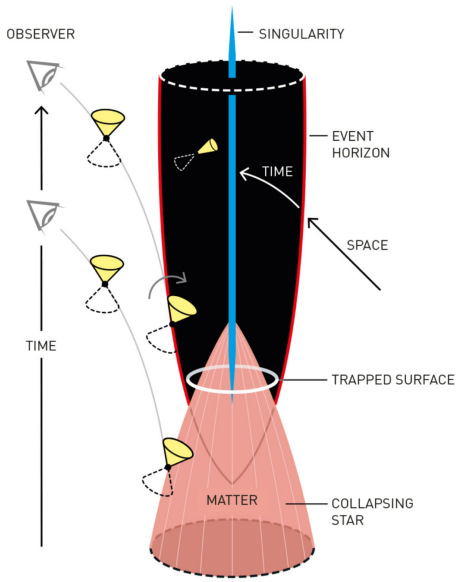


Figure 4.1. Schematic diagram of star collapsing into a black hole Credit Johan Jarnestad/The Royal Swedish Academy of Science.

with $M_\odot$ being the mass of the Sun. Since the interior and the exterior solutions must match, we can understand the behavior at the surface of the star following geodesics of Schwarzschild geometry.[1] For example, we can ask how the light emitted at the surface of the collapsing star would be seen by a distant observer. This light would be redshifted by

$$z = \frac{\Delta\lambda}{\lambda} \propto e^{t/2t_{\rm Sch}} \tag{4.1.3}$$

where the crossing time of the Schwarzschild radius is

$$t_{\rm Sch} = \frac{r_{\rm Sch}}{c} = \frac{2GM}{c^3} \simeq 10 \left(\frac{M}{M_\odot}\right) \mu{\rm s}\,. \tag{4.1.4}$$

The redshift increases exponentially with a characteristic time scale of $2t_{\rm Sch}$. The luminosity of the emitted light will also decay at this rate. Therefore, for an external observer, the light emitted by the star will dim at an incredible high rate for astrophysical standards. For a 1 solar mass star, the characteristic decaying time of the luminosity is only 20μs. A schematic representation of the collapse into a black hole is given in Fig. 4.1.

[1] Recall this was covered in detail in the GR course of block 2 and its lecture notes.

Exercise 4.1: The light from a collapsing star

Consider a comoving distant observer at a fixed spatial position around a collapsing star. Show that the radiation emitted at the surface of the star redshifts exponentially at late times. In other words, derive Eq. (4.1.3).

Literature:

The rest of this section follows Carroll's Chapter 5.8 [2] and Zee's VII.4 [10].

We now wish to study the interior of the star. We will consider the simplest possible composition for the interior: a perfect fluid. Its energy momentum tensor is described by

$$T^{\mu\nu} = (\rho + p)U^\mu U^\nu + p g^{\mu\nu} \,. \tag{4.1.5}$$

where ρ is the energy density and p the pressure of the fluid. U^μ is the local four velocity. We assume also a static, spherically symmetric, interior metric

$$ds^2 = -A(r)dt^2 + B(r)dr^2 + r^2 d\Omega^2 \,. \tag{4.1.6}$$

The fact that we are choosing a static solution implies that the fluid cannot flow, $U^i = 0$. Taking U^μ pointing in the timelike direction, the normalization condition $U_\mu U^\mu = -1$ fully fixes it to $U^0 = A^{-1/2} = -1/U_0$. Altogether explicitly:

$$U_\mu = (-A^{1/2}, 0, 0, 0) \,, \tag{4.1.7}$$

and, as a consequence,

$$T_{\mu\nu} = \mathrm{diag}(A \cdot \rho, B \cdot p, r^2 \cdot p, r^2 \sin^2\theta \cdot p) \,. \tag{4.1.8}$$

With the metric and the energy-momentum tensor one can solve the field equations

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu} \,. \tag{4.1.9}$$

Because of all our assumptions, only the diagonal terms are relevant. In fact, only one of the two angular equations is independent, $G_{\phi\phi} = \sin^2\theta G_{\theta\theta}$, and we are left with three couple equations to solve.[2] As we will see in a second, for us we will only be concerned about the temporal and radial components because we can always use the energy-momentum conservation instead of the angular Einstein field equation. Explicitly, the equations are (we denote primes as derivatives w.r.t the radius $dF/dr = F'$):

$$G_{tt} = \frac{A}{B^2 r^2}\left(B^2 + rB' - B\right) = 8\pi G T_{tt} = 8\pi G A\rho \,, \tag{4.1.10}$$

$$G_{rr} = \frac{1}{r^2}\left(1 - B + r\frac{A'}{A}\right) = 8\pi G T_{rr} = 8\pi G B p \,. \tag{4.1.11}$$

[2] You are welcome to derive the Einstein equations for this metric and energy-momentum tensor by hand! If interested in learning a Mathematica package designed for tensorial calculations I have left a notebook example in the course materials.

Note however that we are saying that there are 3 couple equations, but we have four functions to solve (two from the metric + two from the perfect fluid): $A(r)$, $B(r)$, $p(r)$, $\rho(r)$. The fourth equation arises from the equation of state, $p = p(\rho)$ that defines the properties of the fluid.

Inspired by Schwarzschild geometry, one can define a new radial-dependent mass function $\mathcal{M}(r)$

$$B(r)^{-1} \equiv 1 - \frac{2G\mathcal{M}(r)}{r} \,. \tag{4.1.12}$$

Inserting this definition in the tt equation, we find

$$\frac{d\mathcal{M}(r)}{dr} = 4\pi r^2 \rho(r) \,, \tag{4.1.13}$$

which can be integrate to find

$$\mathcal{M}(r) = 4\pi \int_0^r dr' r'^2 \rho(r') \,. \tag{4.1.14}$$

For a star of radius R, the matching conditions with the exterior metric imply that

$$M = \mathcal{M}(R) = 4\pi \int_0^R dr r^2 \rho(r) \,, \tag{4.1.15}$$

so the mass function $\mathcal{M}(r)$ that we introduced can be interpreted as the mass within a radius r. It is interesting to point out however, that this does not match with the integration of the energy-density ρ over a spatial volume element, $\gamma_{ij} dx^i dx^j$, specifically

$$\tilde{M} = \int_{r<R} \rho(r) \sqrt{\gamma} d^3x = 4\pi \int_0^R \rho(r) B^{1/2}(r) r^2 dr = 4\pi \int_0^R \frac{\rho(r) r^2}{\sqrt{1 - \frac{2G\mathcal{M}(r)}{r}}} dr \,. \tag{4.1.16}$$

The difference between the two is due to the gravitational binding energy arising from the gravitational attraction of the fluid within the star:

$$E_B = \tilde{M} - M > 0 \,. \tag{4.1.17}$$

Note that in GR, the notion of gravitational binding energy is different than from Newtonian gravity.

Back to the field equations, using the radial equation and the energy-momentum conservation $\nabla_\mu T^{\mu\nu} = 0$ (instead of the $\phi\phi$-equation), which for a perfect fluid and a static spherically symmetric metric takes a very simple form. This equation can be derived using the identity for the covariant derivative

$$\nabla_\mu T^{\mu\nu} = \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g} T^{\mu\nu}) + \Gamma^\nu_{\mu\lambda} T^{\mu\lambda} \,, \tag{4.1.18}$$

and recalling the metric compatibility $\nabla_\mu g_{\alpha\beta} = 0$. With this and the perfect fluid's four-velocity properties ($U^i = 0$), one finds

$$\frac{dp}{dr} = -(\rho + p) \frac{d \ln A^{1/2}}{dr} \,. \tag{4.1.19}$$

One can eliminate the $A(r)$ metric component to arrive at the **Tolman–Oppenheimer–Volkoff (TOV) equation**

$$\frac{dp}{dr} = -\frac{G\mathcal{M}(r)\rho(r)}{r^2}\left(1+\frac{p(r)}{\rho(r)}\right)\left(1+\frac{4\pi r^3 p(r)}{\mathcal{M}(r)}\right)\left(1-\frac{2G\mathcal{M}(r)}{r}\right)^{-1} \tag{4.1.20}$$

describing hydrostatic equilibrium. For a given equation of state (EoS) $p = p(\rho)$, the TOV equation together with the interior mass equation (4.1.14) can be used to solve $p(r)$ and $\rho(r)$. These set of couple equations generically require numerical integration except for simplified scenarios as we will check next. Before that, plugging the TOV equation (4.1.20) back to the energy-momentum conservation (4.1.19), we can fully solve for the interior metric

$$\frac{d\ln A^{1/2}}{dr} = \frac{G\mathcal{M}(r)}{r^2}\left(1+\frac{4\pi r^3 p(r)}{\mathcal{M}(r)}\right)\left(1-\frac{2G\mathcal{M}(r)}{r}\right)^{-1} . \tag{4.1.21}$$

We can check that the matching conditions with the exterior geometry are also correct, since at the surface $\mathcal{M}(R) = M$ and $p(R) = 0$ and the above equation can be easily integrated to obtain $A = 1 - \frac{r_{\rm Sch}}{r}$. Finally, it is interesting to point out that in the non-relativistic limit the TOV equation reduces to Newton's equation for stellar interiors:

$$\frac{dp}{dr} = -\frac{G\mathcal{M}(r)\rho(r)}{r^2} . \tag{4.1.22}$$

This equation just corresponds to the balance between the gravitational force and the radial pressure.

4.1.1 A star of constant density

Now that have an equation for the interior of the star, we wish to solve it in order to understand the conditions for collapse. We choose a simple model that allows us to inegrate the TOV equation analytically, that is an **incompressible fluid** of constant density

$$\rho(r < R) = \rho_* , \qquad \rho(r > R) = 0 . \tag{4.1.23}$$

Then, the interior mass is simply

$$\mathcal{M}(r) = \frac{4\pi r^3}{3}\rho_* , \tag{4.1.24}$$

with the density defined by the mass and radius of the star $\rho_* = M/(4\pi R^3/3)$. The TOV equation can be solve to find

$$p(r) = \rho_* \frac{\left(1-\frac{r_{\rm Sch}}{R}\right)^{1/2} - \left(1-\frac{r^2 r_{\rm Sch}}{R^3}\right)^{1/2}}{\left(1-\frac{r^2 r_{\rm Sch}}{R^3}\right)^{1/2} - 3\left(1-\frac{r_{\rm Sch}}{R}\right)^{1/2}} , \tag{4.1.25}$$

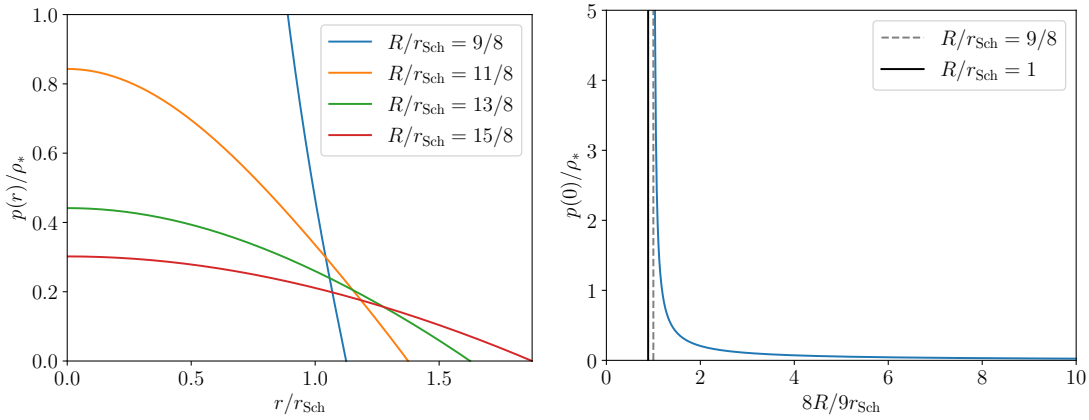


Figure 4.2. On the left, pressure profile in the radial component for a spherically star of radius R made of an incompressible perfect fluid of density ρ_*. On the right, pressure at the center as a function of the star's radius.

whose pressure profile is plotted in the left panel of Fig. 4.2. The pressure increases as one approaches the center of the star $r \to 0$. Moreover, the central pressure $p(0)$ increases as the star becomes more compact. This is plotted in the right panel of Fig. 4.2. In fact, when reaching $R = 9r_{\rm Sch}/8$ the central pressure diverges. This highlights that there is a limit to how compact a constant density star can be in order to have a static solution. In other words, there is a maximum mass that a spherically symmetric star can have in order to have a static solution because for more massive stars it will require a pressure larger than infinity to sustain it! The bound on the mass is then:

$$M < \frac{4}{9}\frac{Rc^2}{G} \simeq 3\left(\frac{R}{10\text{km}}\right) M_\odot \,. \tag{4.1.26}$$

If a shrinking star reaches this critical radius, then it will continue shrinking and eventually collapse into a black hole when $R < r_{\rm Sch}$. Although this important results was derived for a rather specific EoS, Buchdahl's theorem [11] states that this result holds for any "reasonable" EoS. We will see next that real stars in our universe are in fact less compact, satisfying the Buchdahl's theorem. Note, however, that the compactness limit is almost (9/8) the one of a black hole already

4.1.2 A universe within a star

Before moving on to gravitational collapse in realistic stars, we are going to adventure in a small digression. This is motivated by the remarkable similarities between the (simplified) set up that we have considered for the interior of a star and the geometry of the Universe at large scales as described by a (flat) Friedmann-Robertson-Walker (FRW) metric:

$$ds^2 = -dt^2 + a(t)^2 d\vec{x}^2 \,, \tag{4.1.27}$$

where $a(t)$ is the scale factor. In fact, in order to study the background expansion of the Universe one also models matter as a perfect fluid. However, instead of assuming a static, spherically symmetric metric as in the interior of the star, one assumes a homogeneous and isotropic time varying metric.

Let us consider the case of a universe filled with a cosmological constant $\rho = \Lambda$ with EoS $p = -\Lambda$. The field equations then reduce to $R_{\mu\nu} = 8\pi G\Lambda g_{\mu\nu}$ (noting that $R = 32\pi G\Lambda$). A cosmological constant directly leads to a mass profile of

$$\mathcal{M}(r) = \frac{4\pi r^3}{3}\Lambda\,. \tag{4.1.28}$$

The interior metric can be obtained solving Eq. (4.1.21), which after inputting our ansatzs looks like

$$\frac{d\ln A}{dr} = \frac{-2H^2 r}{1-H^2r^2}\,, \tag{4.1.29}$$

where we have defined $H^2 = 8\pi G\Lambda/3$ in resemblance of the Hubble parameter $H = d\ln a/dt$. This equation can be solved to obtain $A(r) = 1 - H^2r^2$. From Eq. (4.1.12) one then obtains $B = 1/A$. Altogether, the spherically symmetric static metric that solves for a cosmological constant is

$$ds^2 = -(1-H^2r^2)dt^2 + (1-H^2r^2)^{-1}dr^2 + r^2d\Omega^2\,. \tag{4.1.30}$$

In a cosmological setup of a universe filled with a cosmological constant solving the Friedman equation

$$\left(\frac{d\ln a}{dt}\right)^2 = \frac{8\pi G}{3}\Lambda = H^2 \tag{4.1.31}$$

leads to a exponentially increasing scale factor $a \propto e^{Ht}$. Back to the FRW metric this is

$$ds^2 = -dt^2 + e^{2Ht}d\vec{x}^2\,, \tag{4.1.32}$$

which is nothing but a de Sitter space-time. Therefore, the interior of a star is mathematically equivalent to an expanding universe.

Exercise 4.2: A ball of dust

Let's consider an even simpler model (perhaps the simplest) for the interior of a star: a spherically symmetric pressureless fluid of constant density, i.e. a "ball of dust". This is also known as the Oppenheimer-Snyder model [12]. Because there are no pressure forces, the dust particles in the surface of the star follow radial geodesics. Following what you have learned so far, study the gravitational collapse of this model. Is it possible to have a static solution? To what cosmology is this model equivalent to?

Explore: black hole formation in the early universe

Although astrophysically we typically only think of stars as dense enough environments to trigger gravitational collapse into a black hole, cosmologically there are other options. In particular, black holes could form directly from the collapse of curvature fluctuations ζ in the early Universe. The fraction of black holes forming β is typically postulated in terms of the probability of having a fluctuation beyond a given threshold

$$\beta = \int_{\zeta_{\rm thr}}^{\infty} p(\zeta) d\zeta \,, \tag{4.1.33}$$

where $p(\zeta)$ is the probability density function. The mass of the black hole that forms is proportional to the size of the causal horizon that is collapsing. Large curvature fluctuations can be produced during inflation. Such fluctuations may collapse upon reentry. Their mass is therefore associated to the epoch in which the fluctuations are generated counted as the number of e-folds N:

$$M_{\rm pbh} \sim 4\pi\gamma \frac{M_{\rm pl}}{H_{\rm inf}} e^{2N} \,, \tag{4.1.34}$$

where $M_{\rm pl}$ is the reduced Planck mass, $H_{\rm inf}$ is the energy scale of inflation, and γ is efficiency parameter encapsulating the details of the gravitational collapse, typically $\gamma \sim 0.2$. Black holes formed in the early Universe are typically referred as **primordial black holes (PBHs)**. The range of masses of PBHs is therefore subject to the details of the concrete inflationary model. There are also other mechanisms different from inflation that could produce PBHs. A famous one is cosmic strings.

4.2 Stellar graveyard: white dwarfs, neutron stars and black holes

In the last section we have studied a generic process of gravitational collapse in a simplified toy model. Now we wish to explore a bit what are the actual stellar remnants that can form in our Universe and what is their interplay with the other fundamental forces apart form gravity. A detailed derivation of such processes goes beyond the scope of this course and here we only aim to draw the main arguments and relevant scales.

Different to other celestial bodies such as planets whose structure is supported by material pressures, stars are sustained by thermonuclear power. Starting from Hydrogen, the core of the stars converts lighter elements into heavier ones by fusion emitting heat that supports its structure. How far the fusion leads the core to climb up the periodic table is very sensitive to the star's mass, rotation, magnetic field and chemical composition. In any case, when the nuclear fuel is used (at most until iron since its binding energy per nucleon decreases, meaning that heavier elements cannot release the extra energy needed to compensate the additional gravity due to the contraction), the

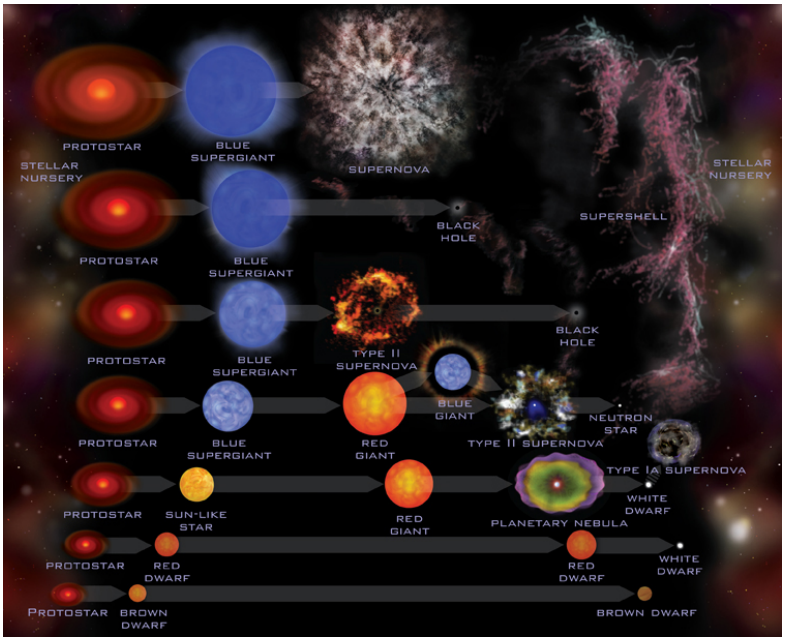


Figure 4.3. Schematic diagram of the different outcomes of stellar evolution. Image credit Chandra.

star compresses and there are two options: **i)** the star reaches equilibrium again due to non-thermal forces or **ii)** the star continues gravitational collapse. The most relevant non-thermal pressure is associated to the Pauli exclusion principle that prevents fermions to occupy the same (quantum) state, also known as Fermi pressure.

When a star is supported by the Fermi pressure of electrons then it is a **white dwarf (WD)**. When a star is supported by the Fermi pressure of neutrons it is a **neutron star (NS)**. Beyond Fermi pressure and nuclear forces there are no other source of energy to prevent the gravitational collapse. Therefore, once one passes the support of NS one end up with a BH. This is why when we think about the remnants of stars, referred sometimes as the **stellar graveyard**, these are WDs, NSs and BHs. A schematic diagram of the different routes of stellar evolution is presented in Fig. 4.3. Something that we did mention though is that through this process of collapse, there could be instabilities triggered that will end up in runaway processes leading to explosion. These are of course the **supernovas**.

Explore: Boson stars

As we have seen the Fermi pressure is a basic ingredient to balance the gravitational pull and form stable stars when the object is composed of fermions. Matter in the standard model of particle physics is made of fermions, so this should cover it all. However, there could be additional fundamental particles of bosonic type.

How then a stable configuration could be form? The key is self-interaction and the simplest example is a complex scalar field [13].

4.2.1 Maximum mass of white dwarfs

Literature:

This section follows Hartle's chapters 12.1 and 24 [14].

We now wish to investigate a bit further the stability of WDs. As explained above, these are stars that are supported by the Fermi pressure of the electrons. The gravity, on the other hand, is driven by the mass of the nuclei ($m_e \ll m_p, m_n$). The balance between these two forces determine the stability of a WD for different configurations. In particular, we are interested in finding what is the maximum mass for which the electron Fermi pressure can no longer counterbalance the gravity, i.e. the maximum mass of a WD.

To get a taste of the problem we are first going to do a rough estimate. We assume a spherically symmetric star of radius R made of A electrons. To ensure neutrality we take A protons. Since electrons cannot occupy the same state, we can think of their wavelength λ as associated to their number density within star, $n_e \sim A/R^3$. This corresponds to a scaling of $\lambda \sim R/A^{1/3}$. The more compact the star is, the larger the frequency and therefore energy and momentum:

$$p \sim \hbar/\lambda \sim A^{1/3}\hbar/R\,. \tag{4.2.1}$$

Assuming that the electrons have been compressed enough to be relativistic, $E = \sqrt{p^2c^2 + m_e^2c^4} \approx pc$, the total Fermi energy is roughly

$$E_F \sim A(pc) \sim A^{4/3}\hbar c/R\,. \tag{4.2.2}$$

This energy is compensated by the gravitational interaction driven by the protons

$$E_G \sim -G(m_pA)^2/R\,. \tag{4.2.3}$$

The total energy is $E_T = E_F + E_G$ and scales as $1/R$. For a sufficient number of electrons A within a radius R then the gravitational energy wins. This determines a critical number of electrons

$$A_{\text{crit}} \sim \left(\frac{\hbar c}{Gm_p^2}\right)^{3/2} \sim 10^{57} \tag{4.2.4}$$

(recall $2\pi\hbar = 6.62 \cdot 10^{-34}$Js, $G = 6.67 \cdot 10^{-11}\text{Nm}^2\text{kg}^2$, $c = 3 \cdot 10^8$m/s and $m_p = 1.7 \cdot 10^{-27}$kg) which in turns defines a critical mass

$$M_{\text{crit}} \sim m_pA_{\text{crit}\sim} \sim M_\odot\,. \tag{4.2.5}$$

(recall $M_\odot = 2 \cdot 10^{30}\text{kg}$) . Therefore, WDs cannot be much heavier than the Sun.

A more detailed analysis follows from solving the TOV equation (4.1.20). Modeling the star with a given EoS, one can find the maximum mass that can be supported. This is precisely what Chandrasekhar discovered [15]. Therefore, the maximum WD mass is known as the **Chandrasekhar limit**.

In order to solve this problem we are only missing to define the EoS of the WD. For a completely degenerate electron gas (see details in [14]) this can be found in the nonrelativistic to be

$$p = \frac{1}{5}(3\pi^2)^{3/2} \left(\frac{\hbar^2}{m_e}\right) n_e^{5/3} , \tag{4.2.6}$$

while in the relativistic one

$$p = \frac{1}{4}(3\pi^2)^{1/3} (\hbar c)\, n_e^{4/3} , \tag{4.2.7}$$

where, again, n_e is the number density of electrons. This is sometimes written in terms of the number of protons Z and nucleons A, to give the energy density ρ

$$n_e = \frac{Z}{A}\frac{\rho}{m_H c^2} , \tag{4.2.8}$$

where m_H is the mass of the hydrogen. Putting together the last two equations one then gets an EoS: $p = p(\rho)$. EoS are sometimes parametrized in terms

$$\gamma \equiv \frac{\rho + p}{p}\frac{dp}{d\rho} , \tag{4.2.9}$$

a dimensionless quantity measuring the stiffness of an EoS. For degenerate fermions in the nonrelativistic limit $\gamma = 5/3$, smoothly transitioning towards $\gamma = 4/3$ as they become relativistic. Larger γ means more increase in pressure with increases in the energy density.

For a given central density ρ_c, the TOV equations can be integrated from the center $p(\rho_c)$ to the radius of the star $p(R) = 0$. Exploring all possible central densities ρ_c, from 0 to ∞, allows to determine the family of spherical stars made of matter with the assumed EoS. The final product is a set of allowed masses $M(\rho_c)$ and radii $R(\rho_c)$ for this type of stars. An example of the solutions of the TOV equations is given in Fig. 4.4. Both the pressure and density steadily decrease from their central value to 0 at the sruface of the star, R. The mass does the opposite: increasing from 0 to M.

For WDs it turns out the maximum mass is $\sim 1.4 M_\odot$ with a radius of 1000 km for densities of about 10^{11}g/cm^3. Roughly, WDs have the mass of the Sun within the size of the Earth. This is a very compact object, but still mostly in the the Newtonian regime,

$$\frac{GM_{\text{wd}}}{c^2 R_{\text{wd}}} \sim \text{few} \times 10^{-3} \tag{4.2.10}$$

For higher densities to 10^{11}g/cm^3, there are other relevant physical phenomena that we study next.

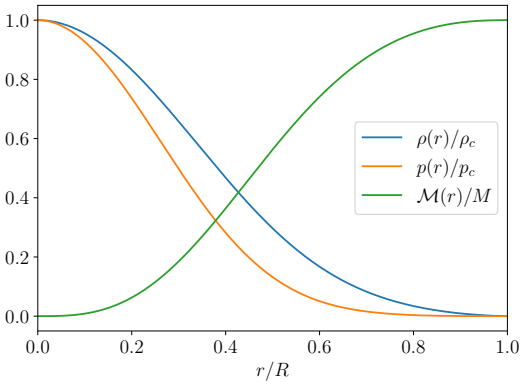


Figure 4.4. The stellar structure for a star with a equation of state with stiffness parameter $\gamma = 5/3$ obtained solving the Tolman–Oppenheimer–Volkoff (TOV) equations (4.1.13) and (4.1.20). M and R are the mass and radius of the star as obtained from solving the TOV equations.

Exercise 4.3: Finding hydrostatic equilibrium

Solve the TOV equations (numerically) to find the maximum mass of stars with different equations of state $p = p_c(\rho/\rho_c)^\gamma$. To do so it is convenient to rewrite the TOV equations in terms of dimensionless variables $\tilde{\rho} = \rho/\rho_c$, $\tilde{p} = p/p_c$, $\tilde{r} = r/R$ and $\tilde{\mathcal{M}} = \mathcal{M}/M$. Note that both M and R need to be obtained from the solutions themselves, but giving an initial guess serves to simplify the numerical implementation.

4.2.2 Maximum mass of neutron stars

We are now going to study the case of stars more dense than WDs. Recall that for WDs we are dealing with a star with free electrons and nuclei (this roughly occurs at $\rho \gtrsim 10^4 \text{g/cm}^3$). As the star becomes more compact, electrons become relativistic ($m_e c^2 \simeq 0.5\text{MeV}$) at about $\rho \sim 10^6 \text{g/cm}^3$. Soon after it becomes energetically favorable to convert electrons and protons into neutrons releasing neutrinos:

$$e^- + p^+ \to n + \nu_e \,. \tag{4.2.11}$$

This occurs at energies $\sim m_n c^2 - m_p c^2 \simeq 1.3\text{MeV}$. As the density increases, the star becomes more and more neutron rich. At around $\rho \sim 4 \cdot 10^{11} \text{g/cm}^3$ the most energetic neutrons become unbound from nuclei. Eventually, all neutrons are free forming a fluid sometimes refered as neutron matter. This is the material of which NSs are made of.

Studying the maximum mass of NSs could be done in a similar fashion to what we have done for WDs. Neutrons are also fermions and therefore are equally subject to the Fermi pressure which mass be compensated by the gravitational pull. Differently though, the gravitational energy is also provided by the neutrons themselves. The main difficulty however of studying NSs is the fact that one could reach densities in which

nuclear interactions become relevant. This is a very active field of research, trying to understand how the matter within NSs behaves.

The first attempts to set this maximum mass were done by Tolman [16] and Oppenheimer–Volkoff [17] in what is now known as the **Tolman–Oppenheimer–Volkoff limit** in analogy to the Chandrasekhar limit. (Much) Later studies including nuclear interactions have set the maximum mass at around $2M_\odot$ for non-rotating NSs and about $3M_\odot$ for rapidly rotating ones [18]. A typical NS has a mass of $1.4M_\odot$ for a radius of about 10km. Therefore

$$\frac{GM_{\rm ns}}{c^2 R_{\rm ns}} \sim \text{few} \times 10^{-1} \tag{4.2.12}$$

which is 100 times the case of WDs. For NSs, relativistic effects at the surface can become important. For scale, a NS has about the mass of the Sun within the size of the urban area of Copenhagen.

Explore: Stellar-origin black hole maximum mass and the pair instability supernovae

We have just seen that there are fundamental processes in nature that prevent white dwarfs and neutron stars to form more massive than a given value. Is there a similar limit for black holes?
In fact, there is. If the core of the star is massive enough it reach energies in which it is possible to produce electron and positrons. This pair production is sourced by high energy gamma rays that no longer help supporting the star against gravity, reducing temporarily the pressure. This photon pressure reduction can makes the star compress. This can lead to a runaway process in which the higher energy of the photons leads to more pair production. Eventually, if this instability continues the star ends up in a supernova in what is known as **pair instability supernova (PISN)** [19]. The explosion ejects so much material that a massive remnant black hole cannot be formed. This process is subject to many details about stellar evolution and the composition of the star, but it is thought that it would prevent stellar-origin black holes to form above $\sim 50M_\odot$. Interestingly, for very heavy stars (which are expected to be rare on the other hand), this instability is insufficient to prevent the collapse of the star into a black hole. Therefore, it is possible that heavier black holes above $\sim 120M_\odot$ could form. This leads to the interesting observational prospect that there is a gap in the mass spectrum of stellar-origin black holes. This is sometimes called the PISN or upper mass gap. This theory is actively being probed with gravitational wave observations.
As an aside, one should note that this only refers to the direct formation of black holes from stars. Black holes of larger masses can form due to accretion or black hole mergers as we will see later. Moreover, in the context of primordial black holes it is also possible to directly collapse curvature perturbations into more massive black holes.

Explore: Is there a black hole minimum mass?

From stellar evolution (recall Fig. 4.3) we have seen that black holes only form for stars more compact that a NS. Does this mean that astrophysically black holes smaller than $1-3M_\odot$ cannot be formed? Well... this depends on the type of matter that the star is made of, but for "standard" configurations (i.e. material within the standard model of particle physics) this is the case.
But still, what is the minimum mass of black hole forming from a star. 2, 3, 4 solar masses? This is a complicated question that requires simulating accurately the collapse of a star. Observationally, there is an interesting fact. With X-ray binaries in the Milky Way (we will talk more about this in 4.3.3) there have been observations of black holes with only $> 5M_\odot$ [20], while neutron star masses are $< 3M_\odot$ as discussed before. This is sometimes referred as the **neutrons star - black hole mass gap** or **lower mass gap** when referring to all the compact binaries collectively. Interestingly, with gravitational waves there has been recently the observation of a compact object with mass between $2.5-4.5M_\odot$ [21], squarely within this purported gap. More observation will tell if there is a new population down there!
On a different note, as argued before, in the early universe it is "easy" to form black holes of any mass. Therefore, detecting sub-solar mass black holes have been though as a smoking gun for primordial black holes. If they are too small, they would however evaporate by emission of Hawking radiation.

4.3 Evidence for black holes

Since any light that falls into a black hole cannot escape, by its own nature black holes are difficult to observe. However, evidence for the existence of black holes can be obtained in several indirect ways. Essentially, the whole game is being able to weight the object and demonstrate that the mass enclosed in such volume can only be explained by a black hole. In this section we wish to make a rapid tour around these different pieces of evidence. This is not meant to be exhaustive or self-contained, but rather an invitation to explore more and a validation of the fact that nowadays we are pretty sure black holes exist in Nature.

4.3.1 Astrometry

Astrometry is the branch of astronomy dedicated to precisely measuring the position and velocities of stars. Within our galaxy, where most precise measurements are possible, observing the motion of stars has turn out to be a prolific way of finding black holes of different masses. Perhaps most well known is the study of "S stars" around Sagittarius A* (SgrA*), the super massive object at the galactic center of the Milky Way.

Before diving into the center of our galaxy, let us refresh some scales. A typical galaxy weights around $10^{12}M_\odot$. It is composed (essentially) of dark matter and stars,

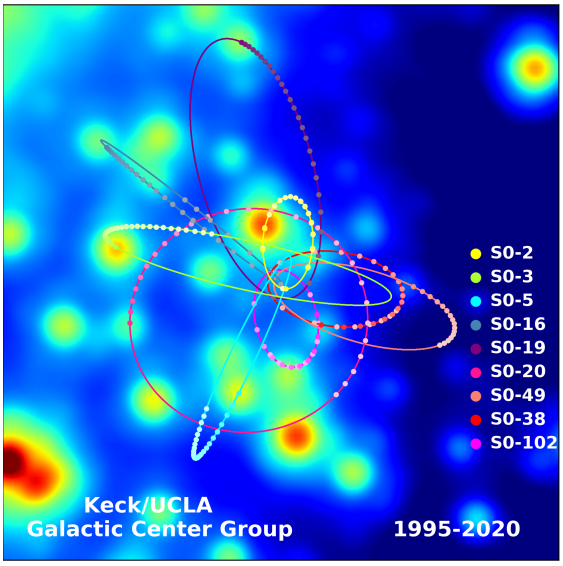


Figure 4.5. The orbits of stars within the central 1.0 × 1.0 arcseconds of our Galaxy. In the background, the central portion of a diffraction-limited image taken in 2015 is displayed. While every star in this image has been seen to move over the past 20 years, estimates of orbital parameters are best constrained for stars that have been observed through at least one turning point of their orbit. The annual average positions for these stars are plotted as colored dots, which have increasing color saturation with time. Also plotted are the best fitting simultaneous orbital solutions. Credit: plot and caption from UCLA Galactic Center Group.

which form a bulge of about $10^{10} M_\odot$. Typical sizes for the bulge (which is the part that we can observe directly!) are around a few kiloparsec (recall 1pc $\simeq 3 \cdot 10^{16}$m). It is thought that typical galaxies host a super massive black hole (SMBH) of between 10^6 and $10^9 M_\odot$. The Schwarzschild radius of such object can be as large as $\sim 10^9$km, but this is only $\sim 10^{-4}$pc. Therefore the size of the SMBH in the center of the galaxy is much smaller than the size of the galaxy. This makes very challenging to resolve such small scales for distant galaxies.

For our own Galaxy, different observatories have been tracking the trajectories of the closest stars to SgrA* for decades, cf. Fig. 4.5. These trajectories are described by Newtonian mechanics so that the velocity at different positions serves to bound the total mass within the orbit. Current observations determine $M_{\mathrm{SgrA}^*} \sim 4 \cdot 10^6 M_\odot$ [22]. These amazing observations were recognized with the Nobel prize in 2020 to Andrea Ghez and Reinhard Genzel as leaders of the competing teams.

Explore: Gaia black holes

Gaia, a European space mission, aims to map a billion stars in our Galaxy. Its precision in positioning stars is unprecedented, making it the perfect instrument for astrometry. In particular, by monitoring the trajectories of so many stars it

is well positioned to find the rare cases in which a star forms a binary with a dormant black hole. These systems are very hard to detect otherwise do to the lack of additional electromagnetic emission by the black hole. There has been so far three black holes discovered in Gaia data [23–25].

4.3.2 Gravitational lensing

GR predicts that both the propagation of massive and massless particles are affected by the gravitational interaction with other objects. This means that in the same way that celestial bodies modify their trajectories according to the curvature of space-time, electromagnetic radiation can be deflected by massive objects which act as **gravitational lenses**. In fact, the deflection of light observed during the solar eclipse of 1919 was an impactful experimental test of GR. In the weak filed limit, the deflection angle produced by a compact lens is given by[3]

$$\Delta\alpha \sim \frac{4GM_L}{bc^2} = 2\frac{r_{\rm Sch}}{b}\,, \tag{4.3.1}$$

where M_L is the mass of the lens and b the impact parameter. For example, for our Sun which has $R_\odot \simeq 7\cdot 10^5$km, a light ray passing by its surface will be deflected by $\sim 10^{-5}$ radians or ~ 2 arcsec.

Such a point lens will also have the effect of magnifying any foreground source. The angular scale at which lensing becomes relevant is

$$\theta_E = \sqrt{\frac{4GM_L}{c^2}\frac{D_{LS}}{D_L D_S}}\,, \tag{4.3.2}$$

where D_L, D_S and D_{LS} are respectively the (angular diameter) distances from the observer to the lens, to the source and between the lens and the source. For short wavelengths compared to the size of the lens, a point lens always produces two images ($\pm$) and their positions and magnifications can be solved analytically, in particular

$$\mu_\pm = \frac{b^2+2}{2b\sqrt{b^2+4}} \pm \frac{1}{2}\,. \tag{4.3.3}$$

When the two images cannot be resolved, an observer is only sensitive to the total flux

$$\mu_{\rm tot} = |\mu_+| + |\mu_-| = \frac{b^2+2}{b\sqrt{b^2+4}}\,. \tag{4.3.4}$$

As $b \to 0$, the magnification diverge. This means that if we are a distant source and a lens passes by then the flux of source is expected to increased significantly over a period of time. It is to be noted that the magnification does not diverge in reality, it is just a consequence of our approximations of small wavelengths and points sources. Those are

[3]Recall this was described in detail in the GR course of block 2 and its lecture notes.

eventually broken in reality due to either the finite size of the wavelength or the finite size of the source, whatever is larger. Still, the magnifications can be very large for a point lens and a point source

$$\mu_{\rm max} \sim \pi t_{\rm Sch}\omega \simeq 4\cdot 10^5 \left(\frac{M_L}{M_\odot}\right)\left(\frac{f}{\rm GHz}\right) , \tag{4.3.5}$$

where f is the frequency of the wave. When the source has a size of $R_{\rm src}$, the maximum magnification is

$$\mu_{\rm max} = \left(1+4\frac{R_E}{R_{\rm src}}\right)^{1/2} , \tag{4.3.6}$$

where $R_E = D_L\theta_E$ is the Einstein radius of the lens.

Since it is typically difficult to know the intrinsic luminosity of a given source, the total magnification is not such a informative parameter. However, the time variability of the magnification is more robust, as this depends more directly on the source-lens configuration and not the knowledge of the source properties. In particular, to know the relevant time scale we want to compute how long the source will be within the Einstein radius of the lens. If we denote v the relative velocity between the lens and the source, this time scale is simply

$$t_\mu = \frac{R_E}{v} \simeq 0.2{\rm yr}\left(\frac{200{\rm km/s}}{v}\right)\left(\frac{M_L}{M_\odot}\right)^{1/2}\left(\frac{D_L}{10{\rm kpc}}\right)^{1/2}\left(\frac{D_{LS}}{D_S}\right) . \tag{4.3.7}$$

Therefore, for solar-mass compact lenses with typical velocities of hundreds of kilometers per second within our galaxy the duration of the magnification is of the order of a fraction of a year. Larger lenses are therefore more challenging to constrain as one needs to monitor the source for longer time. In essence this was the idea of Paczynski, who proposed this as a method to map stellar mass compact objects in our Galaxy [26]. Different surveys have pursued this search such as MACHO, EROS and most recently OGLE. The best candidate to date for an isolated black hole using this lensing method is OGLE-2011-BLG-0462.

Exercise 4.4: Gravitational telescopes

Due to the effect of gravitational lensing, structures in the universe can become "gravitational telescopes" magnifying distant objects. Black hole are among the most efficient lenses. They are also the simplest as they can be described by point lenses. For a universe filled of black holes of the same mass, compute the probability that a source is amplified with a magnification larger than 10.

4.3.3 Electromagnetic emission

Due to the large gravitational attraction around a black hole, any surrounding material that is being accreted can experience significant accelerations, heating up and emitting electromagnetic (EM) radiation. The key to identify such source of radiation as a black

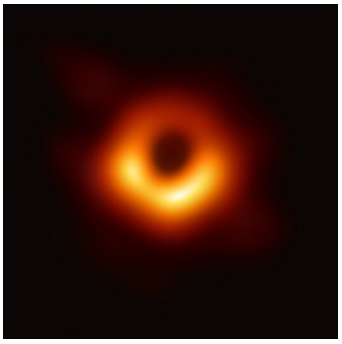


Figure 4.6. Image of the supermassive black hole at the center of our neighbor galaxy, Messier 87, captured by the Event Horizon Telescope [28].

hole is to demonstrate that no other object could source such amount of power in such a small region.

It has been known since the 1960s that there are extremely bright quasi-stellar objects at large distances [27]. These quasi-stellar objects, also known as **quasars**, are particularly surprising because of their small size (this is why initially they were confused with stars) and high luminosity. A Quasar is nowadays understood as very luminous active galactic nucleus (AGN). Quasar luminosities can be 10^4 times brighter than the luminosity of all stars in a galaxy. Therefore, in the most distant cases, only the quasar itself and not the host galaxy is observable. The energy powering these powerful quasars is thought to come from the gravitational binding energy liberated during the accretion as well as the EM extraction of the rotational energy of the BH.

At a much smaller scale, a BH that form in a binary star system can also lead to significant EM emission. Note that 2/3 of all stars are members of binary systems! If one of them undergoes a supernova to end up in a BH, then trajectory of the companion star will be affected and new effects will arise. In particular there could be a periodic lensing of the star. However, more importantly, the BH can accrete the star emitting large amounts of energy. Binary systems discovered in this way are known as **X-ray binaries** because of the observed frequencies. Many X-ray binaries have been identified in the galaxy. However, not all of them contain a BH since, as we have seen before, WDs and NSs are also natural stellar remanents. Therefore, the key for identifying a BH within a X-ray binary is to be able to measure its mass and, in particular, to prove that it is larger than $\gtrsim 3M_\odot$. This can be (partially) done thanks to the Doppler shift of the radial velocity which contains information about the masses in the binary. With additional information about the mass of the star, which can be estimated from its spectrum if it is a common star, then the mass of the potential BH can be inferred.

Explore: Imaging a black holes

Although a direct image of the black hole interior is not possible, one can hope to construct a telescope with enough resolution to see in detail the surroundings of a black hole down to the photon ring, the last stable photon orbit. This is precisely what the Event Horizon Telescope (EHT) is set to do through very-long-baseline interferometry. And, of course, they did it! We have now pictures of black holes, in particular the one in our own Galaxy and in a close neighbor, M87. The image of the later is presented in Fig. 4.6 where one can see the expected ring like structure. The actual size of the inner ring and its relation to the photon ring is an active area of research since its detectability with current telescopes is subject to the details of the accretion disk around the black hole. In any case, this observation already serves to measure the mass of the black hole in a completely independent way. This measurement is most interesting for the black hole in our Galaxy as we have precise astrometric measurements to compare with.

4.3.4 Compact binary coalescence

When two compact objects encounter with each other and collide, they perturb the space-time around them producing gravitational waves (GWs) that travel across the cosmos. These waves encode information about the masses of the binary and its shape is a firm prediction of GR.

The frequency of these waves is correlated with the orbital motion of the binary. As the black holes approach each other they orbit in a faster way. It is easy to guess that the amplitude of the wave will be larger when the black holes are about to merge (and therefore also easier to detect!). For a binary of similar masses, an order of magnitude of the frequency of a GW around merger time can be obtained by considering that is should inversely proportional to the Schwarzschild crossing time of both hole. This rough approximation shows that

$$f \sim \frac{1}{2\pi}\frac{1}{2t_{\rm Sch}} \sim 800{\rm Hz}\left(\frac{10M_\odot}{M}\right) . \tag{4.3.8}$$

Therefore, black holes of tens of solar masses will merge with frequencies of hundreds of Hz. The associated wavelength is much larger than any electromagnetic transient! Similarly, if we expect that spacetime close to merger leads to a distortion of order one, an order of magnitude estimate of the amplitude h leads to

$$h \sim \mathcal{O}(1)\frac{r_{\rm Sch}}{r} \sim 10^{-23}\left(\frac{1{\rm Gpc}}{r}\right)\left(\frac{M}{10M_\odot}\right) . \tag{4.3.9}$$

In words, black holes of tens of solar masses merging at cosmological distances will have very small amplitudes in our detectors.

The waveform of the first GW detected is presented in Fig. 4.7. On the top panel one can see the time domain waveform with its characteristic inspiral, merger and ring-down

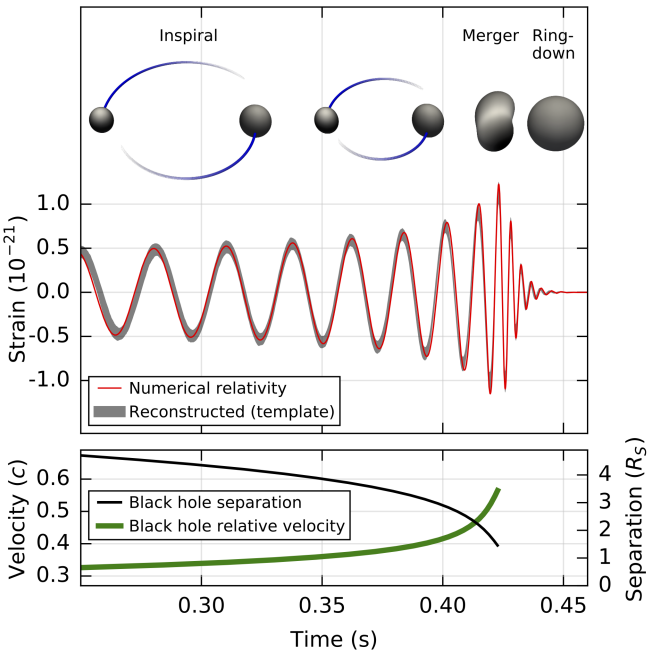


Figure 4.7. GW150914, the first detection of a binary black hole merger [29]. The top panel shows the strain as a function of time. The lower panel shows the relative velocity and separation of the objects as they merge. The velocity is a fraction of the speed of light, c, and the separation just a few Schwarzschild radius, R_S. For more details about the first GW detection, see the science summary.

phases. This was a binary of $\sim 30M_\odot$ merging at ~ 400Mpc. The merger frequency was about 250Hz. On the bottom panel one can see the orbital separation and the velocity. As the black holes are about to merge they are moving almost at the speed of light!

These binary coalescences are fascinating and their study will be the focus of the rest of the course.

CHAPTER

5

When black holes meet each other

The goal of this chapter is to study the gravitational wave (GW) emission produced during the coalescence of two black holes. These GWs are intrinsically small and, consequently, we need to start by reviewing some basic concepts about general relativity (GR) in the limit of small perturbations, i.e. **linearized gravity**. The starting point is to perturb slightly our metric $g_{\mu\nu}$ around a given background. We will call this perturbation $h_{\mu\nu}$ and it must satisfy $|h| \ll 1$. For simplicity we start perturbing around a flat background $\eta_{\mu\nu}$.[1] This has the advantage that there is a well defined separation between the (fixed) background metric and the perturbations. Then, applying this decomposition

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{5.0.1}$$

to the Einstein field equations one arrives at the following equation describing the evolution of h:

$$\Box h_{\mu\nu} = -16\pi G \left(T_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} T \right) , \tag{5.0.2}$$

where $\Box$ is the D'Alembertian operator in flat spacetime: $\Box = \partial^\mu \partial_\mu = -c^{-2}\partial^2/\partial t^2 + \partial^2/\partial \vec{x}^2$. Famously this describes the propagation of GWs.

The linearized Einstein equations admit plane wave solutions. To show that, we focus on the vacuum solution $\Box h_{\mu\nu} = 0$ and make the following ansatz[2]

$$h_{\mu\nu}(x) = \mathrm{Re}\left[A_{\mu\nu}(x) e^{i\theta(x)} \right] . \tag{5.0.3}$$

We further define the wave-vector $k_\mu \equiv \partial_\mu \theta$ and the scalar amplitude A and the polarization vector $\epsilon_{\mu\nu}$ so that $A_{\mu\nu} \equiv A\epsilon_{\mu\nu}$. The polarization tensor is normalized so that $\epsilon^{\mu\nu}\epsilon^*_{\mu\nu} = 1$ and $A = \sqrt{A^*_{\mu\nu} A^{\mu\nu}}$. We will assume that the phase θ varies rapidly compared

[1] We will relax this assumption in 6.1 where we will consider perturbations around a curved background. Note that this was also covered in Chapter 3.

[2] In the GR course of block 2, the initial ansatz was taken to be $h_{\mu\nu}(x) = \mathrm{Re}\left[A_{\mu\nu} e^{ik_\alpha x^\alpha} \right]$, with $A_{\mu\nu}$ and k_α being constant.

to the amplitude A. This can be formally implemented by introducing an expansion parameter ε and substituting $\theta \to \theta/\varepsilon$.[3] Anyhow, at leading order (ε^{-2}) from the wave equations one finds

$$\eta_{\mu\nu} k^\mu k^\nu = 0\,. \tag{5.0.4}$$

This equation also implies that $\nabla_\nu(k_\mu k^\mu) = 2k^\mu \nabla_\nu k_\mu = 0$, which can be rewritten as

$$k^\mu \nabla_\mu k_\nu = 0\,, \tag{5.0.5}$$

using that the derivatives on θ commute (because it is a scalar). In words, GWs travel along null geodesics and propagate at the speed of light. At next to leading order (ε^{-1}) one has from the wave equation

$$2k^\alpha \nabla_\alpha A_{\mu\nu} + A_{\mu\nu} \nabla^\alpha k_\alpha = 0\,, \tag{5.0.6}$$

which can be rewritten in terms of the scalar amplitude and the polarization tensor. Introducing the scalar amplitude is handy because one can realize that

$$\begin{aligned} k^\mu \nabla_\mu (A^2) &= 2A k^\mu \nabla_\mu A \\ &= k^\mu \nabla_\mu \left(A^*_{\alpha\beta} A^{\alpha\beta} \right) = A^2 \nabla^\mu k_\mu\,, \end{aligned} \tag{5.0.7}$$

where the first equality of each line is just (trivially) rewriting the formula and in the last equality of the second line we have use (5.0.6). With this result then (5.0.6) implies

$$\nabla^\mu (A^2 k_\mu) = 0\,, \tag{5.0.8}$$

$$k^\alpha \nabla_\alpha \epsilon_{\mu\nu} = 0\,. \tag{5.0.9}$$

The first equation indicates that the number of gravitons is conserved (in analogy to having a conserved current $\nabla^\mu j_\mu = 0$), while the second equation implies that the polarization tensor is parallel transported. Note that at leading order, (ε^{-1}), the Lorenz gauge condition $\nabla^\mu h_{\mu\nu} = 0$ implies

$$k^\alpha \epsilon_{\alpha\mu} = 0\,, \tag{5.0.10}$$

i.e. the polarization is orthogonal to the rays. The next orders, ε^0 and higher, would solve for the actual propagation. From our initial ansatz one gets

$$\Box A_{\mu\nu} = 0\,, \tag{5.0.11}$$

$$\nabla^\alpha A_{\alpha\mu} = 0\,, \tag{5.0.12}$$

for the wave equation and the gauge condition. Note that we could have also added higher order terms in our initial ansatz for the amplitude, i.e. $h_{\mu\nu} = \left(A_{\mu\nu} + \varepsilon A^{(1)}_{\mu\nu} + \cdots \right) e^{i\theta/\varepsilon}$, and those would modify these last two equations. Moreover, one could have taken an

[3] For further context, this is the eikonal or shortwave expansion that will appear in full glory when solving the GW propagation on a curved background.

alternative, more restrictive setup, in which the amplitude tensor $A_{\mu\nu}$ and the wavevector k_μ are constant. Then, there is no need to distinguish between the variation of the amplitude and the phase because, by definition, only the phase varies. In this case the plane wave solution is $h_{\mu\nu}(x) = A_{\mu\nu}e^{ik_\alpha x^\alpha}$. This is the typical choice in textbooks when introducing GWs for the first time.

Exercise 5.1: Gravitational wave propagation in geometric optics

Using the plane wave ansatz (5.0.3), explicitly derive that GWs propagate in null geodesics, the number of gravitons is preserved and that the polarization tensor is orthogonal to the propagation direction and parallel transported.

The final piece of information about the vacuum propagation of GWs is the number of independent polarizations. Inherited form the metric perturbation, the polarization tensor is symmetric in its indices, $\epsilon_{\mu\nu} = \epsilon_{\nu\mu}$. In four dimensions this means that at most it could have 10 independent components. Taking into account the Lorenz condition (6.1.10), we are left with at most 6 independent components. However, not all these components are independent because there is a residual gauge freedom. In vacuum one can fix the remaining gauge degrees of freedom so that the metric perturbation is purely spatial and traceless, i.e. $\epsilon_{0\mu} = 0$ and $\epsilon^\mu_\mu = 0$ (note that for the traceless part we need vacuum). Altogether, in the **transverse-traceless (TT) gauge** there are only two polarization degrees of freedom:

$$\epsilon_{\mu\nu}(x) = \epsilon_+(x)\hat{\epsilon}^+_{\mu\nu} + \epsilon_\times(x)\hat{\epsilon}^\times_{\mu\nu} \tag{5.0.13}$$

If ones fixes the propagation in the $\hat{z}$ direction then, explicelty,

$$\epsilon_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \epsilon_+ & \epsilon_\times & 0 \\ 0 & \epsilon_\times & -\epsilon_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} . \tag{5.0.14}$$

The TT-gauge is specially convenient because it fully fixes all the gauge freedom and relates nicely with the linearized Riemann tensor

$$R_{i0j0} = -\frac{1}{2}\ddot{h}^{TT}_{ij} . \tag{5.0.15}$$

In vacuum, all the components of the Riemann tensor can be determined from R_{i0j0}. The $+, \times$ polarizations then correspond to the physical propagating degrees of freedom.

Exercise 5.2: Non-radiative degrees of freedom

On a global vacuum spacetime we have seen that the equations of motion only contain two physical propagating degrees of freedom. These are **radiative** modes. On a global spacetime with matter sources, however, this is not generically true.

There could be physical, **non-radiative** degrees of freedom. These modes are characterized by non-wave equations. Schematically, instead of $g_{\alpha\beta}\nabla^\alpha\nabla^\beta\phi_A = \cdots$ one has $\gamma_{ij}\nabla^i\nabla^j\phi_A = \cdots$, where γ_{ij} is the spatial metric. Derive such equations! If you are looking for inspiration, this is nicely discussed in Flanagan and Hughes review [30] as well as Carroll's book [2].

Explore: Massless spin-2 fields

We have seen that in the linearized limit, GR predicts the propagation of two physical degrees of freedom traveling at the speed of light. In the field theory language, interactions are mediated by fundamental particles. Fundamental particles can be either bosons (related to forces) or fermions (related to matter) and they are described (at least) by their mass and spin. Bosons are represented by an integer spin, $s = 0, 1, 2, 3, \cdots$, and fermions by half integers $s = 1/2, 3/2, \cdots$. The particle associated to the gravitational force is normally referred as the graviton. The fact that we have seen that gravity is mediated at the speed of light implies a massless particle. This is also refered as a lon-range forcer. But, what should be the spin of the graviton? Well, this would be a longer digression, but there is beautiful proof by Weinberg [31] showing that for a local, unitary, Lorentz invariant theory only a massless spin-2 field is consistent with the equivalence principle. In short, scalars ($s = 0$) do not couple to photons (since they couple through the trace of the energy-momentum tensor ϕT) and vectors ($s = 1$) have charges of different signs. Massless $s > 2$ particles do not have a conserved tensor (except for total derivatives) with three or more indices $T^{\mu\nu\alpha\cdots}$.

5.1 Gravitational wave generation

The vacuum GW propagation is a good representation of the behavior of the waves at a distance point from their source, in the **far zone**. We now wish to connect this solution with the GW generation in the **near zone**. Using the trace-reversed perturbations

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}g_{\mu\nu}h\,, \tag{5.1.1}$$

which are equivalent to $h_{\mu\nu}$ in the TT-gauge, we can rewrite the field equations as

$$\Box\bar{h}_{\mu\nu} = -16\pi G T_{\mu\nu} + \mathcal{O}(h^2) \equiv -16\pi G \tau_{\mu\nu}\,, \tag{5.1.2}$$

where in the first equality we have explicitly written the higher order terms in the perturbative expansion which can be thought as additional effective sources. The last equation is just a definition of an **effective stress-energy tensor** which satisfies $\partial^\mu\tau_{\mu\nu} = 0$ in the Lorenz gauge, $\partial^\mu\bar{h}_{\mu\nu} = 0$. Note that these field equations are **exact** since they contain all higher order terms. The above equation can be solved directly using the

Green's function, since its solution is well known for a flat-space d'Alembertian operator $\Box$.[4] This is

$$\bar{h}_{\mu\nu}(t,\vec{x}) = 4G\int d^3x' \frac{\tau_{\mu\nu}(t-|\vec{x}-\vec{x}'|,\vec{x}')}{|\vec{x}-\vec{x}'|}\,. \tag{5.1.3}$$

Causality imposes that the solution depends on the **retarded time** $t-|\vec{x}-\vec{x}'|/c$ so that the information takes some time (c is finite!) form point $\vec{x}$ to $\vec{x}'$.

5.1.1 Newtonian limit

We will now compute the metric perturbations generated by a source in the weak-gravity regime and whose internal motion is slow compared to the speed of light. We study the solutions in the far and near zones in order to match the GW signal to the source.

Far zone solution

For a source of size R and a GW with wavelength λ, the far zone is defined as those positions at distance r in which

$$R \ll \lambda \ll r\,. \tag{5.1.4}$$

We are interested in computing the GWs at a point $\vec{x}$ and time t, taking into account the source mass distribution by integrating over all possible positions $\vec{x}'$ within the source of size R. We will define the normal vector $\hat{n}$ and the distance to the point of interest r so that $\vec{x} = r\hat{n}$. Since the distance r is much larger than the size of the source R, we could expand

$$|\vec{x}-\vec{x}'| = r - \vec{x}'\hat{n} + \mathcal{O}\left(\frac{R^2}{r}\right) = r\left(1 - \frac{\vec{x}'\vec{x}}{r} + \mathcal{O}\left(\frac{R^2}{r^2}\right)\right) \tag{5.1.5}$$

Therefore, in the far-field we can approximate at leading order $|\vec{x}-\vec{x}'| \simeq r$. Because the retardation within the source size R will be small, we can also approximate $t-|\vec{x}-\vec{x}'|/c \simeq t-r/c$. Moreover, we have already seen that in the far zone only the spatial components of the metric are relevant, that is, we will only focus on

$$\bar{h}^{ij}(t,\vec{x}) = \frac{4G}{r}\int d^3x' \frac{\tau^{ij}(t-r,\vec{x}')}{|\vec{x}-\vec{x}'|}\,. \tag{5.1.6}$$

The conservation of $\tau_{\mu\nu}$ implies the following identity (you are encouraged to prove this!)

$$\partial_0^2(x^ix^j\tau^{00}) = \partial_k\partial_l(x^ix^j\tau^{kl}) - 2\partial_k(x^i\tau^{jk}+x^j\tau^{ki}) + 2\tau^{ij}\,, \tag{5.1.7}$$

which can be used to obtain

$$\bar{h}^{ij}(t,\vec{x}) = \frac{2G}{c^4 r}\frac{d^2M^{ij}(t-r/c)}{dt^2}\,, \tag{5.1.8}$$

[4] Explicitly, the Green's function $G(x,s)$ of the flat space d'Alembertian $\Box = -c^{-2}\partial_t^2 + \partial_{\vec{x}}^2$ is $G(t,\vec{x}) = -\delta(t-|\vec{x}-\vec{x}'|/c)/4\pi|\vec{x}-\vec{x}'|$ which is obtained from $\Box G(t,\vec{x}) = \delta(\vec{x}-\vec{x}')$.

if we define the quadrupole mass moment

$$M^{ij}(t) = \int d^3x' \tau^{00} x'^i x'^j = \int d^3x' \rho(t, \vec{x}') x'^i x'^j + \mathcal{O}(h^2)\,, \tag{5.1.9}$$

where ρ is the energy-density. Note that upon integration the first two terms in the right hand side of (5.1.7) become boundary terms that can be made arbitrarily small by expanding the integration volume. Equation (5.1.8) is the famous **quadrupole formula** showing that GWs are generated by the quadrupole moment of accelerated energy densities. As before, the $\mathcal{O}(h^2)$ terms account for the gravitational field within the source, i.e. the self-gravity of the system.

The final solution is obtained projecting into the TT-gauge using the transverse projector operator

$$P_{ij} = \delta_{ij} - \hat{n}_i \hat{n}_j\,, \tag{5.1.10}$$

where $\hat{n}_i = x^i/r$ is the unit vector in the propagation direction. This means

$$h^{TT}_{ij}(t, \vec{x}) = \frac{2G}{c^4 r} \frac{d^2 M^{TT}_{ij}(t - r/c)}{dt}\,, \tag{5.1.11}$$

with

$$M^{TT}_{ij} = P_{ik} M^{kl} P_{lj} - \frac{1}{2} P_{ij} P_{kl} M^{kl}\,. \tag{5.1.12}$$

Near zone solution

The next step is to match this far-field solution with the corresponding near zone around the source. In the slow velocity approximation this zone is defined by

$$R \ll r \ll \lambda\,. \tag{5.1.13}$$

This is because the velocity of the source is $v \sim \omega_s R$, where ω_s is the frequency of the source. The radiation frequency will also be of order of magnitude of the source's frequency, $\omega \sim \omega_s$ (we will study the concrete example of an inspiraling binary in §5.2). Since $\lambda = c/(2\pi\omega) \sim (c/v)R$, we obtain that the slow-motion implies $\lambda \ll R$. In other words, in the non-relativistic limit, $v \ll c$, the wavelength of the radiation is much larger than the source itself.

In the Newtonian limit the metric perturbation is fixed by the Newtonian potential[5]

$$\Phi = -\frac{1}{2} h^{00} = -\frac{1}{4}(\bar{h}^{00} + \delta_{ij} \bar{h}^{ij})\,. \tag{5.1.14}$$

As a consequence, within the near zone

$$\Phi(t, x) = -G \int d^3x' \frac{\tau^{00}(t, \vec{x}') + \delta_{ij} \tau^{ij}(t, \vec{x}')}{|\vec{x} - \vec{x}'|}\,, \tag{5.1.15}$$

[5]Recall that in this limit the metric looks like $ds^2 = -(1 + 2\Phi)dt^2 + d\vec{x}^2$. This corresponds to the slow-motion limit of the weak-field limit in which $ds^2 = -(1 + 2\Phi)dt^2 + (1 - 2\Phi)d\vec{x}^2$ (the spatial term is suppressed by $1/c^2$). For a spherically symmetric source, one could also derive this result directly from the Schwarzschild metric in the limit of $2GM/r \ll 1$.

ignoring the retardation effects in the slow motion limit. The term $\delta_{ij}\tau^{ij}$ accounts for the internal stresses. Ignoring them for the moment, as they are small compared to the τ^{00} contributions by $1/c^2$, we can expand the potential in powers of $1/r$ to obtain. To do so first let us note that $|\vec{x}-\vec{x}'| \simeq r - \vec{x}\vec{x}'/r$ with the second term being a small expansion parameter and, as a consequence, the denominator is expanded as

$$\frac{1}{|\vec{x}-\vec{x}'|} = \frac{1}{r} + \frac{x^i x'_i}{r^3} + \frac{(x^i x'_i)(x^j x'_j)}{r^5} + \cdots . \tag{5.1.16}$$

Note that the x^i terms in the numerator and the r^n in the denominator can be taken out of the integral. By defining the different multipole moments, in particular, the monopole, dipole and quadrupole:

$$M \equiv \int d^3x \tau^{00}(x) , \tag{5.1.17}$$

$$D^i \equiv \int d^3x \tau^{00}(x) x^i , \tag{5.1.18}$$

$$Q^{ij} \equiv \int d^3x \tau^{00}(x) \left(x^i x^j - \frac{1}{3} r^2 \delta^{ij} \right) , \tag{5.1.19}$$

we can rewrite the previous expression for the Newtonian potential to obtain

$$\Phi(t,\vec{x}) = -G \left(\frac{M}{r} + \frac{D_i x^i}{r^3} + \frac{3}{2} \frac{Q_{ij} x^i x^j}{r^5} + \cdots \right) . \tag{5.1.20}$$

Note that higher order multipoles would contribute to h_{ij} with higher time derivatives (just using the same trick of the conservation of the stress-tensor to replace $x^i \cdots x^{i_n} \tau^{00}$ with $\partial_t^n \tau^{00}/c^n$ that we used in (5.1.7)) and, therefore, are suppressed in the slow-motion expansion by increasing powers of v/c.

However not even all these low multipoles can contribute to the radiative degrees of freedom. Conservation of energy momentum prevents any contribution from the monopole which accounts from the total mass-energy. Similarly, the dipole cannot contribute because in nearly Newtonian gravity we can always choose a reference frame in the center of mass in which $D_i = 0$. Considering the internal motion would lead to the spin dipole at leading order. However, angular momentum conservation also prevents this term to contribute. Therefore, we are left with the conclusion that only the quadrupole moment contributes at leading order. Because Q_{ij} and I_{ij} only differ by a trace, we can link the near and far zones in the TT-gauge:

$$h_{ij}^{TT}(t,\vec{x}) = \frac{2G}{c^4 r} \frac{d^2 Q_{ij}^{TT}(t - r/c)}{dt^2} , \tag{5.1.21}$$

Again, this is the acclaimed **quadrupole formula**.

5.1.2 Post-Newtonian expansion

We have just seen that in the Newtonian limit the GW radiation is generated by the quadrupole moments of the energy density. This is however a crude approximation as

the compact objects approach to each other. We thus want a systematic approach to compute corrections to this limit. We are interested in including corrections from the slow velocity expansion v/c. Note that for a self-gravitating system of mass M the virial theorem suggests that $(v/c)^2 \sim r_{\rm Sch}/R$ where R is the typical size of the system. We will this make an expansion in a slowly moving, weakly self-gravitating system. This is known as the post-Newtonian (PN) expansion. The PN expansion breaks down as one approaches to the strong gravity regime where $(v/c)^2 \sim r_{\rm Sch}/R \sim 1$.

The whole business of the PN formalism is to make an expansion in a small parameter

$$\epsilon \sim (r_{\rm Sch}/R)^{1/2} \sim v/c\,. \tag{5.1.22}$$

Starting with the metric, one can note that if g_{00} is order ϵ^n, g_{0i} requires ϵ^{n+1} and g_{ij} goes to ϵ^{n+2}. In the Newtonian limit we had $g_{00} = -1 - 2\Phi/c^2$, $g_{0i} = 0$ and $g_{ij} = \delta_{ij}$ going to ϵ^2. The first PN correction will start at ϵ^2 and therefore will require going to ϵ^4. A similar expansion can be applied to the energy-momentum tensor $T^{\mu\nu}$. One can then plug in these expanded quantities to obtain corrections to the perturbed metric $h_{\mu\nu}$.

We have seen that the Green's function solution of $h_{\mu\nu}$ involves the retarded time $t - r/c$. For a generic function of the retarded time, $F(t-r/c)$, the PN expansion implies an expansion in small retardation

$$F(t - r/c) = F(t) - \frac{r}{c}\dot{F}(t) + \frac{r^2}{2c^2}\ddot{F}(t) + \cdots\,. \tag{5.1.23}$$

Each time derivative involves a factor of the frequency $\omega = 2\pi c/\lambda$. Therefore, the small retardation expansion is in fact an expansion in r/λ. In other words, the PN expansion is only valid in the near field.

5.1.3 Post-Minkowskian expansion

We now wish to make an equivalent expansion to the post-Newtonian formalism that is valid in the far zone. For that one should notice that the further we are from the source, the closer we are to flat spacetime. The corrections will be proportional to $r_{\rm Sch}/r \propto G$. Therefore, the post-Minkowskian (PM) expansion is in orders of G.

It is interesting to note that in full GR we can always define a quantity $\mathbb{h}^{\mu\nu}$ as

$$\sqrt{-g}g^{\mu\nu} = \eta^{\mu\nu} + \mathbb{h}^{\mu\nu} \tag{5.1.24}$$

which contains all differences with respect to Minkowski spacetime $\eta^{\alpha\beta}$. The appeal of this quantity is that satisfies the same equations that we were working with before

$$\Box \mathbb{h}_{\mu\nu} = -16\pi G \tau_{\mu\nu}\,, \tag{5.1.25}$$

with $\Box$ the flat space d'Alembertian and the effective stress tensor $\tau_{\mu\nu} = T_{\mu\nu} + \mathbb{t}_{\mu\nu}$ with $\mathbb{t}_{\mu\nu}$ being a pseudo-tensor constructed of $\mathbb{h}$. Therefore, we recover directly the quadrupole formula for $\mathbb{h}^{\mu\nu}$.

The PM formalism makes benefit of this fact to expand

$$\mathbb{h}^{\mu\nu} = \sum_{n=1}^{\infty} G^n \mathbb{h}_n^{\mu\nu} \tag{5.1.26}$$

and solve around $T^{\mu\nu} = 0$ so that

$$\Box \mathbb{h}_{\mu\nu} = -16\pi G \mathbb{t}_{\mu\nu} \,. \tag{5.1.27}$$

This expansion is of course only valid in the far zone. The remaining game is to solve it and then match it with the PN expansion in the near region.

5.1.4 Energy carried by a GW

GWs carry energy and produce a physical effect on nearby particles accelerating them. In the context of GR this means that GWs themselves curve the background through their energy-momentum tensor $T^{\mu\nu}_{\rm gw}(h)$. In fact, we have already seen that GW perturbations $h_{\mu\nu}$ contribute to the energy-momentum tensor at second order in Eq. (5.1.2). We precisely wish to compute this leading order contribution $T^{\mu\nu}_{\rm gw}(h^2)$. The only subtlety is that we need to distinguish the background metric contribution $\bar{g}_{\mu\nu}$ from the GW contribution. In general, those cannot be disentangle. Or in other words, there is not a unique definition of a local energy in GR. The trick will be again on separating the relevant scales, assuming that the typical scale of variation of the background L_B is much larger than the (reduced) wavelength of the GW $\lambda\!\!\!^{-} = \lambda/2\pi$. Explicitly $\lambda\!\!\!^{-} \ll L_B$. This is known as the **short-wave expansion** and will be covered in more detail in §6.1. For the time being, this separation of scales will be mostly only relevant when averring the energy density of the GW over a given region.

Literature:

This section follows mostly Maggiore's section 1.4 [32].

We start by expanding the Einstein field equations in vacuum up to second order in h, that is

$$G_{\mu\nu} = \bar{G}_{\mu\nu}[\bar{g}] + G^{(1)}_{\mu\nu}[h] + G^{(2)}_{\mu\nu}[h] + \cdots = \frac{8\pi G}{c^4} T_{\mu\nu} \,, \tag{5.1.28}$$

where $\bar{G}_{\mu\nu}$ is the Einstein tensor associated to the background metric. $G^{(1)}_{\mu\nu}$ is linear in h (the term we considered before) and $G^{(2)}_{\mu\nu}$ is quadratic in h (the new term). $\bar{G}_{\mu\nu}$ only contains terms associated to the background scale L_B. Therefore, it is a "long" wavelength term. On the contrary, $G^{(1)}_{\mu\nu}$ is only made of "short" wavelength modes. Finally, $G^{(2)}_{\mu\nu}$ can contain both long and short modes.[6] In this way we can split the

[6] For example in a quadratic term $\sim h_{\mu\nu} h_{\alpha\beta}$ each term can have a high wave-vector $\vec{k}_i$ but on opposite directions $\vec{k}_2 = -\vec{k}_1$ to form a low wave-vector quadratic mode.

Einstein equations in long/short modes:

$$\bar{G}_{\mu\nu} = -\left[G^{(2)}_{\mu\nu}\right]^{\text{long}} + \frac{8\pi G}{c^4}\left[T_{\mu\nu}\right]^{\text{long}}, \tag{5.1.29}$$

$$G^{(1)}_{\mu\nu} = \frac{8\pi G}{c^4}\left[T_{\mu\nu}\right]^{\text{short}} - \left[G^{(2)}_{\mu\nu}\right]^{\text{short}}. \tag{5.1.30}$$

The first equation solves for the background metric (including the energy carried by the GWs) and the second equation solves the propagation of the metric perturbations.

The next step is to project into the short/long modes. When those are well separated, the easiest is to introduce an additional scale l between the two:

$$\lambda\!\!\!^{-} \ll l \ll L_B\,. \tag{5.1.31}$$

Then, we will average out the equations over a spatial volume with size l. We will denote this operation as $\langle\cdots\rangle$. Therefore, the background metric will be a solution of

$$\bar{G}_{\mu\nu} = \frac{8\pi G}{c^4}\left(\bar{T}_{\mu\nu} + t_{\mu\nu}\right), \tag{5.1.32}$$

where we have defined $\bar{T}_{\mu\nu} \equiv \langle T_{\mu\nu}\rangle$ and

$$t_{\mu\nu} \equiv -\frac{c^4}{8\pi G}\langle R^{(2)}_{\mu\nu} - \frac{1}{2}\bar{g}_{\mu\nu}R^{(2)}\rangle\,. \tag{5.1.33}$$

To get to the energy carried by a GW we only need to expand the Ricci tensor to second order in h and read of the components. The Ricci tensor is defined in general as

$$R_{\mu\nu} = \partial_\alpha\Gamma^\alpha_{\mu\nu} - \partial_\mu\Gamma^\alpha_{\alpha\nu} + \Gamma^\alpha_{\mu\nu}\Gamma^\beta_{\alpha\beta} - \Gamma^\alpha_{\beta\nu}\Gamma^\beta_{\mu\alpha}\,, \tag{5.1.34}$$

where $\Gamma^\alpha_{\mu\nu}$ is the Christoffel symbol:

$$\Gamma^\alpha_{\mu\nu} = \frac{1}{2}g^{\alpha\beta}\left(\partial_\mu g_{\nu\beta} + \partial_\nu g_{\mu\beta} - \partial_\beta g_{\mu\nu}\right)\,. \tag{5.1.35}$$

Therefore, schematically $\Gamma = \Gamma(\partial g)$ and $R_{\mu\nu} = R_{\mu\nu}(\partial^2 h, \partial h\partial h)$. Since we are interested in the second order terms, we only need to care about $\partial h\partial h$ terms. A great simplification will come from the average scheme we just described. This is because whenever we have the average $\langle\cdots\rangle$ of a space-time derivative term ∂_μ, tha is $\langle\partial_\alpha A^{\beta\cdots\gamma}\rangle$, then this term can be integrated by parts contributing as a boundary term.[7] Boundary terms can then be neglected since their contribution can be made arbitrarily small compared to the bulk by extending the region in which the integral is performed. If we are interested in the GW energy far from the source, we can also approximate the background metric as flat, $\bar{g}_{\mu\nu} \to \eta_{\mu\nu}$. With these tricks and approximation, together with imposing the TT-gauge ($\partial^\mu h_{\mu\nu} = 0$ and $h = 0$) and the vacuum field equations $\Box h_{\alpha\beta} = 0$, then only two terms from $R^{(2)}_{\mu\nu}$ contribute (and can be related by integration by parts) leading to

$$\langle R^{(2)}_{\mu\nu}\rangle = -\frac{1}{4}\langle\partial_\mu h_{\alpha\beta}\partial_\nu h^{\alpha\beta}\rangle\,. \tag{5.1.36}$$

[7] Recall that $\int_V u\partial v dx = [uv]_{\partial V} - \int_V u\partial u dx$. If $u = 1$, then only the boundary term ∂V of v survives.

With this results we can also see that the term $\langle R^{(2)} \rangle$ vanishes, as one can integrate by parts and use the equations of motion. Altogether, we have found

$$t_{\mu\nu} = \frac{c^4}{32\pi G} \langle \partial_\mu h_{\alpha\beta} \partial_\nu h^{\alpha\beta} \rangle \,. \tag{5.1.37}$$

This is the seminal result derived by Isaacson [33, 34].

The GW energy-momentum tensor (5.1.37) is gauge invariant (prove it in the exercise below!). Therefore we can substitute $h_{\mu\nu}$ with the physical modes in the TT-gauge h_+ and $h_\times$. The energy density is then

$$t^{00} = \frac{c^2}{32\pi G} \langle \dot{h}^{TT}_{ij} \dot{h}^{ij}_{TT} \rangle = \frac{c^2}{16\pi G} \langle \dot{h}_+^2 + \dot{h}_\times^2 \rangle \,, \tag{5.1.38}$$

where in the first equality we have defined the dot as $\partial_t = c\partial_0$ and in the second equality we have just expanded in the amplitudes of the two polarizations.

Finally, let us note that because the left hand side of Eq. (5.1.32) is covariantly conserved, the right hand side will be as well. Since far from the source $\bar{T}_{\mu\nu} = 0$ and the metric is approximately flat, one then has the conservation of the GW energy-momentum tensor:

$$\partial^\mu t_{\mu\nu} = 0 \,. \tag{5.1.39}$$

Exercise 5.3: Gauge invariance of the GW energy

Show that the GW energy-momentum tensor (5.1.37) is invariant under a gauge transformation

$$h_{\mu\nu} \to h_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu \,. \tag{5.1.40}$$

Energy flux

Now that we have computed the GW energy-momentum tensor we can compute its energy flux, that is the amount of energy transmitted per unit time and surface area. The GW energy inside a volume V is

$$E_V = \int_V d^3x t^{00} \,. \tag{5.1.41}$$

Its variation over time is therefore

$$\frac{1}{c}\frac{dE_V}{dt} = \int_V d^3x \partial_0 t^{00} = -\int_V d^3x \partial_i t^{0i} = -\int_{\partial V} dA n_i t^{0i} \,. \tag{5.1.42}$$

To obtain the second equality we have used the energy momentum conservation $\partial^\mu t_{\mu\nu} = 0$. The last equality is the boundary term where n^i is the normal to the surface ∂V.

We can then consider a GW propagating in the radial direction outwards. Then $dA = r^2 d\Omega$ and $\hat{n} = \hat{r}$. At sufficiently large distance the wave front is approximately

a plane wave. In that case $h_{ij}^{TT}(t,r) = \frac{1}{r} f_{ij}(t - r/c)$ as we have already seen. Then, $\partial_r h_{ij}^{TT} = -\partial_0 h_{ij}^{TT} = +\partial^0 h_{ij}^{TT}$ and as a consequence

$$t^{0r} = t^{00} \,. \tag{5.1.43}$$

Using this identity we then have the energy time variation within the volume

$$\frac{dE_V}{dt} = -cr^2 \int d\Omega t^{00} \,. \tag{5.1.44}$$

This can then be converted directly into the energy flux that the outward propagating GW carries away

$$\frac{d^2E}{dtd\Omega} = \frac{c^3 r^2}{32\pi G} \langle \dot{h}_{ij}^{TT} \dot{h}_{TT}^{ij} \rangle \,. \tag{5.1.45}$$

This expression can be further related to the quadrupole moment Q_{ij} to obtain

$$\frac{d^2E}{dtd\Omega} = \frac{dP}{d\Omega} = \frac{G}{8\pi c^5} \langle \dddot{Q}_{ij}^{TT} \dddot{Q}_{TT}^{ij} \rangle \,, \tag{5.1.46}$$

which depends on the third time derivatives of the quadrupole moment. Since Q only depends on the propagation direction $\hat{n}$ through the projection into the TT-gauge with the P_{ij} operator, the angular part can be integrated directly (check this!). The final result is that the total radiated power P or, equivalently, the GW luminosity $\mathcal{L}$ is given by

$$\mathcal{L} = \frac{dE}{dt} = \frac{G}{5c^5} \langle \dddot{Q}_{ij} \dddot{Q}^{ij} \rangle \,. \tag{5.1.47}$$

5.2 Inspiral of compact binaries

We have seen that GWs are generated by accelerated energy densities. This can be achieved in many different systems. Astrophysically, and for current detectors, it turns out that the most relevant sources are compact binary coalescences (CBCs). The more compact the objects, the larger their GW emission as they could accelerated more. Therefore, the most relevant sources will be black holes and neutron stars forming binary black hole (BBH), binary neutron star (BNS) and neutorn star black hole (NSBH). We will not enter on how such binaries could form in the first place as this would be a separate course altogether. A brief summary of the landscape of the observations so far will be given in §7. For now, in fact, our sources will be point particles, so we will not care much about them.

5.2.1 Circular orbits

Literature:

This section follows Maggiore's chapter 4 [32] and Creighton & Warren chapter 3 [35].

We will begin by studying a binary system made of two point particles with masses m_1 and m_2 at positions $\vec{r}_1$ and $\vec{r}_2$. Their energy density is then $T^{00}(\vec{x}) = m_1 c^2 \delta^{(3)}(\vec{x} - \vec{r}_1) + m_2 c^2 \delta^{(3)}(\vec{x} - \vec{r}_2)$. In the non-relativistic limit of the two-body problem, the second mass moment can be written as

$$M^{ij} = m_1 r_1^i r_1^j + m_2 r_2^i r_2^j = M_{\rm tot} r_{\rm CM}^i r_{\rm CM}^j + \mu r^i r^j \,, \tag{5.2.1}$$

where we have introduced the center of mass

$$\vec{r}_{\rm CM} = \frac{m_1 \vec{r}_1 + m_2 \vec{r}_2}{m_1 + m_2} \tag{5.2.2}$$

and the relative coordinate $\vec{r} = \vec{r}_2 - \vec{r}_1$. If the system is isolated, then the center of mass is not accelerated and therefore does not contribute to the GW emission. If we locate ourselves in the center of mass frame, in the Newtonian limit the dynamics reduce to a one-body problem. The equivalent one-body has a reduced mass $\mu = m_1 m_2/(m_1 + m_2)$ and equations of motion $\ddot{\vec{r}} = -(GM_{\rm tot}/r^3)\vec{r}$ where $M_{\rm tot} = m_1 + m_2$ is the total mass. By Kepler's law, the orbital frequency ω_s is related to the orbital radius R by

$$\omega_s^2 = GM_{\rm tot}/R^3 \,. \tag{5.2.3}$$

In this frame, the mass quadrupole is then

$$Q^{ij}(t) = \mu \left(r^i(t) r^j(t) - \frac{1}{3} R^2(t) \delta_{ij} \right) \,, \tag{5.2.4}$$

where $R = |\vec{r}|$. It is interesting to note already that if there is periodic motion, for example along a given dimension $z(t) = a \cos \omega_s t$, then

$$Q^{ij}(t) = \frac{2}{3} \mu z(t)^2 \delta^{i3} \delta^{j3} = \frac{1}{3} \mu a^2 (1 + \cos 2\omega_s t) \delta^{i3} \delta^{j3} \,. \tag{5.2.5}$$

Since h_{ij}^{TT} is sourced by the second derivative of the quadrupole moment, $\ddot{Q}_{ij}$, then a non-relativistic source performing a harmonic oscillation of frequency ω_s will produce monochromatic radiation of frequency twice that frequency: $\omega = 2\omega_s$.

For the two point particles, if we fix their motion to the xy-plane and their initial position to the y-axis, the effective one body description can be captured in

$$\begin{aligned} x(t) &= R \cos(\omega_s t + \pi/2) \,, \\ y(t) &= R \sin(\omega_s t + \pi/2) \,, \\ z(t) &= 0 \,. \end{aligned} \tag{5.2.6}$$

Then the second mass moment in the center of mass frame becomes $M^{ij} = \mu x_{\rm CM}^i x_{\rm CM}^j$, and its (non-vanishing) components are

$$\begin{aligned} M_{11}(t) &= \mu x(t)^2 = \mu R^2 \frac{1 - \cos 2\omega_s t}{2} \,, \\ M_{22}(t) &= \mu y(t)^2 = \mu R^2 \frac{1 + \cos 2\omega_s t}{2} \,, \\ M_{12}(t) &= M_{21}(t) = \mu x(t) y(t) = -\frac{1}{2} \mu R^2 \sin 2\omega_s t \,. \end{aligned} \tag{5.2.7}$$

This implies

$$\begin{aligned} \ddot{M}_{11} &= -\ddot{M}_{22} = 2\mu R^2 \omega_s^2 \cos 2\omega_s t \,, \\ \ddot{M}_{12} &= \ddot{M}_{21} = 2\mu R^2 \omega_s^2 \sin 2\omega_s t \,. \end{aligned} \tag{5.2.8}$$

For an observer at a distance r on the z-axis $\ddot{I}_{ij}$ is already in the TT-gauge and therefore

$$h_{ij}^{TT} = \frac{4G\mu R^2 \omega_s^2}{rc^4} \begin{pmatrix} \cos 2\omega_s t & \sin 2\omega_s t & 0 \\ \sin 2\omega_s t & -\cos 2\omega_s t & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{5.2.9}$$

As pointed out before, the GW frequency is twice the orbital frequency. Moreover, the two polarizations, h_+ and $h_\times$ are de-phased by $\pi/2$. It is interesting to rewrite this expression in terms of the orbital velocity $\nu = R\omega_s$, which using Kepler's law can be related to the GW frequency $\omega = 2\pi f$ via

$$\nu = (\pi G M_{\rm tot} f)^{1/3} \,. \tag{5.2.10}$$

ν is also related to the orbital period $\nu = (2\pi G M_{\rm tot}/P)^{1/3}$ and the orbital separation $\nu = (GM/R)^{1/2}$. The observer on the axis of the binary measures the polarizations

$$h_+(t) = \frac{1}{r}\frac{4G\mu}{c^2}\left(\frac{\nu}{c}\right)^2 \cos 2\omega_s t \,, \tag{5.2.11}$$

$$h_\times(t) = \frac{1}{r}\frac{4G\mu}{c^2}\left(\frac{\nu}{c}\right)^2 \sin 2\omega_s t \,. \tag{5.2.12}$$

We can also rewrite this expression in an alternative form introducing the **chirp mass**

$$\mathcal{M}_c = \mu^{3/5} M_{\rm tot}^{2/5} = \frac{(m_1 m_2)^{3/5}}{(m_1+m_2)^{1/5}} \,, \tag{5.2.13}$$

and the GW frequency $f = \omega/2\pi = \omega_s/\pi$ to obtain

$$h_+(t) = \frac{4}{r}\left(\frac{G\mathcal{M}_c}{c^2}\right)^{5/3}\left(\frac{\pi f}{c}\right)^{2/3} \cos 2\pi f t \,, \tag{5.2.14}$$

$$h_\times(t) = \frac{4}{r}\left(\frac{G\mathcal{M}_c}{c^2}\right)^{5/3}\left(\frac{\pi f}{c}\right)^{2/3} \sin 2\pi f t \,. \tag{5.2.15}$$

For a generic observer, we need to generalize slightly the possible orientation and initial position of the binary. This can be done transforming from the **source frame** to the **wave frame** with two rotations

$$R = R_\phi R_\theta = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} . \tag{5.2.16}$$

With this transformation we obtain

$$h_+(t) = \frac{4}{r}\left(\frac{G\mathcal{M}_c}{c^2}\right)^{5/3}\left(\frac{\pi f}{c}\right)^{2/3}\left(\frac{1+\cos^2\theta}{2}\right)\cos(2\pi f t + 2\phi)\,, \tag{5.2.17}$$

$$h_\times(t) = \frac{4}{r}\left(\frac{G\mathcal{M}_c}{c^2}\right)^{5/3}\left(\frac{\pi f}{c}\right)^{2/3}\cos\theta\sin(2\pi f t + 2\phi)\,. \tag{5.2.18}$$

To make this expression more transparent we further introduce the Schwarzschild radius associate to the chirp mass $R_c = 2G\mathcal{M}_c/c^2$ and the reduced GW wavelength $\lambda = c/\omega$ to finally get

$$h_+(t) = \mathcal{A}\left(\frac{1+\cos^2\theta}{2}\right)\cos(2\pi f t + 2\phi)\,, \tag{5.2.19}$$

$$h_\times(t) = \mathcal{A}\cos\theta\sin(2\pi f t + 2\phi)\,, \tag{5.2.20}$$

with the amplitude

$$\mathcal{A} = \frac{1}{2^{1/3}}\left(\frac{R_c}{r}\right)\left(\frac{R_c}{\lambda}\right)^{2/3}\,. \tag{5.2.21}$$

When the binary is **face on/off** $\theta = 0, \pi$, then the amplitude of both polarizations is the same. The wave is then **circularly** polarized. On the opposite limit, when the orbit is **edge-on** $\theta = \pi/2$, then $h_\times$ vanishes and the wave is **linealry** polarized. For any other inclination, the wave is **elliptically** polarized.

For later purposes it will be useful to stop for a second and count how many parameters describe the binary. We have the two masses m_1 and m_2, the distance r, the inclination θ and the initial phase ϕ. Implicitly we also have an initial reference time $t_{\rm ref}$ that we have set to 0 so far.

GW luminosity

With the GW polarizations at hand we can compute the GW luminosity of an inspairaling binary. From Eq. (5.1.46) we have that

$$\frac{d\mathcal{L}}{d\Omega} = \frac{d^2E}{dtd\Omega} = \frac{r^2c^3}{16\pi G}\langle \dot{h}_+^2 + \dot{h}_\times^2\rangle = \frac{2}{\pi}\frac{c^5}{G}\left(\frac{G\mathcal{M}_c\pi f}{2c^3}\right)^{10/3} g(\theta)\,, \tag{5.2.22}$$

where

$$g(\theta) = \left(\frac{1+\cos^2\theta}{2}\right)^2 + \cos^2\theta\,. \tag{5.2.23}$$

Note that this expression is independent of ϕ because $\langle\cos^2(2\pi f t + 2\phi)\rangle = 1/2$. The angular average over the inclination angle is

$$\int d\Omega g(\theta) = 16\pi/5\,. \tag{5.2.24}$$

Altogether the GW luminosity is

$$\mathcal{L} = \frac{32}{5}\frac{c^5}{G}\left(\frac{G\mathcal{M}_c\pi f}{c^3}\right)^{10/3} = \frac{32}{5}\frac{c^5}{G}\eta^2\left(\frac{\nu}{c}\right)^{10}\,, \tag{5.2.25}$$

where we have introduced the symmetric mass ratio

$$\eta = \frac{\mu}{M_{\rm tot}} = \frac{\mathcal{M}_c^{5/3}}{M_{\rm tot}^{5/3}} = \frac{m_1 m_2}{(m_1+m_2)^2}\,. \tag{5.2.26}$$

Exercise 5.4: GW luminosity

Derive the expression for the luminosity emitted in GWs by a binary during the inspiral. For a binary of 30–30$M_\odot$ masses at 250Hz, how does this luminosity compare to the one emitted by the Sun?

The chirp

So far we have seen the GW emission by a compact binary system in a (fixed) circular orbit. We have studied the GWs emitted due to this motion and the energy they carry. Because this energy carried by the GWs has to come from the orbital motion, the orbit has to shrink. This is easy to see from the Newtonian energy of the orbit

$$E_{\rm orbit} = E_{\rm kin} + E_{\rm pot} = -\frac{G m_1 m_2}{2R}\,, \tag{5.2.27}$$

which follows from the virial theorem. Therefore, to compensate the loss of energy, R has to decrease. If R decreases, then the orbital frequency ω_s increases (recall Eq. (5.2.3)). Moreover, this increase of the frequency also enlarges the power radiated in GWs that in turn reduces more the orbit and increases even more the frequency. This eventually leads to the coalescence of the binary. This inspiral process is known as the **chirp** because of the rapid increase of the frequency. Despite this change in the orbital radius, as long as the variation in the frequency is not too large (also known as adiabatic)

$$\dot{\omega}_s \ll \omega_s^2\,, \tag{5.2.28}$$

then we are in the **quasi-circular** regime and our previous formulas apply noting that now the frequency of the GW is a function of time: $\omega \to \omega(t)$.

We can obtain the change in the frequency from the energy radiated. For that we equate the GW luminosity in (5.2.25) to the change in the orbital energy $-dE_{\rm orbital}/dt$. Explicitly:

$$\mathcal{L} = \frac{32}{5}\frac{c^5}{G}\left(\frac{G\mathcal{M}_c \pi f}{c^3}\right)^{10/3} = -\frac{dE_{\rm orbital}}{dt} = \frac{2}{3}\left(\frac{(2\pi)^2 G^2 \mathcal{M}_c^5}{32}\right)^{1/3} f^{-1/3}\dot{f}\,, \tag{5.2.29}$$

where in the last equality we have taken the time derivative of the orbital energy written in terms of the GW frequency f. Simplifying, we arrive at the important equation determining the frequency evolution of a binary system

$$\dot{f} = \frac{96}{5}\pi^{8/3}\left(\frac{G\mathcal{M}_c}{c^3}\right)^{5/3} f^{11/3}\,. \tag{5.2.30}$$

Note that this equation is solely determined by the chirp mass. In other words, at leading order, the chirp mass controls the inspairaling of the binary.

At this level of approximation the frequency of the GW grows until it diverges at a finite time. This defines the **time of coalescence** t_c. Integrating Eq. (5.2.31), one obtains

$$f(t) = \frac{1}{\pi}\left(\frac{5}{256}\frac{1}{t_c - t}\right)^{3/8}\left(\frac{G\mathcal{M}_c}{c^3}\right)^{-5/8} . \tag{5.2.31}$$

This can also be inverted to obtain the time to coalescence starting from a given frequency:

$$t_c - t = \frac{5}{256}\frac{1}{(\pi f)^{8/3}}\left(\frac{G\mathcal{M}_c}{c^3}\right)^{-5/3} . \tag{5.2.32}$$

Recalling that the Schwarzschild time of $1M_\odot$ is 10μs, then

$$t_c - t \simeq 100s\left(\frac{15\text{Hz}}{f}\right)^{8/3}\left(\frac{\mathcal{M}_c}{1.21M_\odot}\right)^{-5/3} . \tag{5.2.33}$$

Therefore, a GW signal from a binary formed of two $1.4M_\odot$ objects will take about 100 seconds to merge starting at 15 Hz. This kind of numbers will be important when understanding the duration of the signals in a GW detector. The general rule of thumb is that the lighter the binary, the longer it will take to merge starting at a fixed frequency.

Exercise 5.5: Number of cycles of a GW

A useful quantity to determine the precision of GW measurements is the number of cycles that a signal is in band in the detector. The longer, the easier it would be to characterize the signal. Noting that for a GW signal with a slowly varying period $T(t)$ the number of cycles is defined by $dN_{\text{cyc}} = dt/T(t) = f(t)dt$, compute the number of cycles that an equal mass binary of $30M_\odot$ will stay of band starting at 10Hz. How does this generalized to a binary of supermassive black holes ($\sim 10^6 M_\odot$) and a detector sensitive down to 0.1mHz?

Phase evolution

Now that we have seen how a binary system chirps, the final step is to compute how this affect the signal itself. Since the frequency is changing, the GW will not be monochromatic in time. For a time evolving wave, the orbital phase needs to be generalized to

$$\Phi(t) = 2\int dt\,\omega_s(t) = \int dt\,\omega(t)\,. \tag{5.2.34}$$

Because the orbital phase $\Phi(t)$ and the orbital radius $R(t)$ are now functions of time,[8] we should in principle take into account their time derivatives when using the quadrupole formula. However, for an adiabatic inspiral, cf. (5.2.28), those terms are negligible.

[8] To see the time dependence of the orbital radius recall the relation between ω_s and R in (5.2.3) and then use the frequency evolution in (5.2.31). You will find that $\dot{\ln}R = -2\dot{\ln}\omega/3 = -1/4(t_c - t)$.

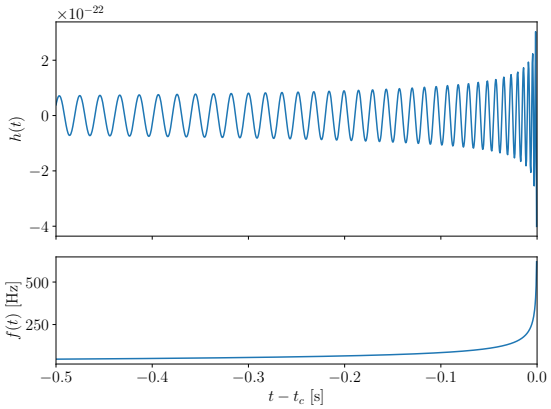


Figure 5.1. Time evolution of the GW strain (top) and frequency (bottom) for an inspiraling binary in the Newtonian limit. This corresponds to a binary with $\mathcal{M}_c = 10M_\odot$ at 1Gpc.

Therefore the only change that we need to apply to our previosu expressions is $\omega \to \omega(t)$ and $\omega t \to \Phi(t)$. That is

$$h_+(t) = \frac{4}{r}\left(\frac{G\mathcal{M}_c}{c^2}\right)^{5/3}\left(\frac{\pi f(t)}{c}\right)^{2/3}\left(\frac{1+\cos^2\theta}{2}\right)\cos(\Phi(t))\,, \tag{5.2.35}$$

$$h_\times(t) = \frac{4}{r}\left(\frac{G\mathcal{M}_c}{c^2}\right)^{5/3}\left(\frac{\pi f(t)}{c}\right)^{2/3}\cos\theta\sin(\Phi(t))\,, \tag{5.2.36}$$

where the phase is given by

$$\Phi(t) = -2\left(\frac{5G\mathcal{M}_c}{c^3}\right)^{-5/8}(t_c - t)^{5/8} + \Phi_0\,. \tag{5.2.37}$$

Here Φ_0 is an integration constant that defines the phase at the time of coalescence $\Phi_0 = \Phi(t_c) = 2\phi_c$. Because the GW frequency increases, then the time domain amplitude also increases as one approaches to the coalescence. We see, once again, that in the Newtonian limit the whole signal is controlled by the chirp mass value. Moreover, if one can measure the frequency evolution to obtain $\mathcal{M}_c$, then the amplitude of the GW polarizations gives a direct measurement of a combination of the distance r and inclination angle θ. If the polarizations can be measured separately, then in principle θ and r can be measure separately. Being able to measure absolute distances is a key distinctive of GW signals over other transients that will be crucial for the cosmological discussion that we will do later. For reference, we include a plot of the time domain signal for an inspiraling binary in Fig. 5.1.

Frequency domain signal

We have just seen what is the time domain waveform of a GW from a binary in the inspiral phase. As in many other branches of signal processing, it can be advantageous to analyze the GW data in Fourier space because the noise sensitivity is better understood

there (we will see this in §7.2). Therefore, we proceed now to compute such frequency domain signal. For that purpose we further rewrite Eqs. (5.2.38) and (5.2.39) to be only a function of time

$$h_+(t) = \frac{1}{r}\left(\frac{G\mathcal{M}_c}{c^2}\right)^{5/4}\left(\frac{5}{c(t_c-t)}\right)^{1/4}\left(\frac{1+\cos^2\theta}{2}\right)\cos(\Phi(t))\,, \tag{5.2.38}$$

$$h_\times(t) = \frac{1}{r}\left(\frac{G\mathcal{M}_c}{c^2}\right)^{5/4}\left(\frac{5}{c(t_c-t)}\right)^{1/4}\cos\theta\sin(\Phi(t))\,. \tag{5.2.39}$$

Schematically we have something like $h(t) = A(t)\cos(\Phi(t))$. We adopt the following Fourier transform convention

$$\tilde{h}(f) = \int_{-\infty}^{\infty} dt h(t) e^{i2\pi ft}\,, \tag{5.2.40}$$

which is common in the GW community. Functions with tildes are in frequency domain. The first thing to notice is that the time domain GW strain is a **real** function, while its Fourier transform in **complex** valued. Therefore, the frequency domain signal must satisfy

$$\tilde{h}(-f) = \tilde{h}^*(f)\,, \tag{5.2.41}$$

where * denotes the complex conjugate. This is made to ensure that the inverse Fourier transform leads to a real function

$$h(t) = \int_{-\infty}^{\infty} df \tilde{h}(f) e^{-i2\pi ft}\,, \tag{5.2.42}$$

as you can directly verify. Eq. (5.2.41) also allows to work only with positive frequencies as the negative ones can be obtained by complex conjugation.

For a wave with a slowly varying amplitude and phase acceleration,

$$\frac{d\ln A}{dt} \ll \frac{d\Phi}{dt}\,, \qquad \frac{d^2\Phi}{dt^2} \ll \left(\frac{d\Phi}{dt}\right)^2\,, \tag{5.2.43}$$

then the Fourier transform can be solved using the **stationary phase approximation (SPA)**. We are going to show this explicitly. The first step is to rewrite the expression in terms of exponential $\int dt A e^{i\phi(t)}$ in which we could identify the stationary points as $d\phi/dt = 0$. Doing so, and focusing on positive frequencies (negative ones are obtained by complex conjugation):

$$\tilde{h}(f>0) = \frac{1}{2}\int dt A(t)(e^{i\Phi(t)} + e^{-i\Phi(t)})e^{i2\pi ft} \simeq \frac{1}{2}\int dt A(t) e^{i(2\pi ft - \Phi(t))}\,, \tag{5.2.44}$$

where in the second equality we have noted that since $\dot{\Phi} = \omega(t) > 0$, then only the second term of the first integrand leads to an stationary point t_* defined by

$$\dot{\Phi}(t_*) = 2\pi f\,. \tag{5.2.45}$$

Note that in finding this stationary point we have used the fact that $e^{\ln A(t)}$ varies slowly (thus the first condition in (5.2.43)). We then proceed with the standard SPA apparatus and expand the exponent to second order to get the first correction around the stationary point:

$$\tilde{h}(f) = \frac{1}{2}A(t_*)e^{i(2\pi f t_* - \Phi(t_*))}\int e^{-\frac{i}{2}\ddot{\Phi}(t-t_*)^2} = \frac{1}{2}A(t_*)\left(\frac{2\pi}{\ddot{\Phi}(t_*)}\right)^{1/2} e^{i(2\pi f t_* - \Phi(t_*) - \pi/4)}\,, \tag{5.2.46}$$

where in the second equality we have used the fact that if we define a new variable $x = (\ddot{\Phi}/2)^2(t - t_*)$ we can just solve the Gaussian integral

$$\int_{-\infty}^{\infty} dx e^{-ix^2} = \sqrt{\pi}e^{.i\pi/4}\,. \tag{5.2.47}$$

The frequency domain signal is typically written in terms of the amplitude and phase

$$\tilde{h}(f) = \tilde{A}(f)e^{i\Psi(f)}\,. \tag{5.2.48}$$

Therefore, the only remaining step is to relate t_* and f. We can do this immediately because we have already solved for $\omega = \omega(t)$ in (5.2.32), therefore

$$t_* - t_c = -\frac{5}{256}\frac{1}{(\pi f)^{8/3}}\left(\frac{G\mathcal{M}_c}{c^3}\right)^{-5/3}\,. \tag{5.2.49}$$

With this expression we find for the phase

$$\Psi(f) = 2\pi f t_c - \Phi_0 - \frac{\pi}{4} + \frac{3}{4}\left(\frac{8\pi G\mathcal{M}_c f}{c^3}\right)^{-5/3}\,. \tag{5.2.50}$$

Note that we have implicitly assumed that the time domain waveform varied as $\cos\Phi(t)$, which corresponds to $h_+(t)$. We should therefore define the frequency domain signal for each polarization:

$$\tilde{h}_+(f) = \bar{A}(f)e^{i\Psi_+(f)}\left(\frac{1+\cos^2\theta}{2}\right)\,, \tag{5.2.51}$$

$$\tilde{h}_\times(f) = \bar{A}(f)e^{i\Psi_\times(f)}\cos\theta\,, \tag{5.2.52}$$

where $\Psi_+(f)$ is given by (5.2.50) and $\Psi_\times(f) = \Psi_+(f) + \pi/2$. The amplitude on the other hand is given by

$$\bar{A}(f) = \frac{1}{\pi^{2/3}}\left(\frac{5}{24}\right)^{1/2}\frac{c}{r}\left(\frac{G\mathcal{M}_c}{c^3}\right)^{5/6} f^{-7/6}\,. \tag{5.2.53}$$

Therefore, we arrive at the important conclusion that the amplitude of the GW in frequency domain scales as $|\tilde{h}| \propto f^{-7/6}$.

Now that we have the frequency domain waveform, we can compute the energy emitted as a function of frequency. Starting from the energy flux as a function of the GW polarizations in (5.1.46) we can write

$$\frac{dE}{dA} = \frac{c^3}{16\pi G}\int_{-\infty}^{\infty} dt \langle \dot{h}_+^2 + \dot{h}_\times^2 \rangle \,. \tag{5.2.54}$$

For practical purposes, the spatial average $\langle \cdots \rangle$ is an average over a few periods. Therefore, it does not play any role within this integral and we can omit it. If we then introduce the definition of the Fourier transform we get

$$\frac{dE}{dA} = \frac{c^3}{16\pi G}\int_{-\infty}^{\infty} df (2\pi f)^2 \left(|\tilde{h}_+|^2 + |\tilde{h}_\times|^2 \right) . \tag{5.2.55}$$

Therefore, the energy emitted per frequency is

$$\frac{dE}{df} = \frac{\pi c^2}{2G} f^2 r^2 \int d\Omega \left(|\tilde{h}_+|^2 + |\tilde{h}_\times|^2 \right) . \tag{5.2.56}$$

If we include next the inspiral waveforms we arrive at

$$\frac{dE}{df} = \frac{\pi^{2/3}}{3G} (G\mathcal{M}_c)^{5/3} f^{-1/3} \,. \tag{5.2.57}$$

Integrating over a maximum frequency $f_{\rm max}$ at which the inspiral is still valid, we can estimate the total radiated energy

$$\Delta E_{\rm rad} \sim \frac{\pi^{2/3}}{2G} (G\mathcal{M}_c)^{5/3} f_{\rm max}^{2/3} \,. \tag{5.2.58}$$

In numerical values, this corresponds to

$$\Delta E_{\rm rad} \sim 4 M_\odot c^2 \left(\frac{\mathcal{M}_c}{27 M_\odot} \right)^{5/3} \left(\frac{f_{\rm max}}{250 {\rm Hz}} \right)^{2/3} , \tag{5.2.59}$$

or, in other words, a 30–30$M_\odot$ binary will emit about 3$M_\odot$ in GWs!

Exercise 5.6: Validity of the stationary phase approximation

In order to determine the validity of the stationary phase approximation, one can compare the prediction for the energy emitted using this approximation or the numerical solution. To that end, solve the energy emitted numerically by starting directly from the time domain signal where no approximation was made. That is, insert (5.2.38) in Eq. (5.2.54) and then solve. Up to what frequency do both results agree?

With the energy emission per unit frequency, Eq. (5.2.57), one can compute the spectrum of a stochastic background of inspiral signals. It is customary to report this

as a function of the energy density per logarithmic frequency normalized to the critical energy density of the universe $\rho_c = 3c^2 H_0/8\pi G$:[9]

$$\Omega_{\rm gw}(f) = \frac{1}{\rho_c}\frac{d\rho(f)}{d\ln f}\,. \tag{5.2.60}$$

The stochastic GW background integrates all contributions from compact binaries of a given class over cosmic time. Therefore, if we have a population of binaries with a merger rate density $\mathcal{R}(z)$, then the above expression translates to

$$\Omega_{\rm gw}(f) = \frac{f}{\rho_c}\int dz \frac{R(z)}{(1+z)H(z)}\frac{dE}{df}\,, \tag{5.2.61}$$

where the redshift factor $1+z$ dependence and the Hubble parameter $H(z)$ appear when changing from integrating in cosmic time to redshift. For inspiraling binaries, there is a characteristic

$$\Omega_{\rm gw}(f)|_{\rm inspiral} \propto f^{2/3} \tag{5.2.62}$$

spectrum.

Before we conclude this section we should note that so far we have studied the inspiral of a binary over flat space time at leading order. As the compact object get close to each other strong gravity effects become important. The Schwarzschild geometry predicts that there is a minimum radius for circular orbits, the **inner most stable circular orbit (ISCO)**. Beyond this point, the two objects plunge into each other. In the test mass limit, when the reduced mass μ is much smaller than the total mass, $\mu \ll M_{\rm tot}$, this is given by

$$r_{\rm ISCO} = \frac{6GM_{\rm tot}}{c^2} = 3r_{\rm Sch}(M_{\rm tot})\,. \tag{5.2.63}$$

This translates into a maximum frequency for the inspiral phase which can be obtained using the Keplerian relation (5.2.3) to obtain a GW frequency

$$f_{\rm ISCO} = \frac{1}{6\pi\sqrt{6}}\frac{c^3}{GM_{\rm tot}}\,. \tag{5.2.64}$$

Exercise 5.7: Inner most stable circular orbit

A compact binary transitions from the inspiral to the merger phase at around the inner most stable circular orbit (ISCO), cf. Eq. (5.2.63). We have derived the inspiral waveform in the frequency domain using the stationary phase approximation (SPA). In the test mass limit, when the reduced mass μ is much smaller than the total mass, $\mu \ll M_{\rm tot}$, determine up to which frequency is the SPA valid. This can be shown by checking if $d\ln A/dt \ll \dot{\Phi}$ and that $\ddot{\Phi} \ll (\dot{\Phi})^2$.

[9] This normalization arises to give a cosmological context to this background that can then be directly compared to the fractional density of photons, neutrinos or any other component in the universe.

Explore: Elliptic orbits

A natural extension of our quasi-circular study is to consider two point particles in a Keplerian elliptic orbit. In this case the orbit is characterized by a semi-major axis a and an eccentricity e. The circular limit is achieved when $e \to 0$ and, as a consequence $a \to R$. The total radiated power is modified to

$$P = \frac{dE}{dt} = \frac{32G^2\mu^2 M_{\rm tot}^3}{5c^5 a^5} f(e) \tag{5.2.65}$$

where

$$f(e) = \frac{1}{(1-e^2)^{7/2}} \left(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4\right) \tag{5.2.66}$$

This is the classical result from Peters & Mathews in 1963 [36]. Importantly, an eccentric binary (rapidly) circularizes due to GW emission.

Exercise 5.8: Hulse-Taylor pulsar

The Hulse-Taylor pulsar is a binary neutron star system in which one of the neutron stars is a pulsar. A pulsar is a highly magnetized rotating neutron star that emits pulses of electromagnetic radiation. Pulsars are fantastic astronomical objects because they serve as clocks. Their EM emission has a reliable periodcity. In this case a clock to time the shrink of the orbit due to the GW emission. The Hulse-Taylor pulsar was the first system of this kind to be discovered. Compute the energy loss by GW of this system. This was the first indirect evidence of the existence of GWs. The discovery of this system received the Nobel Prize in Physics in 1993.

5.2.2 Beyond the Newtonian limit

In our treatment so far we have assumed a Newtonian motion of a binary made of point particles and computed its GW emission. This set of assumptions can be broken in many directions. For example, we could add effects to each particle such spin. We could also compute corrections as one approaches to strong-field gravity near merger. Similarly, each object has a finite size and its composition is relevant to compute tidal effects.

5.3 Full waveforms: inspiral, merger, ringdown

The coalescence of a compact binary is typically divided in three phase: inspiral, merger and ring-down. The inspiral phase is all we have studied so far. It is the situation in which we can obtain analytical solutions for the motion of the binary. These solutions can be obtained to increasing order of precision including higher order corrections. However, eventually, the system reaches a point in which the gravity is at its greatest stage and the strong-filed evolution can only be solved numerically. This is the merger phase.

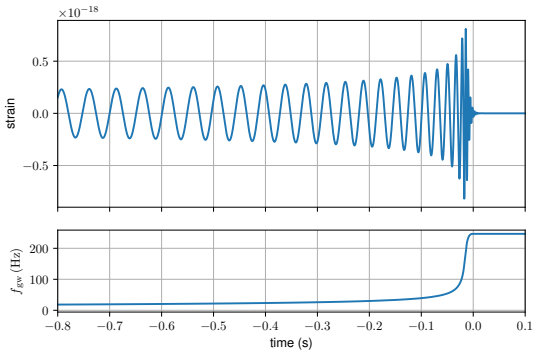


Figure 5.2. Full waveform for a GW including the inspiral, merger and ringdown. The top panel shows the time domain waveform while the bottom panels shows the frequency evolution.

Numerical relativity has demonstrated an enormous strength in providing waveform models, but it is limited by its (still) very large computational cost. Matching the inspiral and merger phase consistently its a whole industry on its own and defines different classes of waveform models used in the community. However, after the objects have merged, the remnant stabilizes and one inevitably reaches a point of small perturbations around a single compact object. For the case of a black holes this can be studied through perturbation theory as you have learned in the first part of the course. This is the ringdown phase.

Exploring the construction of inspiral-merger-ringdown waveforms goes beyond the scope of this course. We can still describe its qualitative features from Fig. 5.2.

CHAPTER

6

Gravitational waves across the cosmos

In the last chapter we have studied how GWs are generated during the coalescence of a compact binary. Our next task is then investigate the propagation of such waves. We have seen already that over flat backgrounds they propagate according to a simple wave equation

$$\Box_{\rm flat}\bar{h}_{\mu\nu} = \partial_\alpha\partial^\alpha h_{\mu\nu} = 0 \tag{6.0.1}$$

and we have even included second order corrections to compute the GW back-reaction and the energy that they carried. We did not pay much attention though to the effect of the background geometry on the GWs themselves and their propagation's properties. This is precisely the goal of this chapter.

6.1 Propagation in curved backgrounds

To study the propagation of GWs on curved space-times we need to be able to distinguish between the GW and the background metric in the first place! This can be done in very different ways depending on the problem at hand. For example, the background could be a smooth function and the perturbations the ripples on top, or the background could have a symmetry (e.g. spherical symmetry) that the perturbations break, or the background could be at a particular state (e.g. equilibrium) that the perturbations deviate from. In §5.1.4 we used an approach in which we introduced a separation of scales between the GW (reduced) wavelength λ and the typical scale of variation of the background L_B, so that they satisfy $\lambda \ll L_B$. For the moment we will **not** make such assumption. The only thing we will demand is that the amplitude of the GW perturbation $\mathcal{A}$ is small, as compared to the typical value of the background $|h_{\mu\nu}| \lesssim |\bar{g}_{\mu\nu}|\mathcal{A}$, where $\bar{g}_{\mu\nu}$ is the background metric. With this condition in mind we will be making the decomposition

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}\,. \tag{6.1.1}$$

Focusing our analysis to linear order in h implies that the inverse metric is given by $g^{\mu\nu} = \bar{g}^{\mu\nu} - h^{\mu\nu}$ with $h^{\mu\nu} = \bar{g}^{\mu\alpha}\bar{g}^{\nu\beta}h_{\alpha\beta}$.

Literature:

The first part of this section follows chapter 35 of Misner-Thorne-Wheleer [37] as well as an unpublished chapter on gravitational wave propagation by Kip Thorne [38] that can be found here.

For the moment we will restrict to propagation in vacuum, therefore, $R_{\mu\nu} = 0$. Expanding the field equations as in §5.1.4, we are therefore interested in the solutions of the linear propagation of the modes:

$$R^{(1)}_{\mu\nu}(h) = 0\,. \tag{6.1.2}$$

Expanding the Ricci tensor to linear order in h is something you already did in the first part of the course. Refreshing you about the notation for the trace-reversed perturbation in a curved background

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\bar{g}_{\mu\nu}h\,, \tag{6.1.3}$$

where the trace is $h = \bar{g}^{\mu\nu}h_{\mu\nu}$ and the Lorenz gauge

$$\nabla^{\mu}\bar{h}_{\mu\nu} = 0\,, \tag{6.1.4}$$

where the covariant derivative is w.r.t. to the background metric, the propagation equation is given by

$$\Box\bar{h}_{\mu\nu} + 2\bar{R}_{\alpha\mu\beta\nu}\bar{h}^{\alpha\beta} = 0\,, \tag{6.1.5}$$

where $\Box$ is the D'Alembertian operator over the background metric and $\bar{R}_{\alpha\mu\beta\nu}$ the Riemann tensor also for $\bar{g}_{\mu\nu}$.[1] This is the main equation that we now want to solve. It carries all the information about the interactions of the linear perturbations with the background. However, by its own construction, it neglects non-linear effects. It is important to emphasize that Eq. (6.1.5) is indeed valid for **any** ratio of $\lambda\!\!\!^{-}/L_B$. In what comes next we will first focus in the short-wave expansion $\lambda\!\!\!^{-} \ll L_B$ in order to understand the cosmological propagation in §6.2. We will then study in §6.3 the full linear propagation (for any wavelength) for a simplified (weak-field) background, i.e. GW lensing.

[1] Note that one could get a similar equation for $h_{\mu\nu}$ noting that, in vacuum, the contributions from $\bar{R}_{\mu\nu}$ have to be sourced by the GWs and, as a consequence, they are of order $(\partial h)^2 \sim \mathcal{A}^2/\lambda\!\!\!^{-2}$. Then, a term $\bar{R}_{\alpha(\mu}h^{\alpha}_{\nu)}$ (where the parenthesis indicates symmetrization of the indices) would be of higher order, $\mathcal{A}^3/\lambda\!\!\!^{-2}$.

Explore: Scattering of GWs with the background metric

For situations in which the curvature scale is smaller than the typical GW wavelength, $L_B \lesssim \lambda$, an interesting phenomenon is the scattering of GWs with the background metric. As you have studied in the firs part of the course, this is particular important effect when computing the normal mode vibrations of black holes [39]. This effect is also important for the production of tails of waves in the near zone [40, 41] and radiative tails in the far zone [42, 43].

6.1.1 Short-wave expansion

We are now going to solve the linearized GW propagation in curved, vacuum space-times (6.1.5) in the limit of $\lambda \ll L_B$, this is known as the **short-wave** expansion or **eikonal** approximation. To make it more convoluted, it is also referred as the **geometric optics** limit and the Wentzel–Kramers–Brillouin (WKB). The truth is that there are so many names because this is a very standard approach to solve the wave propagation appearing in many branches of physics and, most notably, in optics. Besides the possible nomenclature confusion, we will focus on the physical properties that describe the wave propagation in this regime.

We are going to follow a similar approach to the one we used for linearized perturbation around flat space-times (recall (5.0.3)) and expand the metric perturbations in an amplitude and a phase

$$h_{\mu\nu}(x) = \mathrm{Re}\left[\left(A_{\mu\nu}(x) + \varepsilon A^{(1)}_{\mu\nu}(x) + \varepsilon^2 A^{(2)}_{\mu\nu}(x) + \cdots\right) e^{i\theta(x)/\varepsilon}\right] , \tag{6.1.6}$$

where ε is a "dummy" expansion parameter that just help us remember that phase θ oscillates faster than the amplitude and that the amplitude can have higher order corrections $A^{(n)}_{\mu\nu}$. We make the same definitions for the wavevector $k_\mu = \partial_\mu\theta$, scalar amplitude $A = \sqrt{A^*_{\mu\nu}A^{\mu\nu}}$ and polarization vector $\epsilon_{\mu\nu} = A_{\mu\nu}/A$ that satisfies $\epsilon^*_{\mu\nu}\epsilon^{\mu\nu} = 1$. The amplitude tensors ($A_{\mu\nu}$, $A^{(1)}_{\mu\nu}$,...) are all complex valued. As we did before, we solve iteratively in powers of ε. The main difference will be the presence of $2\bar{R}_{\alpha\mu\beta\nu}\bar{h}^{\alpha\beta}$ that effectively behaves as a mass term.[2] At leading order (ε^{-2}) we get

$$\bar{g}_{\mu\nu}k^\mu k^\nu = 0 , \tag{6.1.7}$$

which is equivalent to the flat spacetime propagation changing $\eta \to \bar{g}$. Similarly we get $k^\mu\nabla_\mu k_\nu = 0$ but now for the covariant derivative on the curved background. Therefore, at leading order, GWs still travel along null geodesics and propagate at the speed of light. In the context of geometric optics, null geodesics are also known as **rays**. At next to leading order (ε^{-1}) we obtain the conservation of gravitons and the parallel transport of the polarization tensor:

$$\nabla^\mu(A^2 k_\mu) = 0 , \tag{6.1.8}$$

$$k^\alpha\nabla_\alpha\epsilon_{\mu\nu} = 0 . \tag{6.1.9}$$

[2] In the 1D problem for a scalar field $\phi(x)$ in flat space time then $(-c^{-2}\partial_t^2 + \partial_x^2 + m^2)\phi = 0$.

At this order (ε^{-1}) the Lorenz gauge condition also implies the orthogonality of the polarization along the rays:

$$k^\alpha \epsilon_{\alpha\mu} = 0\,. \tag{6.1.10}$$

So far everything is conceptually the same as in flat space. The next order, ε^0, introduces the first corrections

$$\Box A_{\mu\nu} + 2\bar{R}_{\alpha\mu\beta\nu}A^{\alpha\beta} + i\left(2k^\alpha \nabla_\alpha A^{(1)}_{\mu\nu} + A^{(1)}_{\mu\nu}\nabla^\alpha k_\alpha\right) + k^\alpha k_\alpha A^{(2)}_{\mu\nu} = 0\,, \tag{6.1.11}$$

$$\nabla^\alpha A_{\alpha\mu} - ik^\alpha A^{(1)}_{\alpha\mu} = 0\,, \tag{6.1.12}$$

There are several points to highlight here. The first one is that, in the absence of tensor corrections $A^{(1)}_{\mu\nu}$ and higher, we obtain a curved space-time generalization of the leading order wave propagation:

$$\Box A_{\mu\nu} = 0\,, \tag{6.1.13}$$

where, again, $\Box$ is the D'Alembertian of the background metric. This fully defines the **geometric optics limit** of GWs over curved backgrounds. Beyond geometric optics, the term $2\bar{R}_{\alpha\mu\beta\nu}A^{\alpha\beta}$ would source a (small) correction $A^{(1)}_{\mu\nu}$ (note that because of (6.1.7), the term $k^\alpha k_\alpha A^{(2)}_{\mu\nu}$ vanishes). If $R_{\alpha\mu\beta\nu}$ was to be diagonal in the $+, \times$ polarizations, then this extra term would effectively behave as a mass term. This would introduce dispersion of the waves, i.e. a frequency dependent propagation speed. In general, it can mix the polarization content. Secondly, both the Einstein field equations and the gauge condition beyond leading order in geometric optics (ε^0 and higher) mixes the real and imaginary parts of $A^{(n)}_{\mu\nu}$. This means that beyond geometric optics will correct both the amplitude and phase. Thirdly, and importantly, the gauge condition (6.1.12) implies that beyond geometric optics the polarization tensor is not transverse. Therefore, the polarizations could rotate while propagating.

One of the strengths of the WKB approach is that it allows to systematically solve all higher order corrections. In particular, a term of order (n), $A^{(n)}_{\mu\nu}$, can always be solved in terms of the previous order $(n-1)$:

$$2k^\alpha \nabla_\alpha A^{(n)}_{\mu\nu} + A^{(n)}_{\mu\nu}\nabla^\alpha k_\alpha = i\left(\Box A^{(n-1)}_{\mu\nu} + 2\bar{R}_{\alpha\mu\beta\nu}A^{\alpha\beta}_{(n-1)}\right)\,, \tag{6.1.14}$$

$$k^\alpha A^{(n)}_{\alpha\mu} = -i\nabla^\alpha A^{(n-1)}_{\alpha\mu}\,, \tag{6.1.15}$$

Note also that if $A^{(n-1)}_{\mu\nu}$ is real, then the next correction would be purely imaginary. This would be the case, for example, if one has at leading order a linearly polarized GW.

Exercise 6.1: GW energy-momentum tensor in the short wave expansion

Compute the energy momentum tensor of GWs in the short-wave expansion. What is the leading term? How is it corrected? For doing this please verify first the results presented in the main text for the short-wave expansion.

6.2 Cosmological propagation

The propagation of GWs over curved backgrounds becomes relevant for astrophysical sources such as compact binary coalescences. This is because these mergers occur at cosmological distances (recall the simple estimate (4.3.9)) where the curvature of the background metric is important. Because of their astrophysical nature, these GWs will typically have wavelenghts which are much smaller than the scale at which the cosmological background changes. In particular, for a compact binary its frequency is $f \sim$ few $\times$ 100Hz for a $\sim 10M_\odot$ binary (recall (4.3.8) and the inverse scaling with the mass) or a wavelength of $\lambda = c/f \sim$ few $\times$ 1000km. This typical frequency is much larger than the rate of change of the cosmological background. For present day, the rate of expansion is determined by the Hubble parameter[3]

$$H_0 = h_0 \cdot 100\,\mathrm{km\,s^{-1}\,Mpc^{-1}} \sim h_0 \cdot 10^{-18}\mathrm{Hz}\,, \tag{6.2.1}$$

where current observations indicate that $h_0 \sim 0.7$. The corresponding horizon scale is

$$d_H \equiv \frac{c}{H_0} \sim 3h_0^{-1}\mathrm{Gpc}\,. \tag{6.2.2}$$

Therefore, for this problem the short-wave expansion is totally justified and we aim to obtain the geometric optics solution.

6.2.1 Times and distances in cosmology

At large scales, $\sim$ Gpc, the Universe is approximately homogeneous and isotropic. Therefore, the background metric can be approximately described by a Friedmann-Robertson-Walker (FRW) metric:

$$ds^2 = -c^2dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right]\,, \tag{6.2.3}$$

where $a(t)$ is the scale factor and k the curvature.[4] $k = 0$ indicates a flat Universe (where there is only curvature in the temporal component), $k = +1$ corresponds to a closed Universe and $k = -1$ is an open Universe.

The above metric is expressed in the so-called **comoving** coordinates (t, r, θ, ϕ). Physical distances are therefore stretched by the scale factor. For example, for two points separated by a radial comoving distance r (and fixed angular position), the physical distance is

$$r_{\mathrm{phys}}(t) = a(t)\int_0^r \frac{dr'}{(1-kr'^2)^{1/2}}\,. \tag{6.2.4}$$

[3]Remember that 1 parsec $\simeq 3 \cdot 10^{16}$m and $c \simeq 3 \cdot 10^8$m/s.

[4]The FRW metric is an exact solution when one assumes the **cosmological principle** that states that the Universe is homogeneous and isotropic or, more mathematically, that the spatial metric is maximally symmetric.

For a signal that propagates around null geodesics in the radial direction (again for fixed angles θ, ϕ) between two points separated by the comoving radial distance r and emitted at time $t_{\rm emit}$, it will be observed at time $t_{\rm obs}$ as determined by

$$\int_{t_{\rm emit}}^{t_{\rm obs}} \frac{cdt}{a(t)} = \int_0^r \frac{dr'}{(1-kr'^2)^{1/2}} \,. \tag{6.2.5}$$

If a second signal is emitted $\Delta t_{\rm emit}$ later, it would be observed with a delay $\Delta t_{\rm obs}$ similarly fixed by

$$\int_{t_{\rm emit}+\Delta t_{\rm emit}}^{t_{\rm obs}+\Delta t_{\rm obs}} \frac{cdt}{a(t)} = \int_0^r \frac{dr'}{(1-kr'^2)^{1/2}} \,. \tag{6.2.6}$$

Therefore, equating the left-hand sides of both equations and expanding to linear order in Δt, which is justified when the period of the signal is much smaller than the variation time scale of $a(t)$, one finds

$$\Delta t_{\rm obs} = \frac{a(t_{\rm obs})}{a(t_{\rm emit})} \Delta t_{\rm emit} \,. \tag{6.2.7}$$

This naturally leads to the definition os the **cosmological redshift** z:

$$1 + z = \frac{a(t_{\rm obs})}{a(t_{\rm emit})} \,. \tag{6.2.8}$$

In the same way that time is dilated by the cosmic expansion, the observed frequency $f_{\rm obs}$ will be redshifted compared to the source frequency $f_{\rm src}$:

$$f_{\rm obs} = \frac{f_{\rm src}}{1+z} \,. \tag{6.2.9}$$

When particularized to this FRW background and energy-momentum tensor composed of perfect fluids, the Einstein field equations reduce to the Friedmann equations. For us, the most important one is the one relating the Hubble parameter $H(t) \equiv d\ln a(t) dt$ and the energy density of the perfect fluid and curvature

$$H^2(t) = \left(\frac{d\ln a(t)}{dt}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} \,. \tag{6.2.10}$$

If we define the fractional energy density of a given component X with respect to the critical energy density $\rho_c = 8\pi G/3H_0^2$ as $\Omega_X = \rho_X/\rho_c$, and we particularize to our Universe made of (dark) matter (pressureless fluid $p = 0$, $\rho \propto a^{-3}$), radiation ($p = \rho/3$, $\rho \propto a^{-4}$), cosmological constant ($p = -\rho = -\Lambda$) and curvature ($\rho_k = -3kc^3/8\pi G a^2$), then the above equation takes the simple form

$$H^2(t) = H_0^2 \left(\Omega_m (1+z)^3 + \Omega_r (1+z)^4 + \Omega_\Lambda + \Omega_k (1+z)^2\right) \,, \tag{6.2.11}$$

where we have substituted the scale factor dependence by the redshift. For the purpose of our following studies, the GW propagation from astrophysical compact binaries, then

we can restrict to the late Universe where $\Omega_r \sim 10^{-5}$ has a negligible contribution. Moreover, current observations also indicate that the Universe is spatially flat to high accuracy, $\Omega_k \sim 10^{-3}$, so we are left with two components, the (dark) matter $\Omega_m \sim 0.3$ and the dark energy $\Omega_\Lambda \sim 0.7$ (here presented in the form of a cosmological constant).

In cosmology there are different notions of distance (see [44] for a pedagogical review). Again, for a signal propagating around null geodesics, the comoving distance is obtained from solving (6.2.5). For a flat background this is

$$d_C^{k=0} = \int \frac{cdt}{a(t)} = \int \frac{cdz}{H(z)} = d_H \int \frac{dz}{E(z)}\,, \tag{6.2.12}$$

where in the second equality we have simply changed integration variables and in the last equality we have introduced the normalized Hubble parameter $E(z) \equiv H(z)/H_0$. For general curvatures, the comoving distance is defiend by

$$d_C = \frac{c}{H_0\sqrt{|\Omega_k|}}\mathrm{sink}\left(H_0\sqrt{|\Omega_k|}\int \frac{dz}{H(z)}\right) = \frac{d_H}{\sqrt{|\Omega_k|}}\mathrm{sink}\left(\sqrt{|\Omega_k|}\frac{d_C^{k=0}}{d_H}\right)\,, \tag{6.2.13}$$

where $\mathrm{sink}(x)$ is a shortcut for $\sinh(x)$, x and $\sin(x)$ for $k=+1$, $k=0$ and $k=-1$ respectively. The two other important distances are the angular diameter distance

$$d_A = \frac{d_C}{1+z}\,, \tag{6.2.14}$$

and the luminosity distance

$$d_L = (1+z)d_C = (1+z)^2 d_A\,. \tag{6.2.15}$$

These distances arise naturally when measuring angular separations and fluxes respectively. They are also a direct product of the geometric optics formalism.

6.2.2 GW redshift and damping

After a quick summary of some basic concepts in cosmology, we are ready to study the GW propagation over a cosmological background in the geometric optics approximation. The first thing we learned from the short-wave expansion is that GW will follow null geodesics. Therefore, GWs will suffer cosmological redshifting. While they propagate from the **source's frame** to the **detector's frame**, GWs will be stretched by the cosmic expansion. For a compact binary during the inspiral phase, the detected GW frequency $f_{\rm det}$ will be related to the source frame frequency $f_{\rm src}$, which was given by (5.2.31), as

$$\begin{aligned} f_{\rm det}(\tau_{\rm det}) &= \frac{f_{\rm src}(t_{\rm src})}{1+z} = \frac{1}{(1+z)\pi}\left(\frac{5}{256}\frac{1}{\tau_{\rm src}}\right)^{3/8}\left(\frac{G\mathcal{M}_c}{c^3}\right)^{-5/8} \\ &= \frac{1}{\pi}\left(\frac{5}{256}\frac{1}{\tau_{\rm det}}\right)^{3/8}\left(\frac{G\mathcal{M}_c^{\rm det}}{c^3}\right)^{-5/8}\,, \end{aligned} \tag{6.2.16}$$

where $\tau_{\rm det} = (1+z)\tau_{\rm src}$ is the time to coalescence in the observer's frame. Note that in the last equality we have introduced the **detector frame mass**

$$m_{\rm det} = (1+z)m\,, \tag{6.2.17}$$

which applied to the chirp mass is $\mathcal{M}_c^{\rm det} = (1+z)\mathcal{M}_c$. These masses are also known as redshifted masses and denoted by m_z and $\mathcal{M}_z$ respectively. Interestingly, once we introduce the detector frame masses the formula for the frequency evolution is formally the same if one replaces the time and masses from the source to the detector's frame.

The second thing we learned from geometric optics is that the polarization tensor is parallel transported. Therefore, if we decompose initially the GW into the $+,\times$ polarizations,

$$\bar{h}_{\mu\nu}(x) = h_+(x)\epsilon^+_{\mu\nu} + h_\times(x)\epsilon^\times_{\mu\nu}\,, \tag{6.2.18}$$

those polarization will not change along the propagation. When we apply this into the propagation equation $\Box A_{\mu\nu} = 0$, then we find that the tensor modes decouple:

$$\Box h_{+,\times} = 0\,. \tag{6.2.19}$$

Each of the physical, radiative modes follow the same wave equation. The advantage is that this is a (simpler) scalar wave equation:

$$\Box h_{+,\times} = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}g^{\mu\nu}\nabla_\nu h_{+,\times}) = 0\,, \tag{6.2.20}$$

since we avoid having to deal with the Levi-Civita connection.

We are therefore left with the task of solving the scalar wave equation $\Box h_A = 0$ where $A = +,\times$. It will be convenient to work in conformal time, $\eta = \int dt/a(t)$, so that the FRW metric reads

$$ds^2 = a^2(\eta)[-c^2 d\eta^2 + dr^2 + r^2 d\Omega] \tag{6.2.21}$$

for flat spatial backgrounds. This choice of coordinates is convenient because then this metric is conformally equivalent to flat space $g_{\mu\nu} = a^2(\eta)\eta_{\mu\nu}$. The determinant of the metric is simply $g = -a^8(\eta)c^2r^4\sin^2\theta$ and $\sqrt{-g} = a^4(\eta)cr^2\sin\theta$. Since we have already seen that the GW amplitude decays as the inverse to the distance $1/r$ and the background is conformally equivalent to flat space, we search for spherically symmetric solutions of the wave equation of the form $h_A(\eta, r) = f_A(\eta, r)/a(\eta)r$. With this ansatz, we can solve the wave equation as:

$$\begin{aligned} 0 &= \partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu h_A) \\ &= -c^{-2}\partial_\eta(a^2r^2\partial_\eta h_A) + \partial_r(a^2r^2\partial_r h_A) \\ &= -(af_A)'' - 2\frac{a'}{a}(af_A)' + \partial_r^2(af_A) \\ &= -f_A'' + \frac{a''}{a}f_A + \partial_r^2 f_A\,, \end{aligned} \tag{6.2.22}$$

where primes denote derivatives with respect to $c\eta$. If we look for approximate wave solutions

$$f_A(\eta, r) = e^{\pm i\omega(\eta - r/c)}\,, \tag{6.2.23}$$

we can realize that the term a''/a is subdominant provided that the background varies slowly compared to the wave: $a''/a \ll \omega^2$ (which is indeed satisfied in the short-wave expansion!). Therefore, we arrive at the wave equation $\partial_r^2 f_A - f_A'' = 0$ whose general solution is of the form $f_A(\eta - r/c)$. Altogether we have

$$h_A(\eta, r) \simeq \frac{1}{a(\eta) r} f_A(\eta - r/c)\,. \tag{6.2.24}$$

If we define the conformal time so that in the present $\eta = t$, then

$$h_A(t, r) \simeq \frac{1}{a(t_0) r} f_A(t - r/c)\,, \tag{6.2.25}$$

where t_0 is the present (cosmological time) and we have further approximated that the scale factor is approximately constant during the duration of the signal that is parametrized by t. Therefore, we arrive at the conclusion that the due to the expansion of the Universe the GW amplitude scales with the physical distance $1/r_{\rm phys} = 1/a \cdot r$. In practical terms this means that we can replace $r \to a \cdot r$ in our previous calculations.

We want to investigate how this affects the GWs emitted by a compact binary coalescence. If we focus first in the polarization independent term of the GW amplitude $\bar{A}(t)$, $h_+ = \bar{A}(t)(1 + \cos^2\theta)\cos\Phi(t)/2$ and $h_\times = \bar{A}(t)\cos\theta\sin\Phi(t)$, we have that in the detector's frame (recalling to redshift the frequencies)

$$\begin{aligned} \bar{A}(t_{\rm det}) &= \frac{4}{a(t_0) r}\left(\frac{G\mathcal{M}_c}{c^2}\right)^{5/3}\left(\frac{\pi(1+z) f_{\rm det}(t_{\rm det})}{c}\right)^{2/3} \\ &= \frac{4}{d_L(z)}\left(\frac{G\mathcal{M}_z}{c^2}\right)^{5/3}\left(\frac{\pi f_{\rm det}(t_{\rm det})}{c}\right)^{2/3}\,. \end{aligned} \tag{6.2.26}$$

We find that GWs scale inversely with the luminosity distance $d_L = (1+z)a(t_0)r$. For the rest the expression is formally the same if we replace the chirp mass by its redshifted version. Then, it is easy to realize that the same will happen for the phase evolution

$$\Phi(\tau_{\rm det}) = -2\left(\frac{5G\mathcal{M}_z}{c^3}\right)^{-5/8}\tau_{\rm det}^{5/8} + \Phi_0\,. \tag{6.2.27}$$

This concludes our derivation of the GW signal from a cosmological compact binary merger.

Exercise 6.2: GW damping by the Hubble friction

If one introduces the spatial wave vector k, and looks for wave solutions $h_A \propto e^{i\vec{k}\vec{x}}$,

the cosmological wave propagation can also be written as

$$h_A'' + 2\mathcal{H}h_A' + c^2k^2a^2h_A = 0\,, \tag{6.2.28}$$

where $\mathcal{H} \equiv a'/a$ is the Hubble parameter in terms of conformal time. We can also rewrite this equation in terms of comoving time t:

$$\ddot{h}_A + 3H\dot{h}_A + c^2k^2h_A = 0\,, \tag{6.2.29}$$

where $\dot{h} = c^{-1}\partial h/\partial t$. From this point of view, the expansion of the Universe $H > 0$ introduces a damping of the GWs, which is sometimes referred as the **Hubble friction**. Solve this equation using the WKB approximation. What is the leading order solution? How is it corrected?

Explore: Scalar-Vector-Tensor decomposition

We have seen that in geometric optics the tensor polarizations of the GW decouple from each other. This is a consequence of the polarization tensor being parallel transported along null geodesics. But what happens if we have additional fields? Could they couple to the GWs over a cosmological backgrounds? The answer is that, at linear order, a GW could only couple with other additional tensor modes. This is because in cosmological perturbation theory, the symmetries of the FRW lead to a decoupling of the scalar, vector and tensor modes. This is known as the **scalar-vector-tensor decomposition**. It is not difficult to prove!

Explore: GWs from the early Universe

So far we have focused on the case in which the GWs have an astrophysical origin and, therefore, their wavelength is much smaller than the curvature scale. We then applied the short-wave expansion. However, in the early Universe, the GWs generated could have had comparable (or even larger!) scales to the background curvature. This is precisely the case of the tensor perturbations generated during inflation, the period of rapid accelerated expansion in the first instants of the Universe. During inflation, the scale factor $a(t)$ grows exponentially. Therefore both the background connection $\bar{\Gamma}^{\mu}_{\nu\alpha}$ and the Riemann tensor $\bar{R}_{\mu\alpha\nu\beta}$ in (6.1.5) rapidly vary. This can lead to parametric amplifications of the tensor perturbations $h_{\mu\nu}$ that could source a stochastic GW background. From the quantum mechanical perspective, defining a quantum field theory on a curved spacetime implies that, in general, there is not a well defined notion of vacuum state and, therefore, of number of particles. The parametric amplification can then be interpreted as a large particle production. This fascinating phenomenon was first pointed out by Schrodinger in 1939 [45] and later developed in full details by many others. Note that the same occurs for the scalar perturbations (inflaton modes) produced

during inflation. Vector modes however rapidly decay during inflation. Quantum fluctuations during inflation in the inflaton and metric fields can then be stretched to cosmological scales, producing the inhomegeneities in the spatial distribution of density perturbations (matter and radiation) that serve as seeds for all the structures in the Universe.

6.3 Gravitational lensing of gravitational waves

In the previous section we have solved the GW propagation under the short-wave approximation in the context of a cosmological background in which the curvature scale is much larger than the GW wavelength. We now wish to consider another example in which we solve the curved spacetime propagation for wavelengths that could be comparable to the curvature scale. In full generality this is a very difficult problem to solve and, for that reason, we restrict to the limiting of **weak-field** gravity.

If we are interested in the weak-field limit, we can expand the background metric around flat space-time[5]

$$\bar{g}_{\mu\nu} = \eta_{\mu\nu} + \delta g_{\mu\nu} \tag{6.3.1}$$

and take $|\delta g_{\mu\nu}| \ll 1$. As you are already guessing, this is the same formalism of linearized gravity that we used to derive the GW propagation! As we discussed at the beginning of §5, the metric perturbation can have at most 10 components that after fixing the gauge reduce to 6. We now take a slightly different route and decompose δg_{ij} into two scalar modes Φ and Ψ, a vector mode w_i and a traceless tensor s_{ij}:

$$\delta g_{00} = -2\Phi\,, \tag{6.3.2}$$

$$\delta g_{0i} = w_i\,, \tag{6.3.3}$$

$$\delta g_{ij} = 2s_{ij} - 2\Psi\delta_{ij}\,. \tag{6.3.4}$$

This defines a general metric for linearized gravity. Note that w_i is a spatial vector and, therefore, contains 3 degrees of freedom. s_{ij} is defined as traceless, symmetric, spatial tensor and, as a consequence, only has 2 degrees of freedom. In this sense Ψ represents the trace of the spatial metric perturbations.

We now restrict to the case of a static source of mass density ρ, whose energy-momentum tensor is given by $T_{\mu\nu} = \mathrm{diag}(\rho, 0, 0, 0)$. If we introduce this ansatz into the Einstein field equations, one finds that the tensor and vector modes vanish and that the scalar modes are equal to each other $\Phi = \Psi$ and follow a non-radiative equation:

$$\nabla^2\Phi = 4\pi G\rho\,, \tag{6.3.5}$$

which is nothing but the Poisson equation describing Newtonian mechanics. In this case the weak-field metric reduces to

$$ds^2 = -(1+2\Phi)dt^2 + (1-2\Phi)d\vec{x}^2\,. \tag{6.3.6}$$

[5]For the derivation of the weak-field limit I follow Carroll's chapter 7 [2].

This is the metric that we are going to consider as a background for the GW propagation.

In order to solve the curved space-time propagation of GWs in (6.1.5), we need to compute $\Box$ and $\bar{R}_{\alpha\mu\beta\nu}$ around the background under consideration. We have already seen that for a weak-field, static, mass density only the scalar modes are excited. In terms of the Riemann tensor, this means that the only non-zero components are $\bar{R}_{j0k0} = -\bar{R}_{j00k} = \partial_j\partial_k\Phi$. In other words, $\bar{R}_{\alpha\mu\beta\nu}$ will not couple with the GWs which over an approximately flat metric (a very good approximation far away from the mass density) are h_{ij}^{TT}.

We are therefore just left with computing the d'Alembertian operator on the GWs. The main difference from before is that we will not be expanding the equations in the short-wave limit $\lambda\!\!\!^{-} \ll L_B$ (where as before L_B indicates the typical scale of the background curvature), although the weak-gravity approximation will bring other simplifications. In the weak-field limit of a static mass density source the Christoffel symbols are given by

$$\Gamma^0_{0i} = \Gamma^i_{00} = \partial_i\Phi\,, \tag{6.3.7}$$

$$\Gamma^i_{jk} = \delta_{jk}\partial_i\Phi - \delta_{ik}\partial_j\Phi - \delta_{ij}\partial_k\Phi\,, \tag{6.3.8}$$

where in both equations we are exploiting the fact that at linear order in Φ we raise indices with the flat metric $\eta_{\mu\nu}$ (or in other words the inverse metric is given by $\delta g^{\mu\nu} = \eta^{\mu\alpha}\eta^{\nu\beta}\delta g_{\alpha\beta}$). Therefore, for the spatial indices we do not care if they are up or down. The Christoffel symbols are therefore of order $\Gamma \sim \partial\Phi \sim |\Phi|/L_B$. This means that we will neglect terms of order Γ^2. We would like to decompose the metric perturbations into a scalar amplitude (describing the shape of the waveform) and a tensorial part (describing the polarization content) in order to study their wave equations separately, if possible. Our ansatz is $h_{\mu\nu} = h_A(x)\epsilon_{\mu\nu}(x)$, where, in principle, both the amplitude and polarization tensor could vary. The full propagation equation is then

$$\begin{aligned} 0 = \Box h_{\mu\nu} &= \Box(h_A)\epsilon_{\mu\nu} + h_A\Box(\epsilon_{\mu\nu}) \\ &\sim \left(\partial^2 h_A + \Gamma\partial h_A\right)\epsilon + h_A\left(\partial^2\epsilon + \Gamma\partial\epsilon + \partial\Gamma\epsilon + \Gamma^2\epsilon\right) \\ &\sim \left(\partial^2 h_A + \Gamma h_A\right)\epsilon + h_A\left(\partial^2\epsilon + \Gamma\partial\epsilon\right) + \mathcal{O}\left(\Phi^2, \Phi/L_B^2\right) \end{aligned} \tag{6.3.9}$$

where in the second line we have schematically written the type of terms that appear in the expression. We are interested in the weak-field limit in which the waves travel far away from the mass density and the Newtonian approximation is valid. If we think of a spherically symmetric mass density then $\Phi \sim r_{\rm Sch}/r$ and we are in the limit of $r \gg r_{\rm Sch}$. Terms that involve derivatives of this potential will therefore be highly suppressed. For example from the Christoffel symbols we get $\Gamma \sim \partial\Phi \sim r_{\rm Sch}/r^2 \ll 1/r_{\rm Sch}$. In other words, weak-field limit implies that L_B is large. This is why in the third line we neglect higher order terms and keep the leading order (small) corrections $\sim \Phi/L_B$ (note that we have not assumed anything about $\lambda\!\!\!^{-}$). Altogether, we see that the trajectory of the wave can be slightly deflected and the polarizations slightly rotated. Because GW detectors are much more sensitive to the phase evolution of the wave than its sky location and polarization content (we will learn about this in chapter 7), we will focus on h_A.

Given the above discussion, our task now is to solve

$$\begin{aligned} 0 = \Box h_A &= \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}g^{\mu\nu}\partial_\nu h_A\right) \\ &= g^{00}\partial_0^2 h_A + \frac{1}{\sqrt{-g}}\sum_{i=1,2,3}\partial_i\left(\sqrt{-g}g^{ii}\partial_i h_A\right) \\ &\simeq (-1+2\Phi)\partial_0^2 h_A + \sum_i\left((1+2\Phi)\partial_i^2 h_A + 4\Phi\partial_i\Phi\partial_i h_A\right) \\ &\simeq (-1+2\Phi)\partial_0^2 h_A + (1+2\Phi)\nabla^2 h_A\,, \end{aligned} \tag{6.3.10}$$

where in the first equality we have used the same identity that we used when solving the cosmological propagation, cf. (6.2.20), and in the second line we have imposed that Φ is static. Noting that $g^{00} = -1/(1+2\Phi)$, $\sqrt{-g} = (1+2\Phi)^{1/2}(1-2\Phi)^{3/2}$ and $\sqrt{-g}g^{ii} = (1+2\Phi)^{1/2}(1-2\Phi)^{1/2} = \sqrt{(1-4\Phi^2)}$, and expanding to leading order in Φ (e.g. $g^{00} \simeq -1+2\Phi$), we can obtain the third line. We have already argued that $\partial\Phi \sim \Phi/L_B$ terms are small, so then we get the last equality, where ∇^2 is the flat space Laplacian operator. At linear order in Φ, we can rewrite the equation as

$$0 = \nabla^2 h_A - (1-4\Phi)\partial_0^2 h_A\,. \tag{6.3.11}$$

Because the potential is static it is useful to bring this to Fourier space $h_A(t,\vec{x}) = \int d\omega \tilde{h}(\omega,\vec{x})e^{-i\omega t}$ so that

$$\left(\nabla^2 + \omega^2\right)\tilde{h}_A = 4\Phi\omega^2\tilde{h}_A\,. \tag{6.3.12}$$

This is a Helmholtz equation commonly appearing in Mathematics and Physics. We are looking to solve the effect of the Newtonian potential on the wave propagation. Therefore, it is common to define an **amplification function** F that encodes this information:

$$F(\omega,\vec{x}) \equiv \tilde{h}_A(\omega,\vec{x})/\tilde{h}_{A,0}(\omega,\vec{x})\,, \tag{6.3.13}$$

where $h_{A,0}(\omega,\vec{x})$ is the solution of the wave equation when $\Phi = 0$. Without the potential, the wave will not be lensed and therefore can be chosen to be a spherical wave $\tilde{h}_{A,0}(r = |\vec{x}|) \propto e^{i\omega r}/r$. Plugging these definitions into the above equation we obtain

$$\begin{aligned} 0 &= \left(\nabla^2+\omega^2\right)F\tilde{h}_{A,0} - 4\Phi\omega^2 F\tilde{h}_{A,0} \\ &= \nabla^2(F)\tilde{h}_{A,0} + 2\partial^i F\partial_i\tilde{h}_{A,0} + F\left(\nabla^2+\omega^2\right)\tilde{h}_{A,0} - 4\Phi\omega^2 F\tilde{h}_{A,0} \\ &= \left(\partial_r^2 F + \frac{1}{r^2}\nabla_\theta^2 F\right)\tilde{h}_{A,0} + 2i\omega\partial_r(F)\tilde{h}_{A,0} - 4\Phi\omega^2 F\tilde{h}_{A,0}\,, \end{aligned} \tag{6.3.14}$$

where in the first line we have simply introduced our new function F. In the second line we have applied the chain rule and in the third line we have used that $\tilde{h}_{A,0}$ is a solution of the propagation without a lens and it is spherically symmetric. Because of the spherical symmetry, we have also expanded the Laplacian operator into its radial r and angular components (θ,ϕ):

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \tag{6.3.15}$$

where ∇^2_θ is the angular Laplacian operator on the 2D sphere encoding the last two terms. Note that there are two terms $(2/r)\tilde{h}_{A,0}\partial_r F$ of opposite signs that cancel each other. We see that the remaining equation for the amplification factor is

$$\partial_r^2 F + \frac{1}{r^2}\nabla^2_\theta F + 2i\omega\partial_r F = 4\Phi\omega^2 F\,. \tag{6.3.16}$$

To further simplify this equation we note that because of the weak-field limit, the deflection angles will be small. Therefore, if we situate the origin at the source, and the lens at the polar axis, an observer at (r_o, θ_o, ϕ_o) will have $\theta_o \ll 1$. This means that we can simplify $\sin\theta \simeq \theta$ and effectively look at $\vec{\theta} = \theta(\cos\phi, \sin\phi)$ as a two-dimensional vector in a flat plane. This also means that our angular Laplacian simplifies to $\nabla^2_\theta \to \partial^2_\theta = \partial^2_\theta + \theta^{-1}\partial_\theta + \theta^{-2}\partial^2_\phi$.

If we assume that the wave travels far from the center of the lens at a distance D, far compared to the wavelength ($D \gg \lambda\!\!\!^{-}$), then $\partial_r^2 F$ will be subdominant compared to $2i\omega\partial_r F$. This is equivalent to say that $\lambda\!\!\!^{-}$ is small compared to the scale at which F varies ($\omega \gg |\partial_r \ln F|$). This has the advantage that then we get partial differential equation to first order in r, which can then be interpreted as a Schrödinger like equation:

$$i\partial_r F = -\frac{1}{2\omega r^2}\partial^2_\theta F + 2\Phi\omega F\,, \tag{6.3.17}$$

where r is the equivalent to time, and Φ to the 'time' dependent potential. To make more explicit the comparison with the usual Schrödinger equation, $i\hbar\frac{\partial}{\partial t}\Psi(\vec{x},t) = \hat{H}\Psi(\vec{x},t) = -\frac{\hbar^2}{2m}\nabla^2\Psi(\vec{x},t) + V(\vec{x})\Psi(\vec{x},t)$, we can rewrite the above expression

$$i\frac{1}{\omega}\partial_r F = -\frac{1}{2\omega^2 r^2}\partial^2_\theta F + 2\Phi F\,, \tag{6.3.18}$$

where the quantization (or wave-like behavior) come from the fact ω is finite. The Lagrangian giving rise to the classical equations of motion is

$$\mathcal{L}(\vec{\theta}, \dot{\vec{\theta}}, r) = \frac{r^2}{2}|\dot{\vec{\theta}}|^2 - 2\Phi(r, \vec{\theta})\,, \tag{6.3.19}$$

where to make the analogy even further we denote derivatives w.r.t. r as $\dot{\vec{\theta}} \equiv d\vec{\theta}/dr$. One can then use the path integral formulation of quantum mechanics to write the formal solution[6]

$$F(\omega, \vec{x}_o) = \int \mathcal{D}\vec{\theta}\exp\left[i\omega\int_0^{r_o} dr\mathcal{L}(\vec{\theta}, \dot{\vec{\theta}}, r)\right]\,, \tag{6.3.20}$$

at the observers position $\vec{x}_o$. Note that here $1/\omega$ is playing the role of $\hbar$, so that the classical limit $\hbar \to 0$ is achieved when $\omega \to \infty$. Therefore, the analogous of the quantum limit in lensing is known as **wave optics**. When taking the frequency as a common denominator in the exponential, then we can think of the rest as a time delay surface.

[6]The path integral formulation of gravitational lensing is presented pedagogically in Nakamura & Deguchi [46].

The kinetic and potential terms of the Lagrangian then correspond to the **geometric** and **Shapiro delays** respectively. The geometric time delay carries the lag associated to the change in the trajectory

$$t_{\rm geo} = \int \frac{r^2|\dot{\vec{\theta}}|^2}{2} dr \,, \tag{6.3.21}$$

while the Shapiro delays accounts for the gravitational potential

$$t_\Phi = -\frac{2}{c^3}\int \Phi dr \,, \tag{6.3.22}$$

where we have added back the factors of c.

Next we are going to use the **thin lens approximation**, in which the extent of the lens in the radial direction is small compared to the distance to the lens from the observer r_L and the distance between the source and the lens r_{LS}. Similarly, the distance from the observer to the source r_S will be large. Under this approximation then the potential only contributes effectively at a single distance and we can substitute

$$\psi(\vec{\theta}_L) = \frac{2}{c^3}\int \Phi(r,\vec{\theta}) dr \,. \tag{6.3.23}$$

Because the lens is thin compared to the paths that the wave travels before and after, the phase contribution is going to be dominated by a constant vector $\vec{\theta}(r) \approx \vec{\theta}_l$. With this, the geometric time delay simplifies to

$$t_{\rm geo} \approx \frac{r_L r_S}{2r_{LS}c}|\vec{\theta}_L - \vec{\theta}_S|^2 \,, \tag{6.3.24}$$

where $\vec{\theta}_S$ is the angular position of the source from the observer perspective, which is equivalent to our previously defined angle of the observer from the source perspective $\vec{\theta}_o$ to linear order. The thin lens approximation also makes that we can simplify the integral over all paths $\mathcal{D}\vec{\theta}$ to a 2D integral $d^2\theta_L$ so that

$$F(\omega,\vec{\theta}_S) \approx \frac{\omega}{2\pi i}\frac{r_L r_S}{c r_{LS}}\int d^2\theta_L \exp\left[i\omega\left(\frac{r_L r_S}{2c r_{LS}}|\vec{\theta}_L - \vec{\theta}_S|^2 - \psi(\vec{\theta}_L)\right)\right] \,, \tag{6.3.25}$$

where the normalization is such that in the absence of a potential, $\Phi = 0$, then $F = 1$. After a somewhat long detour we have found that the wave propagation around a lens in the weak-field limit can be solved in terms of a a Kirchhoff diffraction integral.

In the cosmological context, we have seen that time is dilated, then we will need to replace $t \to (1+z_L)t$. Similarly, distances will be substituted by angular diameter distances $r_L \to D_L$. Therefore we will have

$$F(\omega,\vec{\theta}_S) = (1+z_L)\frac{D_L D_S}{c D_{LS}}\frac{\omega}{2\pi i}\int {\rm d}^2\theta \;\exp\Big[i(1+z_L)\omega t_d(\vec{\theta},\vec{\theta}_S)\Big] \,, \tag{6.3.26}$$

for later convenience we introduce the time scale associated to the angular diameter distances $\tau_D \equiv (1+z_L)D_L D_S/cD_{LS}$. The time delay t_d is determined by

$$(1+z_L)t_d(\vec{\theta},\vec{\theta}_S) = \frac{\tau_D}{2}|\vec{\theta} - \vec{\theta}_S|^2 + (1+z_L)t_\Phi \,. \tag{6.3.27}$$

Lensing typically occurs at a reference angular scale, θ_*. If we redefine the coordinates appropriately

$$\vec{x} \equiv \vec{\theta}/\theta_* \,, \quad \vec{y} \equiv \vec{\theta}_S/\theta_* \,, \quad w \equiv \tau_D \theta_*^2 \omega \,, \tag{6.3.28}$$

then the amplification factor takes a more compact form

$$F(w, \vec{y}) = \frac{w}{2\pi i} \int \mathrm{d}^2 x \;\; \exp[iwT_d(\vec{x}, \vec{y})] \,, \tag{6.3.29}$$

in terms of the dimensionless frequency w and time delay $T_d \equiv t_d/\tau_D\theta_*^2$.

It is interesting to connect this path integral formulation with **Fermat's principle**. Among all the possible paths that the wave could take, Fermat's principle states that the time delay is extremized when the path match the actual rays.[7] As we have seen before, in the geometric optics limit the rays are defined as the null geodesic where the wave propagates. Therefore, in this limit, the amplification factor is telling us that there will be distinct rays dominating the integral. This is equivalent to having **multiple images** of the original source. However, in general, the diffraction integral also implies that multiple paths could interfere around the rays leading to **diffraction**. We will study these two regimes next.

6.3.1 Multiple images

The amplification factor F accounts for the time delay associated for all possible lensed paths from the source to the observer, which in general can lead to complicated interference or diffraction patterns. There is a limit, however, in which the integral in F has a highly oscillatory exponent. In this case the integral is dominated by its stationary points

$$\left.\frac{\partial t_d}{\partial \vec{\theta}}\right|_{\vec{\theta}=\vec{\theta}_j} = 0 \,, \tag{6.3.30}$$

and lensing is characterized by having distinct images at the locations $\vec{\theta}_j$. For sufficiently strong lenses, multiple images are produced. In this limit, the amplification factor can be solved using the stationary phase approximation (SPA). The derivation is completely analogous to what we did in §5.2.1 to derive the frequency domain inspiral waveform. Therefore, we leave the derivation as an exercise. The result is

$$F(\omega, \vec{\theta}_j) \approx \sum_j \sqrt{|\mu_j|} \exp\left[i\omega t_j - i\mathrm{sign}(\omega) n_j \frac{\pi}{2}\right] \,, \tag{6.3.31}$$

where the magnifications are obtained from the determinant of the Hessian matrix evaluated at the stationary points

$$\mu(\theta_j) = 1/\det(T_{ab}(\theta_j)) \,, \tag{6.3.32}$$

[7] The connection between Fermat's principle and gravitational lensing and the subsequent derivation of the properties of the images from the time delay surface is nicely presented in the seminal paper of Blandford & Narayan [47].

where

$$T_{ab} \equiv \tau_D^{-1} \partial^2 t_d / \partial\theta_a \partial\theta_b \,. \tag{6.3.33}$$

Lensing is described by a set of j images each arriving at a different time $t_j \equiv t_d(\vec{\theta}_j)$, with a different magnification $\mu_j \equiv \mu(\vec{\theta}_j)$ and Morse phase $n_j = 0, 1, 2$ for type I, II and III images respectively. The different Morse phases corresponds to the different type of solutions of the lens equation (6.3.30): minimum, saddle point and maximum. The different phases come into play because in the SPA we look at the leading order quadratic correction to the phase. Therefore, the sign of T_{ab} around the stationary points matters when solving the Gaussian integral. Note that because we are considering both positive and negative frequencies, the term $\text{sign}(\omega)$ is only there to ensure that the lensed signal in time domain is real. We could avoid this term by restricting to positive frequencies and imposing that $F^*(\omega) = F(-\omega)$. Apart from the phase shift of type II images that can induce waveform distortions on signals with multiple frequency components,[8] lensing is achromatic in the SPA or, in other words, the amplification factor changes the waveform in a frequency independent way. Explicitly in terms of our lensed time domain waveform

$$h_L(t) = \int df \tilde{h}_L(f) e^{-i2\pi f t} \,, \tag{6.3.34}$$

where the frequency domain is given by

$$\begin{aligned} \tilde{h}_L(f) &= F(f, \vec{\theta}) \tilde{h}(f) \\ &\approx \sum_j \sqrt{|\mu_j|} \exp\left[i\omega t_j - i \text{sign}(\omega) n_j \frac{\pi}{2} \right] \frac{A}{d_L} e^{i\Phi(t_c, \phi_c)} \\ &= \sum_j \frac{A\sqrt{|\mu_j|}}{d_L} e^{i\Phi(t_c + t_j, \phi_c + n_j \pi/4)} \\ &\equiv \sum_j \frac{A}{\tilde{d}_{L,j}} e^{i\Phi(\tilde{t}_{c,j}, \tilde{\phi}_{c,j})} \,. \end{aligned} \tag{6.3.35}$$

In the multiple image regime for the leading quadrupolar radiation, lensing produces multiple images each of them with a different effective distance, time of coalescence and phase:

$$\tilde{d}_{L,j} = d_L / \sqrt{|\mu_j|} \,, \tag{6.3.36}$$

$$\tilde{t}_{c,j} = t_c + t_j \,, \tag{6.3.37}$$

$$\tilde{\phi}_{c,j} = \phi_c + n_j \pi / 4 \,. \tag{6.3.38}$$

The SPA is satisfied when the arrival time difference between the stationary points is larger than the duration of the signal, i.e.

$$|\Delta t_d \cdot \omega| \gg 1 \,. \tag{6.3.39}$$

[8] In chapter 5 we focused on the leading quadrupolar radiation. However, higher order modes are also present on GWs signals. Each of them are characterized by a different frequency, this is why the frequency independent phase could play a role

It is to be noted that the geometric optics (or eikonal) approximation, determined by the wavelength of the wave being smaller than the size of the lens curvature, $\lambda/\mathcal{R} \ll 1$, encompasses the SPA. However, there are source-lens configurations in geometric optics where the SPA is broken, namely at the caustics where the Hessian matrix T_{ab} becomes singular. At those locations a single, highly magnified image is formed and the full diffraction integral needs to be solved. The other relevant limit in lensing is that of strong gravity, where $\Phi \sim 1$, and our initial weak-field assumption is broken. In practice, astrophysically this only occurs very close to extremely dense objects such as black holes, whose characteristic length scale is given by the Schwarszchild radius $r_{\rm Sch} = 2GM/c^2$. If we model such black hole as a point lens, its characteristic lensing scale is no other than the Einstein radius

$$R_E \approx \theta_E D_L = \sqrt{\frac{2(1+z_L) r_{\rm Sch}}{c\tau_D}} D_L = \sqrt{\frac{\tau_M}{\tau_D}} D_L \,, \tag{6.3.40}$$

where in the last equality we have introduced the dilated Schwarszchild diameter crossing time $\tau_M = D_{\rm Sch}/c$. The ratio of the two scales is then

$$\frac{R_E}{D_{\rm Sch}} = \sqrt{\frac{D_L D_{LS}}{D_{\rm Sch} D_S}} \,. \tag{6.3.41}$$

In other words, weak-field lensing occurs at scales much larger than the strong gravity regime unless the source is very close to the lens, $D_{LS} \ll D_L, D_S$, as it is the case of the photons near the super massive black holes imaged by the Event Horizon Telescope. For typical astrophysical lensing situations in which $D_{LS} \approx D_L \approx D_S$, then

$$\frac{R_E}{D_{\rm Sch}} \approx 10^{10} \left(\frac{10 M_\odot}{(1+z_L)M}\right)^{1/2} \left(\frac{D_{LS}}{1{\rm Gpc}}\right)^{1/2} \,, \tag{6.3.42}$$

and strong gravity can be safely neglected for all practical purposes.

Exercise 6.3: Gravitational lensing in the stationary phase approximation

Use the stationary phase approximation (SPA) to obtain the multiple image limit of the diffraction integral.

6.3.2 Diffraction

Whenever $|\Delta t_d \cdot \omega|$ is not a large number, then we need to solve the diffraction integral in full generality. This is in general a complicated problem and many different methods have been developed. However, for some simple lenses one can obtain analytic solutions. This is the case of a point mass whose expression can be derived in terms of Laguerre polynomials $L_n(z)$ (more details will follow!)

$$F(w,y) = \nu^{1-\nu} e^{2\nu \log\theta_E} \Gamma(\nu) L_\nu(-\nu y^2) \,, \tag{6.3.43}$$

where $\nu = -iw/2$. One can check that this is equivalent to the standard hypergeometric functions except for the $e^{2\nu\log\theta_E}$ term. The result is plotted in Fig. 6.1.

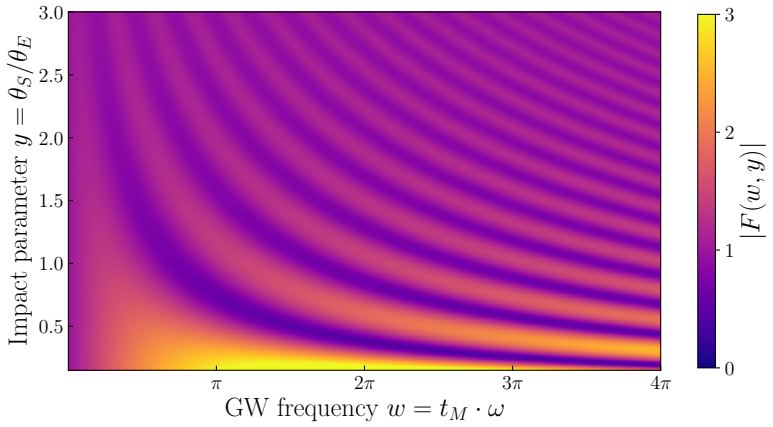


Figure 6.1. Amplification factor $F(w, y)$ for a point lens as a function of the dimensionless frequency w and dimensionless impact parameter y.

6.4 Standard siren cosmology

Gravitational waves from compact binary coalescences are unique cosmic messengers to learn about the Universe's history and fundamental components. A few characteristics are outstanding in the quest for GW cosmology:

1. Compact binaries merge at cosmological distances.
2. Signals are understood from first principles (solving Einstein field equations).
3. GWs travel unaltered through the Universe, except for gravitational lensing.
4. GW wavelengths are of astrophysical scale.

As we saw in §4.3.4, current GW detectors are already sensitive to sources at cosmological distances that merged a few billion years back in time. This is only going to improve with upcoming detectors as schematically depicted in Fig. 6.2. Although the origin of the observed binary black holes is still uncertain, if those originate from stars, their merger rate is expected to follow the star formation rate with some time delay. Thus, binaries are expected to merge with an increasing rate up to redshifts of ~ 2, decreasing afterwards. If other populations of black holes are present, for example originating from the direct collapse of population III stars, mergers at higher redshifts would be possible.

The fact that GWs waveforms are known from theory – numerically solving Einstein's equations – is unusual in astronomy and the aftermath of the "simplicity" of black holes. For most transients, the emission is highly complex due to the interplay of multiple physical phenomena and scales.

As we have seen in this chapter, GW's amplitude is inversely proportional to the luminosity distance to the source. Therefore, because GW detectors are directly sensitive to the GW strain (as we will see in the next chapter), GW signals from binary

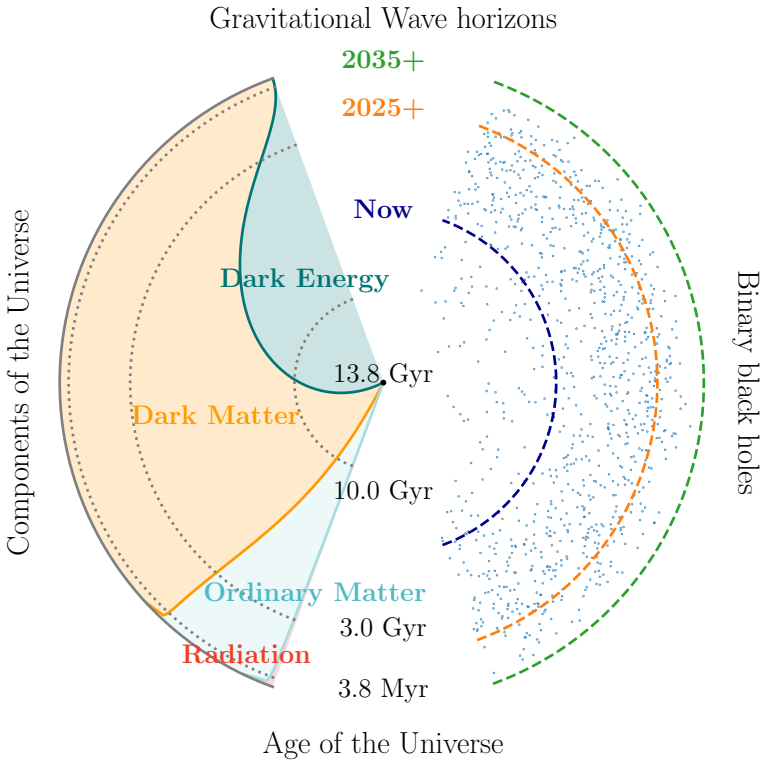


Figure 6.2. Schematic diagram of the potential of GW observations to probe the evolution of the Universe. The radial coordinate represents the time from the present (center) to the beginning of the Universe (edge). On the right side of the diagram, distribution of stellar-mass BBH mergers (blue dots) and their associated horizon distance for different GW detectors (dashed lines). On the left side of the diagram, fraction of each Universe's component (dark energy, dark matter, ordinary matter and radiation) as a function of time, where a fraction of 1 at a given time is represented by the full arc. Although at present time we are observing a Universe dominated by dark energy, future GW detectors will observe an ancient Universe dominated by dark matter. Figure and caption reproduced from [48].

coalescences provide a direct **absolute** measure of distance. This is something rare in cosmology. Normally, distances are only measured **relative** to other quantities. For example, certain type of supernovae, those of type Ia, have standarizable luminosities. Then, if properly calibrated, the relative distance can be extracted from the inverse square law. Such supernovae are referred as **standard candles**. The absolute distance can then be obtained going down the "cosmic distance ladder", to find SN Ia in galaxies with other extragalactic distance indicators (for example cepheids with a well measured period-luminosity relation or stars in the tip of the red giant branch) which can then be matched to absolute distance measures with parallax. Similarly, when we think of the large scale structures of the Universe, the distribution of galaxies carry the imprint of the primordial sound waves in the matter and radiation hot plasma just before recom-

bination, i.e. baryon acoustic oscillations. By measuring such angular scale in a galaxy survey, and with knowledge of the sound horizon at recombination, one effectively obtains a **standard ruler** to measure distances in these surveys. In analogy to these other cosmological probes, GWs are referred as **standard sirens**.

CHAPTER

7

The new era of Gravitational Wave Astronomy

In the last two chapters we have studied how GWs are generated and how they propagate across the cosmos. We focused on the theory and the general physical principles behind each phenomena. Maybe this is all we could do pre-2015, but now GW astronomy is a reality and therefore it is more important than ever to not only understand the theory but also how we analyze the data and what kind of observations we are getting. Of course, we are just at the beginning of this new era in Physics and Astronomy and there is still plenty to discover. Therefore, in this chapter we will first focus on the detectors in §7.1, then on how to analyze their data in §7.2.

7.1 GW observatories

As GWs propagate they carry energy and displace positions of test masses. This characteristic displacement affect celestial bodies as well as apparatus on Earth. If you have a precise way to track the position of your favorite test mass then you have a GW detector. The technology that makes possible the detection of GWs is fascinating, but goes beyond the scope of this course. For us, the most important thing will be to characterize the sensitivity of such detectors and their interactions with GWs.

Current GW detectors are interferometers. A schematic representation of the LIGO detectors is shown in Fig. 7.1. The current detector network is composed of the two LIGO detectors (Handford and Livingston in the United States), Virgo (in Europe) and KAGRA (in Japan). There is also a smaller operating detector in Germany, GEO. It is currently being built another detector in India.

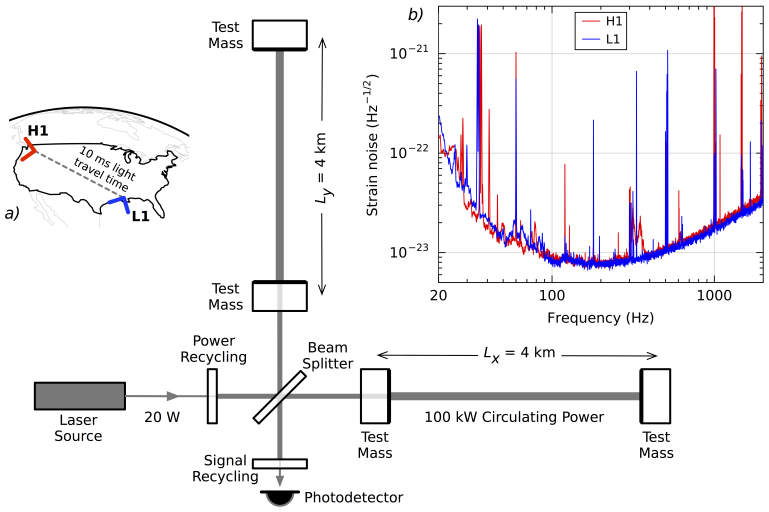


Figure 7.1. Schematic representation of the LIGO detectors [29]. Panel *a*) shows the location of the LIGO detector: Handford in Washington (H1) and Livingston in Lousiana (L1). The time travel distance between the two is 10 ms. Panel *b*) plots the detectors sensitivity around the time of the detection of the first event GW150914. Finally, the central panel shows the key components of the 4 kilometer long interferometers at each location. For more details about this figure see the science summary.

7.1.1 GW detectors

The output of any GW detector is a time series of data $d(t)$. In general this data may contain a signal $s(t)$ and noise $n(t)$:

$$d(t) = s(t) + n(t)\,. \tag{7.1.1}$$

We have already studied how some GW signals look like. For example, we have computed the time domain waveform a compact binary coalescence in the inspiral phase. The purpose of the next section will be to understand how we can disentangle such signals from the noise. However, in order to do so we need to characterize first the noise itself, which will be intrinsic property defining each detector. Moreover, the noise itself determines the minimum value of the GW that could be measured.

Noise spectral density

The noise at a GW detector could be complicated. There are many people who dedicate their careers and heroic efforts to characterize them. Here, we will take simplified approach and make the following assumptions:

1. the noise is **random** process,
2. the noise is **stationary**,
3. the noise is **ergodic**,

4. the noise has **zero-mean**,

5. the noise is **Gaussian** .

The first point implies that we need to characterize each noise realization by a stochastic quantity. The stationarity of the noise implies that we only care about the difference in noise at different times and that the Fourier components are uncorrelated. Stationary noise is defined by its **autocorrelation function**

$$R(\tau) \equiv \langle n(t)n(t+\tau)\rangle \,, \tag{7.1.2}$$

which only depends in the time difference $\tau = t - t'$ because the noise is statistically the same at any given time. The ergodicity of the noise implies that we can exchange ensemble averages, an average over many possible realizations of the probability density function $p_n(n)$ at an arbitrary time t,

$$\langle n\rangle \equiv \int n p_n(n) dn \,, \tag{7.1.3}$$

with time averages that we have direct access to

$$\langle n\rangle = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} n(t)dt \,. \tag{7.1.4}$$

This ensamble average is then performed over a time interval T. The Fourier transform will then have a resolution of $\Delta f = 1/T$. The zero-mean is just a convention:

$$\langle n(t)\rangle = 0 \,. \tag{7.1.5}$$

Finally, that the noise is Gaussian (also referred as white noise), implies that it can be fully characterized by its two-point function:

$$\langle \tilde{n}^*(f)\tilde{n}(f')\rangle = \frac{1}{2} S_n(f)\delta(f-f') \,. \tag{7.1.6}$$

The **noise spectral density or power spectral density (PSD)** $S_n(f)$ then fully characterized the noise. It is measured in units of Hz^{-1}. Also, because the noise is a real function, its Fourier transform satisfies $\tilde{n}(-f) = \tilde{n}^*(f)$, which then implies $S_n(-f) = S_n(f)$. Although the above equation formally diverges when $f = f'$, the fact that we are taking averages over a time interval T implies that

$$\langle |\tilde{n}(f)|^2\rangle = \frac{1}{2} S_n(f) T \,, \tag{7.1.7}$$

where we have replaced the delta function by its integral form $\delta(x) = \int_{-\infty}^{\infty} dk e^{ikx}/2\pi$ on a limited range $[-T/2, T/2]$. Again, $S_n(f)$ will have a resolution of $\Delta f = 1/T$. Note that the factor of $1/2$ is conventional in order to represent the noise in the (physical) range of positive frequencies:

$$\langle n^2(t)\rangle = R(0) = \int_0^{\infty} df S_n(f) \,. \tag{7.1.8}$$

If we have Gaussian noise on a time interval T, if we take N samples at regular intervals Δt, then each noise sample n_j with $j = 0, \cdots, N-1$ is an independent Gaussian random variable and the joint probability of obtaining $\{n_j\}$ is

$$p_n(\{n_j\}) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N \exp\left[-\frac{1}{2\sigma^2}\sum_{j=0}^{N-1} n_j^2\right] , \tag{7.1.9}$$

where σ is the variance. If we make the sampling intervals smaller, $\Delta t \to 0$, then we can approach to the continuum limit

$$\lim_{\Delta t\to 0} \exp\left[-\frac{1}{2\sigma^2\Delta t}\sum_{j=0}^{N-1} n_j^2\Delta t\right] = \exp\left[-\frac{1}{S_n}\int_0^T n^2(t)dt\right] \approx \exp\left[-\int_{-\infty}^{\infty}\frac{|\tilde{n}(f)|^2}{S_n(f)}df\right] . \tag{7.1.10}$$

Then, the continuum probability of a noise realization $n(t)$ is

$$p_n[n(t)] \propto \exp\left[-2\int_0^{\infty}\frac{|\tilde{n}(f)|^2}{S_n(f)}df\right] , \tag{7.1.11}$$

where we have reduced the integral to positive frequencies.

As said before, real noise in current GW detectors certainly deviates from this idealized scenario. In particular it is non-Gaussian and non-stationary. However, for the purpose of our discussion these assumptions will be enough to characterize the main features of the data analysis process.

Antenna pattern functions

In our presentation of the GW generation and propagation in chapters 5 and 6, we were always thinking of how the GWs are produced in the **source frame** and how that signal was seen by a distant observer, sometimes know as the **wave frame**. For example, we took into account that the observer might not be aligned with the propagation direction and, as a consequence, each polarization will carry a different factor due to the projection from the inclination between each reference frame. GW detectors themselves will be sensitive to different parts of the GW signal is different ways. Therefore, we need to define a **detector frame** and study the projection of a GW signal into them. Fig. 7.2 summarizes the different definitions and conventions that we will be using. Note that those vary from reference to reference making the process of matching results and fully defining all the GW parameters maximally confusing. In any case for our purpose, the very concrete definitions will not be relevant and we will focus on the important physical and observational implications.

Literature:

This section follows mostly Maggiore's Chapter 7.2 [32].

In general, we have seen that a given GW propagating in the direction $\hat{n}$ can be written as a sum over its polarizations A:

$$h_{ij}(t,\vec{x}) = \sum_A \epsilon_{ij}^A(\hat{n}) \int_{-\infty}^{\infty} df \tilde{h}_A(f) e^{-2\pi i(t-\hat{n}\vec{x}/c)}\,, \tag{7.1.12}$$

where ϵ_{ij}^A are the polarizations tensors and we have included the retardation from the finite propagation speed. However, since ground-based detectors have sizes ($\sim$4km) much smaller than the waves they detect ($\lambda = c/2\pi f \sim 1000$km), then $\hat{n}\vec{x}/\lambda \ll 1$ and we can neglect this delay and simplify the above expression to

$$h_{ij}(t,\vec{x}) = \sum_A \epsilon_{ij}^A(\hat{n}) h_A(t)\,. \tag{7.1.13}$$

The detector will measure a **strain amplitude** h that will be a projection D^{ij} of each polarization

$$h(t) = \sum_A D^{ij}\epsilon_{ij}^A(\hat{n}) h_A(t) = \sum_A F_A(\hat{n}) h_A(t)\,, \tag{7.1.14}$$

where F_A are the **detector antenna pattern functions**. Since in GR GWs only produce $+,\times$ polarizations, we will focus only in

$$h(t) = h_+(t)F_+(\hat{n}) + h_\times(t)F_\times(\hat{n})\,, \tag{7.1.15}$$

although in general the detector could also be sensitive to other polarizations. The functions $F_{+,\times}$ therefore tell us how a GW detector reacts to the $+$ and $\times$ for a signal arriving from the sky position $\{\theta,\varphi\}$. Their particular functional form depends on the geometry of the detector itself. For current L-shaped ground-based interferometers their antenna pattern functions are

$$F_+(\theta,\varphi) = \frac{1}{2}\left(1+\cos^2\theta\right)\cos 2\varphi\,, \tag{7.1.16}$$

$$F_\times(\theta,\varphi) = \cos\theta \sin 2\varphi\,. \tag{7.1.17}$$

The transformation between the wave and detector frames is fully fixed by three Euler angles. However, the antenna function above are defined only by two: $\{\theta,\varphi\}$. This is because we were implicitly assuming that the polarization basis defined in the plane perpendicular to $\vec{n}$ was the same between both frames. In general this need not be the case and there is a third Euler angles ψ, known as the **polarization angle**, that rotates between both frames in the transverse plane R_ψ. A rotation of ψ changes the polarization tensor by

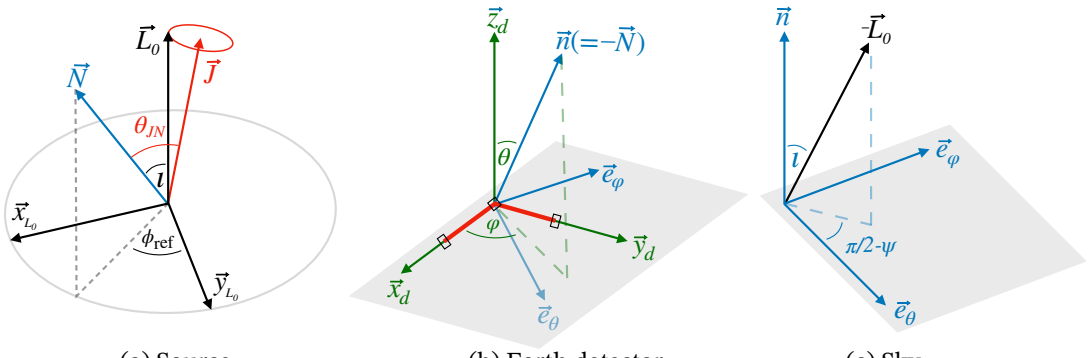
$$\epsilon_{ij}^+(\vec{n}) \to \epsilon_{ij}^+(\vec{n})\cos 2\psi - \epsilon_{ij}^\times(\vec{n})\sin 2\psi\,, \tag{7.1.18}$$

$$\epsilon_{ij}^\times(\vec{n}) \to \epsilon_{ij}^+(\vec{n})\sin 2\psi + \epsilon_{ij}^\times(\vec{n})\cos 2\psi\,, \tag{7.1.19}$$

which effectively rotates by 2ψ due to the spin-2 nature of the polarization tensors. Then, the antenna pattern functions will become

$$F_+(\theta,\varphi,\psi) = F_+(\theta,\varphi,0)\cos 2\psi - F_\times(\theta,\varphi,0)\sin 2\psi\,, \tag{7.1.20}$$

$$F_\times(\theta,\varphi,\psi) = F_+(\theta,\varphi,0)\sin 2\psi + F_\times(\theta,\varphi,0)\cos 2\psi\,. \tag{7.1.21}$$



(a) Source (b) Earth detector (c) Sky

Figure 7.2. Summary figure of our frame conventions. The source frame (a) defines the coordinate system in which the intrinsic parameters of the binary are defined: masses, spins and phase. It is anchored to the orbital angular momentum at the reference frequency $\vec{L}_0$. The Earth detector frame (b) serves to define the time of arrival and the position of the sky of the binary event for a fiducial detector at the center of the Earth, in order to compute the antenna response function of each detector. The sky frame (c) defines the remaining extrinsic parameters, the inclination ι and the polarization angle ψ. See text for further details. Figure and caption reproduced from [49].

Note that the definition of ψ changes between different communities. The LIGO software uses a different convention to the above equation, but it can be easily obtained redefining the rotation angle: $\psi \to \pi/2 - \psi$. This is the convention used in Fig. 7.2.

Altogether this means that when analyzing a given GW signal $h(t)$, this signal will depend on three additional parameters $\{\theta, \varphi, \psi\}$ **extrinsic** to the properties of the signal itself.

There are two simplifications to the above derivation. First, we were implicitly assuming a detector at the center of Earth. When analyzing real data we need to project to the actual position of each detector on the surface of the Earth. Second, the Earth rotates and therefore we need to anchor the sky positions to the Greenwich Mean Sidereal Time (gmst) of the observation. This is done most commonly by astronomers defining the right ascension (ra) and declination:

$$\text{ra} = \varphi + \text{gmst}\,, \tag{7.1.22}$$

$$\text{dec} = \pi/2 - \theta\,. \tag{7.1.23}$$

Sky localization

With a single GW detector it is very difficult to localize a given source. This is because we cannot break the degeneracies between all the parameters describing the binary. If we have multiple detectors, then we can use them to triangulate the sky position of the source. The trick is that GWs travel at the speed of light and we know the fixed positions of the detectors on the surface of the Earth $\vec{r}_d$. Then, the arrival time at each detector will be

$$t_d = t_c - \vec{n} \cdot \vec{r}_d / c\,. \tag{7.1.24}$$

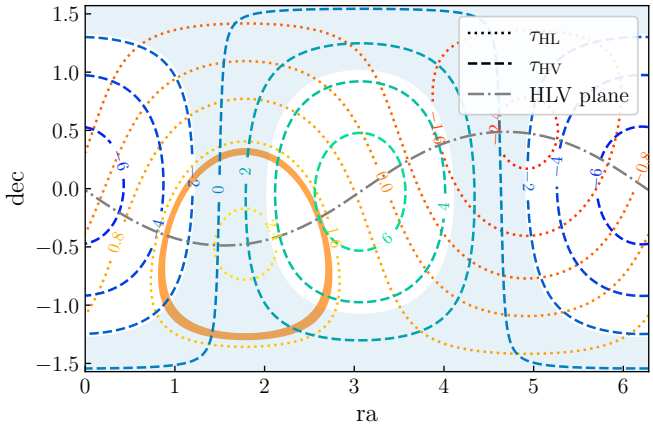


Figure 7.3. Time delay phase contour lines as a function of the right ascension (ra) and declination (dec) for GW150914. Dotted lines indicate contours for Hanford and Livingston (τ_{HL}), while dashed lines are for Hanford and Virgo (τ_{HV}). The shaded orange and blue regions correspond to the 95% CL from the reconstructed time delays, τ_{HL} and τ_{HV}respectively. The dash-dotted line indicates the plane defined by the position of the HLV detectors. The intersection of both shaded regions occurs in reflection symmetric positions above and below the detector's plane and correspond to a bimodality in the localization from time delays. Figure and caption reproduced from [49].

Similarly, the arrival time difference between two detectors d_1 and d_2 is

$$\Delta t_{d_1 d_2} = \vec{n} \cdot \vec{r}_{d_1 d_2}/c\,, \tag{7.1.25}$$

where $\vec{r}_{d_1 d_2} = \vec{r}_{d_2} - r_{d_2}$. Each pair of detectors then measures an arrival time difference that constraints one angle $\vec{n}\cdot\vec{r}_{d_1 d_2}$, which defines a ring in the sky. With three detectors, two angles are constrained, $\vec{n} \cdot \vec{r}_{d_1 d_2}$ and $\vec{n} \cdot \vec{r}_{d_1 d_3}$, defining two rings in the sky that intersect at two points where the event localization is possible. These two possible localization regions correspond to a reflection symmetry of the time delays across the hemispheres delineated by the plane defined by the position of the three detectors, i.e. distinguished by the sign of $\vec{n} \cdot (\vec{r}_{d_1 d_2} \times \vec{r}_{d_1 d_3})$.

7.2 Data analysis

As said before, the output of a GW detector is a time series data $d(t)$. In the absence of a signal $s(t)$ the data is just noise $n(t)$. We have seen in §7.1.1 that the noise can be characterized by its spectral density. We are going to consider now the more exciting case in which the data has both a signal and some noise component. Our task is therefore disentangling one from each other. In the limit in which the signal is much larger than the noise, $|s(t)| \gg |n(t)|$, then this is a simple task as one is literally seeing the signal on top of the noise. However, recall that the GW strain at the detector is very small... In fact, in current observing scenarios we are always in the opposite regime in which

the noise amplitude is larger than the signal itself. Does this mean that we cannot detect GWs? No! We can detect GWs because of a very unique property of this type of transient:[1] we understand the shape of a GW (at least for compact binary coalescences) from first principles. Therefore, we can use our waveform models as **templates** to **filter** the data and dig up the signal from the noise that **matches** our template.

In this section we will study the method to extract the signal from the data in §7.2.1, **matched filtering**, then we will learn how to characterize the signal and perform **parameter estimation** in §7.2.2, and, finally, we will cover how to study signals collectively in **population analyses** in §7.2.3.

7.2.1 Matched filtering

Literature:

This section follows the derivation of Maggiore 7.3 [32].

The idea of matched filtering is very simple in theory. If you have a data set $d(t)$ with a GW signal $s(t)$ for which you know the signal (in other words you have a model for $s(t)$), then, if you wait long enough, you will detect the signal. More explicitly, you can convolve the data with your model signal over an observing time T then

$$\frac{1}{T}\int_0^T dt\, d(t)s(t) = \frac{1}{T}\int_0^T dt\, s^2(t) + \frac{1}{T}\int_0^T dt\, n(t)s(t)\,, \tag{7.2.1}$$

where both the signal and the noise are typically oscillating functions. The first term on the right hand side is positive definite and, at late times, grows linearly in T. On the other hand, the second term has two uncorrelated terms that, over late times, scale as $T^{1/2}$ as a random walk. Then, we have

$$\frac{1}{T}\int_0^T dt\, d(t)s(t) \sim s_0^2 + \left(\frac{\tau_0}{T}\right)^{1/2} n_0 s_0\,, \tag{7.2.2}$$

where s_0 and n_0 are the characteristic amplitudes of the signal and noise, and τ_0 is the characteristic time of the signal. A signal is then detected when

$$s_0 > \left(\frac{\tau_0}{T}\right)^{1/2} n_0 \tag{7.2.3}$$

and the first term dominates over the second one. This is achieved if the signal amplitude is large, but also if one observe many oscillations and $\tau_0/T \ll 1$.

[1] For most (if not all!) other transients we do not have fundamental physics understanding of how their signals look like. We learn about them empirically. The main difference is that the physics of a compact binary coalescence is relatively "simple". Black holes are just described by their mass and spin and general relativity does the rest. For other explosive transients there are many other physical properties involved: temperature, chemical composition, magnetic fields...

We now want to formalize this matched filtering process and find the optimal filter that maximizes the **signal-to-noise ratio (SNR)**. If we filter our data with a function $K(t)$, the filtered data is

$$\hat{d} = \int_{-\infty}^{\infty} dt\, d(t)K(t)\,. \tag{7.2.4}$$

The SNR is defined as the ratio of the expected value of the filtered data when the signal is present, $S = \langle \hat{d}(t)\rangle_{s\neq 0}$, with the root mean square of the filtered data without signal, $N = \sqrt{\langle \hat{d}^2(t)\rangle_{s=0} - \langle \hat{d}(t)\rangle^2_{s=0}}$. In Fourier space we have that the signal is

$$S = \langle \hat{d}(t)\rangle_{s\neq 0} = \int_{-\infty}^{\infty} dt\, s(t)K(t) = \int_{-\infty}^{\infty} df\, \tilde{s}(f)\tilde{K}^*(f)\,, \tag{7.2.5}$$

where in the second equality we have used that the noise has zero mean, and in the last one we have moved to frequency domain and exploit that $K(t)$ is a real function.[2] Similarly, the noise is

$$\begin{aligned} N^2 &= \langle \hat{d}^2(t)\rangle_{s=0} - \langle \hat{d}(t)\rangle^2_{s=0} = \langle \hat{d}^2(t)\rangle_{s=0} \\ &= \int_{-\infty}^{\infty} dt dt' K(t)K(t')\langle n(t)n(t')\rangle \\ &= \int_{-\infty}^{\infty} df \frac{1}{2} S_n(f)|\tilde{K}(f)|^2\,, \end{aligned} \tag{7.2.6}$$

where, again, we have used in the second equality that we have zero-mean noise and in the third line we have used that the noise is Gaussian and defines by the power spectral density (PSD). Altogether we have that

$$S/N = \frac{\int_{-\infty}^{\infty} df\, \tilde{s}(f)\tilde{K}^*(f)}{\sqrt{\int_{-\infty}^{\infty} df \frac{1}{2} S_n(f)|\tilde{K}(f)|^2}}\,. \tag{7.2.7}$$

We now just need to find the adequate $K(t)$ that maximizes S/N. To do so it is useful to introduce the **noise-weighted inner product** of two real function $a(t)$ and $b(t)$

$$\begin{aligned} (a|b) &\equiv \mathrm{Re}\left[\int_{-\infty}^{\infty} \frac{\tilde{a}^*(f)\tilde{b}(f)}{S_n(f)/2}\right] \\ &= 4\mathrm{Re}\left[\int_{0}^{\infty} \frac{\tilde{a}^*(f)\tilde{b}(f)}{S_n(f)}\right]\,, \end{aligned} \tag{7.2.8}$$

where in the second line we have used that time domain reality implies $\tilde{a}^*(f) = \tilde{a}(-f)$ and $S_n(f) = S_n(-f)$. If we define a function

$$\tilde{u}(f) = \frac{1}{2} S_n(f)\tilde{K}(f)\,, \tag{7.2.9}$$

[2] Remember that our convention is $g(t) = \int df \tilde{g}(f)e^{-2\pi i f t}$ and $\tilde{g}(f) = \int dt g(t)e^{2\pi i f t}$.

the SNR is

$$S/N = \frac{(u|s)}{\sqrt{(u|u)}}\,. \tag{7.2.10}$$

Therefore, it is now clear that the optimal filter will define a function u that is parallel to s so that $(u|s)$ is maximal. Mathematically:

$$\tilde{K}(f) \propto \frac{\tilde{s}(f)}{S_n(f)}\,. \tag{7.2.11}$$

We have found (maybe not surprisingly in perspective) that the optimal filter is the signal itself. For a given GW signal h, the optimal SNR is therefore

$$\rho^2_{\rm opt} = (h|h) = 4{\rm Re}\left[\int_0^\infty df \frac{|\tilde{h}(f)|^2}{S_n(f)}\right]\,. \tag{7.2.12}$$

A common proxy for determining whether a signal is observable by a given detector is if its SNR is larger than 8. This threshold can be set in terms of a desired false alarm probability as we will see later. It is important to note that the optimal filter (7.2.11) is independent of the amplitude. Therefore, the overall GW amplitude is a parameter that we will be able to marginalize over.

Finally, let us note that with the definition of this inner product the probability density function of a noise realization $n(t)$ in Eq. (7.1.11) simplifies to

$$p_n[n(t)] \propto e^{-(n|n)/2}\,. \tag{7.2.13}$$

Exercise 7.1: Noise-weighted inner product

Show that the noise-weighted inner product (7.2.8) can also be written as

$$(a|b) = \int_{-\infty}^{\infty} df \frac{\tilde{a}^*(f)\tilde{b}(f) + \tilde{a}(f)\tilde{b}^*(f)}{S_n(f)}\,. \tag{7.2.14}$$

Exercise 7.2: Signal-to-noise ratio

Demonstrate that the noise-weighted inner product $(a|b)$ of two real functions $a(t)$ and $b(t)$ defined as Eq. (7.2.8) can be equivalently written in time domain as

$$(a|b) = \int_{-\infty}^{\infty} dt \hat{a}(t)\hat{b}(t)\,,$$

where $\hat{a}(t)$ is the inverse Fourier transform of the frequency domain noise-weighted signal $\tilde{\hat{a}} = \tilde{a}(f)/\sqrt{S_n(f)/2}$.

Mismatch

Although we might known the general form of the GW signal h, it is not always that we know the concrete parameters or we might have different hypothesis for the concrete model itself. In those situations, a useful quantity to measure the similarity between two models h_1 and h_2 in light of the detector under consideration is the **match**:

$$\mathcal{M}(h_1, h_2) = \frac{(h_1|h_2)}{\sqrt{(h_1|h_1)(h_2|h_2)}}\,. \tag{7.2.15}$$

The match has the advantage that it is insensitive to the overall amplitude of each template. It is sometimes referred as the **fitting factor**. The degree of difference, or **mismatch**, is simply:

$$\epsilon(h_1, h_2) = 1 - \mathcal{M}(h_1, h_2)\,. \tag{7.2.16}$$

It is to be noted that with the above definition the match can range from 1 to -1, taking the limiting cases of two equal templates, $\mathcal{M}(h, h) = 1$, and two completely off-phase, $\mathcal{M}(h, -h) = -1$. This translates into a mismatch spanning from 0 to 2. Because when we project into the detector antenna pattern functions the signal can pick an overall phase, it is common to maximize the match over phase. Similarly, one is typically interested in looking for the best match across different potential arrival times of the template. This can be achieved simply by multiplying one of the signals by $e^{i(2\pi f t_0 + \phi_0)}$ and compute the maximum match for a range of t_0 and ϕ_0:

$$\begin{aligned}
\mathcal{M}_{\max}(h_1, h_2) &= \max_{t_0,\phi_0}\left[\frac{(h_1|h_2 e^{i(2\pi f t_0+\phi_0)})}{\sqrt{(h_1|h_1)(h_2 e^{i(2\pi f t_0+\phi_0)}|h_2 e^{i(2\pi f t_0+\phi_0)})}}\right] \\
&= \max_{t_0,\phi_0}\left[\frac{e^{i\phi_0}(h_1|h_2 e^{2i\pi f t_0})}{\sqrt{(h_1|h_1)(h_2|h_2)}}\right] \\
&= \max_{t_0}\left[\frac{\left|(h_1|h_2 e^{2i\pi f t_0})\right|}{\sqrt{(h_1|h_1)(h_2|h_2)}}\right]\,.
\end{aligned} \tag{7.2.17}$$

In the second line we have noted that the extra complex exponential does not affect the denominator and that the constant phase ϕ_0 can be taken out of the integral. Therefore, in the third line, we have exchange the phase marginalization with an absolute value. With this definition, the match now is limited to be between 0 and 1. As a consequence, the minimum mismatch is also limited to the same range, simply using

$$\epsilon_{\min}(h_1, h_2) = 1 - \mathcal{M}_{\max}(h_1, h_2)\,. \tag{7.2.18}$$

Horizon distance

For a given SNR detection threshold $\rho_{\rm thr}$ it is easy to compute the maximum (luminosity) distance that a GW could be observed because the SNR is inversely proportional to d_L, $\rho \propto 1/d_L$. The horizon distance is

$$d_L^{\rm hor} = \frac{\rho_{\rm opt}(d_L = 1{\rm Mpc})}{\rho_{\rm thr}}{\rm Mpc}\,. \tag{7.2.19}$$

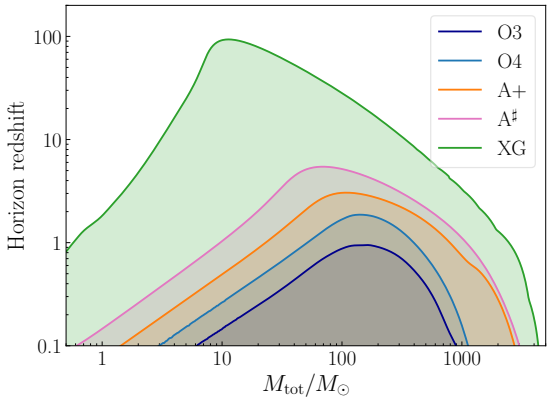


Figure 7.4. Horizon redshift for different past and future ground-based detectors. O3 and O4 refers to the third and fourth observing runs of LIGO. A+ is the design sensitivity of LIGO detectors. $A^\sharp$ is a planned upgrade of final LIGO. XG refers to next-generation detectors such as Einstein Telescope or Cosmic Explorer.

This horizon distance can also be translated into a horizon redshift for a fixed cosmology. A few examples for different past and future ground-based detectors are presented in Fig. 7.4. The shape of these curves is determined by the detectors sensitivities at different frequencies.

SNR time series

In our calculation of the optimal SNR we have assumed that we knew the arrival time of the signal so that we could match a template at that precise reference time $t_{\rm ref}$ when the signal enters the interferometer bandwidth at a reference frequency $f_{\rm ref}$. In reality, however, we do not have that information. Therefore, in principle, we should compute the inner product of the data with our template at every different time step. If this step is ∼ms, then in a year ($\sim 3\cdot 10^7$s) we would need 10^{10} evaluations. Luckily we do not need to make so many computations. The unknown arrival time just makes a time translation of our template

$$h(t, t_{\rm ref}) = h(t - t_{\rm ref})\,, \tag{7.2.20}$$

which in Fourier space is simply

$$\tilde{h}(f, t_{\rm ref}) = \tilde{h}(f) e^{i2\pi f t_{\rm ref}}\,. \tag{7.2.21}$$

The scalar product of this signal with the data defines a time series for the SNR:

$$\rho^2_{\rm opt}(t_{\rm ref}) = (d|h(t, t_{\rm ref})) = 4{\rm Re}\left[\int_0^\infty df \frac{\tilde{d}^*(f)\tilde{h}(f)}{S_n(f)} e^{i2\pi f t_{\rm ref}}\right]\,. \tag{7.2.22}$$

Note that this expression is just the Fourier transform of $\tilde{d}^*(f)\tilde{h}(f)/S_n(f)$, so we get the whole time series with a single Fourier transform. The maximum of this time series defines the value of $t_{\rm ref}$.

Detector network

When we have a set of detectors, we can also define a **network SNR** as

$$\rho_{\rm ntw} = \sqrt{\sum_i \rho_i^2}\,. \tag{7.2.23}$$

Note that since each detector is located at different positions on Earth, they will carry different projections of the antenna pattern functions $F_{+,\times}$. In a two detector network, a proxy for a GW detection is $\rho_{\rm ntw} > 12$.

It is also important to note that a network of detectors play a crucial role in discarding spurious noise fluctuations that mimic a GW in a given detector. This is because since the distance between detectors is known, a GW traveling at the speed of light arriving first at one detector should arrive at the others within a time window. Therefore one can look for **coincident** triggers in multiple detectors. If two detectors are co-aligned and not too far apart, as the two LIGO detectors, then one would expect their antenna pattern functions to be similar and, therefore, if both detectors have similar sensitivity, they should see the same waves.

Matched filtering statistics

In our definition of the SNR we restricted to the expectation value of the filtered data $\langle\hat{d}\rangle$, we now wish to compute the full distribution of SNR

$$\rho = \frac{\hat{d}}{N}\,, \tag{7.2.24}$$

which is sometimes referred as the **observed** SNR to be distinguished from the **optimal** SNR $\rho_{\rm opt}$. The main difference is that now we will have to take into account the noise fluctuations:

$$\hat{d} = \int_{-\infty}^{\infty} dt\,(s(t) + n(t))\,K(t)\,. \tag{7.2.25}$$

In the absence of a signal, ρ is then a random variable characterized by the statistical properties of the noise. Because we have zero mean $\langle n(t)\rangle = 0$ and because we are normalizing by its own root mean square, the distribution of SNR without signal is a Gaussian of variance 1 centered around 0:

$$p(\rho|s=0) = \frac{1}{\sqrt{2\pi}} e^{-\rho^2/2}\,. \tag{7.2.26}$$

With a signal with optimal SNR $\rho_{\rm opt}$, then the observed SNR distribution is

$$p(\rho|\rho_{\rm opt}) = \frac{1}{\sqrt{2\pi}} e^{-(\rho-\rho_{\rm opt})^2/2}\,, \tag{7.2.27}$$

which again has unit variance but it is centered around the optimal SNR. Therefore, in practice, if one want to simulate observed SNRs, one simply need to scatter around the optimal value with variance 1.

If we are defining a detection as a observed SNR above a given value $\rho_{\rm th}$, then it is important to note that the noise itself could have fluctuations which are larger than this value. This defines the **false alarm probability (FAP)**:

$$p_{\rm FA} = \int_{\rho_{\rm th}}^{\infty} d\rho \frac{1}{\sqrt{2\pi}} e^{-\rho^2/2} = 2{\rm erfc}(\rho_{\rm th}/\sqrt{2})\,, \tag{7.2.28}$$

where $\mathrm{erfc}(z)$ is the complementary error function. Similarly, there is a probability that a real GW event does not pass the threshold and there is a **false negative**:

$$p_{\rm FN} = \int_{-\infty}^{\rho_{\rm th}} d\rho \frac{1}{\sqrt{2\pi}} e^{-(\rho-\rho_{\rm opt})^2/2}\,, \tag{7.2.29}$$

where we have noted that ρ is not positive definite because it is proportional to h.

7.2.2 Parameter estimation

Once we believe that there is a signal in our data, the next step is to infer the parameters that characterize it. This is necessary for two main reasons. First, in our matched filtering analysis we assumed that we knew the signal $s(t)$. However, in general, we only have models or **templates** that are characterized by a set of parameters $h(t|\theta)$. For example, for a compact binary coalescence, at leading order, the main parameters of the source are the distance, chirp mass, arrival time, reference phase, to which we need to add the extrinsic parameters defining the relation between the source's and detector's frames: inclination, sky position and polarization angle. Therefore, we need to estimate the parameters that describe the signal. The second reason is that in those parameters is precisely where all the science is hidden. By measuring the parameters of a GW signal we can learn about its astrophysical origin, the cosmological time when it was generated and whether the signal follows the predictions of general relativity.

Bayesian statistics

Given a candidate GW event, we want to reconstruct its most probable parameters for our assumed model of the signal. We can compute such distributions using Bayesian statistics. Let's review the basics!

Given a set S with subsets $A, B, \cdots$, we define a probability P as a real function that satisfies that **i)** for every subset the probability is positive $P(A) \geqslant 0$, **ii)** for disjoint subsets, $A \cap B = 0$, the probability is additive $P(A \cup B) = P(A) + P(B)$, and **iii)** for the whole set $P(S) = 1$. The probability of A given B, the **conditional probability** $P(A|B)$ is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}\,, \tag{7.2.30}$$

where $P(A \cap B)$ is the **joint probability** of both A and B being true. Since the definition of the conditional probabilities implies $P(A \cap B) = P(A|B)P(B)$ and $P(B \cap A) = P(B|A)P(A)$, and the intersection is commutative, $A \cap B = B \cap A$, we find that

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}\,, \tag{7.2.31}$$

which is known as **Bayes' theorem**. Note that the term in the denominator could also be written as

$$P(B) = \sum_i P(B|A_i)P(A_i) \tag{7.2.32}$$

for a set of subsets A_i such that their union is the full set $\cup_i A_i = S$. Therefore, the denominator is just a normalization function. In the jargon of Bayesian statistics $P(A|B)$ is the **posterior probability** of A being true given B that is determined in terms of the **prior probability** $P(A)$, the **likelihood** of B being true given A and normalized by the **evidence** $P(B)$.

If we interpret A as the hypothesis/parameters and B as the data, then Bayes' theorem tell us that

$$P(\text{hypothesis}|\text{data}) \propto P(\text{data}|\text{hypothesis})P(\text{hypothesis})\,. \tag{7.2.33}$$

Note that in Bayesian statistics we can define a **confidence level (CL)** from our posterior probability that defines the range of the subset in which we believe the hypothesis is true given the data at some probability level.

Another useful statistical concept is the **marginal probability**. Given two sets A and B forming a joint distribution, the marginal distribution of A can be obtained taking the expectation value of the conditional probability of A given B, $P(A|B)$, over the distribution of B:

$$P(A) = \int P(A|B)P(B)dB\,. \tag{7.2.34}$$

Interestingly, the concept of marginalization is ubiquitous not only in Mathematics, but also in Physics. Whenever we do not know about a physical phenomena, for example a theory of fundamental interactions at high energies, we can **integrate out** or marginalize over all the degrees of freedom that are not relevant for the problem under consideration. This is the basis of statistical mechanics and effective field theories.

Bayesian inference

After reviewing the basics of Bayesian statistics, we are ready to use this framework to infer the properties of a signal. As in our previous discussions, we will be assuming that the noise is stationary and Gaussian. Therefore, we know that in the absence of a signal it is described by a normal distribution with variance determined by the noise weighted inner product $p(n(t)) \propto e^{-(n|n)/2}$, cf. Eq. (7.2.13). If our data contains a signal then the noise can be written as

$$n(t) = d(t) - s(t)\,. \tag{7.2.35}$$

Our hypothesis is that the signal $s(t)$ is described by a model $h(t;\theta)$ with parameters $\theta = \{\theta_i\}$ so that $s(t) = h(t;\theta_t)$, where θ_t are the true values. Therefore, if the hypothesis is true, the likelihood of the data $\Lambda(d|\theta)$ is described by the fact that if we subtract the signal to the data we should recover Gaussian noise:

$$\begin{aligned}\Lambda(d|\theta) &\propto \exp\left[-\frac{1}{2}(d-h(\theta)|d-h(\theta))\right] \\ &= \exp\left[(d|h(\theta)) - \frac{1}{2}(h(\theta)|h(\theta)) - \frac{1}{2}(d|d)\right]\,.\end{aligned} \tag{7.2.36}$$

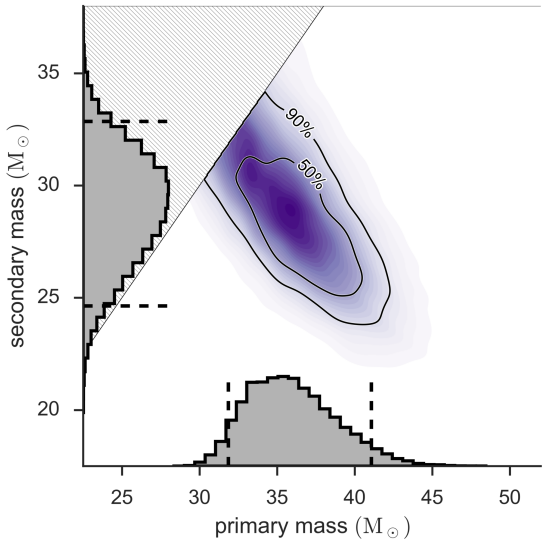


Figure 7.5. Inferred component masses for the first detection, GW150914. More information in the science summary.

According to Bayes' theorem to get the posterior distribution of the parameters given the data $p(\theta|d)$, we only need to multiply by our prior knowledge on the distribution of the parameters $p(\theta)$:

$$p(\theta|d) \propto p(\theta) \exp\left[(d|h(\theta)) - \frac{1}{2}(h(\theta)|h(\theta))\right], \tag{7.2.37}$$

where we have reabsorbed the $(d|d)$ term in the exponent into the overall normalization. This solves the problem of inferring the values of the parameters θ consistent with the data.

Although the definition of the posterior distribution for the GW parameters in Eq. (7.2.37) seems innocuous, its actual maximization is rather nontrivial since a GW waveform models are defined in a high dimensional parameter space. In practice one needs to solve this problem numerically and there are a plethora of methods to efficiently solve this maximization. Within the GW community a common software that includes many of these methods is bilby. Once the posterior distribution is found, this is typically presented in a **corner plot** where all possible 2D sections of the high-dimensional posterior distribution are plotted together with the 1D marginal distributions. As an example, the inferred component masses of the first event are plotted in Fig. 7.5.

Maximum likelihood estimator

Let us consider a case in which the prior distribution is flat. This is sometimes referred as **uninformative** priors, although really in Bayesian statistics there is no such concept since we always need to make a choice for the prior. Anyways, in this situation the maximization of the posterior distribution is simply the maximization of the likelihood.

We are looking for the **maximum likelihood** values $\theta_{\rm ML}$. In order to maximize $\Lambda(d|\theta)$ it is easier to take the logarithm to focus in its exponent:[3]

$$\log \Lambda(d|\theta) = (h(\theta)|d) - \frac{1}{2}(h(\theta)|h(\theta))\,. \tag{7.2.38}$$

The maximum values $\theta_{\rm ML}$ are simply given by the solution of

$$(\partial_i h(\theta)|d) - (\partial_i h(\theta)|h(\theta)) = 0. \tag{7.2.39}$$

The errors of the parameters $\Delta\theta^i$ are then given by the posterior distribution around the point $\theta_{\rm ML}$.

If we have a generic signal of unknown amplitude a and other parameters $\lambda = \{\lambda_i\}$, $h(t;\theta) = a h_a(t;\lambda)$, the log-likelihood is given by

$$\log \Lambda(d|a,\lambda) = a(h_a(\lambda)|d) - \frac{a^2}{2}(h_a(\lambda)|h_a(\lambda))\,. \tag{7.2.40}$$

Therefore, we can obtain the maximum likelihood value of the amplitude analytically

$$a_{\rm ML}(d) = \frac{(h_a|d)}{(h_a|h_a)}\,. \tag{7.2.41}$$

If we substitute this expression back into the likelihood, then we find that for the rest of parameters

$$\log \Lambda(d|\lambda) = \frac{1}{2}\frac{(h_a|d)^2}{(h_a|h_a)}\,. \tag{7.2.42}$$

Therefore, the maximization just entails finding the maximum of the noise-weighted inner product of the data with the normalized template $\hat{h} = h_a/(h_a|h_a)^{1/2}$. In other words, the maximum likelihood is equivalent to the value that maximizes the SNR, which was our defining goal of the matched filtering process.

Measurement uncertainty in the high-SNR limit

The measurement uncertainty of a parameter $\Delta\theta^i$ around a given value $\hat{\theta}$ is determined by the shape of the posterior distribution around that point. In general the posterior distribution can be highly complicated having multiple peaks and non-Gaussian features. We are going to restrict here to the case in which the SNR is high and, as a consequence, the errors $\Delta\theta^i$ are small. Such errors will be determined by the curvature of $p(\theta|d)$ around $\hat{\theta}$. Because the SNR is high we can consider our data to be informative. In other words, the likelihood is peaked around $\hat{\theta}$ and the prior approximately constant around the relevant range.[4] Then, the inference is fully determined by the likelihood and choose

[3] This is very advantageous numerically because then we convert products of probabilities into sums of logarithms and also because the numbers become more tractable, e.g. $\log(2839289829) = 9.45$.

[4] The formalism can also be generalized to other types of priors, but we will not consider this here for simplicity.

our reference point in parameter space as the maximum likelihood $\hat{\theta} = \theta_{\rm ML}$. We can then expand $\Lambda(d|\theta)$ around

$$\theta^i = \theta^i_{\rm ML} + \Delta\theta^i \,, \tag{7.2.43}$$

where $\Delta\theta^i$ is small. Since at the maximum likelihood point the first derivative of Λ vanishes by definition, the leading contribution will be quadratic in the uncertainties:

$$p(\theta|d) \propto \exp\left[-\frac{1}{2}\Gamma_{ij}\Delta\theta^i\Delta\theta^j\right] \,, \tag{7.2.44}$$

where Γ_{ij} is the **Fisher information matrix** defined by

$$\Gamma_{ij} = (\partial_i\partial_j h|h-d) + (\partial_i h|\partial_j h) \approx (\partial_i h|\partial_j h) \,, \tag{7.2.45}$$

evaluated at $\theta_{\rm ML}$. The approximate equality comes from the use of the high-SNR limit in which the amplitude of the noise $|n| = |d-h|$ is small compared to $|h|$. As it can be seen directly, the Fisher matrix is symmetric. The inverse of the Fisher matrix defines the **covariance**

$$C_{ij} = (\Gamma^{-1})^{ij} = \langle \Delta\theta^i \Delta\theta^j \rangle \,, \tag{7.2.46}$$

which defines the expectation values of the errors and their correlations. In particular, the diagonal terms define the variances

$$\sigma_i^2 = C_{ii} = (\Gamma^{-1})^{ii} \,, \tag{7.2.47}$$

which are the squares of the standard deviations.

In order to exemplify this procedure of estimating the measurement uncertainties we can consider a toy model waveform $h(t)$ whose Fourier transform is described by a constant amplitude and a phase, $\tilde{h}(f) = Ae^{i\phi}$, that we want to infer. In order to compute the Fisher matrix we first compute the derivatives of the waveform w.r.t each variable:

$$\frac{\partial\tilde{h}}{\partial \ln A} = \tilde{h} \,, \tag{7.2.48}$$

$$\frac{\partial\tilde{h}}{\partial\phi} = i\tilde{h} \,. \tag{7.2.49}$$

Then, if we order the matrix components 1,2 as $\{\ln A, \phi\}$ we have

$$\Gamma_{11} = (\partial_{\ln A} h|\partial_{\ln A} h) = (h|h) = \rho^2 \,, \tag{7.2.50}$$

$$\Gamma_{12} = \Gamma_{21} = (\partial_{\ln A} h|\partial_\phi h) = (h|ih) = 0 \,, \tag{7.2.51}$$

$$\Gamma_{22} = (\partial_\phi h|\partial_\phi h) = (ih|ih) = \rho^2 \,, \tag{7.2.52}$$

which in matrix for reads

$$\hat{\Gamma} = \rho^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \,. \tag{7.2.53}$$

We find that amplitude and phase are uncorrelated, $C_{12} = 0$, and their variance scale inversely with the SNR:

$$\sigma_{\ln A} = \sigma_\phi = 1/\rho \,. \tag{7.2.54}$$

This means that for this toy model and a SNR of 10, we should expect an error of 10% in the fractional amplitude and of 0.1 radians in the phase.

Marginalizing over extrinsic parameters

When performing parameter estimation we do not need to include all the parameter as we can estimate some of them directly. We have already seen that both the arrival time and the amplitude that give the maximum SNR can be obtained separately and therefore they do not affect the matched filtering. After marginalizing over the amplitude and time, the likelihood was

$$\log \Lambda(d|\lambda) = \frac{1}{2}\frac{(h|d)^2}{(h|h)}\,, \tag{7.2.55}$$

where $\lambda = \{\lambda_i\}$ encapsulates all the parameters but the reference time and amplitude. We now want to show that something similar happens to the phase.

If we split our template as

$$h(t) = h_c(t)\cos\varphi + h_s(t)\sin\varphi\,, \tag{7.2.56}$$

to extract the phase term φ, then the log-likelihood is

$$2\log\Lambda = \frac{\left((h_c|d) + (h_s|d)\tan\varphi\right)^2}{(h_c|h_c) + (h_s|h_s)\tan^2\varphi + 2(h_c|h_s)\tan\varphi}\,. \tag{7.2.57}$$

This expression can be maximized directly in terms of $\tan\varphi$ and get the marginalized $\log\Lambda$. It is however more intuitive if we define a new basis $\{h_p, h_q\}$ that is orthonormal w.r.t. the scalar product: $(h_p|h_q) = 0$. It can be defined introducing two angles ϕ_p and ϕ_q as

$$h_p = h_c\cos\phi_p + h_s\sin\phi_p\,, \tag{7.2.58}$$

$$h_q = h_c\cos\phi_q + h_s\sin\phi_q\,. \tag{7.2.59}$$

This expression can be inverted to

$$h_c = \frac{h_p\sin\phi_q - h_q\sin\phi_p}{\cos\phi_p\sin\phi_q - \cos\phi_q\sin\phi_p}\,, \tag{7.2.60}$$

$$h_s = \frac{-h_p\cos\phi_q + h_q\cos\phi_p}{\cos\phi_p\sin\phi_q - \cos\phi_q\sin\phi_p}\,, \tag{7.2.61}$$

where note that the denominator is a common factor. After some trigonometry one finds that the likelihood of the remaining parameters is

$$2\log\Lambda = \frac{(h_p|d)^2}{(h_p|h_p)} + \frac{(h_q|s)^2}{(h_q|h_q)}\,. \tag{7.2.62}$$

This means that the maximization with respect to the rest of the variables is equivalent to the maximization of two matched filters h_p and h_q that are added in quadrature.

Exercise 7.3: Marginalization over phase

Following the steps described above, demonstrate that indeed the likelihood after marginalizing over phase and amplitude is given by Eq. (7.2.62).

Goodness-of-fit test

Literature:

This section follows the discussion of Ref. [50] where a pedagogical introduction to this concept was included.

The GW likelihood of a template h_T described by a set of parameters θ_T given a time series d follows a χ^2 distribution

$$\Lambda(d|h_T) \propto e^{-\chi^2}\,, \tag{7.2.63}$$

where

$$\begin{aligned}\chi^2 = 2\log\Lambda = (d-h|d-h) = (d|d) - 2(h_T|d) + (h_T|h_T)\\ \simeq (h|h) - 2(h_T|h) + (h_T|h_T) + (n|n)\,.\end{aligned} \tag{7.2.64}$$

Note that in the second line we have assumed that there is a real signal h in the data, $d = h + n$, and that the noise is uncorrelated with the signal and templates.

Our objective is to compare the fit of our template h_T to the true signal h. The increase in the χ^2 of the template with respect to the truth is given by

$$\Delta\chi^2 = \chi^2_{\rm template} - \chi^2_{\rm truth} = (h|h) - 2(h_T|h) + (h_T|h_T)\,, \tag{7.2.65}$$

where the common noise term vanishes. In the case of a perfect fit, $h_T = h$, then $\Delta\chi^2 = 0$. Larger $\Delta\chi^2$ will then indicate worse fits. If we have a family of templates $\{h^i_T\}$, their difference in $\Delta\chi^2$ will tell us how well each template fits the data.

It is interesting to note that we can rewrite the $\Delta\chi^2$ in terms of the mismatch ϵ defined in (7.2.16). That is:

$$\Delta\chi^2 = (h|h) - 2(1-\epsilon)\sqrt{(h_T|h_T)(h|h)} + (h_T|h_T)\,. \tag{7.2.66}$$

If the difference between the signal and the template is small, $(h|h) \approx (h_T|h_T)$ this simplifies to

$$\Delta\chi^2 \simeq 2\epsilon\rho^2_{\rm opt}\,, \tag{7.2.67}$$

where we have reintroduced the definition of the optimal SNR: $\rho^2_{\rm opt} = (h|h)$. The usefulness of this expression is that ϵ is basically only a statement of the shape of the signals and, therefore, one can scale its value to the SNR of the event in the detector to determine if a distortion in the signal could be detected. In the frequentist interpretation, an improvement of $\Delta\chi^2 = X^2$ from minimizing over a single parameter corresponds to

an $X\sigma$ preference for adding that parameter. This implies that an improvement in $\Delta\chi^2/\rho^2 \sim Y^2 \approx 2\epsilon$ corresponds to $(Y\rho) \approx \sqrt{2\epsilon}\rho$ in units of σ. In other words, with a high-SNR event, even a fractionally small change in the template can give a significant preference for an additional parameter. In a Bayesian framework, $\Delta\chi^2$ provides the likelihood ratio for the improvement in the posterior probability. This is in general then weighted with the prior, and as a consequence, it can be interpreted as parameter constraints for a flat prior.

7.2.3 Population analyses

Analyzing events individually allow us to learn about their properties and the type of signals that describe them. One can for instance learn about the masses and distances of the event, or if general relativity is a good description of the waveform. However, this is not all the information we can extract! If we consider all events collectively we can learn about their population properties. For example, we could learn about the population model that describes the distribution of source masses and redshifts $p_{\rm pop}(m_1, m_2, z)$.

It is very important to emphasize that in general, the population of sources is not going to be the same as the detected population. This is because our detectors are more or less sensitives depending on the properties of the source. This is a common factor to any astronomical instrument. There is always a **selection function**. However, what is special about GWs is that because we have models for our signals, which in fact are predicted by fundamental principles (namely by GR), we can simulate events inside the noise (referred in the GW community as **injections**) and therefore resolve the selection function precisely. With this selection function at hand we can the "undo" the selection effects and constrain directly the source population. This is very rare in Astronomy and should not be underestimated!

Hierarchical Bayesian inference

We are interested in consider a situation in which we have a set of $N_{\rm obs}$ observations that give us the data $\{d_i\}$ for $i = 1, \cdots, N_{\rm obs}$. For each of this observations we have already we already have the posterior distribution of the parameters $\theta = \{\theta_i\}$ for each event: $p(\theta|d_i)$. Our task is to infer the properties of the population model that predicts such parameters θ_i observed in the data $\{d_i\}$. In the jargon of statistics, the set of parameters that describe the population model are referred as **hyper-parameters** $\lambda = \{\lambda\}$. The framework to infer them is known as **hierarchical Bayesian inference**. The name eludes to the fact that there is a hierarchy between the underlying population model that predicts the event parameters that are then detected. Therefore, the inference is hierarchical process in which we need first $p(\theta|d_i)$ before getting $p(\lambda|\{d_i\})$.

Literature:

A nice review of this formalism is given in Mandel, Farr & Gair 2018 [51].

Following Bayes' theorem, the posterior probability of the hyper-parameters is given by

$$p(\lambda|\{d_i\}) \propto p(\lambda)p(\{d_i\}|\lambda) = p(\lambda)\prod_{i=1}^{N_{\rm obs}} \frac{p_{\rm pop}(\theta_i|\lambda)}{\int d\theta p_{\rm pop}(\theta|\lambda)}\,, \tag{7.2.68}$$

where in the second equality we have expressed the likelihood of all the events given the population model as the product of the likelihood of each event data given the model λ. This assumes that all detections are independent, which is a safe assumption for GW observations. Moreover, this expression assumes that all events in the population are detected (there is no selection effect) and that the measurement uncertainties are negligible, so that we evaluate the likelihood in the numerator $p_{\rm pop}(\theta|\lambda)$ at a single value θ_i. These two last assumptions are too restrictive and we need to relax them.

Selection effects and measurement uncertainties

As we have seen before, GW detectors are characterized by a given power spectral density (PSD) $S_n(f)$. We have seen that different waveform models and parameters can lead to different signal-to-noise ratio (SNR). In particular, the SNR scales with the amplitude $\rho \sim h$. For example, at leading order we have that the strain of an inspiraling binary scales as $h \propto \mathcal{M}_z^{5/3}/d_L$. Therefore, massive and nearby events are easier to detect than light and far away sources. We will generically define our **selection function** $p_{\rm det}(\theta)$ as the probability of detecting given set of parameters for a given waveform model $h(\theta)$ and detector noise $S_n(f)$. $p_{\rm det}(\theta)$ is a probability and therefore ranges from 0 to 1.

With selection effects the likelihood of the data set given the population model becomes

$$p(\{d_i\}|\lambda) = \prod_{i=1}^{N_{\rm obs}} \frac{p_{\rm pop}(\theta_i|\lambda)p_{\rm det}(\theta_i)}{\int d\theta p_{\rm pop}(\theta|\lambda)p_{\rm det}(\theta)} = \prod_{i=1}^{N_{\rm obs}} \frac{p_{\rm pop}(\theta_i|\lambda)}{\int d\theta p_{\rm pop}(\theta|\lambda)p_{\rm det}(\theta)}\,. \tag{7.2.69}$$

In the second equality we have used the fact that we are looking at the values of the parameters of the detected events and, as a consequence, $p_{\rm det}(\theta_i) = 1$ by definition.

The final step is then to include measurement uncertainties, because we do not have perfect measurements of each parameter. For example, for a given event the likelihood of the data given the population model is

$$p(d_i|\lambda) = \int d\theta p(\theta|d_i)p_{\rm pop}(\theta|\lambda)\,, \tag{7.2.70}$$

which is obtained simply by marginalizing over the posterior distribution of the parameters $p(\theta|d_i)$ that is obtained from he parameter estimation. Therefore, altogether the likelihood of the data set given the population model with selection effects and uncertainties is

$$p(\{d_i\}|\lambda) = \prod_{i=1}^{N_{\rm obs}} \frac{\int d\theta p(\theta|d_i)p_{\rm pop}(\theta|\lambda)}{\int d\theta p_{\rm pop}(\theta|\lambda)p_{\rm det}(\theta)}\,. \tag{7.2.71}$$

Putting this together with the prior and the evidence we obtain the posterior distribution for λ:

$$p(\lambda|\{d_i\}) = \frac{p(\lambda)}{p(\{d_i\})} \prod_{i=1}^{N_{\rm obs}} \frac{\int d\theta p(\theta|d_i) p_{\rm pop}(\theta|\lambda)}{\int d\theta p_{\rm pop}(\theta|\lambda) p_{\rm det}(\theta)}\,, \tag{7.2.72}$$

which is the expression we were looking for.

In practical terms, the posterior distribution of the parameters θ_i for a given event d_i is provided in terms of a set of posterior samples $\{\theta_i^j\}$ for $j = 0, \cdots, S_i$. Using, again, Bayes' theorem we have that

$$p(d_i|\theta) = \frac{p(\theta|d_i)p(d_i)}{p(\theta)}\,, \tag{7.2.73}$$

where $p(\theta)$ is the parameter estimation prior on θ. The integral in the numerator of the likelihood of the data set given the population model can then be written as an expectation value

$$\int d\theta p(\theta|d_i) p_{\rm pop}(\theta|\lambda) = \frac{1}{S_i}\sum_j^{S_i} p_{\rm pop}(\theta^j|\lambda)\frac{p(d_i)}{p(\theta)} = p(d_i)\left\langle \frac{p_{\rm pop}(\theta|\lambda)}{p(\theta)} \right\rangle_{\theta_j}\,. \tag{7.2.74}$$

Plugging this back to the general expression, and noting that $p(\{d_i\}) = \prod_i p(d_i)$, we obtain

$$p(\lambda|\{d_i\}) = p(\lambda) \prod_{i=1}^{N_{\rm obs}} \frac{1}{\int d\theta p_{\rm pop}(\theta|\lambda) p_{\rm det}(\theta)} \left\langle \frac{p_{\rm pop}(\theta_i|\lambda)}{p(\theta_i)} \right\rangle_{\theta_j}\,, \tag{7.2.75}$$

where the average is over the posterior samples of each event.

Including the rate information

So far we have only payed attention to inferring the properties of our population model λ that describe our set of $N_{\rm obs}$ given by $\{d_i\}$. However, we have not used the fact that the number of detections itself also has useful information. To extract that information we need to know how many detections we should expect for our population model.

To obtain the expected number of detections we first define $\mathcal{R}(z)$ as the merger per comoving volume, that is, the number of mergers per unit time of the source frame per unit volume. Then, the number of mergers producing GWs, $N_{\rm gw}$, as seen by the observer is

$$\frac{dN_{\rm gw}}{dt_{\rm det}} = \int \frac{\mathcal{R}(z)}{(1+z)}\frac{dV_c}{dz}dz\,. \tag{7.2.76}$$

Here we have simply used the definition of $\mathcal{R}(z)$ and transformed the time to the observer's frame, $dt_{\rm det} = (1+z)dt_{\rm src}$. For convenience we integrate in redshift instead of volume.

The number of GW detections will be given by the above expression if we include the selection function $p_{\rm det}(z, \theta)$. Because the selection function depends in general in all

the parameters describing the binary we need to marginalize over those (note that we have already separated the distance from the rest of parameters). Doing so we obtain

$$\frac{dN_{\rm det}}{dt_{\rm det}} = \int \frac{\mathcal{R}(z)}{(1+z)} \frac{dV_c}{dz} p(\theta) p_{\rm det}(z,\theta) dz d\theta \,. \tag{7.2.77}$$

Given the rate of detections $N_{\rm det}$, the probability of observing $N_{\rm obs}$ is described by a Poisson distribution

$$p(N_{\rm obs}|N_{\rm det}) \propto e^{-N_{\rm det}} (N_{\rm det})^{N_{\rm obs}} \,. \tag{7.2.78}$$

Then, the posterior distribution of the population parameters λ and the rate N is given by

$$p(\lambda, N|\{d_i\}) \propto p(\lambda) p(N) e^{-N_{\rm det}} (N_{\rm det})^{N_{\rm obs}} \prod_{i=1}^{N_{\rm obs}} \frac{1}{\int d\theta p_{\rm pop}(\theta|\lambda) p_{\rm det}(\theta)} \left\langle \frac{p_{\rm pop}(\theta_i|\lambda)}{p(\theta_i)} \right\rangle_{\theta_j} \,. \tag{7.2.79}$$

At this point it is interesting to note that the denominator of the product can be written as the fraction of mergers that are detected with respect to the total rate N:

$$\int d\theta p_{\rm pop}(\theta|\lambda) p_{\rm det}(\theta) = N_{\rm det}/N \equiv \xi \,. \tag{7.2.80}$$

By its own definition, ξ is a quantity that does not depend on N. This implies that the $(N_{\rm det})^{N_{\rm obs}}$ term in front of the product can be simplified with the denominators when put all together:

$$p(\lambda, N|\{d_i\}) \propto p(\lambda) p(N) N^{N_{\rm obs}} e^{-N_{\rm det}} \prod_{i=1}^{N_{\rm obs}} \left\langle \frac{p_{\rm pop}(\theta_i|\lambda)}{p(\theta_i)} \right\rangle_{\theta_j} \,. \tag{7.2.81}$$

If we marginalize this expression using a logarithmic prior in N, $p(N) \propto 1/N$, then the above expression can be simplified to

$$\begin{aligned} p(\lambda|\{d_i\}) &\propto \int dN p(\lambda) N^{N_{\rm obs}-1} e^{-N\xi} \prod_{i=1}^{N_{\rm obs}} \left\langle \frac{p_{\rm pop}(\theta_i|\lambda)}{p(\theta_i)} \right\rangle_{\theta_j} \\ &\propto p(\lambda) \xi^{-N_{\rm obs}} \prod_{i=1}^{N_{\rm obs}} \left\langle \frac{p_{\rm pop}(\theta_i|\lambda)}{p(\theta_i)} \right\rangle_{\theta_j} \,, \end{aligned} \tag{7.2.82}$$

where in the second line we have used that

$$\begin{aligned} \int dN N^{N_{\rm obs}-1} e^{-N\xi} &= \xi^{-N_{\rm obs}} \int \frac{dN}{N} e^{-N\xi} (N\xi)^{N_{\rm obs}} \\ &= \xi^{-N_{\rm obs}} \int dN_{\rm det} e^{-N_{\rm det}} (N_{\rm det})^{N_{\rm obs}-1} \\ &= \xi^{-N_{\rm obs}} \Gamma(N_{\rm obs}) = \xi^{-N_{\rm obs}} (N_{\rm obs}-1)! \end{aligned} \tag{7.2.83}$$

Here in the first line we have simply multiply and divide by $\xi^{N_{\rm obs}}$. In the second line we have redefining the integration variables, $N_{\rm det} = \xi N$. In the third line we have introduced the definition of the gamma function. Therefore, when marginalizing over N we recover our previous expression.

List of exercises

Bibliography

[1] H. Reall, *Part 3 Black Holes*, 2020.

[2] S. M. Carroll, *Spacetime and Geometry: An Introduction to General Relativity*. Cambridge University Press, 7, 2019. Based on gr-qc/9712019.

[3] R. H. Boyer and R. W. Lindquist, *Maximal analytic extension of the Kerr metric*, *J. Math. Phys.* **8** (1967) 265.

[4] T. Harmark, *General Relativity and Cosmology*, 2023.

[5] R. M. Wald, *General Relativity*. Chicago Univ. Pr., Chicago, USA, 1984.

[6] "Black Hole Perturbation Toolkit." (bhptoolkit.org).

[7] M. P. Ryan, *Teukolsky equation and Penrose wave equation*, *Phys. Rev. D* **10** (1974) 1736–1740.

[8] R. M. Wald, *On perturbations of a Kerr black hole*, *J. Math. Phys.* **14** (1973), no. 10 1453–1461.

[9] B. F. Whiting, *Mode Stability of the Kerr Black Hole*, *J. Math. Phys.* **30** (1989) 1301.

[10] A. Zee, *Einstein Gravity in a Nutshell*. Princeton University Press, New Jersey, 5, 2013.

[11] H. A. Buchdahl, *General relativistic fluid spheres*, *Phys. Rev.* **116** (Nov, 1959) 1027–1034.

[12] J. R. Oppenheimer and H. Snyder, *On continued gravitational contraction*, *Phys. Rev.* **56** (Sep, 1939) 455–459.

[13] R. Ruffini and S. Bonazzola, *Systems of selfgravitating particles in general relativity and the concept of an equation of state*, *Phys. Rev.* **187** (1969) 1767–1783.

[14] J. B. Hartle, *Gravity: An introduction to Einstein's general relativity*. 2003.

[15] S. Chandrasekhar, *The maximum mass of ideal white dwarfs*, *Astrophys. J.* **74** (1931) 81–82.

[16] R. C. Tolman, *Static solutions of einstein's field equations for spheres of fluid*, *Phys. Rev.* **55** (Feb, 1939) 364–373.

[17] J. R. Oppenheimer and G. M. Volkoff, *On massive neutron cores*, *Phys. Rev.* **55** (Feb, 1939) 374–381.

[18] V. Kalogera and G. Baym, *The maximum mass of a neutron star*, *Astrophys. J. Lett.* **470** (1996) L61–L64, [astro-ph/9608059].

[19] W. A. Fowler and F. Hoyle, *Neutrino Processes and Pair Formation in Massive Stars and Supernovae*, *Astrophys. J. Suppl.* **9** (1964) 201–319.

[20] W. M. Farr, N. Sravan, A. Cantrell, L. Kreidberg, C. D. Bailyn, I. Mandel, and V. Kalogera, *The Mass Distribution of Stellar-Mass Black Holes*, *Astrophys. J.* **741** (2011) 103, [arXiv:1011.1459].

[21] **LIGO Scientific, Virgo,, KAGRA, VIRGO** Collaboration, A. G. Abac et al., *Observation of Gravitational Waves from the Coalescence of a 2.5–4.5 M Compact Object and a Neutron Star*, *Astrophys. J. Lett.* **970** (2024), no. 2 L34, [arXiv:2404.04248].

[22] A. M. Ghez et al., *Measuring Distance and Properties of the Milky Way's Central Supermassive Black Hole with Stellar Orbits*, *Astrophys. J.* **689** (2008) 1044–1062, [arXiv:0808.2870].

[23] K. El-Badry et al., *A Sun-like star orbiting a black hole*, *Mon. Not. Roy. Astron. Soc.* **518** (2023), no. 1 1057–1085, [arXiv:2209.06833].

[24] K. El-Badry et al., *A red giant orbiting a black hole*, *Mon. Not. Roy. Astron. Soc.* **521** (2023), no. 3 4323–4348.

[25] **Gaia** Collaboration, P. Panuzzo et al., *Discovery of a dormant 33 solar-mass black hole in pre-release Gaia astrometry*, arXiv:2404.10486.

[26] B. Paczynski, *Gravitational microlensing by the galactic halo*, *Astrophys. J.* **304** (1986) 1–5.

[27] M. Schmidt, *3C 273 : A Star-Like Object with Large Red-Shift*, *Nature* **197** (1963), no. 4872 1040.

[28] **Event Horizon Telescope** Collaboration, K. Akiyama et al., *First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole*, *Astrophys. J. Lett.* **875** (2019) L1, [arXiv:1906.11238].

[29] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott et al., *Observation of Gravitational Waves from a Binary Black Hole Merger*, *Phys. Rev. Lett.* **116** (2016), no. 6 061102, [arXiv:1602.03837].

[30] E. E. Flanagan and S. A. Hughes, *The Basics of gravitational wave theory*, *New J. Phys.* **7** (2005) 204, [gr-qc/0501041].

[31] S. Weinberg, *Photons and Gravitons in S-Matrix Theory: Derivation of Charge Conservation and Equality of Gravitational and Inertial Mass*, *Phys. Rev.* **135** (1964) B1049–B1056.

[32] M. Maggiore, *Gravitational Waves. Vol. 1: Theory and Experiments*. Oxford University Press, 2007.

[33] R. A. Isaacson, *Gravitational Radiation in the Limit of High Frequency. I. The Linear Approximation and Geometrical Optics*, *Phys. Rev.* **166** (1968) 1263–1271.

[34] R. A. Isaacson, *Gravitational Radiation in the Limit of High Frequency. II. Nonlinear Terms and the Ef fective Stress Tensor*, *Phys. Rev.* **166** (1968) 1272–1279.

[35] J. D. E. Creighton and W. G. Anderson, *Gravitational-wave physics and astronomy: An introduction to theory, experiment and data analysis*. 2011.

[36] P. C. Peters and J. Mathews, *Gravitational radiation from point masses in a Keplerian orbit*, *Phys. Rev.* **131** (1963) 435–439.

[37] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*. W. H. Freeman, San Francisco, 1973.

[38] K. S. Thorne, *Gravitational Radiation: A New Window onto the Universe*. unpublished, 1989.

[39] W. H. Press, *Long Wave Trains of Gravitational Waves from a Vibrating Black Hole*, *Astrophys. J. Lett.* **170** (1971) L105–L108.

[40] R. H. Price, *Nonspherical perturbations of relativistic gravitational collapse. 1. Scalar and gravitational perturbations*, *Phys. Rev. D* **5** (1972) 2419–2438.

[41] R. H. Price, *Nonspherical Perturbations of Relativistic Gravitational Collapse. II. Integer-Spin, Zero-Rest-Mass Fields*, *Phys. Rev. D* **5** (1972) 2439–2454.

[42] E. W. Leaver, *Spectral decomposition of the perturbation response of the Schwarzschild geometry*, *Phys. Rev. D* **34** (1986) 384–408.

[43] L. Blanchet, *Radiative gravitational fields in general relativity. 2. Asymptotic behaviour at future null infinity*, *Proc. Roy. Soc. Lond. A* **409** (1987) 383–399.

[44] D. W. Hogg, *Distance measures in cosmology*, astro-ph/9905116.

[45] E. Schrodinger, *The proper vibrations of the expanding universe*, *Physica* **6** (July, 1939) 899–912.

[46] T. T. Nakamura and S. Deguchi, *Wave Optics in Gravitational Lensing*, *Prog. Theor. Phys. Suppl.* **133** (1999) 137–153.

[47] R. Blandford and R. Narayan, *Fermat's principle, caustics, and the classification of gravitational lens images*, *Astrophys. J.* **310** (1986) 568–582.

[48] H.-Y. Chen, J. M. Ezquiaga, and I. Gupta, *Cosmography with next-generation gravitational wave detectors*, *Class. Quant. Grav.* **41** (2024), no. 12 125004, [arXiv:2402.03120].

[49] J. M. Ezquiaga, W. Hu, and R. K. L. Lo, *Identifying strongly lensed gravitational waves through their phase consistency*, *Phys. Rev. D* **108** (2023), no. 10 103520, [arXiv:2308.06616].

[50] J. M. Ezquiaga, W. Hu, M. Lagos, M.-X. Lin, and F. Xu, *Modified gravitational wave propagation with higher modes and its degeneracies with lensing*, *JCAP* **08** (2022), no. 08 016, [arXiv:2203.13252].

[51] I. Mandel, W. M. Farr, and J. R. Gair, *Extracting distribution parameters from multiple uncertain observations with selection biases*, *Mon. Not. Roy. Astron. Soc.* **486** (2019), no. 1 1086–1093, [arXiv:1809.02063].