Lecture notes Relativity Theory II (Winter 2017)

March 5, 2017

Harald Pfeiffer

CITA, University of Toronto

Revision History

- **Jan 15, 2016**: Add Chapter 4 (Event Horizons).
  Correct missing factor 1/2 in Eqs. (2.31)–(2.34). Correct typo in Eq. (2.36).
- **Jan 29**: Add first part of Chapter 5 (post-Newtonian approximation)
- **Feb 2**: Add more post-Newtonian, up to Eq. (5.71).
- **Feb 5**: Insert discussion of Keplerian 2-body problem (Eqs. (5.5)—(5.12) and the surrounding text). Add Section 5.3.
- **Feb 8**: Fix missing factor 1/5 in Eq. (5.87). Add Chapter 6.
- **Feb 16**: Add bullet-point about Spatial metric, namely Eq. (6.14). This bullet was accidentally omitted in the earlier version of the notes. Correct typo in Eq. (6.40).
- **Feb 26**: Add Chapters 7 and 8.
- **Mar 1**: Add image caption to Fig. 2.5
  Add Chapter 9
- **Mar 5**: Add item 7 on page 59. Use lowercase m (instead of M) to denote total mass in Chapter 8. Chapter 9 significantly expanded and reordered.

Contents

Chapter 1

Introduction

This course serves as in-depth introduction to the gravitational wave discoveries that happened in 2015, and new discoveries that –we hope– are going to happen soon: Coalescing binary black holes were observed, and we hope on coalescing neutron stars in the near future. As such, you will learn about the objects that are colliding: black holes and neutron stars, and the basics of gravitational waves. Because gravitational waves are only emitted in *dynamical and non-symmetric* situations, we will pay significant emphasis on how to treat General Relativity in such generic circumstances, where few (or no) symmetries are present, and the standard GR-I analytical techniques are not applicable.

Textbooks

There is no good text-book that covers all material to be discussed (hey, GW discoveries are still too new!). We will refer to the following sources:

- Hobson, Efstathiou, Lasenby: "General Relativity: An Introduction for Physicists" (Cambridge University Press)
- Carroll, "Spacetime and Geometry: An Introduction to General Relativity", Pearson (early version of notes available at `https://arxiv.org/abs/gr-qc/9712019`).
- Townsend, "Black holes (Lecture Notes)" `https://arxiv.org/abs/gr-qc/9707012`
- The 2nd half of the course (taught by Prof. Chris Thomson) will also refer to Shapiro & Teukolsky: "Black holes, white dwarfs and neutron stars"

Chapter 2

Geodesics around Schwarzschild & Kerr black holes

Geodesics describe orbits of test-masses, here, around a black hole. Binary black hole orbits inherit many features that are already present in the simpler geodesic motion. We will perform an in-depth analysis using the effective potential method. This method utilizes *constants of motion* to simplify the problem so much that all qualitative features of the solutions (and quite a few quantitative features) can be understood with little calculations.

Some of this lecture re-cap material of GR-I, but not all of you took GR-I.

2.1 Geodesics, space-time symmetries and constants of motion

We parameterize the geodesic as a curve

$$x^a(\lambda), \qquad \lambda \in \mathbb{R}. \tag{2.1}$$

The **velocity** is

$$u^a = \frac{dx^a}{d\lambda} = \dot{x}^a. \tag{2.2}$$

For **affine parameter** λ, the geodesic equation reads

$$u^a \nabla_a u^b = 0. \tag{2.3}$$

It holds that

$$-\dot{x}^a \dot{x}^b g_{ab} = \kappa \equiv \begin{cases} 1, & \text{time-like geodesics (massive particles)} \\ 0, & \text{null geodesics (massless particles)} \end{cases} \tag{2.4}$$

Proof: The derivative of $g_{ab}\dot{x}^a\dot{x}^b$ along the curve is

$$\begin{aligned} u^a \nabla_a \left(g_{bc} u^b u^c \right) &= u^a \left(u^b u^c \nabla_a g_{bc} + g_{bc} u^b \nabla_a u^c + g_{bc} u^c \nabla_a u^b \right) \\ &= 0 + 2 g_{bc} u^b u^a \nabla_a u^c \\ &= 0. \end{aligned}$$

We see that $g_{ab}\dot{x}^a\dot{x}^b$ =const along the geodesic. Rescale λ by a constant factor to achieve Eq. (2.4)

Equation (2.4) is our first constant of motion. To obtain more, we need to consider symmetries of the spacetime. Symmetries are described by Killing vectors. A **Killing vector** k^a satisfies $\nabla_{(a}k_{b)} \equiv \frac{1}{2}\left(\nabla_a k_b + \nabla_b k_a\right) = 0$. It follows:

$$\begin{aligned} u^a\nabla_a(k_b u^b) &= k_b u^a \nabla_a u^b + u^a u^b \nabla_a k_b \\ &= 0 + u^a u^b \nabla_{(a}k_{b)} \\ &= 0. \end{aligned}$$

Here, we have used the geodesic equation to eliminate the first term. Adding symmetrization in $\nabla_a k_b$ in the second term is possible because it is multiplied by the symmetric tensor $u^a u^b$. Then Killing's equation causes the second term to vanish.) Therefore, a Killing-vector induces a constant of motion along the geodesic:

$$\dot{x}^a\nabla_a\left(k_b\dot{x}^b\right) = 0 \quad \Rightarrow \quad k_a\dot{x}^a = \text{const} \tag{2.5}$$

2.2 Schwarzschild spacetime

The Schwarzschild black hole in Schwarzschild coordinates (t, r, θ, ϕ) is given by

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1}dr^2 + r^2\left(d\Omega^2 + \sin^2\theta d\phi^2\right), \tag{2.6}$$

where we have used units such that $G = 1, c = 1$. Note that the term "Schwarzschild" is used with two meanings: "Schwarzschild spacetime" refers to the whole manifold which solves Einstein's equations. "Schwarzschild coordinates" are just one of many coordinate-systems one can use to describe this manifold.

The geodesic has coordinates $x^a = [t, r, \theta, \phi]$, and velocity $u^a = \dot{x}^a = [\dot{t}, \dot{r}, \dot{\theta}, \dot{\phi}]$. Note that overdot indicates the derivative with respect to the affine parameter, e.g. $\dot{t} \equiv dt/d\lambda$.

We rotate the coordinates such that the orbit is in the equatorial plane: $\theta = pi/2$, $\dot{\theta} = 0$, $\sin\theta = 1$.

Time-translation Killing vector
$(\partial_t)^a = [1, 0, 0, 0]$. Lowering the index yields $(\partial_t)_a = g_{ab}(\partial_t)^b = \left[-\left(1 - \frac{2M}{r}\right), 0, 0, 0\right]$, and therefore, the associated constant of motion is

$$-\left(\partial_t\right)_a\dot{x}^a = \left(1 - \frac{2M}{r}\right)\dot{t} = \text{const} \equiv E. \tag{2.7}$$

This is the energy per unit-restmass of the orbiting body.

Rotational Killing vector
$(\partial_\phi)^a = [0, 0, 0, 1]$ (rotation axis normal to orbital plane). Lowering the index yields $(\partial_\phi)^a = g_{ab}(\partial_\phi)^b = [0, 0, 0, r^2]$, and the associated constant of motion is

$$\left(\partial_\phi\right)_a\dot{x}^a = r^2\dot{\phi} = \text{const} \equiv L, \tag{2.8}$$

the angular momentum of the test-body per unit-mass.

Now let's combine the three constants of motion that we have discovered: Spelling out Eq. (2.4) explicitly for the Schwarzschild metric (recall $\theta = \pi/2$) yields

$$-\kappa = -\left(1-\frac{2M}{r}\right)\dot{t}^2 + \left(1-\frac{2M}{r}\right)^{-1}\dot{r}^2 \qquad + r^2\dot{\phi}^2. \tag{2.9}$$

Multiplying by $(1-2M/r)$, and replacing the terms with $\dot{t}^2$ and $\dot{\phi}^2$ by E and L, respectively, gives:

$$-\kappa\left(1-\frac{2M}{r}\right) = -E^2 + \dot{r}^2 + \left(1-\frac{2M}{r}\right)\frac{L^2}{r^2}. \tag{2.10}$$

Solving for $\dot{r}^2$, we find:

$$\boxed{\frac{1}{2}\dot{r}^2 + V(r) = \frac{1}{2}E^2,} \tag{2.11}$$

where

$$\boxed{V(r) = \frac{1}{2}\kappa - \kappa\frac{M}{r} + \frac{L^2}{2r^2} - \frac{ML^2}{r^3}} \tag{2.12}$$

is called the **effective potential**. Equations (2.11) and (2.12) are extremely powerful to understand the structure and qualitative properties of geodesics around a Schwarzschild black hole

Equation (2.11) is identical to the equation that governs mechanical motion in a 1-dimension poential. E.g., taking a time-derivative of Eq. (2.11) yields

$$\dot{r}\ddot{r} + V'(r)\dot{r} = 0, \tag{2.13}$$

where $V'(r) = \partial V/\partial r$. Division by $\dot{r}$ yields the well-known second order ODE for 1-D motion,

$$\ddot{r} = -V'(r). \tag{2.14}$$

The only differences are: (i) in the GR-case, derivatives $\dot{r}$ are derivatives with respect to the affine parameter of the geodesic (i.e. proper time of the test-mass for time-like geodesics). This changes the *timing* of how the orbit is traversed, but not the *shape* of the orbit. And, (ii) the energy *squared* appears on the right-hand-side of Eq. (2.11). This is just what happens. However, because E also contains the rest-mass of the particle this is still quite similar to standard Newtonian mechanics. Let's separate out the rest-mass and write $E = 1 + E_{\rm newtonian}$ (recall that our E is the energy per unit-rest mass, so the rest-mass is just 'one'). We now have

$$\frac{1}{2}E^2 - \frac{1}{2} = E_{\rm newtonian} + \frac{1}{2}E^2_{\rm newtonian} \approx E_{\rm newtonian}. \tag{2.15}$$

2.2.1 Geodesics of massive particles: $\kappa = 1$

Let's get going!

$$V(r) = \frac{1}{2} - \frac{M}{r} + \frac{L2}{2r^2} - \frac{L^2 M}{r^3}, \qquad L = r^2\phi. \tag{2.16}$$

The terms on the right-hand-side represent, in order:

- rest-mass
- Newtonian gravitational attraction
- Centrifugal repulsion
- A genuine GR term, which increases the gravitational attraction

2.2.1.1 Graph and qualitative features

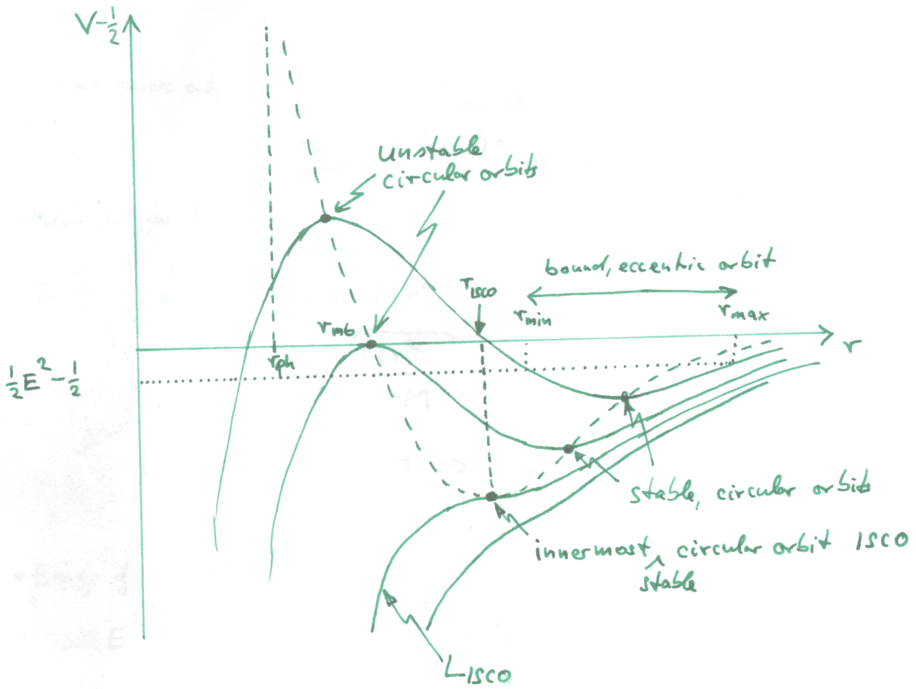


Figure 2.1: Effective potential for a Schwarzschild black hole. Plotted are several curves of $V(r)$ for different values of L.

Figure 2.1 plots several effective potential contours. From these contours we learn:

- Each line represents $V(r)$ at a fixed value of L, with smaller values of L being further down.
- For given value of $E^2 < 1/2$ (e.g. dotted horizontal line), the radial motion is bounded between with radius $r_{\rm min} \le r \le r_{\rm max}$, i.e. a bound eccentric orbit.
- *Minima* of V correspond to circular orbits at that given angular momentum L.

- There is an *innermost stable circular orbit*, $r_{\rm ISCO}$.
- *unstable circular orbits* correspond to maxima of V. Bound, unstable, circular orbits (i.e. $E^2/2 < 1/2$) exist for $r_{\rm mb} < r \le r_{\rm ISCO}$, i.e. closer to the black hole.
- $r_{\rm mb}$ (mb="marginally bound") is also the closest distance any orbiting test-mass can come to the black hole, no matter how eccentric its orbit is.
- Unbound, unstable circular orbits correspond to maxima of V at positive values of $V - 1/2$. They occur at radii $r_{\rm photon} < r < r_{\rm mb}$.
- $r_{\rm photon}$ (ph='photon') is also the closest distance any geodesic can approach the BH and still escape, no matter the values of E and L.

2.2.1.2 Circular orbits

- For given L, circular orbits for massive particles exist at radii r_c satisfying

$$0 = \frac{\partial V}{\partial r}(r_c) = \frac{M}{r_c^2} - \frac{L^2}{r_c^3} + \frac{3ML^2}{r_c^4} \quad \Rightarrow \quad r_c = \frac{L^2}{2M} \pm \sqrt{\frac{L^4}{4M^2} - 3L^2}. \tag{2.17}$$

  The +-sign in Eq. (2.17) corrresponds to stable circular orbits, the −-sign to unstable orbits. Note that the square-root only exists if $L \ge L_{\rm ISCO} = \sqrt{12}M$. This is the minimum angular momentum of a test-body orbiting a Schwarzschild black hole.

- Solve Eq. (2.17) for L:

$$L^2(r_c - 3M) = Mr_c^2 \quad \Rightarrow \quad L = \sqrt{\frac{Mr_c^2}{r_c - 3M}} \tag{2.18}$$

  The angular momentum $L \to \infty$ as the radius $r \to 3M = r_{\rm photon}$.

- The Energy of a circular orbit is:

$$E^2 = 2V = 1 - \frac{2M}{r} + \frac{L^2}{r^2} - \frac{L^2 M}{r^3}. \tag{2.19}$$

  Substitute Eq. (2.18) and simplify:

$$E^2 = 1 - \frac{M}{r}\frac{r - 4M}{r - 3M}. \tag{2.20}$$

  Because of the factor $r - 4M$, the energy $E^2 = 1$ when $r = r_{\rm mb} = 4M$. We now have all three characteristic radii determined:

$r_{\rm ISCO} = 6M$	Innermost stable circular orbit
$r_{\rm mb} = 4M$	smallest radii achievable by eccentric, bound orbits
$r_{\rm photon} = 3M$	smallest radius achievable by geodesics not falling into the BH (also only radius for a circular null geodesic, hence the name)

- Accretion disks end at $r_{\rm ISCO}$. Particles at the inner edge have energy per unit rest-mass of

$$E_{\rm ISCO} = \sqrt{2V(r_{\rm ISCO})} = \sqrt{8/9} \approx 0.943. \tag{2.21}$$

  The difference to 1 (i.e. about 5.7% of the restmass) represents gravitational binding energy, which was emitted during the slow inspiral of the particle to $r_{\rm ISCO}$.

2.2.1.3 Small eccentricity – periastron advance

Let's Taylor expand around a circular orbit. We write $r(t) = r_c + \delta r(t)$ and substitute into Eq. (2.14).

$$\delta\ddot{r} + V'(r_c + \delta r) = 0. \tag{2.22}$$

Taylor expansion yields $V'(r_c + \delta r) = V'(r_c) + V''(r_c)\delta r = V''(r_c)\delta r$, because $V'(r_c) = 0$, by virtue of r_c being a circular orbit. Therefore, we obtain the equation for a harmonic oscillator:

$$\delta\ddot{r} + V''(r_c)\delta r = 0. \tag{2.23}$$

As for any harmonic oscillator, the oscillation frequency ω_r satisfies

$$\omega_r^2 = V''(r_c) \tag{2.24}$$

$$= -\frac{2M}{r_c^3} + \frac{3L^2}{r_c^4} - \frac{12ML^2}{r_c^5} \tag{2.25}$$

$$= -\frac{2}{r_c}\left(\frac{M}{r_c^2} - \frac{L^2}{r_c^3} + \frac{3ML^2}{r_c^4}\right) + \frac{L^2}{r_c^4} - \frac{6ML^2}{r_c^5}, \tag{2.26}$$

where the second equality just rearranges some terms. We do this, because the quantity inside the parentheses equals $V'(r_c)$, which vanishes because r_c is a circular orbit. In the remaining terms, we substitute in $L = r_c^2\dot{\phi}$ (this is Eq. 2.8), and thus obtain

$$\omega_r^2 = \dot{\phi}^2 - \frac{6M}{r_c}\dot{\phi}^2. \tag{2.27}$$

Therefore, the ratio of radial and angular frequencies are

$$\frac{\omega_r}{\dot{\phi}} = \sqrt{1 - \frac{6M}{r_c}} \approx \left(1 - \frac{3M}{r_c}\right), \tag{2.28}$$

where the last approximation is valid for $r/M \gg 1$.

The radial oscillations have *lower* frequency than the orbital frequency. Therefore, during one radial oscillation period, the phi-motion proceeds through *more* than one full orbit: We have established **periastron advance**. Per *radial* oscillation period, the periastron advances by

$$\Delta\phi = 2\pi\left(1 - \sqrt{1 - \frac{6M}{r_c}}\right) \approx 2\pi\frac{3M}{r_c} = 6\pi\frac{GM}{c^2 r_c} \tag{2.29}$$

This result is only valid in the limit of small eccentricity, but still, we could obtain this result with remarkable simple calculations, compared to explicitly integrating the geodesic equations.

2.2.1.4 Large eccentricity – Zoom-Whirl orbits

Going back to Figure 2.1, let's consider values for E and L, such that (i) the orbit is bounded ($E^2 < 1/2$) and (ii) the effective potential contour has its maximum (the unstable circular orbit) just *barely* larger than the energy E. This situation is plotted in Figure 2.2.

By our careful choice of effective potential contour and energy, we have achieved that $r_{\min}$ of the orbit is very close to the maximum of $V(r)$. Therefore, near $r_{\min}$, the gradient

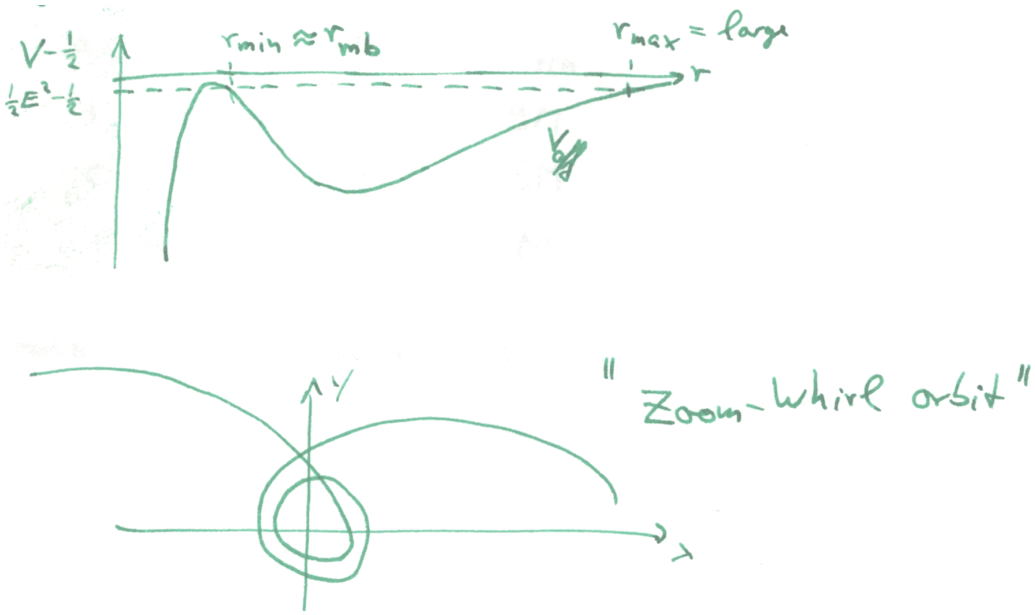


Figure 2.2: Zoom-whirl orbits. **Top:** a typical effective-potential contour that yields to zoom-whirl orbits. **Bottom:** illustration of the resulting orbital trajectory.

$V'(r)$ is very small. By Eq. (2.14), therefore $\ddot{r}$ near apastron will be very small, and therefore, the the test-mass will spend a large amount of time near $r_{\rm min}$. The test-mass still orbits the central black hole with an orbital frequency $\dot{\phi}$, which is not exceptionally small, and therefore it is possible that the test-mass performs *several* orbits near $r_{\rm min}$ before moving away toward $r_{\rm max}$. This is indicated in the lower panel of Fig. 2.2.

2.2.2 Photon trajectories $\kappa = 0$

For photons which have $\kappa = 0$, Eq. (2.11) simplifies to

$$\frac{1}{2}\dot{r}^2 + \frac{1}{2}\left(1 - \frac{2M}{r}\right)\frac{L^2}{r^2} = \frac{1}{2}E^2. \tag{2.30}$$

It is not yet obvious, but it is possible to *eliminate one of the two parameters* L, E from this equation. To do this, we first divide by L^2:

$$\frac{1}{2}\left(\frac{dr}{Ld\lambda}\right)^2 + \frac{1}{2r^2}\left(1 - \frac{2M}{r}\right) = \frac{E^2}{2L^2}. \tag{2.31}$$

Next, we note that for null-trajectories, the affine parameter λ can be freely rescaled ($ds^2 = 0$, and multiplication by an overall constant does not change this). We use this freedom to redefine the affine parameter $\lambda' = L\lambda$. Furthermore, we define

$$b \equiv \frac{L}{E}. \tag{2.32}$$

Figure 2.3: Photon-Trajectory approaching the black hole from infinity with an offset of b relative to a straight line. The incoming particle has momentum $p = E/c$. Its angular momentum (relative to the center of the black hole is $L = bp = Eb/c$. Setting $c = 1$ yields Eq. (2.32).

The meaning of b is illustrated in Fig. 2.3: For an unbound particle (or photon) approaching the black hole from large distance, b is the *impact parameter*, i.e. the distance by which the particle would miss the black hole, if it moved on a straight line.

With these two changes, Eq. (2.31) becomes

$$\frac{1}{2}\left(\frac{dr}{d\lambda'}\right)^2 + \frac{1}{2r^2}\left(1 - \frac{2M}{r}\right) = \frac{1}{2b^2}. \tag{2.33}$$

As promised, individual appearance of E and L has been eliminated from Eq. (2.33), and the effective potential depends only on the geometric parameter b. This means that photon-trajectories are *independent of the energy of the photon.* The properties of photon-trajectories are determined by the effective potential

$$V_{\rm null}(r) = \frac{1}{2r^2}\left(1 - \frac{2M}{r}\right), \tag{2.34}$$

which is plotted in Figure 2.4.

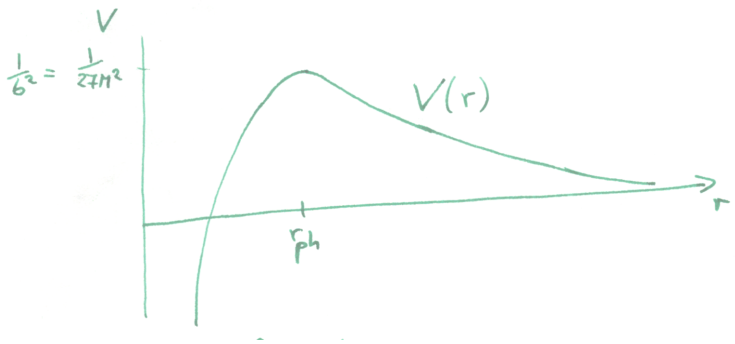


Figure 2.4: Effective potential $V_{\rm null}$ for null geodesics, Eq. (2.34).

$V_{\rm null}(r)$ is qualitatively very different from the effective potential for massive particles. The main differences are:

- No local minima of $V_{\rm null}$, and therefore no bound stable photon orbits around the black hole.

- $V_{\rm null}(r)$ has only one maximum at $r = r_{\rm photon} = 3M$. This maximum corresponds to a critical impact parameter of

$$b_{\rm crit} = \sqrt{\frac{1}{V_{\rm null}(r_{\rm photon})}} = \sqrt{27}M. \tag{2.35}$$

- At the maximum, i.e. $r = r_{\rm photon}$ there is one unstable circular orbit for photons (the existence of this orbit explains the name "photon-radius" $r_{\rm photon}$).
- Geodesics that approach the black hole with impact parameter $b > b_{\rm crit}$ will approach the black hole until they encounter the potential $V_{\rm null}$ (cf. Fig. 2.4. They will reverse direction and move away to infinity – I.e. the photon is scattered by the black hole.
- Photons with $b < b_{\rm crit}$ will pass *over* the maximum of $V_{\rm null}(r)$ in Fig. 2.4, and will reach $r = 0$ – i.e. these photons are *captured* by the black hole.
- Therefore, the cross-section of a Schwarzschild black hole to photon-capture is $\sigma = \pi b_{\rm crit}^2 = 27\pi M^2$.

2.3 Kerr Black holes

To recap, the Kerr-metric in **Boyer-Lindquist coordinates** is given by

$$\begin{aligned} ds^2 = &-\left(1 - \frac{2Mr}{\rho^2}\right) dt^2 - \frac{2Mar\sin^2\theta}{\rho^2}\left(dtd\phi + d\phi dt\right) \\ &+ \frac{\rho^2}{\Delta} dr^2 + \rho^2 d\theta^2 + \left[r^2 + a^2 + \frac{2Mra^2\sin^2\theta}{\rho^2}\right] d\phi^2, \end{aligned} \tag{2.36}$$

where

$$\Delta = r^2 - 2Mr + a^2, \qquad \rho^2 = r^2 + a^2\cos^2\theta.$$

The parameters have the following meaning:

- M mass
- aM angular momentum; $a/M \equiv \chi$ dimensionless spin. To avoid naked singularities, the spin must not exceed 1, i.e. $\chi \leq 1$.
- The event horizon is at $\Delta = 0$, i.e. $r_\pm = M \pm \sqrt{M^2 - a^2}$.

2.3.1 Equatorial Geodesics

For equatorial orbits, the analysis techniques for geodesics in Schwarzschild are still applicable, although more tedious. The results are qualitatively similar, but the numerical values vary with spin. This calculation has been done in **Bardeen, Press, Teukolsky,**

ApJ 178, p.347 (1972):

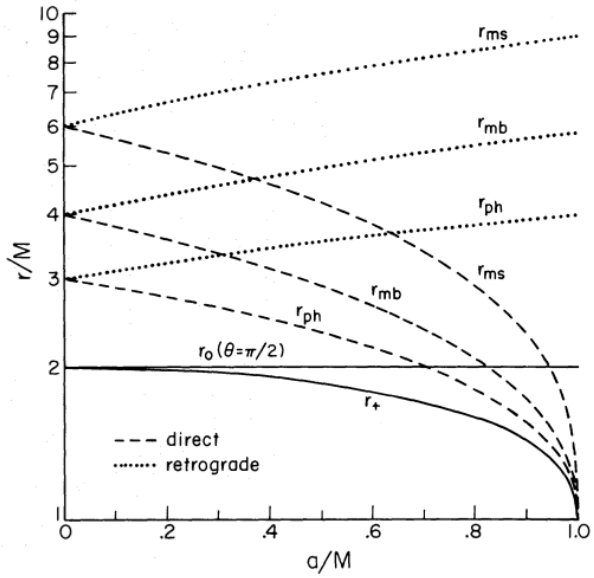


FIG. 1.—Radii of circular, equatorial orbits around a rotating black hole of mass M, as functions of the hole's specific angular momentum a. Dashed and dotted curves (for direct and retrograde orbits) plot the Boyer-Lindquist coordinate radius of the innermost stable (ms), innermost bound (mb), and photon (ph) orbits. Solid curves indicate the event horizon (r_+) and the equatorial boundary of the ergosphere (r_0).

Note that Bardeen et al use the symbol $r_{\rm ms}$ for $r_{\rm ISCO}$. The general tendency is that for co-rotating orbits, all characteristic radii become smaller, whereas for counter-rotating orbits, the characteristic radii become larger. The limiting values for extremal black holes ($\chi = 1$) are summarized in Table 2.1.

Spin a/M	r_+	$r_{\rm ISCO}$	$1 - E_{\rm ISCO}$	$r_{\rm mb}$	$r_{\rm photon}$
-1 (counter-rotating)	M	9M	3.8%	$(3+\sqrt{2})M$	4M
0	2M	6M	5.7%	4M	3M
+1 (co-rotating)	M	M	42%	M	M

Table 2.1: Important radii for equatorial geodesics co- or counter-rotating around Kerr black holes.

For extremal spin and co-rotating orbits, $r_{\rm ISCO} = r_{\rm mb} = r_{\rm ph} = r_+ = M$. It appears all these orbits are coincident with the horizon. This appearance, however, is deceptive, and the radii are at very different physical locations. To see this, we compute the radial proper-separation between the horizon r_+ and a radius $r_0 > r_+$, by setting $dt = d\theta = d\phi = 0$ in Eq. (2.36):

$$ds^2 = \frac{\rho^2}{(r-r_-)(r-r_+)} dr^2, \tag{2.37}$$

$$\Rightarrow \quad \Delta s(r_+, r_0) = \int_{r_+}^{r_0} \frac{\rho}{\sqrt{(r-r_-)(r-r_+)}} dr \tag{2.38}$$

For a *non-extremal* black hole, $a < M$ and $r_- < r_+$. Therefore, the integrand in Eq. 2.38

diverges as $1/\sqrt{r-r_+}$, as $r \to r_+$. A $1/\sqrt{x}$-singularity integrates to a finite value. However, as $a \to M$, r_- approaches r_+, and the integrand diverges as $1/x$ and the integral Eq. (2.38) diverges for any value $r_0 > r_+$. It is this divergence that counteracts the fact that $r_{\rm ISCO}$, $r_{\rm mb}$ and $r_{\rm photon}$ all approach the same *coordinate* value.

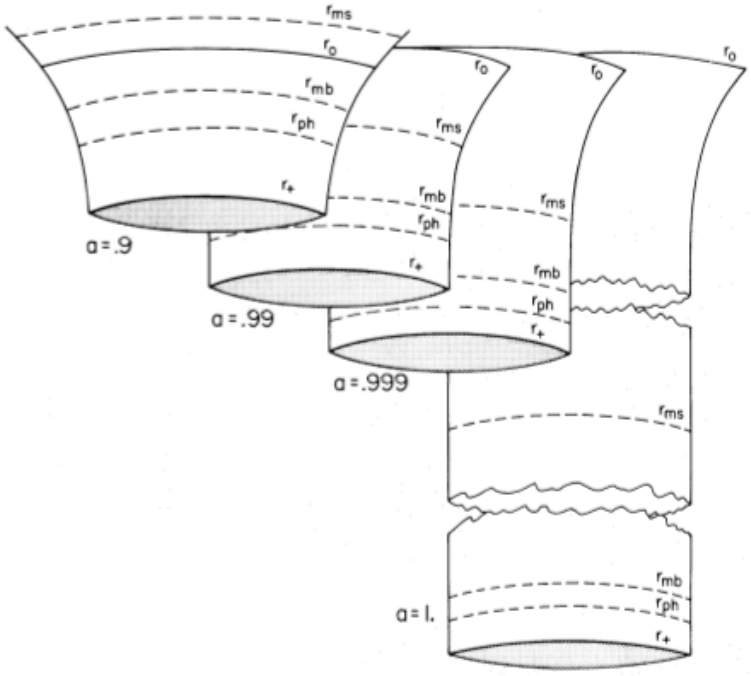


FIG. 2.—Embedding diagrams of the "plane" $\theta = \pi/2$, t = constant, for rotating black holes with near-maximum angular momentum. Here a denotes the hole's angular momentum in units of M. The Boyer-Lindquist radial coordinate r determines only the circumference of the "tube." When $a \to M$, the orbits at $r_{\rm ms}$, $r_{\rm mb}$, and $r_{\rm ph}$ all have the same circumference and coordinate radius, although—as the embedding diagram shows clearly—they are in fact distinct.

2.3.2 Non-equatorial geodesics

Counting the symmetries of Kerr, one finds three constants of motion along geodesics: the 4-velocity $g_{ab}u^a u^b = -1$, the energy E and z-component of angular momentum, i.e. the equivalent of Eq. (2.8) for Kerr. Equatorial orbits remain equatorial, so this suffices for integrability (i.e. given values of the constanst, and a point in the equatorial plane, in general the velocity $\dot{x}^a$ of the geodesic is completely determined).

Kerr also possesses a second rank Killing tensor, which results in a *fourth* constant of motion, the Carter constant. Therefore, non-equatorial orbits are also integrable. Such orbits have three characteristic frequencies, $\dot{\phi}$, ω_r and ω_θ.

Non-equatorial geodesics show a very rich variety of structure. Usually the frequencies are incommensurate, and the geodesic is volume-filling. Sometimes, the three frequencies are commensurate and the geodesic closes after a finite number of orbits ("resonance").

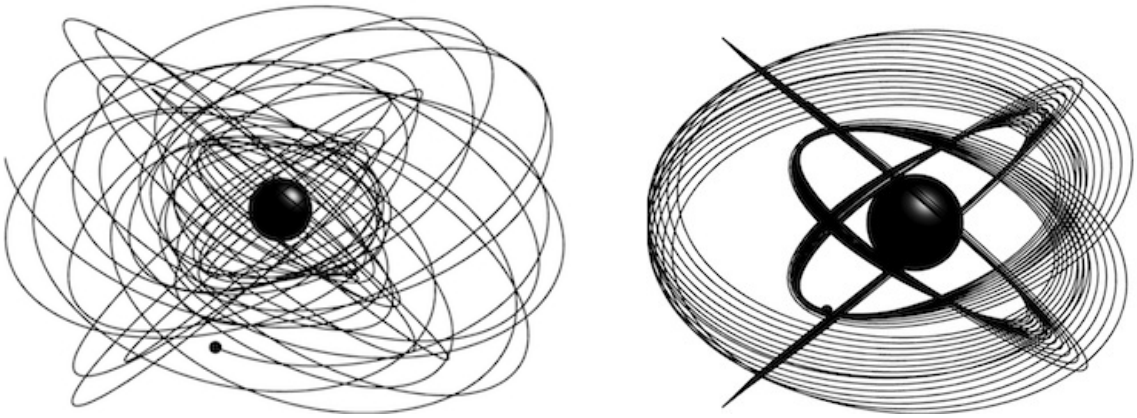


Figure 2.5: Illustration of the geodesic orbit of a small body around a spinning Kerr black hole. The left image shows the generic case, where the three characteristic frequencies $\dot{\phi}, \omega_r, \omega_\theta$ are not commensurable, and the trajectory is volume-filling. The right image shows a resonance, where the orbits repeat after a certain time. (Images are snapshots from a movie by Steve Drasco, available at `http://www.tapir.caltech.edu/~sdrasco/animations/`)

Chapter 3

Useful coordinate systems

3.1 Schwarzschild spacetime

3.1.1 Schwarzschild coordinates

The starting point are **Schwarzschild coordinates**

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2, \tag{3.1}$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$, and where we have used units such that $G = 1, c = 1$. Note that the term "Schwarzschild" is used with two meanings: "Schwarzschild spacetime" refers to the whole manifold which solves Einstein's equations. "Schwarzschild coordinates" are just one of many coordinate-systems one can use to describe this manifold.

The radial coordinate r is an **areal radius** coordinate. This means that the surface area of a $t = \text{const}, r = \text{const}$ sphere in Schwarzschild coordinates has surface area $4\pi r^2$, i.e. the same formula as in Euclidean space.

As $r \to \infty$, ds^2 approaches Euclidean metric. Schwarzschild is **asymptotically flat**.

3.1.2 Kerr-Schild coordinates

Schwarzschild coordinates have a coordinate singularity at $r = 2M$. Changing the time-coordinate via $t \to \bar{t} = t + 2M \ln(r/(2M) - 1)$ results in **Kerr-Schild coordinates**:

$$ds^2 = -d\bar{t}^2 + dr^2 + r^2 d\Omega^2 + \frac{2M}{r}(d\bar{t} + dr)^2 \tag{3.2}$$

These coordinates are very useful, because they're simple and extend easily to Kerr.

- Coordinate singularity at $r = 2M$ gone.
- Physical singularity at $r = 0$.

Let us compute the radial null-directions. Ignoring the angular part by setting $d\Omega \to 0$, Eq. (3.2) becomes:

$$0 = ds^2 = (dr + d\bar{t}) \left(\left(1 + \frac{2M}{r}\right) dr - \left(1 - \frac{2M}{r}\right) d\bar{t} \right) \tag{3.3}$$

Therefore, $dr = -d\bar{t}$ (incoming light-rays), or $dr = (1 - 2M/r)(1 + 2M/r)^{-1} d\bar{t}$ (outgoing lightrays). For $r < r_H \equiv 2M$, outgoing lightray goes to smaller radius: **Event Horizon**

3.1.3 Kruskal coordinates

Sometimes called "Kruskal-Szekres coordinates". These coordinates cover the maximal extension of the Schwarzschild spacetime, i.e. the complete manifold (actually quite remarkable that this can be done with a single coordinate system).

From Schwarzschild coordinates, first go to double null coordinates u, v. Then extend the range of these coordinates *through* the horizon by defining $u' = -e^{-u/(4M)}$, $v' = e^{-v/(4M)}$. Combined, this is

$$v' = \left(\frac{r}{2M} - 1\right)^{1/2} e^{(r+t)/4M} \tag{3.4}$$

$$u' = -\left(\frac{r}{2M} - 1\right)^{1/2} e^{(r-t)/4M} \tag{3.5}$$

u' v' are still null-coordinates (outgoing and incoming, respectively). Return to a spacelike and one time-like coordinate by defining

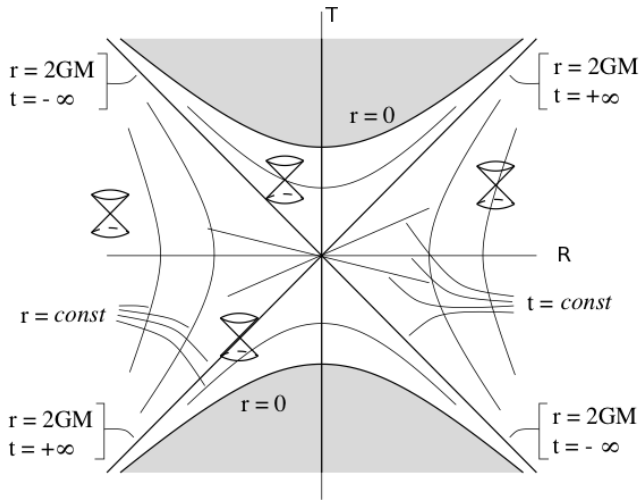
$$T = \frac{1}{2}(v' + u'), \qquad R = \frac{1}{2}(v' - u'). \tag{3.6}$$

When the dust settles, the metric becomes

$$ds^2 = \frac{32M^3}{r} e^{-r/(2M)}(-dT^2 + dR^2) + r^2 d\Omega^2, \tag{3.7}$$

where r is defined implicitly by the solution of

$$T^2 - R^2 = \left(1 - \frac{r}{2M}\right) e^{r/(2M)} \tag{3.8}$$



Kruskal Diagram, taken from p. 188 of gr-qc/9712019

- Lightcones at "45 degrees"

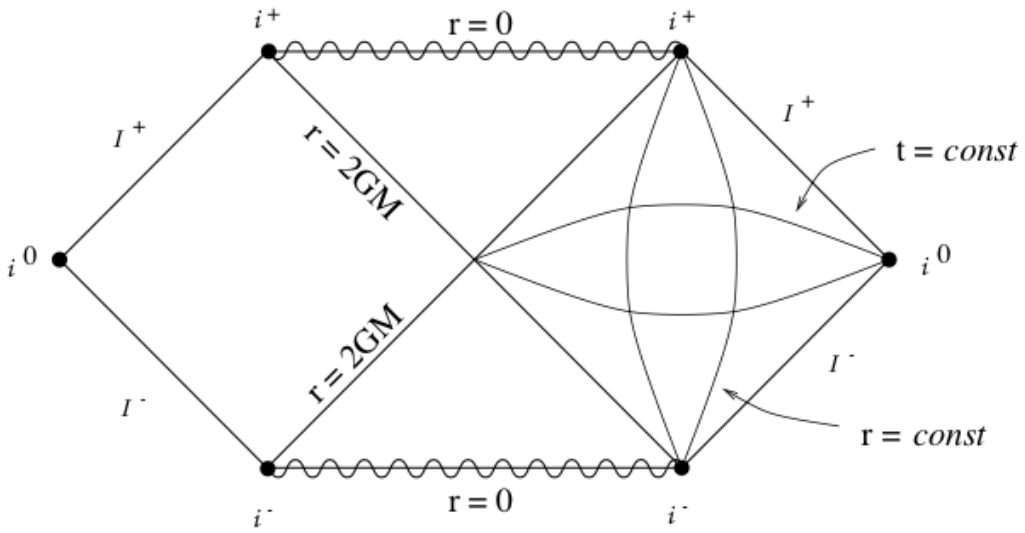
- t=const hypersurfaces are straight lines
- r=const hyperboloids
- Future singularity $r = 0$ at $T^2 - R^2 = 1$

3.1.4 Carter-Penrose diagram for Schwarzschild

Also called "Conformal Diagram", and short "Penrose diagram". We'd like to go one step further beyond the Kruskal diagram, by pulling in infinite distance to a finite coordinate value. Returning to the (u', v') coordinates, define

$$v'' = \arctan(v'), \qquad u'' = \arctan(u'), \tag{3.9}$$

so that $-\infty < u', v' < +\infty$ is mapped to $-\pi/2 < u', v' < \pi/2$.



Penrose diagram for Schwarzschild, taken from p. 205 of gr-qc/9712019.

- Lightcones at "45 degrees"
- *spacelike infinity* — i^0
- *future null infinity* — $\mathcal{I}^+$ "scri-plus"
- *future timelike infinity* — i^+ (separate from the singularity)
- Two asymptotically flat ends

The curves r =const are tangent to the time-translation symmetry.

3.2 Kerr spacetime

3.2.1 Kerr-Schild coordinates

$$g_{ab} = \eta_{ab} + 2H l_a l_b \tag{3.10}$$

H is a scalar function, l_a is null, $l^a = g^{ab} l_b$ and $l^a l_a = 0$. Substitute into Eq. (3.10),

$$0 = g_{ab} l^a l^b = \eta_{ab} l^a l^b + 2H l_a l^a l_b l^b = \eta_{ab} l^a l^b.$$

Hence, l^a is null with respect to *both* g_{ab} and η_{ab}. In Cartesian coordinates $(t, \vec{x})$, for spin vector $\vec{a}$ (3-D vector)

$$H = \frac{M r^3}{r^4 + (a \cdot \vec{x})}, \tag{3.11}$$

$$l_a = [1, \vec{l}\,], \quad \vec{l} = \frac{r\vec{x} - \vec{a} \times \vec{x} + (\vec{a} \cdot \vec{x})\vec{a}/r}{r^2 + a^2}. \tag{3.12}$$

Here, vector-operations are the usual 3-dimensional flat-space operations, which we have used for compactness. The quantity r is the Boyer-Lindquist radial-coordinate, which is related to $\vec{x}$ by

$$r^4 - \left(\vec{x}^2 - \vec{a}^2\right) r^2 - (\vec{a} \cdot \vec{x})^2 = 0 \tag{3.13}$$

The horizon $r = r_+$ maps to an ellipsoid in Kerr-Schild coordinates.

Boosts: Because the Kerr-Schild form is written in terms of tensorial objects, coordinate transformations can be computed by acting on the individual objects. For instance, a Lorentz-transformation

$$x^a \to x^a \Lambda_a{}^b \tag{3.14}$$

is represented by one constant matrix $\Lambda_a{}^b$. Under this transformation:

$$\eta_{ab} \to \eta_{ab} \tag{3.15}$$

$$l_a \to \Lambda_a{}^b l_b \tag{3.16}$$

and the form of the metric Eq.(3.10) is unchanged. By choosing $\Lambda_a{}^b$ to be a Lorentz boost, one obtains moving, spinning black holes; this is exploited, e.g. in numerical relativity. (Of course, all quantities have to be expressed as function of the transformed coordinates $x^a \Lambda_a{}^b$).

Chapter 4

Event Horizon

4.1 Intro & Examples

The event horizon is the boundary between regions of space-time that are in causal contact with "infinity", i.e. regions from which light can escape to large distance. For stationary (i.e. time-independent) space-times, one can find the event horizon by a local analysis just near the black hole at one instant in time. For time-varying space-times, this is no longer possible:

To decide whether null-rays from a certain point in space-time can reach infinity, we must know what happens in the future of the spacetime, because the null-rays must travel through those future regions of the space-time. Therefore, the event horizon is a *global property* of the entire spacetime.

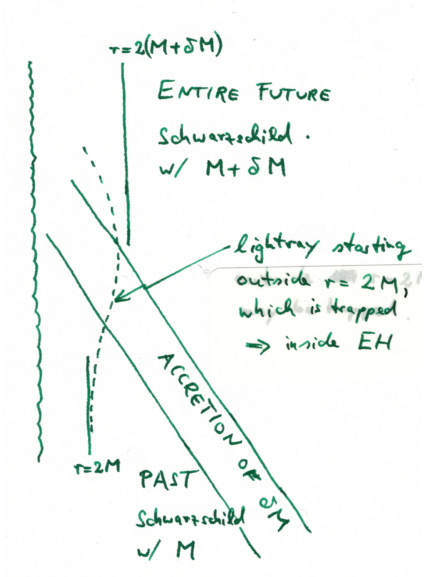


Figure 4.1: Illustration of geodesics for a spherically symmetric, accreting black hole. The dashed curve represents a radially outgoing null-ray which starts outside the $r = 2M$ surface of the original black hole, but is trapped by the enlarged final black hole.

Figure 4.1 illustrates that the event horizon can be at places different than one would navively expect: We start with a Schwarzschild black hole of mass M. If nothing were to happen in this space-time, the event horizon would be at $r = 2M$. However, in Fig. 4.1, we now have a spherical shell of matter accreting onto the black hole.

Because the configuration is spherically symmetric, it suffices to consider radially moving lightrays when determining the event horizon. One such lightrau Figure 4.1 plots one suchlightray that starts slightly outside the original $r = 2M$ surface and propagates outward. Due to the matter falling into the black hole at later time, this light-ray may be captured by the enlarged black hole. If so, this light ray was always inside the event horizon of this spacetime.

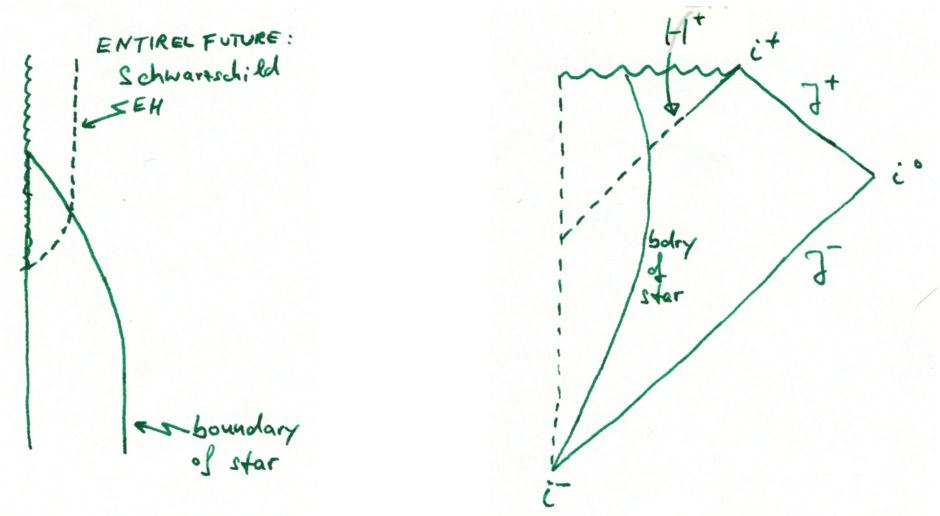


Figure 4.2: Event horizon in a spherical collapse of a star. Left: "traditional" Schwarzschild-like coordinates. Right: The same situation in a conformal diagram (Carter-Penrose diagram).

4.2 Formal Definitions

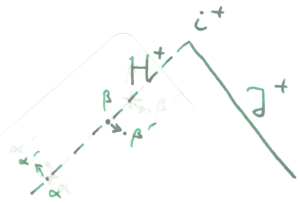
Let's define the event horizon more formally. This discussion follows Townsend, `https://arxiv.org/abs/gr-qc/9707012`, Chapter 2.6. We define

- **Causal past of an event** P, $J^-(P)$: All events in the past of P, which can be connected to P with time-like or null curves (any curve, need not be geodesics).
- **Causal past of a set** $\mathcal{S}$: All events that lie in $J^-(P)$ of at least one event of $\mathcal{S}$.
- $J^-(\mathcal{I}^+)$: All events from which one can reach future null infinity by a null- or time-like curve, i.e. all events from which one can escape to infinity.
- **Event Horizon:** Boundary of $J(\mathcal{I}^+)$.

4.3 Properties of the Future Event Horizon

1. **No two points in H^+ are timelike separated**.

   Assume points α and β are timelike separated with both α, β on the closure of $J^-(\mathcal{I}^+)$, and with β in the future of α. Then one could move β by an infinitesimal distance into *the interior* of $J^-(\mathcal{I}^+)$ and one could move α by an inifinitesimal distance to the *outside* of $\bar{J}^-(\mathcal{H}^+)$, while α' and β' still timelike separated. Now we have:
   $\alpha' \notin J^-(\mathcal{I}^+)$, but $\alpha' \in J^-(\beta') \in J^-(\mathcal{I}^+)$.
   This is a contradiction.



2. **H^+ is a null hypersurface...**

4.3.1 Interlude – Null surfaces

Before continuing with properties of event horizons, we need to explain some features of null hypersurfaces. A **null hypersurface** is a surface, which has a surface-normal which is null l is null: $l^2 = 0$. Such surfaces are surprising beasts.

To make our discussion more precise, consider a smooth function $S(x)$ of the space-time coordinates x, and consider surfaces S=const.

- A *one-form* l_a that is normal to S=const–surfaces is given by

$$l_a = \frac{\partial S}{\partial x^a}. \tag{4.1}$$

  Given a surface defined as a level-set (as here, S=const) the normal one-forms can be defined even if no metric is present.[1]

- Raising the index to obtain a normal-*vector* requires an (inverse) metric g^{ab}:

$$l^b = g^{ab}\frac{\partial S}{\partial x^a}. \tag{4.2}$$

- The surface S=const is *null*, if everywhere on the surface $l^a l_a = 0$. Let's call this surface $\mathcal{N}$.

- *Tangent vectors* t^a are vectors normal to l_a, that is $t \cdot l = t^a l_a = 0$. Note that l^a itself satisfies $l \cdot l = 0$. Therefore, l^a is **both the surface normal and a tangent vector** to $\mathcal{N}$!

[1] We could also introduce an arbitrary normalization constant $\tilde{f}(x)$, and write the normal-one form as $l_a = \tilde{f}(x)\partial_a S$. The normalization slightly complicates the following calculation, but does not change the result (cf. Townsend chapter 2.3.5).

- Because l^a is tangent to $\mathcal{N}$, there is a curve $x^a(\lambda)$ *within* $\mathcal{N}$ such that $l^a = dx^a/d\lambda$. We now show that this curve $x^a(\lambda)$ is a geodesics. We do so by confirming l^a satisfies the geodesic equation:

$$\begin{aligned} l^b\nabla_b l_a &= l^b\nabla_b\partial_a S \\ &= l^b\left(\partial_b\partial_a S - \Gamma^c_{ab}\partial_c S\right) \\ &= l^b\left(\partial_a\partial_b S - \Gamma^c_{ba}\partial_c S\right) \\ &= l^b\nabla_a\partial_b S \\ &= l^b\nabla_a l_b \\ &= \frac{1}{2}\nabla_a\left(l^b l_b\right). \end{aligned}$$

  Although $l^b l_b = 0$ on $\mathcal{N}$, the last term may be non-zero, because $l^b l_b$ is allowed to be non-zero away from $\mathcal{N}$. However, because $l^b l_b$ =const within $\mathcal{N}$, for any tangent vector t^a to $\mathcal{N}$, it must hold that

$$t^a\nabla_a\left(l^b l_b\right) = 0. \tag{4.3}$$

  The one-form $\nabla_a\left(l^b l_b\right)$ is therefore orthogonal to every tangent vector. Therefore, it must be parallel to the normal l_a, i.e.

$$\nabla_a\left(l^b l_b\right) \propto l_a. \tag{4.4}$$

  We therefore have that

$$l^b\nabla_b l_a \propto l_a, \tag{4.5}$$

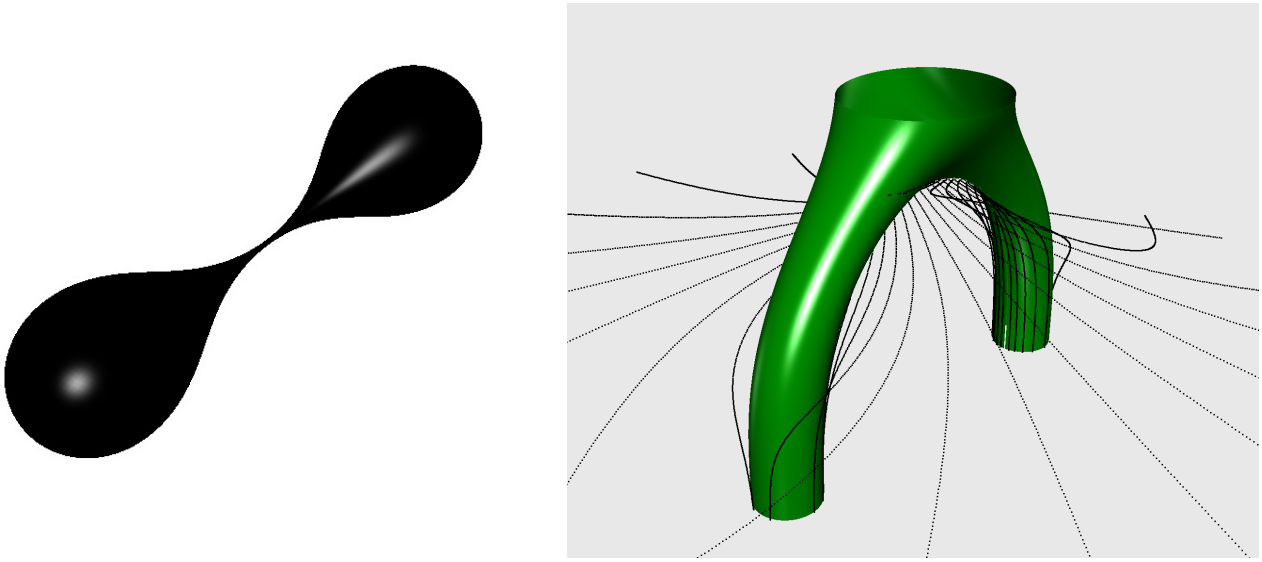
  i.e. $x^a(\lambda)$ is a geodesic, albeit in possibly non-affine parameterization.

- Therefore, **a null-surface is foliated by null-geodesics.** These null-geodesics are called *generators*.

4.3.2 Properties of Future Event Horizon cont'd

2. ... H^+ is a null hypersurface, **foliated by null-geodesics, called "generators".**
3. **Generators of H^+ may have past end-points.** I.e. if the geodesic is continued further into the past, it is no longer on H^+.
   Example: In the stellar collapse shown in Fig. 4.2, new geodesics enter the horizon at the point where it first forms.
4. **Generators of H^+ do not have future end-points.**
   This was shown by Penrose, who used non-timelike arguments like in EH property 1 above, to show that geodesics on H^+ can always be continued further on H^+. (see Townsend, Chapter 2.6)
5. $\Rightarrow$ **Geodesics can join the horizon H^+, but they cannot leave it.**

4.4 Example: Head-On collision of two black holes



From arXiv/0809.2628

4.5 Area theorem

The area theorem states that the area of an event horizon cannot decrease.

This theorem is central to the identification of BH-area as the analog of entropy in black hole thermodynamics. It also makes it impossible for black holes to bifurcate, and it limits the amount of energy that can be extracted from a rotating black hole.

Let us sketch the ingredients into a proof of the area theorem.

Null-geodesic congruences

A congruence is a set of curves, such that exactly one curve passes through each point. If all curves are null-geodesics, then this is a null-geodesic congruence. The congruence can be written as

$$x^a(\lambda; y^\alpha), \tag{4.6}$$

where λ is an affine parameter along each geodesic, and y^α parameterizes the different geodesics. The tangent to the geodesics is

$$k^a = \frac{\partial x^a}{\partial \lambda}. \tag{4.7}$$

For the geodesic congruence, one can define *expansion* θ, *shear* $\hat{\sigma}_{ab}$ and *twist* $\hat{\omega}_{ab}$ (see Caroll, Appendix F or Townsend Chapter 6.1). The expansion, in particular yields the change in the area-element a spanned by nearby geodesics:

$$\frac{da}{d\lambda} = \theta a. \tag{4.8}$$

The expansion is governed by the **Raychaudhuri equation** for null-congruences,

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \hat{\sigma}_{ab}\hat{\sigma}^{ab} + \hat{\omega}_{ab}\hat{\omega}^{ab} - R_{ab}k^a k^b. \tag{4.9}$$

Negative expansion diverges in finite affine length

We need a bound on the right-hand-side of Eq. (4.9), so let's consider the various terms:
(i) As a square, $\hat{\sigma}_{ab}\hat{\sigma}^{ab} \geq 0$.
(ii) By Frobenius theorem, $\hat{\omega}_{ab} = 0$ for surface-forming congruences (as for the horizon we are interested in).
(iii) Using Einstein's equations, the last term becomes

$$R_{ab}k^a k^b = 8\pi G\left(T_{ab} - \frac{1}{2}T g_{ab}\right)k^a k^b = 8\pi G T_{ab}k^a k^b \geq 0 \tag{4.10}$$

The last inequality requires that the matter satisfies the null-energy condition, $T_{ab}k^a k^b \geq 0$ or the slightly stronger weak-energy condition $T_{ab}v^a v^b \geq 0$ for all time-like vectors v^a.

With the inequalities (i) to (iii), Eq. (4.9) implies

$$\frac{d\theta}{d\lambda} \leq -\frac{1}{2}\theta^2. \tag{4.11}$$

Assume $\theta(\lambda)$ becomes negative somewhere along a geodesic, say, $\theta(0) = \theta_0 < 0$. It follows:

$$\frac{d\theta}{d\lambda} \leq -\frac{1}{2}\theta^2 \;\;\Rightarrow\;\; \frac{d}{d\lambda}\left(\theta^{-1}\right) > \frac{1}{2} \;\;\Rightarrow\;\; \theta^{-1} \geq \frac{1}{2}\lambda + \frac{1}{\theta_0} \;\;\Rightarrow\;\; \theta(\lambda) \leq \frac{\theta_0}{1+\lambda\theta_0/2}. \tag{4.12}$$

Therefore, θ will reach $-\infty$ within affine length $\lambda \leq |\theta_0|/2$.

Negative expansion implies caustics

Equation (4.8) implies

$$\frac{d\ln a}{d\lambda} = \theta. \tag{4.13}$$

Therefore, if $\theta \to -\infty$, then the area element $a \to 0$, i.e. nearby geodesics converge onto each other in *finite* affine length, a caustic.

Caustics forbidden on event horizons

Finally, Penrose has shown that the null-generators of event horizons must not have caustics. (if they had a caustic, one can argue via infinitesimal deformations that such geodesics cannot be on the event horizon).

The previous arguments imply that the expansion of null-generators of event horizons must satisfy

$$\theta \geq 0. \tag{4.14}$$

Therefore, each area-element of the event horizon satisfies

$$\frac{da}{d\lambda} \geq 0, \tag{4.15}$$

which implies the area-theorem,

$$\frac{dA_{\rm EH}}{d\lambda} \geq 0. \tag{4.16}$$

Chapter 5

post-Newtonian Approximation

My discussion follows Norbert Strautmann "General Relativity", Springer.

The goal of post-Newtonian approximation is to study the interaction of multiple isolated bodies under mutual gravitational interaction. The first results in this area were obtained by Einstein, Droste and de Sitter in the first years after GR was formulated. The research field is still very active today, with the major current effort spent on computing the 4-PN corrections for non-spinning bodies, and 3-PN corrections for spinning bodies.

Post-Newtonian theory is one of the corner stones for the complete solution of the general relativistic two-body problem. It covers the inspiral-phase of the two bodies (e.g. black holes or neutron stars). The last orbits before merger, and merger of the two objects, and the post-merger evolution must be treated with full numerical simulations of Einstein's equations.

5.1 Effacement Property

In Newtonian gravity, the motion of extended bodies can be approximated by the motion of their centers of mass. In this way, very complicated equations of the motion of extended bodies (e.g. fluids like Jupiter) can be simplified to a small set of ordinary differential equations.

In GR, the same property holds: the overall motion of extended bodies (this also includes black holes – after all, there are no point-masses in GR) turns out to be very insensitive to the internal structure of the bodies.

To motivate this property, consider bodies which would be spherical in isolation. Denote the characteristic size of each body by L, and their characteristic separation by R.

Tidal effects of the other body/bodies will induce an ellipticity of each body of order

$$\varepsilon \sim \frac{\text{tidal gravity}}{\text{self-gravity}} \sim \frac{GML/R^3}{GM/L^2} = \left(\frac{L}{R}\right)^3 . \tag{5.1}$$

This results in a quadrupole-moment of the body of order

$$Q \sim \varepsilon\, M\, L^2, \tag{5.2}$$

which results in an additional force onto other bodies of

$$\delta F \sim \frac{G\,Q\,M}{R^4}. \tag{5.3}$$

The ratio of this additional force to the leading Newtonian force is

$$\frac{\delta F}{F} \sim \frac{G\varepsilon M^2 L^2/R^4}{GM^2/R^2} = \varepsilon \left(\frac{L}{R}\right)^2 = \left(\frac{L}{R}\right)^5. \tag{5.4}$$

If the bodies are well-separated, $L \ll R$, this is indeed a very small perturbation.

5.2 post-Newtonian gravitational field

post-Newtonian theory combines two distinct perturbative expansions that have to be joined for the complete solution. Near the bodies, one expands in the characteristic velocity of the bodies, v/c, the post-Newtonian approximation per se. Far away in the wave-zone, one expands in powers of Newton's constant G, the so-called post-Minkowski expansion.

Let us sketch the near-field solution. We assume that bodies are widely separated (i.e. $R \gg R_S = 2GM/c^2$), and slowly moving (i.e. $v \ll c$). Let's call the small parameter

$$\epsilon \sim \left(\frac{R_S}{2R}\right)^{1/2} = \left(\frac{GM}{c^2 R}\right)^{1/2} \sim \frac{v}{c} \sim \left(\frac{p}{\rho}\right)^{1/2} \ll 1. \tag{5.5}$$

The middle equality, $GM/R \sim v^2$ arises by the virial theorem, which asserts under weak assumptions that the kinetic energy is comparable to the potential energy for a system of point masses in equilibrium.

The relative sizes of the terms can be easily derived for two point masses in circular orbits in Newtonian gravity. Let the masses be m_1 and m_2, and denote their separation by R. The masses are separated from the center of mass by the distances

$$r_1 = \frac{m_2}{m} r, \qquad r_2 = \frac{m_1}{m} r, \tag{5.6}$$

where $m = m_1 + m_2$ is the total mass.

The virial theorem reads

$$E_{\rm kin} = -\frac{1}{2} E_{\rm pot}. \tag{5.7}$$

Using the velocities of body 1 and 2 in the form $v_1 = \Omega r_1 = m_2 \Omega r/m$ and $v_2 = \Omega r_2 = m_1 \Omega r/m$, we find quickly Kepler's 3rd law:

$$\frac{1}{2} m_1 v_1^2 + \frac{1}{2} m_2 v_2^2 = \frac{G m_1 m_2}{2r}, \tag{5.8}$$

$$\Rightarrow \quad \left(m_1 m_2^2 + m_2 m_1^2\right) \frac{\Omega^2 r^2}{m^2} = \frac{G m_1 m_2}{r}, \tag{5.9}$$

$$\Rightarrow \quad \Omega^2 r^3 = Gm. \tag{5.10}$$

Dividing Eq. (5.10) by $c^2 r$, we find

$$\gamma \equiv \frac{Gm}{c^2 r} = \left(\frac{\Omega r}{c}\right)^2 = \left(\frac{v}{c}\right)^2, \tag{5.11}$$

where $v = \Omega r$ is the magnitude of the relative velocity of the two point-masses (because we assumed a circular orbit, here, v is constant). Multiplying Eq. (5.10) by Ω/c^3, we find

$$x \equiv \left(\frac{Gm\Omega}{c^3}\right)^{2/3} = \left(\frac{\Omega^3 r^3}{c^3}\right)^{2/3} = \left(\frac{v}{c}\right)^2. \tag{5.12}$$

γ and x defined in Eqs (5.11) and (5.12) define two common parameters in post-Newtonian expressions. They are particularly useful for the two-body problem in circular orbits, which we disucss below in Sec. 5.3. Until then, Eqs. (5.11) and (5.12) merely serve as a demonstration of the equalities indicated in Eq. (5.5).

5.2.1 Perturbative expansion

We will now make an Ansatz in powers of ϵ, and attempt to solve for consecutive orders in ϵ. In practice, we will stop at the first non-trivial order past Newtonian gravity, which turns out to be $O(\epsilon^2) = O(v^2/c^2)$. We assume the metric g_{ab} takes the form

$$g_{00} = -1 + {}^{(2)}g_{00} + {}^{(4)}g_{00} + \ldots, \tag{5.13a}$$

$$g_{0i} = \quad {}^{(3)}g_{0i} + \ldots, \tag{5.13b}$$

$$g_{ij} = \delta_{ij} + {}^{(2)}g_{ij} + \ldots. \tag{5.13c}$$

Here, the superscripts ${}^{(n)}$ denote terms of order ϵ^n, which are successively smaller with increased value of n. The even- and odd-ness of the different components of g_{ab} can be motivated by the behavior under time-reversal $t \to -t$, where only the space-time cross-terms of the metric (i.e. g_{0i}) change sign. However, ultimately, the ansatz is only justified after the fact when a consistent solution is found.

Having a perturbative metric of the form Eqs. (5.13), one next calculates various derived quantities also as a perturbative series. One has to be careful to keep all terms at the relevant orders, and to discard all higher-order terms. Precisely what order of terms is necessary isn't obvious ahead of time; again, the final result justifies which precise terms we keep.

As long as one is sufficiently careful, the calculations are straightforward, but tedious. We only give a few exemplary intermediate results.

- For the **inverse metric**:

$$g^{00} = -1 + {}^{(2)}g^{00} + {}^{(4)}g^{00}, \tag{5.14}$$
$$\ldots$$

  where

$${}^{(2)}g^{00} = -{}^{(2)}g_{00}, \tag{5.15}$$

$${}^{(4)}g^{00} = -{}^{(4)}g_{00} - \left({}^{(2)}g_{00}\right)^2. \tag{5.16}$$

  Note the general trend ere: The lowest perturbative terms are linear, whereas higher order terms may involve products of lower-order terms.

- For the **Christoffel-symbols**:

$$\Gamma^i_{00} = {}^{(2)}\Gamma^i_{00} + {}^{(4)}\Gamma^i_{00}, \tag{5.17}$$
$$\dots$$

where

$${}^{(2)}\Gamma^i_{00} = -\frac{1}{2}{}^{(2)}g_{00,i}, \tag{5.18}$$
$${}^{(4)}\Gamma^i_{00} = -\frac{1}{2}{}^{(4)}g_{00,i} + {}^{(3)}\, g_{0i,0} + \frac{1}{2}{}^{(2)}g_{ij}\, {}^{(2)}g_{00,j}. \tag{5.19}$$

To keep expressions reasonably compact, we are using the notation that a sub-script $,i$ denotes a partial derivative with respect to x^i. Note that because velocities are assumed to be much smaller than the speed of light, $v/c \ll 1$, a time-derivative increases the order by 1:

$$\frac{\partial}{\partial t} \sim \epsilon \frac{\partial}{\partial x^i} \sim \epsilon \frac{1}{R}. \tag{5.20}$$

Because a time-derivative increases the order of a term, indeed all terms in Eq. (5.16) are $O(\epsilon^4)$.

- For the **Ricci-tensor**:

$$R_{00} = {}^{(2)}\, R_{00} + {}^{(4)}\, R_{00} + \dots \tag{5.21}$$
$$\dots$$

Substituting in the expansion of the Christoffel symbols, and then the Christoffel symbols in terms of the metric, one finds

$${}^{(2)}R_{00} = {}^{(2)}\, \Gamma^i_{00,i} = -\frac{1}{2}{}^{(2)}g_{00,ii} = -\frac{1}{2}\Delta g_{00} \tag{5.22}$$
$${}^{(4)}R_{00} = {}^{(4)}\, \Gamma^i_{0,i} - {}^{(3)}\, \Gamma^i_{0i,0} + {}^{(2)}\, \Gamma^i_{00}\, {}^{(2)}\Gamma^j_{ij} - {}^{(2)}\, \Gamma^i_{00}\, {}^{(2)}\Gamma^0_{0i} = -\frac{1}{2}\Delta\, {}^{(4)}g_{00} + {}^{(3)}\, g_{0i,0i} + \dots \tag{5.23}$$

$\dots$

We have used the symbol Δ for the Laplacian. As can be seen from Eqs. (5.22) and (5.23), the Laplacian acting on the metric perturbation appears quite commonly.

5.2.2 post-Newtonian Gauge conditions

At this stage, we have quite lengthy expressions for the components of the Ricci tensor R_{ab} in terms of the perturbations of the metric g_{ab}. To simplify the expressions, we use the typical trick in GR, namely to choose a coordinate gauge, in which the expressions simplify.

A good gauge is given by:

$$g_{0j,j} - \frac{1}{2}g_{jj,0} = O(\epsilon^5), \tag{5.24}$$
$$g_{ij,j} - \frac{1}{2}\left(g_{jj} - g_{00}\right)_{,i} = O(\epsilon^4). \tag{5.25}$$

This essentially constrains some of the lower-order terms of the metric, namely:

$$ {}^{(3)}g_{0j,j} - \frac{1}{2}{}^{(2)}g_{jj,0} = 0, \tag{5.26a} $$

$$ {}^{(2)}g_{ij,j} - \frac{1}{2}{}^{(2)}g_{jj,i} + \frac{1}{2}{}^{(2)}g_{00,i} = 0. \tag{5.26b} $$

Using these gauge-conditions, one can simplify the expressions for the Ricci-tensor to

$$ {}^{(2)}R_{00} = -\frac{1}{2}\Delta^{(2)}g_{00}, \tag{5.27a} $$

$$ {}^{(4)}R_{00} = -\frac{1}{2}\Delta^{(4)}g_{00} + [\text{terms quadratic in } {}^{(2)}g_{00} \text{ and } {}^{(2)}g_{ij}] \tag{5.27b} $$

$$ {}^{(3)}R_{0i} = -\frac{1}{2}\Delta^{(3)}g_{0i} - \frac{1}{4}{}^{(2)}g_{jj,0i} + \frac{1}{2}{}^{(2)}g_{ij,0j}, \tag{5.27c} $$

$$ {}^{(2)}R_{ij} = -\frac{1}{2}\Delta^{(2)}g_{ij}. \tag{5.27d} $$

5.2.3 Einstein's equations

We'll write Einstein's equations in the form

$$ R_{ab} = 8\pi G\left(T_{ab} - \frac{1}{2}g_{ab}T^{c}{}_{c}\right). \tag{5.28} $$

From the stress-energy tensor, we'll need:

$$ T^{00} = {}^{(0)}T^{00} + {}^{(2)}T^{00} + \ldots, \tag{5.29a} $$

$$ T^{0i} = {}^{(1)}T^{0i} + \ldots, \tag{5.29b} $$

$$ T^{ij} = {}^{(2)}T^{ij} + \ldots. \tag{5.29c} $$

${}^{(0)}T^{00}$ is the rest-mass, ${}^{(2)}T^{00}$ includes, for example, the kinetic energy $\frac{1}{2}\rho v^2$.

5.2.4 Solving Einstein's equations

Substituting Eqs. (5.27) into Eq. (5.28), we arrive at Laplace equations for that are sourced by the energy-monentum tensor. Substituting in Eqs. (5.29), it turns out that the 00 and ij Einstein-equations reduce to essentially the same equation:

$$ \Delta^{(2)}g_{00} = -8\pi G\,{}^{(0)}T^{00}, \tag{5.30} $$

$$ \Delta^{(2)}g_{ij} = -8\pi G\,{}^{(0)}T^{00}\,\delta_{ij}. \tag{5.31} $$

Both equations are solved by the Newtonian potential ϕ:

$$ \Delta\phi = 4\pi G\,{}^{(0)}T^{00}, \tag{5.32a} $$

$$ \phi = -G\int \frac{{}^{(0)}T^{00}(t,\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}', \tag{5.32b} $$

$$ {}^{(2)}g_{00} = -2\phi, \tag{5.32c} $$

$$ {}^{(2)}g_{ij} = -2\phi\,\delta_{ij}. \tag{5.32d} $$

In fact, this is exactly the Newtonian limit of the Einstein equation, and indeed, Eqs. (5.27a) and (5.27d) are the lowest-order at which non-trivial terms arise.

We can now substitute Eqs. (5.32) back into Eqs. (5.27b) and (5.27c). Then all terms are determined except of the leading $\Delta^{(4)} g_{00}$ and $\Delta^{(3)} g_{0i}$. Still having Laplace-equations, we expect the solution to be given in terms of additional potentials ψ and $\vec{\xi}$. These two potentials appear at higher order in ϵ and are the first post-Newtonian potentials.

Let's do ${}^{(4)}R_{00}$ first. Substitute ϕ into Eq. (5.27b), and equate to the appropriate $O(\epsilon^4)$ terms of the 00-component of Eq. (5.28). After some algebra, one finds

$$\Delta \left({}^{(4)}g_{00} + 2\phi^2 \right) = -8\pi G \left({}^{(2)}T^{00} + {}^{(2)}T^{ii} \right). \tag{5.33}$$

This is solved by

$$\Delta\psi = 4\pi G \left({}^{(2)}T^{00} + {}^{(2)}T^{ii} \right), \tag{5.34a}$$

$$\psi = -G \int \frac{{}^{(0)}T^{00}(t,\vec{x}') + {}^{(2)}T^{ii}(t,\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x}', \tag{5.34b}$$

$${}^{(4)}g_{00} = -2\phi^2 - 2\psi. \tag{5.34c}$$

The $0i$-components of Einsteins equations –cf. Eq. (5.27c)– are somewhat more involved. It turns out, they require two extra potentials. Perhaps as expected, the first potential is a vector-potential, written $\vec{\xi}$ or ξ^i:

$$\Delta\xi^i = 16\pi G \, {}^{(3)}T^{0i}, \tag{5.35a}$$

$$\xi^i = -4G \int \frac{{}^{(3)}T^{0i}(t,\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x}'. \tag{5.35b}$$

(Here, as in all other equations, we insert pre-factors into the definitions for later convenience). However, one also needs one further scalar potential χ, to absorb extra terms with derivatives of ϕ:

$$\Delta\chi = \phi, \tag{5.36a}$$

$$\chi = -\frac{G}{2} \int \left|\vec{x}-\vec{x}'\right| \, {}^{(2)}T^{00}(t,\vec{x}') d\vec{x}'. \tag{5.36b}$$

In terms of the potentials ξ^i and χ, we have

$${}^{(3)}g_{0i} = \xi_i + \chi_{0i}. \tag{5.37}$$

At this point, we have the complete post-Newtonian metric in terms of the potentials ϕ, ψ, ξ_i, χ. We can now also express the Christoffel-symbols in terms of these potentials.

5.2.5 Conservation of T^{ab}

If we rewrite the gauge condition (5.26a) in terms of potentials, one finds

$$\xi_{i,i} + 4\phi_{,0} = 0. \tag{5.38}$$

Taking the Laplacian,

$$4\frac{\partial}{\partial t}\Delta\phi + \frac{\partial}{\partial x^i}\left(\Delta\xi_i\right) = 0, \tag{5.39}$$

and substituting Eqs. (5.32) and (5.35), we get

$$\frac{\partial^{(0)}T^{00}}{\partial t} + \frac{\partial^{(1)}T^{0i}}{\partial x^i} = 0. \tag{5.40}$$

This is the leading-order conservation of the energy-momentum.

Therefore, we see that the post-Newtonian potential $\vec{\xi}$ implies the leading order energy-momentum conservation. The next-higher order conservation of T^{ab} does *not* follow, but must be separately imposed. Specifically, the $O(\epsilon^2)$ terms of $\nabla_b T^{ab} = 0$ yield

$$\frac{\partial^{(1)}T^{i0}}{\partial t} + \frac{\partial^{(2)}T^{ij}}{\partial x^j} = -\frac{\partial\phi}{\partial x^i}{}^{(0)}T^{00}. \tag{5.41}$$

This is the force-law of Newtonian gravity.

5.2.6 Interlude: Equations of motion of test particle

Our results so far allow to write down two useful intermediate results. The first one is the motion of test particles in the post-Newtonian metric.x

Having the solution for the metric, Eqs. (5.32), (5.34), (5.35) and (5.36), we can also write out the equation of motion for a test-particle, i.e. a geodesic.

One could either just write down the Geodesic equation itself. Or, one can start from the extremal principle for geodesics

$$\delta\int\left(\frac{d\tau}{dt}\right)dt = 0. \tag{5.42}$$

Writing $v^a = dx^a/dt = [1, dx^i/dt] = [1, \vec{v}]$, we have

$$\left(\frac{d\tau}{dt}\right)^2 = -g_{ab}v^a v^b. \tag{5.43}$$

There is a great freedom in writing down a Lagrangian, as it merely needs to be a monotoninc function in $d\tau/dt$, in order for the extremality condition (5.42) to yield geodesics.

One natural choice is $\mathcal{L} = 1 - d\tau/dt$, because it reduces to the Newtonian Lagrangian in the Newtonian limit. Substituting in $dx^a/dt = [1, \vec{v}]$, multiplying out terms, and truncating at the order ϵ^4 yields

$$\mathcal{L} = 1 - \frac{d\tau}{dt} = \frac{1}{2}v^2 - \phi + \frac{1}{8}v^4 - \frac{1}{2}\phi^2 - \psi - \frac{3}{2}\phi v^2 + v^i\left(\xi_i + \chi_{0i}\right). \tag{5.44}$$

The Euler-Lagrange equations read

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial v^i}\right) - \frac{\partial\mathcal{L}}{\partial x^i} = 0. \tag{5.45}$$

Substitute in Eq.(5.44), and discard higher-order terms. Furthermore, one higher-order term contains $\dot{v}^i$ itself. In this term, substitute in the Newtonian acceleration. Eventually, one finds

$$\frac{d\vec{v}}{dt} = -\nabla\left(\phi + 2\phi^2 + \psi\right) - \frac{d\vec{\xi}}{dt} - \frac{\partial^2}{\partial t^2}\nabla\chi + \vec{v}\times(\nabla\times\vec{\xi}) + 3\vec{v}\frac{\partial\phi}{\partial t} + 4\vec{v}(\vec{v}\cdot\nabla\phi) - v^2\nabla\phi. \quad (5.46)$$

The very first term on the right-hand side represents Newtonian gravity. All other terms are the 1-PN corrections.

5.2.7 Interlude: Asymptotic form of stationary fields

The second intermediate result concerns the asymptotic form of the metric for stationary matter distributions.

Consider the integrals for the potential, for instance Eqs. (5.32):

$$\phi = -G\int \frac{{}^{(0)}T^{00}(t,\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}'. \quad (5.47)$$

If hte matter is limited to a region in space, say $|\vec{x}'| < L$, and if one is interested in the value of the potential far away, i.e. $|\vec{x}| = e \gg L$, then one can make a multi-polar expansion,

$$\frac{1}{|\vec{x}-\vec{x}'|} = \frac{1}{r} + \frac{\vec{x}\cdot\vec{x}'}{r^3} + \ldots. \quad (5.48)$$

Substituting this form for $1/|x-x'|$ back into Eq. (5.32), one finds

$$\phi = -\frac{G^{(0)}M}{r} - \frac{G\,\vec{x}\cdot{}^{(0)}\vec{D}}{r^3} + O(1/r^3) \quad (5.49)$$

The impact of the matter onto the far-field is reduced to a monopole and a dipole term,

$${}^{(0)}M = \int {}^{(0)}T^{00}\; d^3\vec{x}, \quad (5.50)$$

$${}^{(0)}D^i = \int x^i\,{}^{(0)}T^{00}\; d^3\vec{x}. \quad (5.51)$$

The same game can be played for the other potentials. One can simplify in the time-independent case somewhat, and substitute into the formulae that give the metric in term of the potentials. One can further shift the origin of the coordinate system to eliminate the dipole term. The end result is:

$$g_{00} = -1 + \frac{2GM}{r} - 2\frac{G^2M^2}{r^2} + O(1/r^3), \quad (5.52)$$

$$g_{0i} = -2G\epsilon_{ijk}\frac{J^j x^k}{r^3} + O(1/r^3), \quad (5.53)$$

$$g_{ij} = \left(1 + \frac{2GM}{r}\right)\delta_{ij} + O(1/r^2), \quad (5.54)$$

where M and J^i are the total mass and angular momentum of the matter distribution:

$$M = \int \left({}^{(0)}T^{00} + {}^{(2)}T^{00} + {}^{(2)}T^{ii} \right) \, d^3\vec{x}, \tag{5.55}$$

$$J^i = \int \epsilon_{ijk} x^{j\,(1)}T^{0k} \, d^3\vec{x}. \tag{5.56}$$

Therefore, the far-field metric for a stationary matter distribution depends only on mass and angular momentum of the matter.

5.2.8 Point-particle limit: Einstein-Infeld-Hoffman Equations

Let us now return to our main objective, to figure out how multiple bodies move under their mutual gravitational interaction when we include *post-Newtonian* corrections.

The the first post-Newtonian order, in which we work here, one can show that the internal structure of the individual bodies does not matter (e.g. Damour, The problem of motion in Newtonian and Einsteinian gravity. In 300 Yeaes of Gravitation, Cambridge University Press).

We can replace the individual bodies by **point-particles**, i.e. the matter distribution of each body is replaced by a Dirac delta-function. The stress-energy tensor of several point-particles labeled by $A = 1, 2, \ldots$ then takes the form

$$T^{ab}(\vec{x}, t) = \frac{1}{\sqrt{-g}} \sum_A m_A \int \frac{dx^a_A}{d\tau_A} \frac{dx^b_A}{d\tau_A} \, \delta^4 \left(x - x_A(\tau_A) \right) d\tau_A. \tag{5.57}$$

The integral of the A-th particle is along its world-line $x^a_A(\tau_A)$. At each instant of time, t, the integral only has support at the position of the particle at that time, $x^i_A(t)$. The integration over proper time collapses the delta-function in time by the factor factor

$$\int \delta(t - t_A(\tau_A)) d\tau_A = \left(\frac{dt}{d\tau_A} \right)^{-1}. \tag{5.58}$$

Replacing the $d/d\tau_A$-derivatives by d/dt derivatives adds two further factors of $dt/d\tau_A$, so that overall, we have

$$T^{ab}(\vec{x}, t) = \frac{1}{\sqrt{-g}} \sum_A m_A \frac{dx^a_A}{dt} \frac{dx^b_A}{dt} \left(\frac{d\tau_A}{dt} \right)^{-1} \delta^3 \left(\vec{x} - \vec{x}_A(t) \right). \tag{5.59}$$

We can now substitute in the expansion of the metric, Eqs. (5.13), into the term $\sqrt{-g}$, and substitute in further the solution in terms of the PN-potentials. This yields

$$-g = 1 - 4\phi + O(\epsilon^4), \qquad \frac{1}{\sqrt{-g}} = 1 + 2\phi + O(\epsilon^4). \tag{5.60}$$

Furthermore, $d\tau_A/dt$ can be read off from Eq. (5.44),

$$\frac{d\tau_A}{dt} = 1 - \mathcal{L} = 1 - \frac{1}{2} v_A^2 + \phi + O(\epsilon^4). \tag{5.61}$$

Overall, the stress-energy tensor becomes

$$ {}^{(0)}T^{00} = \sum_A m_A \delta^{(3)}(\vec{x} - \vec{x}_A), \tag{5.62a} $$

$$ {}^{(2)}T^{00} = \sum_A m_A \left(\phi + \frac{1}{2} v_A^2 \right) \delta^{(3)}(\vec{x} - \vec{x}_A), \tag{5.62b} $$

$$ {}^{(1)}T^{0i} = \sum_A m_A v_A^i \delta^{(3)}(\vec{x} - \vec{x}_A), \tag{5.62c} $$

$$ {}^{(2)}T^{ij} = \sum_A m_A v_A^i v_A^j \delta^{(3)}(\vec{x} - \vec{x}_A). \tag{5.62d} $$

Note that ${}^{(2)}T^{00}$ depends on the Newtonian potential ϕ. That's ok. We can solve for ϕ from the lower-order terms first, and then substitute into ${}^{(2)}T^{00}$.

Ok, let's solve for the potentials! From Eq. (5.32),

$$ \phi = -G \sum_A \frac{m_A}{|\vec{x} - \vec{x}_A|}. \tag{5.63} $$

From Eq. (5.34),

$$ \nabla\psi = 4\pi G \left({}^{(2)}T^{00} + {}^{(2)}T^{ii} \right) \tag{5.64} $$

$$ = 4\pi G \sum_A m_A \left(\phi'_A + \frac{3}{2} v_a^2 \right) \delta^{(3)}(\vec{x} - \vec{x}A). \tag{5.65} $$

Here, the first difficulty occurs. The potential ϕ in Eq. (5.65) will have to be evaluated at the particle locations $\vec{x}_A$. However, the term $m_A/|\vec{x} - \vec{x}_A|$ is singular there. By careful renormalization of a finite-size body, one can show that one can simply leave out the offending term. Defining

$$ \phi'_A := -G \sum_{B \neq A} \frac{m_B}{|\vec{x} - \vec{x}_B|}, \tag{5.66} $$

we therefore find

$$ \psi = -G \sum_A \frac{m_A \phi'_A}{|\vec{x} - \vec{x}'|} - \frac{3}{2} G \sum_A \frac{m_A v_A^2}{|\vec{x} - \vec{x}_A|}. \tag{5.67} $$

Finally, from (5.35) and (5.36):

$$ \xi^i = -4G \sum_A \frac{m_A v_A^i}{|\vec{x} - \vec{x}'|}, \tag{5.68} $$

$$ \chi = -\frac{G}{2} \sum_A m_A |\vec{x} - \vec{x}'_A|. \tag{5.69} $$

And evaluating the partial derivatives in $\chi_{,0i}$, we finally find the space-time-components of the metric:

$$ g_{0i} = \xi_i + \chi_{,0i} = -4G \sum_A \frac{m_A}{|\vec{x} - \vec{x}_A|} \left(7 v_A^i + (\vec{v}_A \cdot \hat{n}_A) \hat{n}_A \right). \tag{5.70} $$

Here, $\hat{n}_A = (\vec{x} - \vec{x}_A)/|\vec{x} - \vec{x}_A|$.

At this point, we have computed the metric that arises from a collection of point particles. We can now begin to use it. Let us begin to write down the geodesic equation for particle A, i.e. the equation of motion m_A would follow in the limit $m_A \to 0$. To do so, we need to evaluate Eq. (5.44) at the position $\vec{x}_A$. While doing so, one can again show that singular terms can be neglected, i.e. we will remove the A-term from sums whenever it is singular. One finds

$$\begin{aligned}\mathcal{L}_A = &\frac{1}{2}v_A^2 + \frac{1}{8}v_A^4 + G\sum_{B\neq A}\frac{m_B}{r_{AB}} - \frac{G^2}{2}\sum_{B\neq A}\sum_{C\neq A}\frac{m_B m_C}{r_{AB}r_{AC}} \\ &- G^2\sum_{B\neq A}\sum_{C\neq B}\frac{m_B m_C}{r_{AB}r_{BC}} + \frac{3G}{2}v_a^2\sum_{B\neq A}\frac{m_B}{r_{AB}} + \frac{3G}{2}\sum_{B\neq A}\frac{m_B v_B^2}{r_{AB}} \\ &- \frac{G}{2}\sum_{B\neq A}\frac{m_B}{r_{AB}}\left(7\vec{v}_A\cdot\vec{v}_B + (\vec{v}_A\cdot\hat{n}_{AB})(\vec{v}_B\cdot\hat{n}_{AB})\right). \end{aligned} \tag{5.71}$$

Here we have introduced the short-hand $r_{AB} = |\vec{x}_A - \vec{x}_B|$ and $\hat{n}_{AB} = (\vec{x}_A - \vec{x}_B)/r_{AB}$.

We are almost there!

Given the test-particle Lagrangian Eq. (5.71), we can now guess how the total Lagrangian for an N-body system must look like. It must be symmetric under exchange of any two particles. And the portion for the A-th particle must reduce to Eq. (5.71) in the limit $m_A \to 0$.

The answer is

$$\begin{aligned}\mathcal{L} = &\frac{1}{2}\sum_A m_A v_A^2 + \frac{1}{8}\sum_A m_A v_A^4 + \frac{G}{2}\sum_A\sum_{B\neq A}\frac{m_A m_B}{r_{AB}} + \frac{3G}{2}\sum_A\sum_{B\neq A} m_A v_A^2\frac{m_B}{r_{AB}} \\ &- \frac{G}{4}\sum_A\sum_{B\neq A}\frac{m_A m_B}{r_{AB}} - (7\vec{v}_A\cdot\vec{v}_B + (\vec{v}_A\cdot\hat{n}_{AB})(\vec{v}_B\cdot\hat{n}_{AB})) - \frac{G^2}{2}\sum_A\sum_{B\neq A}\sum_{C\neq A}\frac{m_A m_B m_C}{r_{AB}r_{AC}} \end{aligned} \tag{5.72}$$

Finally, evaluating the Euler-Lagrange equations, we find the **Einstein-Infeld-Hofmann equations**, which determine the motion of N point particles in the 1-st post-Newtonian approximation:

$$\begin{aligned}\frac{d\vec{v}_A}{dt} = &- G\sum_{B\neq A}\frac{\vec{x}_{AB}}{r_{AB}}\Bigg[1 - 4G\sum_{C\neq A}\frac{m_C}{r_{AC}} + G\sum_{C\neq B} m_C\left(-\frac{1}{r_B C} + \frac{\vec{x}_{AB}\cdot\vec{x}_{BC}}{2r_{BC}^3}\right) \\ &\qquad + v_A^2 - 4\vec{v}_A\cdot\vec{v}_B + 2v_B^2 - \frac{3}{2}\left(\frac{\vec{v}_B\cdot\vec{x}_{AB}}{r_{AB}}\right)^2\Bigg] \\ &- \frac{7G^2}{2}\sum_{B\neq A}\sum_{C\neq B}\frac{m_B}{r_{AB}}\frac{m_C\vec{x}_{BC}}{r_{BC}^3} + G\sum_{B\neq A} m_B\frac{\vec{x}_{AB}}{r_{AB}^3}(4\vec{v}_A - 3\vec{V}_B)\cdot(\vec{v}_A - \vec{v}_B) \end{aligned} \tag{5.73}$$

5.3 The two-body problem

Let us now specialize to two bodies, with our aim being to describe systems of binary compact objects that lead to gravitational wave emission.

Going to the center of mass frame, Eqs. (5.73) reduce to

$$\frac{d\vec{v}}{dt} = -\frac{Gm}{r^2}\hat{n}\left(1-(4+2\nu)\frac{Gm}{r}+(1+3\nu)v^2-\frac{3\nu}{2}\left(\vec{v}\cdot\hat{n}\right)^2\right)+\frac{Gm}{r^2}\left(4-2\nu\right)\left(\vec{v}\cdot\hat{n}\right)\vec{v}. \tag{5.74}$$

Here,

$$\vec{v} = \vec{v}_1 - \vec{v}_2 \qquad \text{relative velocity} \tag{5.75}$$

$$r = |\vec{x}_1 - \vec{x}_2| \qquad \text{separation} \tag{5.76}$$

$$\hat{n} = (\vec{x}_1 - \vec{x}_2)/r \qquad \text{unit-vector pointing from } m_2 \text{ to } m_1 \tag{5.77}$$

$$m = m_1 + m_2 \qquad \text{total mass} \tag{5.78}$$

$$\mu = \frac{m_1 m_2}{m} \qquad \text{reduced mass} \tag{5.79}$$

$$\nu = \frac{\mu}{m} = \frac{m_1 m_2}{m^2} \qquad \text{symmetric mass-ratio} \tag{5.80}$$

Equation (5.74) must be invariant under the exchange of the particles $1 \leftrightarrow 2$. Therefore, it is natural that the mass-ratio of the two bodies enters Eq. (5.74) only through the *symmetric mass-ratio* ν. For a binary of highly unequal masses $m_1 \ll m_2$ or $m_1 \gg m_1$, the symmetric mass-ratio reduces to the usual mass-ratio, independent of which of the two bodies is the more massive one:

$$\nu \approx \frac{m_{\rm small}}{m_{\rm big}} \tag{5.81}$$

For equal-mass systems, $m_1 = m_2 = m/2$, the symmetric mass-ratio takes the value $1/4$.

Equation (5.74) forms the basis of general relativistic motion in a wide variety of systems. The corresponding metric g_{ab} as given in Eqs. (5.32), (5.34), (5.35) and (5.36) can be used to describe propagation effects. For instance:

- Relativistic binary pulsars (e.g. the Hulse-Taylor pulsar)
- Solar system tests of general relativity
- Gyroscope precession
- Relativistic time-delays (most notably the global positioning system GPS)

All these systems share a very weak gravitational field. For instance, for the Earth's motion around the sun: $\epsilon = v/c \sim 10^{-4}$. For the Hulse-Taylor pulsar: $\epsilon \sim 10^{-3}$. Note from Eq. (5.74) that corrections enter at the *square* $\left(\frac{v}{c}\right)^2 \sim \epsilon^2$. The next-higher corrections enter at $\left(\frac{v}{c}\right)^4 \sim \epsilon^4$. This implies that Newtonian gravity is accurate in the solar system to 10^{-8}, and 1-PN theory to about 10^{-16}.

Weak-field applications of GR will be discussed in the second half of the course by Prof. Thompson. For now, we want to chart our way to the sources of gravitational waves for LIGO.

Nomenclature:
(1) *post-Newtonian order* counts the number of squares of v/c. Terms $O(\epsilon^2) \sim v^2/c^2$ are called "1-PN". Terms $O(\epsilon^5)$ are called "2.5PN".
(2) post-Newtonian order is generally counted relative to the leading (Newtonian) term. The leading order term in Eq. (5.74) is $GM/r^2 \sim \epsilon^2$. This sets the reference, and so the corrections in Eq. (5.74) are called "1-PN" terms.

5.3.1 Binary black holes – inadequacy of low-order PN

The 1-PN equations (5.74) have been quite a large effort to derive. Unfortunately, they are nowhere near accurate enough for LIGO purposes. We'll give two reasons why much more accurate results are needed.

In the 2nd homework set, you computed the gravitational wave energy flux F of two point-masses in a circular orbit:

$$F_{\rm GW} = \frac{32 G\mu^2 r^4 \Omega^6}{5c^5} \tag{5.82}$$

If energy-conservation holds in post-Newtonian theory, then this energy-loss must be derivable from the equation of motion Eq. (5.74), through a term of the form "force times velocity":

$$F_{\rm GW} \sim m \frac{d\vec{v}}{dt} \cdot \vec{v}. \tag{5.83}$$

Substituting $v = \Omega r$ into the formula for $F_{\rm GW}$, we find

$$\frac{32 G\mu^2 r^4 \Omega^6}{c^5} = \frac{32 G\mu^2 v^6}{c^5 r^2} \sim vm\frac{Gm}{r^2}\left(\ldots + \left(\frac{v}{c}\right)^n + \ldots\right) \tag{5.84}$$

On the right-hand-side, we have schematically highlighted the post-Newtonian term of PN-order $n/2$. To have the powers of velocity v balance in Eq. (5.84), the lowest-order PN-term that can contribute to the GW energy loss must have $n = 5$, i.e. 2.5PN-orders beyond Newtonian gravity, or an extra factor of $(v/c)^3$ smaller than the 1-PN corrections we have so laboriously computed in Eq. (5.74)!

Hence, if we were to derive the *inspiral* of two point-masses under their gravitational wave emission, we would have to go to 2.5PN order of the Newtonian-like equations of motion to find *any inspiral at all*!

Moreover, as we show below, the leading order GW-driven inspiral is woefully inadequate for LIGO needs. One needs to go at least 3-PN orders beyond leading order radiation reaction. That means, if we were to compute gravitational waveforms for LIGO based on Newtonian-like ODE's for the motion of the black holes (like Eq. (5.74)), one would need to have the ODE's at least to PN-order $2.5 + 3 = 5.5$, i.e. $(v/c)^{11}$. This is impractical.

5.3.2 Energy Balance

Problem set 2 outlines the important trick how one can achieve the inspiral motion of two orbiting masses, without having to go to excessive post-Newtonian order.

The trick is called energy balance: The energy removed from the binary through gravitational waves must be supplied by the total energy of the binary.

For a Newtonian circular binary, the total energy is

$$E = E_{\rm kin} + E_{\rm pot} = \frac{1}{2}E_{\rm pot} = -\frac{Gm_1m_2}{2r}. \tag{5.85}$$

This can be easily shown from our little calculation of Newtonian circular orbits, Eqs. (5.10)–(5.12), and we have used in the middle equality the Virial theorem, once again.

Now, assuming that the GW-energy loss is equal to the change in the energy of the binary, we have

$$F_{\rm GW} = -\frac{dE}{dt} = -\frac{dE}{dr}\frac{dr}{dt}. \tag{5.86}$$

This equation is called the **energy-balance equation**. We know the flux $F_{\rm GW}$, we know the energy $E(r)$, so we can derive from the energy-balance equation an equality for the inspiral rate dr/dt. This is your homework, so here, we will proceed slightly differently.

To proceed, we consider the inspiral in terms of *orbital frequency* Ω. Instead of using Ω directly, the variable x defined in Eq. (5.12) turns out to be more convenient.

In $F_{\rm GW}$, we replace distance r by Ω via Kepler's 3rd law, and replace further by x. One finds

$$F_{\rm GW} = \frac{32c^5\nu^2}{5G}x^5. \tag{5.87}$$

In the (Newtonian) energy of the binary, we likewise replace r by x, to find:

$$E = -\frac{1}{2}\mu c^2\, x. \tag{5.88}$$

Using energy-balance, we find:

$$F_{\rm GW} = -\frac{dE}{dt} = -\frac{dE}{dx}\frac{dx}{dt} = \frac{1}{2}\mu c^2\frac{dx}{dt} \qquad \Rightarrow \qquad \frac{dx}{dt} = \frac{64c^3\nu}{5G\,m}x^5. \tag{5.89}$$

We shall use this expression to work out how many orbits an inspiraling binary will perform between frequencies x_1 and $x_2 > x_1$.

We define the orbital phase Φ via

$$\frac{d\Phi}{dt} = \Omega. \tag{5.90}$$

Φ changes by 2π each orbit, so if we know Φ, we also know the number of orbits. To compute the phase as a function of frequency, we use the product rule:

$$\frac{d\Phi}{dx} = \frac{d\Phi}{dt}\frac{dt}{dx} = \Omega\frac{64c^3\nu}{5G\,m}x^5 = \frac{5}{64}\frac{1}{\nu}x^{-7/2}. \tag{5.91}$$

Integrating yields

$$\Phi(x) = -\frac{1}{32}\frac{1}{\nu}x^{-5/2} = -\frac{1}{32}\frac{1}{\nu}\left(\frac{Gm\Omega}{c^3}\right)^{-5/3} \tag{5.92}$$

The number of orbits between x_1 and x_2 are then simply

$$N(x_1, x_2) = \frac{\Phi(x_2) - \Phi(x_1)}{2\pi}. \tag{5.93}$$

Equation (5.92) is important, and we'll need some time to discuss it.

- There is a *negative* power of frequency. That means that in the early inspiral (at lower frequencies), there is an increasingly larger number of orbits. No real surprise, GW emission is weak at large separation i.e. small frequencies, and therefore the binary does many orbits there. In fact, arbitrarily many as the frequency approaches 0.

- There is a *negative* power of symmetric mass-ratio. This implies that large mass-ratio systems ($\nu \ll 1$) will execute increasingly more orbits at each frequency as the mass-ratio becomes more and more extreme. Equal-mass binaries spiral in quite rapidly. Extreme mass-ratio systems spiral in much more slowly. Black Hole–neutron star binaries (where the mass-ratio is expected to be somewhere between 5 and 15) will execute more orbits in LIGO's frequency band than a binary black hole of same total mass with mass-ratio near unity. *Extreme mass ratio inspirals* where the more massive object is a super-massive black hole with mass $m_1 \sim 10^6 M_\odot$ have mass-ratios of $m_2/m_1 \sim 10^5$. Such systems will spend $100,000$'s of orbits in LISA's sensitivity band. Because such systems spiral in so slowly, the geodesic approximation becomes quite good, and linking us back to our initial discussion of geodesics (see page 16)

- Finally, Eq. (5.92) is only the leading order term of a post-Newtonian expansion which happens to take the form

$$\Phi(x) = -\frac{1}{32}\frac{1}{\nu}x^{-5/2}\left(1 + (\ldots)x + (\ldots)x^2 + (\ldots)x^{5/2}\ln(x) + (\ldots)x^3 + \ldots\right). \quad (5.94)$$

  The leading-order behavior here –$x^{-5/2}$– arises through the lowest-order radiation reaction, i.e. 2.5PN in the Newtonian-like equations of motions (5.74). The 1-PN correction in Eq. (5.92) –schematically denoted by $(\ldots)x$— therefore corresponds to 3.5PN in the Newtonian-like equations of motion.

  LIGO detects gravitational waves by matched filtering, which, essentially, is a phase-coherent integration of the detector output. This requires the waveform-templates to be phase-accurate to within ~ 1 GW-cycle. For this to be the case in the low-frequency regime, we must know all negative powers of x in Eq. (5.94), i.e. at least up to 2.5-PN order past the leading order, corresponding to 5-PN in the Newtonian-like equations of motion. This argument for 2.5-PN order in phase is very rough, it turns out higher orders are yet better.

  However, note that energy-balance avoids the need to use the Newtonian-like equations of motion at such excessively high PN-order (they aren't known there anyway). We have it replaced by the following four ingredients:

  1. The assumption that binaries move on circular or
     bits [you will investigate this assumption in Problem set number 3]
  2. The assumption of energy-balance [Note: for those PN-orders where the Newton-like equations or motions are known, energy-balance is confirmed to hold]
  3. The GW-energy flux $F_{\rm GW}(x)$ as a function of frequency x
  4. The total energy of the inspiraling binary $E(x)$ as a function of frequency x.

  Here $F_{\rm GW}(x)$ and $E(x)$ are both PN-expansions, which are known to 3.5PN order.

5.3.3 High-order post-Newtonian results

A vast amount of high-order post-Newtonian results have been computed. The review article "Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries" by Luc Blanchet summarizes them nicely:

Living Reviews in Relativity December 2014, 17:2 `http://link.springer.com/article/10.12942/lrr-2014-2`

For your entertainment, here are pointers to the higher-order PN versions of some of the equations we have encountered:

- The Newtonian-like equation of motions, Blanchet Eq. 203, pages 79-82.
- The center-of-mass Newtonian-like equations of motion: Blanchet Eqs. 219, 220 (pages 91-92).
- The total energy $E(x)$ of a circular binary as a function of orbital frequency, Blanchet Eq. 232, page 95
- GW energy flux $F_{\rm GW}(x)$, Blanchet Eq. 314, pages 122-123. The GW energy-flux includes some extra effects, which we have not yet mentioned: In the quadrupole formula we have assumed that there is direct propagation from the source to the observer. This manifested itself by evaluating the quadrupole moments simply *at the suitable retarded time*. However, at higher PN orders a new effect shows up: The outgoing radiation is scattered at the curved background space-time. Such scattering can happen anywhere in the space-time, and therefore, the corresponding contribution to the energy-flux involves an integral over the entire past trajectory of the binary. This is called a "tail-term". At higher PN order there are also tails-of-tails.

Chapter 6

Numerical Relativity

At sufficiently high post-Newtonian order, post-Newtonian theory is capable of representing the inspiral of two compact objects to sufficient accuracy for LIGO. Near merger of two inspiraling black holes, however, their separation will approach the sum of their Schwarzschild radii:

$$r = r_{S,1} + r_{S,2} = \frac{2Gm_1}{c^2} + \frac{2Gm_2}{c^2} = \frac{2Gm}{c^2}. \tag{6.1}$$

Substituting this distance into Kepler's 3rd law yields

$$\frac{v}{c} = \frac{1}{\sqrt{2}} \sim 1. \tag{6.2}$$

This can be most easily read off Eq. (5.11). We see that near the merger of two black holes, the orbital velocities become so large that the perturbative expansion in $\epsilon = v/c$ no longer is a *perturbative* expansion. Moreover, the basis assumption of post-Newtonian theory –the existence of two discrete point-masses– does not apply during the merger of the two bodies.

Therefore, we need entirely different techniques to explore what happens when two black holes (or two neutron stars) collide: Numerical relativity.

Numerical solutions of Einstein's equations have a long history. The first attempt to simulate two colliding black holes was made in 1964 by Hahn & Lindquist, "The two=body problem in geometrodynamics" Annals of physics 29, 304 (1964). The break-through on black-hole simulations was accomplished by Frans Pretorius in 2005.

Good references:

- Baumgarte & Shapiro: "Numerical Relativity", Cambridge University Press
- Baumgarte & Shapiro, Phys.Rept.376:41-131,2003, arXiv:gr-qc/0211028.
- For 3+1 decomposition, see also the PhD thesis of Gregory Cook "Initial Data for the Two-Body Problem of General Relativity" (1990) `http://users.wfu.edu/cookgb/Thesis2side.pdf`

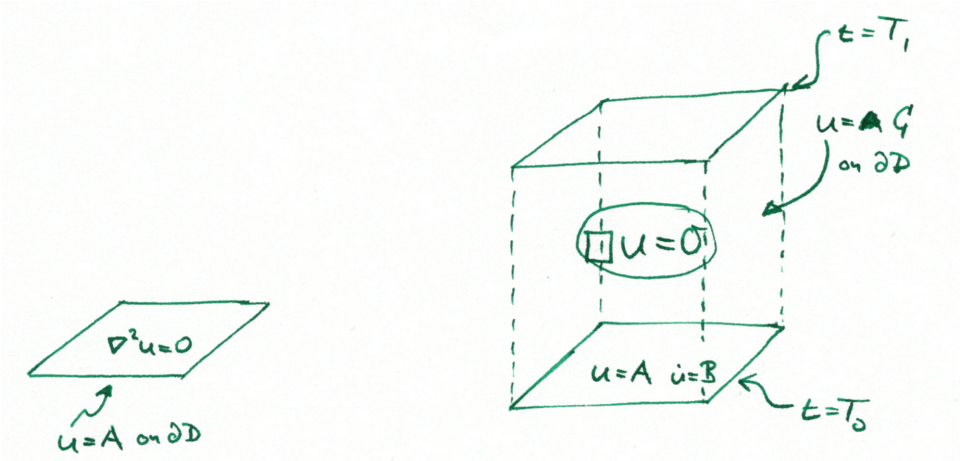


Figure 6.1: Illustration of elliptic and hyperbolic problems. **Left:** Elliptic problems only deal with space, and they need boundary conditions on $\mathcal{D}$. **Right:** Hyperbolic problems involve both space and time; they need initial conditions (at $t = T_1$) and boundary conditions (at $\vec{x} = \partial\mathcal{D}$).

6.1 Recap: partial differential equations

There are two important types of partial differential equations that we will need going forward: Elliptic and hyperbolic ones.

These two types of equations differ radically in their mathematical properties, and therefore also in the numerical strategies that are used to solve them. We need both types when solving Einstein's equations, so we start by briefly introducing them.

6.1.1 Elliptic equations

The prototypical elliptic equation is the Laplace equation or Poisson equation:

$$\nabla^2 u(x) = \rho(x), \qquad x \in \mathcal{D}, \tag{6.3a}$$

where the domain $\mathcal{D}$ and the right-hand-side $\rho(x)$ are given. Elliptic equations like Eq. (6.3a) cover only space and do not involve time, as illustrated in the left panel of Fig. 6.1. Therefore the domain $\mathcal{D}$ on which the equation is being solved is typically $\mathcal{D} \subseteq \mathbf{R}^3$ (I use dimension 3 for concreteness, but this discussion applies to any dimension). If $\mathcal{D}$ has boundaries, then we need *one* boundary condition on each boundary, for instance a *Dirichlet* boundary condition:

$$u(x) = A(x), \qquad x \in \partial\mathcal{D}, \tag{6.3b}$$

or a *von Neumann* boundary condition that uses the derivative along the normal to the boundary:

$$\partial_n u(x) = B(x), \qquad x \in \partial\mathcal{D}. \tag{6.3c}$$

If $\mathcal{D}$ extends to infinity, then we need suitable fall-off conditions there, e.g.

$$u \to 0, \qquad \text{as } x \to \infty \tag{6.3d}$$

(or any other asymptotic condition that makes sense for the problem under study). We already encountered the Poisson-equation when solving Newtonian and post-Newtonian potentials.

Elliptic equations can be more complicated, and they can also consist of a set of multiple equations, rather than just a single equation. For instance, instead of the simple $\nabla^2 = \partial_x^2 + \partial_y^2 + \partial_z^2$, there could be a generic 2-nd order differential operator

$$A^{ij}(x)\frac{\partial^2 u(x)}{\partial x^i \partial x^j} = \dots . \tag{6.4}$$

The coefficient matrix A^{ij} can also be spatially dependent.

If the matrix A^{ij} is positive definite, then the equation is called elliptic. In that case, it requires boundary conditions (as indicated above). If the right-hand-side is sufficiently simple (e.g. zero, or linear in u with a positive sign), then there exists precisely one solution.

6.1.2 Hyperbolic equations

Hyperbolic equations are *evolutionary* equations, that determine how a spatial function changes in time. The typical example is the wave equation:

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}\right) u(\vec{x},t) = S(\vec{x},t), \qquad (\vec{x},t) \in \mathcal{D} \times [T_0, T_1] \tag{6.5a}$$

The solution $u(\vec{x},t)$ of a hyperbolic equation is a function of space and time, as indicated on the right side of Fig. 6.1. Usually, such equations are defined on a certain spatial domain $\mathcal{D} \subseteq \mathbf{R}^3$ and for a certain time-interval $[T_1, T_2] \subseteq \mathbf{R}$.

Hyperbolic equations require *initial conditions* that give the field-values at the initial time T_0:

$$u(\vec{x}, T_0) = A(\vec{x}). \tag{6.5b}$$

$$\frac{\partial u}{\partial t}(\vec{x}, T_0) = B(\vec{x}). \tag{6.5c}$$

Furthermore, hyperbolic equations typically require *boundary conditions*. In its simplest form they can be written as

$$u(\vec{x},t) = C(\vec{x},t), \qquad \vec{x} \in \partial\mathcal{D}, t \in [T_0, T_1]. \tag{6.5d}$$

Note that initial conditions are applied at $t = T_0$ on a *spacelike* boundary, whereas boundary conditions are imposed on $\vec{x}$ =const *time-like* boundaries.

Equation (6.5a) also has the form of Eq. (6.4), if one uses a 4-d coefficient matrix $A = \mathrm{diag}(1, -1, -1, -1)$. In terms of this coefficient matrix, the difference between elliptic and hyperbolic boils down to the signature of the oefficient matrix: If A is symmetric and positive definite (i.e. has only positive eigenvalues), then we have an elliptic equation. If precisely one of the eigenvalues has the opposite sign then the problem is hyperbolic (and the corresponding eigenvector plays the role of time).

6.1.3 First order symmetric hyperbolic systems

By introducing new variables, it is possible to reduce 2nd order hyperbolic equations to *first order form*. For instance, for a scalar wave Eq. (6.5a), one can define the first derivative and the gradient,

$$\vec{\Phi} \equiv \nabla u, \tag{6.6}$$

$$\Pi \equiv \frac{\partial u}{dt}. \tag{6.7}$$

Then, the wave-equation takes the form

$$\frac{\partial \Pi}{\partial t} = \nabla \cdot \vec{\Phi} + S, \tag{6.8}$$

$$\frac{\partial u}{\partial t} = \Pi, \tag{6.9}$$

$$\frac{\partial \Phi}{\partial t} = \nabla \Pi. \tag{6.10}$$

If we label the vector of variable $u^\alpha = (u, \Pi, \Phi_k)$, then the first-order representation takes the form

$$\frac{\partial u^\alpha}{\partial t} + \sum_{k,\beta} A^{\alpha k}_{\beta} \frac{\partial u^\beta}{\partial x^k} = R^\alpha. \tag{6.11}$$

The coefficient-matrices $A^{\alpha k}_{\beta}$ and the right-hand-side can depend on space, time and the fundamental fields u^α, however, they are not allowed to depend on derivatives $\partial_k u^\alpha$.

There is a vast amount of mathematical results that concerns first order hyperbolic equations of the form Eq. (6.11). For instance:

- Whether a system of the form (6.11) is hyperbolic depends on properties of the matrices ${A^\alpha}_\beta{}^k$. Specifically, if the matrices ${A^\alpha}_\beta$ are symmetric for each k, then the system is hyperbolic.
- Information propagates only at finite speed(s), and those speeds can be determined from the matrices $A^{\alpha k}_{\beta}$.
- What boundary conditions must be imposed at time-like boundaries can also be determined from the $A^{\alpha k}_{\beta}$.

6.2 General Relativity

Our goal is to find numerically a space-time metric g_{ab} that satisfies Einstein's equations

$$G_{ab}[g_{ab}] = 8\pi T_{ab}. \tag{6.12}$$

This requires several steps. The first one is to "undo" the great achievement of Einstein, namely the unification of space and time into space-time. It turns out that computers cannot deal with covariant space-time; rather they want explicit and unambiguous instructions of what to calculate at any step in the "computational time".

6.2.1 Introduce "time" explicitly – Foliation

Foliate space-time by space-like hypersurfaces Σ_t, labelled by a continuously and monotonically increasing real value, **the time** t. This is illustrated in Fig 6.2.

From the choice of Σ_t a whole host of further objects follow:

- **Future-pointing unit normal n^a**
  A one-form normal to Σ_t is given by $dt = \partial_a t$. By rescaling

  $$n_a = -\alpha \partial_a t, \quad n^a = -\alpha g^{ab} \partial_b t, \quad n^a n_a = -1, \tag{6.13}$$

  where the scale factor α is given by $\alpha = (-g^{ab}\partial_a t \partial_b t)^{-1/2}$. The function α is called the **lapse function**. It measures the separation of nearby hypersurfaces Σ_t in the direction normal to the hypersurfaces.

- **Spatial metric** The space-time metric induces a **spatial metric** γ_{ab} on the Σ_t hypersurfaces,

  $$\gamma_{ab} = g_{ab} + n_a n_b, \qquad \gamma^a{}_b = \delta^a{}_b + n^a n_b, \qquad \gamma^{ab} = g^{ab} + n^a n^b. \tag{6.14}$$

  The mixed upper-lower metric $\gamma^a{}_b$ removes components orthogonal to Σ_t, i.e. it projects into the hypersurfaces.

  $$\gamma_{ab} = g_{ab} + n_a n_b, \qquad \gamma^a{}_b = \delta^a{}_b + n^a n_b, \qquad \gamma^{ab} = g^{ab} + n^a n^b. \tag{6.15}$$

  The mixed upper-lower metric $\gamma^a{}_b$ removes components orthogonal to Σ_t, i.e. it projects into the hypersurfaces.

- **Adopted coordinates and spatial tensors**
  Write space-time coordinates as $x^a = (t, x^i)$, where $i, j, \ldots = 1, 2, 3$ are coordinates within each Σ_t. Spatial tensors "live within Σ_t", this means each component is orthogonal to the normal n^a:

  $$T^a{}_b n_a = 0, \quad T^a{}_b n^b = 0. \tag{6.16}$$

  For spatial tensors, the spatial components carry all information, so we'll write $T^i{}_j$ instead of $T^a{}_b$.

- **3-D differential geometry within Σ_t**

  The space-time metric induces a **spatial metric** γ_{ij} on the Σ_t hypersurfaces. Consider two points x^a and $x^a + dx^a$ within the same Σ_t. Because they are in the same hypersurface, dx^a is purely spatial, and we can write it as dx^i.

  Their separation is

  $$ds^2 = g_{ab}x^a dx^b \equiv \gamma_{ij}dx^a dx^j, \tag{6.17}$$

  where the second equality defines the spatial metric γ_{ij}.

  Along with it there are spatial Christoffel-symbols

  $${}^{(3)}\Gamma^i{}_{jk} = \frac{1}{2}\gamma^{il}\left(\gamma_{jl,k} + \gamma_{lk,j} - \gamma_{jk,l}\right), \tag{6.18}$$

  a spatial covariant derivative operator

  $$D_i T^j{}_{k\ldots} = \partial_i T^j{}_{k\ldots} + {}^{(3)}\Gamma^j{}_{il}T^l{}_{k\ldots} - {}^{(3)}\Gamma^l{}_{ik}T^j{}_{l\ldots} - \ldots, \tag{6.19}$$

  (tensors to be differentiated must be spatial, otherwise Leibnitz's rule fails), and the spatial Riemann tensor

  $$\left(D_i D_j - D_j D_i\right) W_k = {}^{(3)}R^l_{ijk} W_l. \tag{6.20}$$

- **Extrinsic curvature K_{ij}**

  The spatial metric γ_{ij} determines the *intrinsic* geometry of the hypersurface, but does not contain information about how Σ_t is embedded in the surrounding space-time. The latter information is contained in the extrinsic curvature tensor, which can be defined in several equivalent ways:

  $$K_{ab} = \gamma_a{}^c \gamma_b{}^d \nabla_{(c} n_{d)} \tag{6.21}$$

  $$= -\frac{1}{2}\mathcal{L}_n \gamma_{ab}. \tag{6.22}$$

  The first of these two definitions is quite intuitive: If the unit-normal n_a changes direction as one moves along Σ_t (i.e. n_a has non-zero gradient), then the extrinsic curvature is non–zero.

  The Lie-derivative along n^a that appears in the second definition is, roughly-speaking, a "time-derivative". Therefore, the extrinsic curvature can be thought of, roughly speaking, as the time-derivative of the spatial metric.

6.2.2 Choose space-time coordinates — 3+1 split

So far, we've considered only *one* hypersurface Σ_t, on which we chose spatial coordinates x^i. We now need to decide how to relate spatial coordinates on neighboring hypersurfaces $\Sigma_{t+\delta t}$. For this purpose, introduce a **time-vector field t^a**

$$t^a = \alpha\, n^a + \beta^a, \tag{6.23}$$

and define the spatial coordinates x^i such that integral curves of t^a connect points with the *same* spatial coordinates x^i.

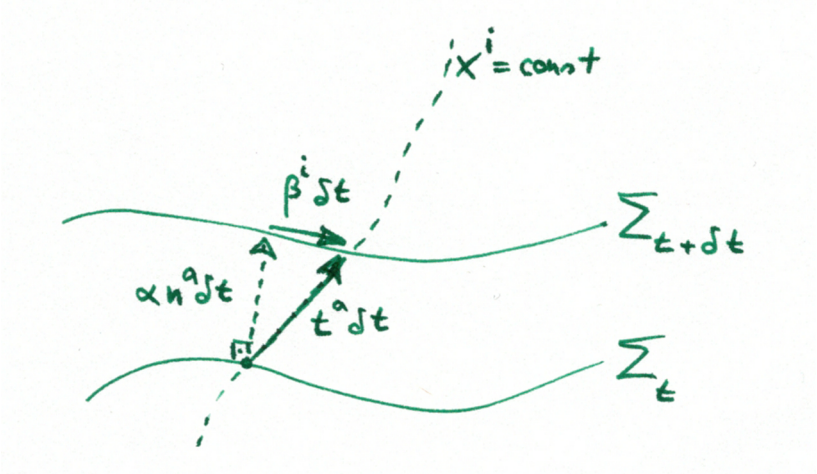


Figure 6.2: 3+1 decomposition of space-time.

This vector field has some component $\alpha\, n^a$ normal to Σ_t, and a component β^a parallel to Σ_t. The spatial vector β^a is called **shift vector**. In these special coordinates, the space-time metric takes the form

$$ds^2 = -\alpha^2 dt^2 + \gamma_{ij}\left(dx^i + \beta^i dt\right)\left(dx^j + \beta^j dt\right) \tag{6.24}$$

that is

$$g_{ab} = \begin{pmatrix} -\alpha^2 + \beta^i\beta_i & \beta_i \\ \beta_j & \gamma_{ij} \end{pmatrix}. \tag{6.25}$$

The upper space–time metric can be written as

$$g^{ab} = \begin{pmatrix} -\alpha^{-2} & \alpha^{-2}\beta^i \\ \alpha^{-2}\beta^j & \gamma^{ij} - \alpha^{-2}\beta^i\beta^j \end{pmatrix}. \tag{6.26}$$

Note that the spatial metric γ_{ij} is given by the spatial components of the space-time metric g_{ab}. However, note that the spatial inverse metric γ^{ij} is *not* equal to the spatial compontents of the inverse space-time metric g^{ab}.

6.2.3 Decompose Einstein's equations — ADM–equations

Having described the nomenclature, we can now look at Einstein's equations, $G_{ab} = 8\pi T_{ab}$. The Riemann-tensor can be split into pieces normal and tangential to Σ_t. In the preferred 3+1 coordinates, this boils down to choosing specific components. It turns out that the *spatial* derivatives inside the space-time Riemann tensor can be rewritten in terms of the spatial (3-D) Riemann tensor within Σ_t. The normal derivatives (along n^a) can be rewritten in terms of the extrinsic curvature K_{ij}, cf. Eq. (6.22). The derivatives along n^a can be re–expressed as time-derivatives and a Lie-derivative along the shift-vector using

$$\partial_t = \mathcal{L}_t = a\mathcal{L}_n + \mathcal{L}_\beta.$$

(The first equality is an identity for Lie-derivatives taken along coordinate basis-vectors. The second identity arises because of $t^a = \alpha n^a + \beta^a$).

For the time-time, time-space and space-space components of the Einstein equations, one finds:

$$ {}^{(3)}R + K^2 - K_{ij}K^{ij} = -16\pi\rho \tag{6.27} $$

$$ D_j K^j{}_i - D_i K = \quad 8\pi j_i \tag{6.28} $$

$$ \partial_t K_{ij} = -D_i D_j \alpha + \alpha \left({}^{(3)}R_{ij} - 2K_{ik}K^k{}_j + K\, K_{ij} \right) - \alpha\, 8\pi \left(S_{ij} - \frac{1}{2}\gamma_{ij}(S-\rho) \right) + \beta^k D_k K_{ij} + K_{ik} D_j \beta^k + K_{kj} D_i \beta^k. \tag{6.29} $$

Here $K = \gamma^{ij} K_{ij}$ is the trace of the extrinsic curvature, and ρ, j_i and S_{ij} the matter density, momentum density, and stress density arising from the stress-energy tensor. The final crucial equation is the definition of the extrinsic curvature, Eq. (6.22), from which it follows that

$$ \partial_t \gamma_{ij} = -2\alpha K_{ij} + D_i \beta_j + D_j \beta_i \tag{6.30} $$

Eqs. (6.27)–(6.30) are referred to as the ADM–equations. (named after Arnowitt, Deser, Misner, 1962, who wrote down a slightly different set of equations. The precise set of Eqs. (6.27)–(6.30) was first published by J. W. York in 1979)

Let's discuss the structure of the ADM–equations by comparison with Maxwells Equations:

$$ \vec{\nabla} \cdot \vec{E} = 4\pi\rho_e \tag{6.31} $$

$$ \vec{\nabla} \cdot \vec{B} = 0 \tag{6.32} $$

$$ \partial_t \vec{E} = \vec{\nabla} \times \vec{B} - 4\pi \vec{J} \tag{6.33} $$

$$ \partial_t \vec{B} = -\vec{\nabla} \times \vec{E} \tag{6.34} $$

Maxwell's equations (as well as Einstein's equations) contain equations *without time-derivatives* (Eqs. (6.31), (6.32) and Eqs. (6.27), (6.28)). These are **constraint equations**, restricting the set of allowable field-configurations *at a given time t*. Then there are **evolution equations**, prescribing how the fields are changing in time (Eqs. (6.33), (6.34) and Eqs. (6.29), (6.30)).

Hence, in principle, solvinging Einstein's equations mirrors solving Maxwell's equations. First, find initial data that satisfies the constraint equations. Second, evolve forward in time.

One important difference lies in the functions α and β^i. They appear in the evolution equations, but there aren't equations that determine them! These functions are freely specifiable; they determine the coordinate system in which the solution is being constructed.

However, there are three important differences between Maxwell's Equations and the ADM-equations:

1. Maxwell's equations are linear, whereas Einstein's equations are highly non-linear (there are squares of extrinsic curvature, and the Ricci-tensors appearing in the ADM-equations is a highly nonlinear combination of the metric components and their derivatives).

2. The Maxwell constraint equations (6.31) and (6.32) are easily solvable, because of their simple form. In contrast, the Einstein constraint equations as written in Eq. (6.27) and (6.28) are of *no known mathematical type*. We will have to first rewrite them in a form in which it becomes clear how to solve these equations.

3. The Maxwell evolution equations (6.33) and (6.34) are well-posed hyperbolic evolution equations. In contrast, the ADM evolution equations as written in Eqs. (6.29) (6.30) are *not hyperbolic*, and therefore naively coding these as they are written will lead to computer codes that will fail. We will have to first rewrite the evolution equations in a form which allows implementation on computers. This is a highly non-trivial comment: It took decades until the importance of hyperbolicity was noted, and several more years until correct and useful hyperbolic formulations were developed.

6.3 Constructing initial data

Let's tackle the first problem. How can we rewrite the initial data constraint equations (6.27) and (6.28) into equations that we know how to solve?

6.3.1 Conformal transformation

The constraint equations (6.27) and (6.28) are coupled non-linear partial differential equations, of no specific mathematical form. They can be transformed into *elliptic equations* by a suitable variable transformation. One separates from the spatial metric an overall conformal factor,

$$\gamma_{ij} = \psi^4 \bar{\gamma}_{ij}. \tag{6.35}$$

The Ricci-scalar in Eq. (6.27) can now be expressed in terms of the Ricci-scalar $\bar{R}$ of $\bar{\gamma}_{ij}$ and covariant derivatives with respect to $\bar{\gamma}_{ij}$,

$$R = \psi^{-4}\bar{R} - 8\psi^{-5}\bar{D}^2\psi.$$

Substituting back into the Hamiltonian constraint Eq. (6.27) gives an equation of the form

$$\bar{D}^2\psi = \dots. \tag{6.36}$$

This is a non-linear elliptic equation for ψ. Furthermore, one makes an Ansatz for K_{ij} in terms of a vector-potential, and arrives at elliptic equations for the vector-potential. The point of view is now:

1. Choose $\bar{\gamma}_{ij}$ and some further "free data" that enter the ansatz for K_{ij}

2. Solve Eq. (6.36) and elliptic equations for the vector-potential.

3. Assemble the full physical initial data γ_{ij} and K_{ij}

6.3.2 BBH initial data

Puncture–initial data (Brandt, Bruegmann, PRL 78 (1997), p. 3606), makes many simplifying assumptions, to reduce the complexity of the numerical problem. Very widely used for BBH evolutions, because simple. Major deficit: Can only deal with black holes of spins $S/M^2 \lesssim 0.9$.

Conformal–thin sandwich initial data (Cook, Pfeiffer, PRD 70 (2004), p.104016) preserves generality, allows construction of very high spins (Lovelace et al, 2008). Not as widely used. Major deficit: More complicated equations to solve.

All BBH initial data are constructed with so-called **spectral methods**. One expands the solution in basis-functions (e.g. spherical harmonics), and solves for the expansion coefficients. Two codes: For **puncture data**, Ansorg's code is the standard work-horse (Ansorg. "A multi-domain spectral method for initial data of arbitrary binaries in general relativity". Classical and Quantum Gravity (2007) vol. 24 pp. 1). For **conformal–thin sandwich initial-data**, my code continues to be state-of-the-art (Pfeiffer et al. "A multidomain spectral method for solving elliptic equations." Computer Physics Communications (2003) vol. 152 pp. 253). Although Trevor Vincent (a graduate student at CITA) is developing a superior code.

6.4 Evolution equations

BH evolutions have four unique features:

- The Constraints must be preserved
  Analytically, vanishing constraints $C = 0$ are preserved under evolutions. Numerically, small errors tend to exponentially amplify:
  $$\frac{dC}{dt} \propto C \quad \Rightarrow \quad C \sim e^{\lambda t}$$
- One has to deal with singularities inside black holes
- One must choose coordinates "on the fly"
  In principle any coordinate system *should* do. But one must avoid coordinate singularities, ensure the coordinates cover the region of space-time of interest.
- Analytical properties of the equations. One has to work hard to find a formulation that is well-behaved. Specifically, writing out ${}^{(3)}R_{ij}$ in terms of derivatives of the metric, Eq. (6.29) becomes
  $$\partial_t K_{ij} = -\frac{\alpha}{2}\gamma^{kl}\partial_k\partial_l\gamma_{ij} + \alpha\gamma_{k(i}\partial^{(3)}_{j)}\Gamma^k + \ldots, \tag{6.37}$$
  where ${}^{(3)}\Gamma^k \equiv^{(3)} \Gamma^k{}_{ij}\gamma^{ij}$. The first term alone would be fine and would result in a hyperbolic wave–equation. The second term destroys hyperbolicity.

... and of course, the equations are very complicated, and will require supercomputers to solve...

6.4.1 BSSN equations

(Baumgarte & Shapiro 1999, based on work by Nakamura et al 1987)

Promote ${}^{(3)}\Gamma^k$ **to evolved variables**, with equal standing to γ_{ij} and K_{ij}. Now the second term in Eq. (6.37) has only first derivatives of evolved variables, and is no longer important for the mathematical analysis. (There are some additional changes of variables, and one ends up evolving about 20 variables.)

Choose very specific coordinate conditions (Γ–driver). These coordinate-conditions were basically found by experimentation, but they work very well. The BH singularity turns out to be separated by an infinitely long Einstein-rosen bridge, and so its effects are mitigated.

constraint damping (similar to Generalized Harmonic, see next section)

Numerics: finite-differences, adaptive mesh refinement, and supercomputers

6.4.2 Generalized Harmonic coordinates

One evolves directly the space-time metric, and uses a specific coordinate choice to simplify Einstein's equations. Nice exposition in Lindblom et al, Classical and Quantum Grabity, 23 (2006) p. S447.

Let's consider the vacuum Einstein's equations, $R_{ab} = 0$. The highest derivative terms in the Ricci-tensor have the form:

$$0 = R_{ab} = -\frac{1}{2} g^{cd} \partial_c \partial_d g_{ab} + \nabla_{(a} \Gamma_{b)} + \dots \tag{6.38}$$

The first term on the right hand side is nice (hyperbolic, well-posed, wave-equation). The second term on the right hand side destroys hyperbolicity. To get rid of this second term, choose coordinates such that Γ_b equals an a priori chosen function H_b:

$$\Gamma_b \equiv -H_b. \tag{6.39}$$

Now the nasty second term turns into first derivatives of a known function.

Harmonic coordinates set $H_b \equiv 0$. This has been used since 1920's to analyze Einstein's equations. "Generalized Harmonic Coordinates" allow for non-zero H_b (Friedrich, 1987).

Constraint damping: We have assumed that $\Gamma_b \equiv -H_b$ to re-write the evolution equations. However, the Γ_b are functions of the evolved variables g_{ab}, and due to numerical errors, there is no guarantee that this relation continues to hold. One modifies the evolution equations:

$$0 = R_{ab} = -\frac{1}{2} g^{cd} \partial_c \partial_d g_{ab} - \nabla_{(a} H_{b)} + \tau \left[t_{(a} \left(H_{b)} + \Gamma_{b)} \right) - \frac{1}{2} g_{ab} t^c (H_c + \Gamma_c) \right] + \dots, \tag{6.40}$$

where $\tau > 0$ is a free parameter. If $H_b + \Gamma_b = 0$, the square-bracket vanishes. In this case, Eq. (6.40) reduces to the Einstein equations, and we solve indeed the problem of interest.

The interesting aspect arises when $H_b + \Gamma_b$ becomes non-zero (for whatever reasons). If $H_b + \Gamma_b \neq 0$, then one can show that for perturbations of Minkowski space,

$$\partial_t \left(H_b + \Gamma_b \right) \sim -\tau \left(H_b + \Gamma_b \right). \tag{6.41}$$

This means, the constraint violations are exponentially suppressed on a time-scale τ.

Black hole excision: The interior of the BH is causally disconnected from the exterior. Place an artifical boundary somewhat inside the BH, and excise the region inside.

Numerics: Frans Pretorius (Princeton) and Luis Lehner (Perimeter/U Guelph) use finite differences, adaptive mesh refinement. The Caltech/Cornell/CITA group uses multidomain spectral methods.

6.5 Some Results of BBH simulations

6.5.1 BBH Kicks

Asymmetric mergers result in asymmetric emission of gravitational waves, resulting in a change in linear momentum of the remnant BH.

RIGHT: BH-kicks for non-spinning binaries, kicks of 100km/sec due to mass-ratio. From Gonzalez et al, PRL 98 (2007) p. 091101 (University of Jena, Germany).

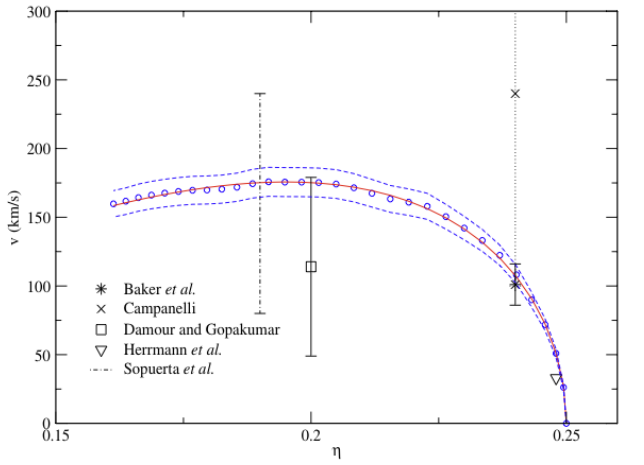


FIG. 2 (color online). The kick velocity as a function of mass ratio, with an error of ±6% indicated by the dotted lines. We also indicate previous numerical results from Baker *et al.* [30], Campanelli [29], and Herrmann *et al.* [24], and the analytic estimates of Damour and Gopakumar [14] and Sopuerta *et al.* [15].

For **spinning binaries** kicks up to 5000km/sec (!!) due to spins.

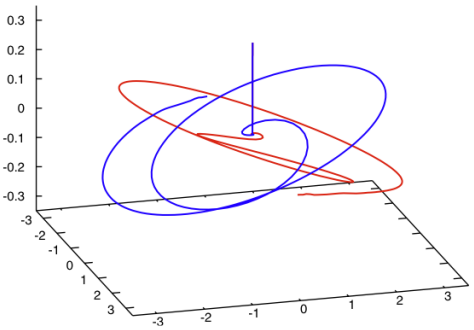
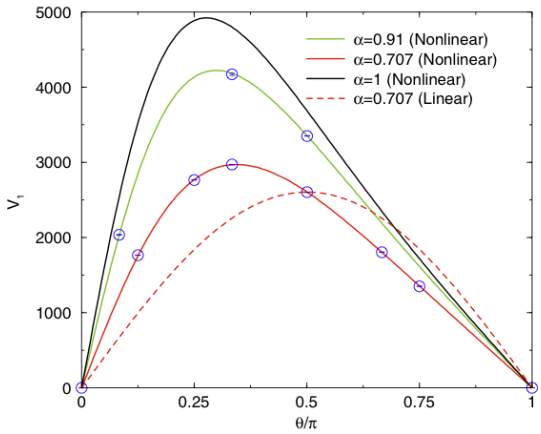


FIG. 4 (color online). The three-dimensional trajectories of the punctures showing the orbital precession and the final recoil for the SP2 configuration. Note that the scale of the z axis is 1/10 that of the x and y axes.

Campanelli et al, 2007. Orbital trajectories of a BH-BH with spins initially tangent to the orbital plane, and pointing in opposite directions. $v_{\rm kick} \sim 4000 km/sec$.



Lousto & Zlochower, 2011: Spins having an angle θ with the initial angular momentum. Components of spins within orbital plane anti-parallel. $\theta = \pi/2$ corresponds to plot on left.

6.5.2 Final Mass & Spin

Left: Rezzolla et al, ApJ 674 (2008), L29 (Albert Einstein Institute, Germany). Right: Hemberger et al, PRD 88, 064014 (2013) arXiv:1305.5991

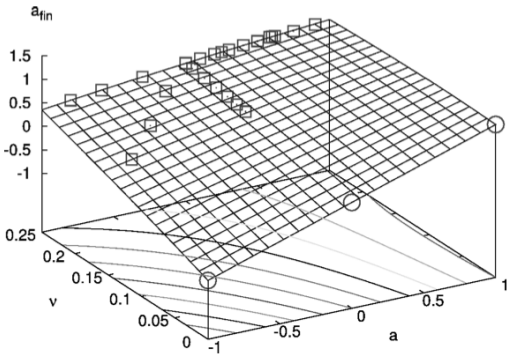


FIG. 1.—Global dependence of the final spin on the symmetric mass ratio and on the initial spins as predicted by expression (5). Squares refer to numerical estimates while circles to the EMRL constraints. [*See the electronic edition of the Journal for a color version of this figure.*]

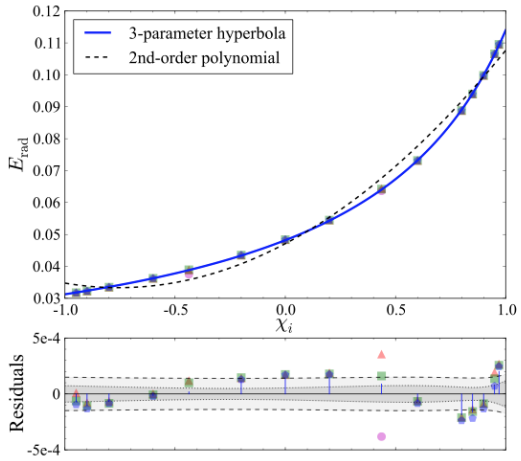
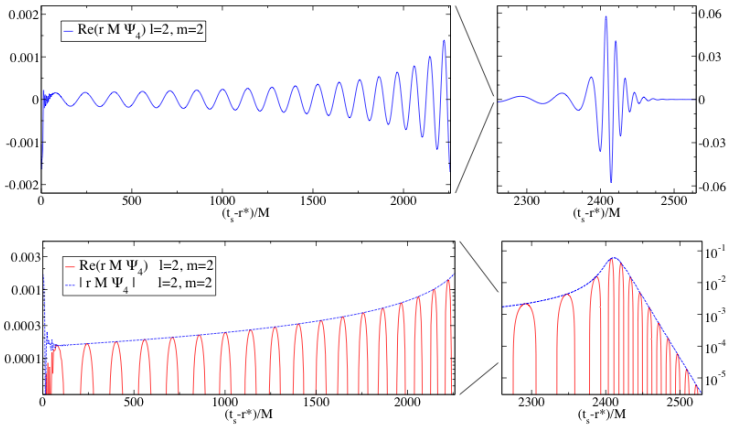


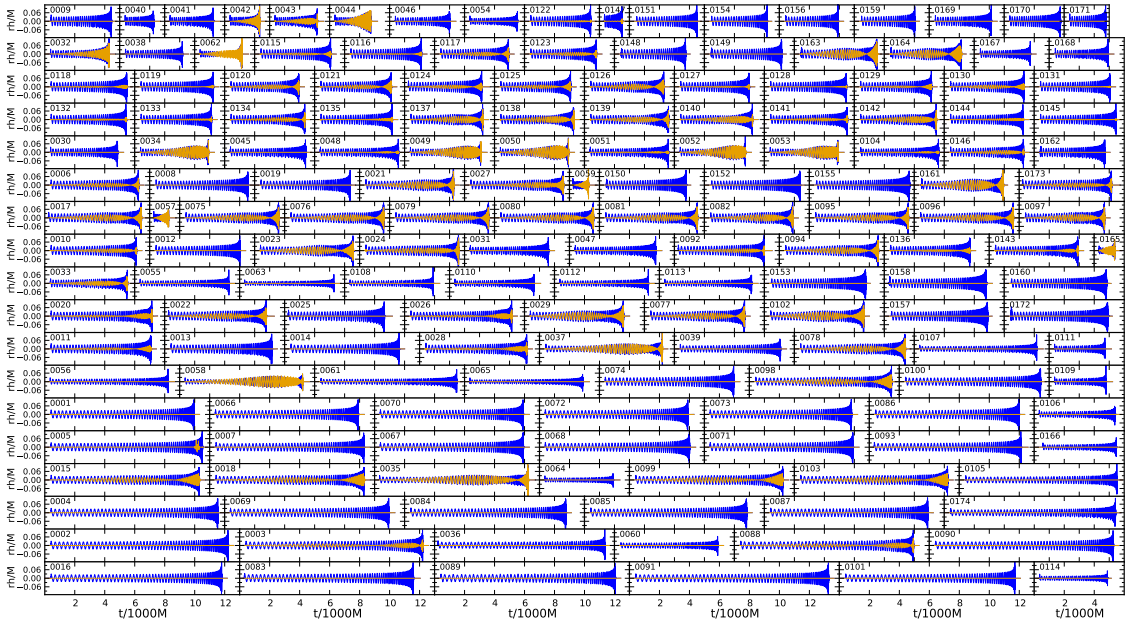
FIG. 8: In the top panel, we plot our preferred fitting formula (solid line), the hyperbolic function in Eq. (9), and a comparison with a second-order polynomial (dashed line) for $E_{\rm rad}$ as a function of χ_i. Our data points are plotted as polygons, where more sides indicates higher resolution level. In the bottom panel, we plot the residuals of the hyperbolic function. We indicate our fit parameter (dotted line) and total prediction (dashed line) uncertainties (defined in Appendix A).

6.5.3 Waveforms for gravitational wave detectors

These simulations need to cover many orbits before merger, and are therefore much more costly than shorter simulations investigating mergers. To the right, one example from Chu et al, arXiv:0909.1313 (Caltech/Cornell/CITA), for an equal mass BBH with antialigned spins



Below, a large exploration of BBH parameter space to date, led by CITA Postdoc Abdul Mroue (Mroue et al, PRL 111, 241104 (2013), arXiv:1304.6077):



6.6 BBH Open Challenges

- Explore parameter space. At least seven parameters: mass-ratio and spin-vectors of each BH. Possibly eccentricity relevant. Note that $3^7 = 2000$, and with each simulation costing 10,000's CPU-hours, one needs 1000's of CPU's for years...
- Near extremal spins
- Large mass-ratios, $M_1/M_2 \gtrsim 10$. Computationally much more expensive.
- Eccentric binaries: Increase dimensionality of parameter space further.

Chapter 7

Black hole quasi-normal modes

Immediately after the horizons of two black holes merge, one has one larger, very deformed black hole. This deformed black hole settles down into a quiescent Kerr black-hole. At some point the deformations have decayed enough to allow a perturbative treatment around the final Kerr black hole.

During this perturbative ringdown, the perturbations decompose into different modes labeled by three integers, (l, m, n). By convention, the first two (l, m) represent the angular shape of the perturbation-mode, in terms of spheroidal harmonics $Ylm(\theta, \phi)$. The third integer controls the radial shape of the pertubation-mode, with $n = 0$ being the most important one, and $n > 0$ being ever more quickly damped (n counts the number of radial zeros in the eigenmode).

Each mode exhibits a time-behavior which is oscillating and exponentially damped:

$$h \propto \exp\{-t/\tau_{lmn}\}\, sin(\omega_{lmn} t + \phi_{lmn}) \tag{7.1}$$

Let's define a quality-factor Q_{lnm} of the oscillations as

$$Q_{lmn} = \omega_{lmn}\tau_{lmn}/2. \tag{7.2}$$

(We include the factor 1/2 for consistency with Figure 7.1.) The meaning of Q is as follows: Within Q periods, $\Delta T = Q/\omega$, the amplitude of the quasi-normal mode decays by a factor

$$e^{-\Delta T/\tau} = e^{-Q/(\omega\tau)} = e^{-2}. \tag{7.3}$$

Therefore, we see that $Q_{lmn}/2$ corresponds to the number of GW cycles per e-folding-time of the decaying perturbations.

The modes are also a function of the remnant BH spin. Figure 7.1 plots the frequencies of the modes, and their quality-factors as a function of BH spin. There are quite a few important points to note:

1. As already mentioned the 220 mode is of highest importance for BBH mergers: It is preferentially excited, because two black holes merging naturally represent a quadrupole deformation. This is the top-most black solid curve in Fig. 7.1, on which we henceforth focus.

2. Black holes are *extremely strongly damped*: There are only O(1) periods per e-fold decay.

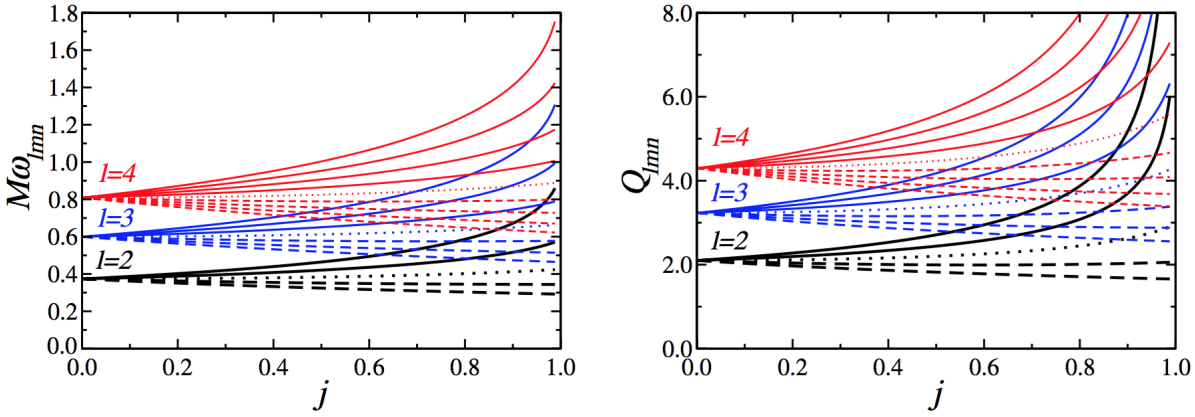


FIG. 5 (color online). Frequency f_{lmn} (left) and quality factor Q_{lmn} (right) for the fundamental modes with $l = 2, 3, 4$ and different values of m. Solid lines refer to $m = l, \ldots, 1$ (from top to bottom), the dotted line to $m = 0$, and dashed lines refer to $m = -1, \ldots, -l$ (from top to bottom). Quality factors for the higher overtones are lower than the ones we display here.

Figure 7.1: Frequencies (left) and decay-times of quasi-normal mode perturbations of a Kerr black hole. The x-axis shows the dimensionless angular momentum of the black hole, $j = S/M^2$, and the y-axis shows frequency and quality-factor for a variety of modes. The most important mode is the $2, 2, 0$ mode, shown as the top-most solid line labeled $l = 2$. Figure from Berti, Cardoso & Will, Physical Review D **73**, p. 064030 (2006).

3. The frequency of the quasi-normal modes is mildly dependent on the spin, with the (220) mode roughly doubling in frequency for an extremal Kerr black hole.

4. The quality-factor is a steeper function of spin: For highly spinning BH's, the ringdown is prolonged, and passes through several times more cyles per e-folding. In fact, the quality factor diverges as the BH spin becomes extremal.

5. $M\omega_{lmn}$ is a *dimensionless frequency*, i.e. it is a number without units (assuming $G = c = 1$). Substituting back factors of G and c, as well as the solar-mass $M_\odot$, we find the physical frequency to be

$$f_{lmn} = \frac{\omega_{lmn}}{2\pi} = \frac{1}{2\pi} \frac{M\omega_{lmn}}{M/M_\odot \, \frac{GM_\odot}{c^3}} = (M\omega_{lmn}) \left(\frac{M}{M_\odot}\right)^{-1} 32\text{kHz}. \qquad (7.4)$$

   As usual in vacuum GR, the frequency scales *inversely* to the mass. For a remnant BH mass $M \sim 60 M_\odot$ (as for GW150914), and remnant spin $j \sim 0.7$, we have for the 220 mode, $f_{\text{ringdown}} \sim 250$Hz.

6. By BH uniqueness theorems, every single (lmn)-ringdown mode can depend only on the black hole mass M and dimensionless spin $j = S/M^2$. Therefore: (a) if one knows already M and S, then measurement of even a single BH ringdown mode is a test of GR. (b) if one does not know yet M and S, then measurement of *two* modes provides a test of GR.

7. Alternative theories of gravity often have identical BH solutions to GR. However, the ringdown frequencies (which are determined by perturbations around the stationary solutions) tend to be different. As such, measurement of ringdown frequencies constitutes an important part of testing general relativity using gravitational waves.

Chapter 8

Compact binary coalescence: Summary of qualitative features

We've now completed a pass through most important aspects of compact object binaries. Let's recap:

- With compact object, we refer to either black holes, or objects only a factor of a few times larger than their Schwarzschild-radius, i.e. neutron stars.
- Such binaries end their inspiral when their separation is roughly comparable to the inner-most stable circular orbit of a black hole of the same mass, i.e. at separation

$$r \sim 6\frac{Gm}{c^2}. \tag{8.1}$$

  Using Kepler's third law, and the fact that the dominant GW emission is at twice the orbital frequency Ω, we find a GW frequency near merger of

$$f_{\rm GW,merger} \sim \frac{2\Omega}{2\pi} \sim \frac{1}{\pi}\sqrt{\frac{Gm}{r^3}} \sim \frac{1}{\pi 6^{3/2}}\frac{1}{GM_\odot/c^3}\left(\frac{m}{M_\odot}\right)^{-1} \sim \left(\frac{m}{M_\odot}\right)^{-1} 4\text{kHz}. \tag{8.2}$$

  Because of the assumptions, this estimate is only good to a factor of a few.

  - For neutron stars, $f_{\rm GW,merger}$ ∼kHz.
  - Stellar mass black holes ($m = 10..100M_\odot$), $f_{\rm GW,merger} \sim 40...400$Hz
  - Intermediate-mass black holes $m \sim 100..10^4 M_\odot$, $f_{\rm GW,merger} \sim 0.4...40$Hz
  - Small super-massive black holes with $m \sim 10^6 M_\odot$, $f_{\rm GW,merger}$ ∼mHz
  - Large super-massive black holes with $m \sim 10^9 M_\odot$, $f_{\rm GW,merger} \sim \mu$Hz

  These frequencies immediately dictate which instrument can observe each of these systems: Ground-based GW detectors/space-based detectors (LISA)/pulsar-timing arrays.

- The *length of the observable inspiral* depends on the frequency range $[f_{\rm low}, f_{\rm high}]$ of the detector. For LIGO, $f_{\rm low} \sim 20$Hz and the merger-frequencies for stellar-mass

compact-object binaries are within LIGO's frequency band. For instance, the number of orbits of a inspiral is (to leading post-Newtonian order) given by Eq. (5.92):

$$N \approx \frac{-\Phi(x_{\rm low})}{32\pi} = \frac{1}{64\pi\nu}\left(\frac{Gm\pi f_{\rm low}}{c^3}\right)^{-5/3} \sim \frac{3000}{\nu}\left(\frac{f_{\rm low}}{20{\rm Hz}}\frac{m}{M_\odot}\right)^{-5/3}. \tag{8.3}$$

In the first equality, we neglected $\Phi(x_2)$, assuming that the end-frequency $f_2 \gg f_{\rm low}$. Binary Neutron stars, therefore, proceed through about 1000 orbits in LIGO's frequency band. Comparable mass ($\nu \approx 1/4$) BBH spend anywhere between ~ 100 down to only a few inspiral-orbits in band, as the BBH mass varies from $m = 10M_\odot$ to $100M_\odot$.

At fixed total mass, un-equal mass systems have a *longer* inspiral, owning to the factor $1/\nu$. Moving the symmetric mass-ratio ν into the parenthesis, we find

$$N \sim 3000\left(\frac{f_{\rm low}}{20{\rm Hz}}\frac{m\nu^{3/5}}{M_\odot}\right)^{-5/3}. \tag{8.4}$$

Therefore, the length of the inspiral depends only on the *chirp-mass* $M_c = m\nu^{3/5} = (m_1 m_2)^{3/5}/m^{1/5}$.

Chapter 9

Gravitational Wave Detectors

Gravitational waves are being searched for in different frequency bands:

- $f = 10\text{Hz} \cdots \text{kHz}$ – ground-based interferometers (existing & successful detections)
- $f \sim$ mHz – space-based detectors (LISA – planned)
- $f \sim$ nHz (periods of $\sim$ years) – pulsar-timing arrays (ongoing searches for GW)

9.1 Ground-based GW interferometers

Because a GW passing through Earth stretches and squeezes space in alternating directions, the typical detection strategy employs kilometer-sized interferometers. The most important one is LIGO, as shown in Fig. 9.1.

9.1.1 Existing and planned detectors

Interferometric GW detectors began operation in earnest around the year 2000, with initial LIGO in the U.S., Virgo in Italy, TAMA300 in Japan and Geo in Germany. These operators were improved until about 2010, and searched for gravitational waves at ever increasing sensitivity. This first generation of gravitational wave detectors successfully reached its goal of a sensitivity that makes GW detection *plausible*. Unfortunately, no GWs were detected.

Starting around 2010, the LIGO detectors were shut down for a significant improvement to the *Advanced LIGO* detectors, with the goal of a 10-fold sensitivity improvement. Over the globe, the following *second-generation* GW detectors are in operation, or under construction:

- The two LIGO detectors in the U.S. (already operating)
- Virgo in Italy (to come online in coming months)
- Kagra in Japan ($\sim$ 2018)
- LIGO-India in India ($\gtrsim$ 2020)

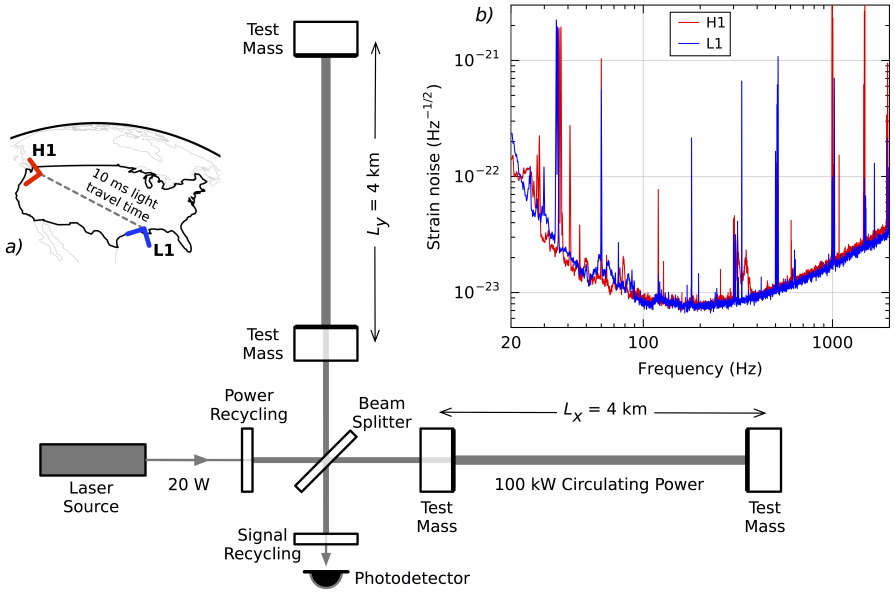


Figure 9.1: Schematic of the LIGO Interferometers and the noise curves during the first Advanced LIGO observing run (O1).

9.1.2 Science Targets

Ground-based GW detectors are sensitive from frequencies $f \sim 10Hz$ to a few KhZ. This sensitivity range determines their science targets:

- Inspiraling BBH binaries: Visible to a couple of Gpc.
- Inspiraling BH-NS binaries: Visible to several 100 Mpc.
- Inspiraling NS-NS binaries: Visible to $\sim$ 80Mpc (currently), $\sim$ 200Mpc at Advanced LIGO design
- core-collapse supernovae. It is unclear how much GW's are emitted, so it is unclear how far these can be seen. Possibly only visible within the Milky Way; since there is a super-nova only once every $\sim$ 100years in the Milky Way, it may therefore be quite unlikely to be observed. However, if seen, entirely complementary information to electro-magnetic observations of supernovae: GW's generated by the matter flows near the centre can escape unimpeded, and offer so a direct means to study the central engine of core-collapse supernovae.
- Individual rotating Neutron stars with an asymmetry. These are *transient sources* which are always active, albeit at much smaller GW amplitude. Current non-detection translates into upper limits as of how elliptic known radio-pulsars can maximally be.
- A stochastic background of gravitational waves. (upper limits are set)

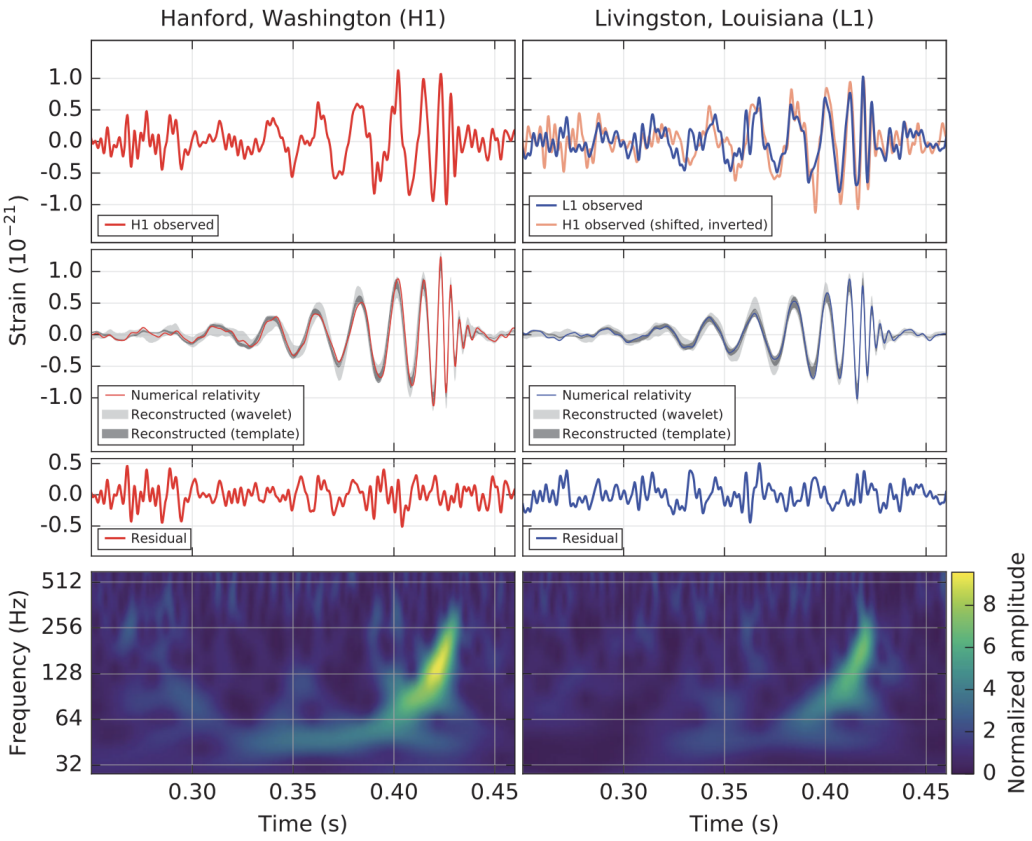


Figure 9.2: GW150914: Measured GW strain, and comparison to numerical solutions of GR. Top: Detector measurement; reconstructed waveforms & comparison with numerical simulations; residuals; bottom: time-frequency plot of excess energy in both detectors.

9.1.3 First GW discoveries

As you all know, LIGO made their first observation of gravitational waves in September, 2015. Two further GW events were observed later in 2015. Figure 9.2 summaries the discovery measurement of Sep 14, 2015. It is a clear chirp-signal, as expected for a coalescing compact object binary. The data are consistent in both detectors, and consistent with direct numerical solutions of GR (see chapter 6).

9.1.4 Parameter estimation

LIGO estimates the parameters of the coalescing binary by comparing the *observed* gravitational waves with the *predicted* gravitational waveforms from post-Newtonian theory combined with numerical simulations.

While not perfect, we can use our intuition from post-Newtonian theory (especially HW3) to gain insights into what parameters LIGO can measure well, and what other parameters not so well. Typically, we will find that there are *correlations* between different measured quantities. Some of the most important measurements of GW150914 are shown in Fig. 9.3. of the degeneracies:

- Masses m_1 and m_2 are reasonably well measured.
- Distance D_L and binary orientation θ_{JN} are *degenerate*. If the binary is face-on or face-off (i.e. $\theta_{JN} \approx 0^\circ$ or $\approx 180^\circ$), then the binary must be at a large distance of $D_L \approx 500$Mpc. If, in contrast, the binary is seen edge-on ($\theta_{JN} \approx 90^\circ$), then the distance is smaller, $D_L \approx 200$Mpc. This particular degeneracy arises because the GW emission is *anisotropic*. Specifically, for a Newtonian binary in the quadrupole-approximation:

$$h_+ = -\frac{1}{2}\,\frac{4\mu r^2\Omega^2}{D_L}\left(1+\cos^2\theta_{JN}\right)\cos 2\Omega t \tag{9.1}$$

$$h_\times = -\frac{4\mu r^2\Omega^2}{D_L}\cos\theta_{JN}\,\sin 2\Omega t. \tag{9.2}$$

$$\tag{9.3}$$

  The important factors are the $\cos\theta_{JN}$ terms, that encode the directional dependence of the emitted gravitational waves. h_+ is diminished by a factor of 2 for GW-emission parallel to the orbital plane ($\theta_{JN} = 90^\circ$), whereas $h_\times$ entirely vanishes in this case.[1] By the way, the sign-change of $h_\times$ incdicates that face-off binaries with $\Omega_{JN} > 90^\circ$ are circularly polarized with the *opposite sense of rotation* than face-on binaries.
- Finally, the bottom panel of Fig. 9.3 summarizes the spin results. As you have shown in HW3, spins enter the post-Newtonian inspiral waveforms at comparably high order. And the first combination of spins that enters is a mass-weighted sum of the spin-components *aligned* with the orbital angular momentum. There are a few

[1] The $\sin(2\Omega t)$, $\cos(2\Omega t)$ terms merely encode the oscillatory nature of gravitational waves during the inspiral.

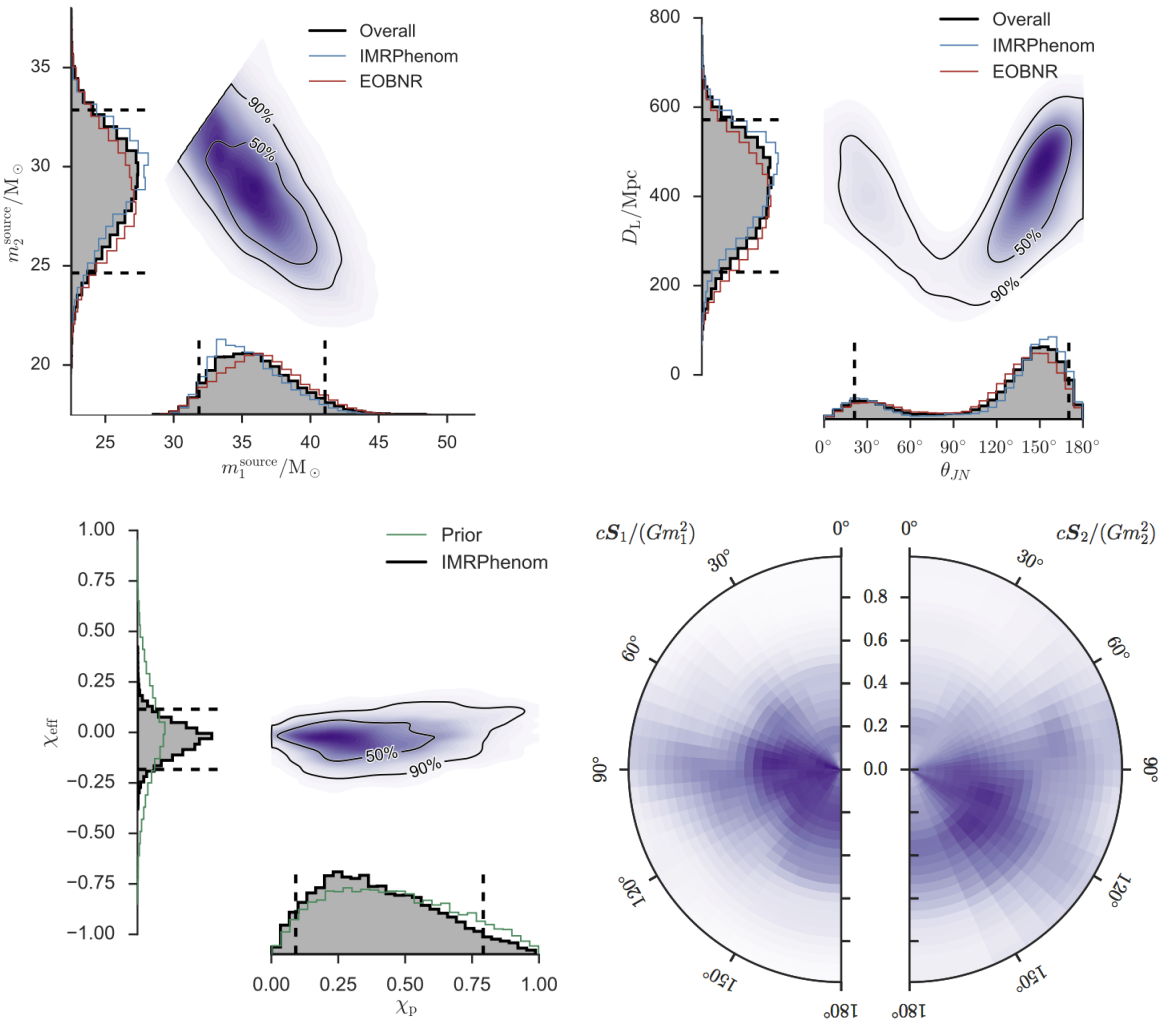


Figure 9.3: Parameter-estimation results for GW150914. Each of the four panels shows in blue the probability density contours in 2-dimensional subspaces of all parameters that characterize the binary. Top-Left: component masses m_1, m_2. Top-right: Inclination θ_{JN} of the orbital plane relative to the line-of-sight to the binary and distance D_L to the binary. Bottom left: Two spin parameters, one characterizing the spin-components parallel to the orbital angular momentum ($\chi_{\rm eff}$), and one characterizing the spin-components orthogonal to the orbital angular momentum (χ_p). Bottom-right: Probability density of the spin-direction of both BH spins. (Figure from `https://arxiv.org/abs/1602.03840`)

ever so slightly different variants of this mass-weighted sum that appear at different places in post-Newtonian theory. Fig. 9.3 considers

$$\chi_{\rm eff} \equiv \frac{c}{G}\left(\frac{\vec{S}_1}{m_1 m} + \frac{\vec{S}_2}{m_2 m}\right)\cdot \hat{L}. \tag{9.4}$$

As is apparent from the bottom left panel of Fig. 9.3, this combination of spins is measured to quite good accuracy: $\chi_{\rm eff} = -0.06 \pm 0.2$. Hence we know that the sum of the spin-components paralllel to L is near zero, but we have little information about the individual spins!

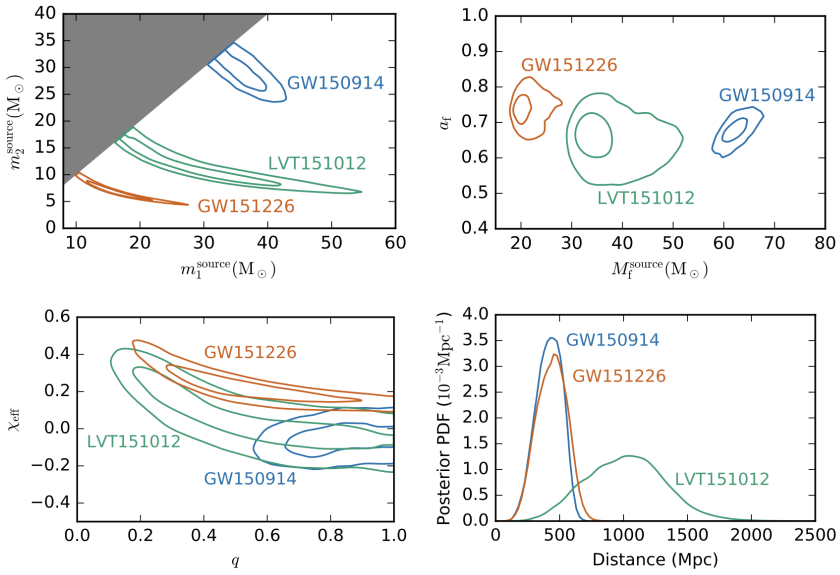


Figure 9.4: Parameter-estimation of the BBH systems observed in the first Advanced LIGO Science Run (O1). (Figure from `https://arxiv.org/abs/1606.04856`)

Figure 9.4, finally, summarizes the three BBH detections of the first observing run of Advanced LIGO, which ran from September 2015 to January 2016. We see that the three detections are at different masses, with total mass ranging from 20 to 65 solar masses. As the total mass becomes lower, ever more inspiral cycles are in band. In the inspiral the waveforms are dominated by the value of the chirp mass,

$$M_c = \frac{(m_1 m_2)^{3/5}}{m^{1/5}}, \tag{9.5}$$

as you have shown in HW3, with other parameters (mass-ratio, spins) less important. Therefore, at low total mass, the $m_1 - m_2$ contours become thinner and more elongated along a M_c =const line, and the mass-ratio is less well measured.

LVT151012 is a weaker signal arriving from larger distance. It was so weak that it cannot be confidently identified as a true gravitational wave. Rather, it only has a $\sim 80\%$ probability of being a real signal, and 20% probability of being a noise-artefact. Therefore,

it is named 'LVT' for LIGO-Virgo-trigger, rather than with a 'GW' moniker that indicates a confident detection.

The lower signal-to-noise ratio of LVT151012 leads to the larger confidence countours in Fig. 9.4.

9.1.5 Testing GR

The first GW observations of coalescing BBH also allow to test GR in the genuinely strong-field, highly dynamic regime. There are three types of tests in particular, which are radically different from tests of GR available before.

- First, one can simply compare the measurement with the prediction of GR; as the third row of panels of Fig. 9.2 demonstrates, the residual of such a comparsion is consistent with noise. This is also confirmed with a more sophisticated analysis than a mere plot.

- BBH systems allow to measure the BH masses and spins of the inspiraling black holes during the inspiral. Furthermore, because GW150914 was so high total mass, the ringdown had significant signal-to-noise ratio. Therefore, one could also measure the quasi-normal-mode ringdown frequency and decay-rate (ω_{220} and τ_{220} in the language of chapter 7). If GR holds, then GR predicts the ringdown properties from the inspiral properties. However, to reach the ringdown, the system has to pass through the genuinely non-linear, asymmetric, and dynamic merger regime. If the evolution in this regime differs from GR, then the remnant properties will almost certainly be different than those predicted by GR. The left panel of Fig. 9.3 demonstrates that the ringdown QNM frequencies are consistent with those expected if the entire signal is perfectly modeled by GR.

- Speed of propagation of gravitational waves. General relativity predicts that gravitational waves propagate with the speed of light c, irrespective of their frequency. This is codified by the dispersion relation

$$E^2 = p^2 c^2 \tag{9.6}$$

for gravitational waves. If the graviton had a non-zero mass $m_g > 0$, then one would expect a dispersion relation like

$$E^2 = p^2 c^2 + m_g^2 c^4. \tag{9.7}$$

The speed of propagation of a gravitational wave is $v_g = pc/E$, so that

$$\left(\frac{v_g}{c}\right)^2 = \frac{p^2 c^2}{E^2} \approx 1 - \frac{m_g^2 c^4}{p^2 c^2} = 1 - \frac{h^2 c^2}{\lambda_g^2 E^2}, \tag{9.8}$$

where in the second equality we have Taylor-expanded the denominator, and in the last equality we substituted in the Compton-wavelength of the graviton,

$$\lambda_g = \frac{h}{m_g c}. \tag{9.9}$$

Equation (9.8) indicates that high-frequency (large energy E) gravitational waves would propagate *faster* than low-frequency waves if the graviton-mass is non-zero. Therefore, the gravitational wave signal reaching Earth (as measured in the top-panels of Fig. 9.2 would be distorted, because the high-frequency merger waveforms would arrive *earlier* on Earth than in pure GR where $v_g = c$ for all frequencies.

The agreement with NR (which uses strictly $v_g = c$) in Fig. 9.2 shows that such dispersive effects must be quite small, with a change in travel-time to Earth $\delta T \ll 0.01$s. However, the travel-time is immense, $T = D_L/c \approx 1.3 \times 10^9yr\sim 4 \times 10^{16}$s. The combination of short duration of the GW signal (i.e. very small δT) and the immense distance to the source (i.e. very large T), therefore immediately bounds the propagation speed

$$\left|\frac{v_g}{c} - 1\right| \lesssim \frac{\delta T}{T} \sim 10^{-19}. \tag{9.10}$$

The right panel of Fig.) 9.5 performs a more careful analysis in terms of the Compton wavelength λ_g.

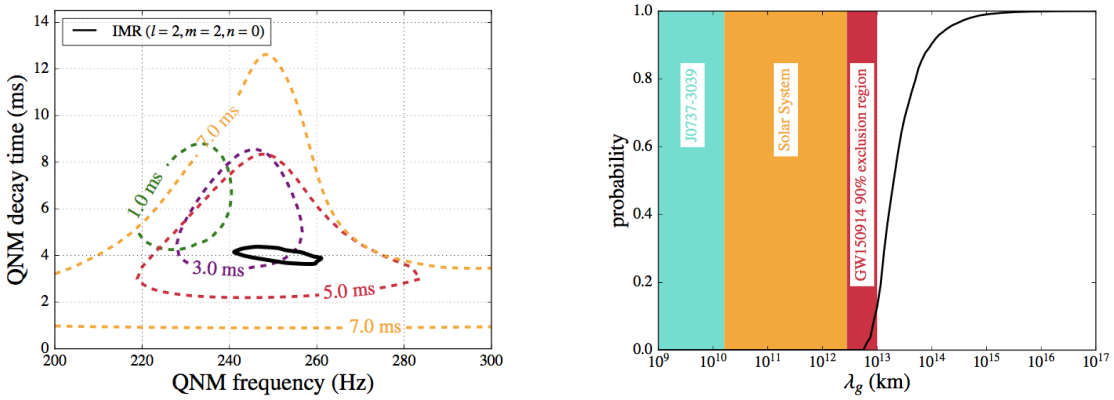


Figure 9.5: Tests of General Relativity using GW150914. **Left:** measured from the ringdown signal starting a certain time-offset after the peak of the waveform. The thick black curve is the measurement if the entire inspiral-merger-ringdown waveform is utilized. **Right:** Bound on the Componton-wavelength of the graviton λ_g. The colored areas are excluded by pulsar-timing ("J0737-3039"), solar-system measurements, and by GW150914. (Figures from `https://arxiv.org/abs/1602.03841`)

9.2 LISA

The Laser Interferometer Space Antenna is a space-based mission that consists of three satellites on orbits such that the three satellites always form approximately an equilateral triangle. This mission is led by the European Space Agency, for a planned launch around 2030.

A lot of information is available from the LISA website, `ww.elisascience.org`, most notably the mission proposal `https://www.elisascience.org/files/publications/LISA_L3_20170120.pdf`

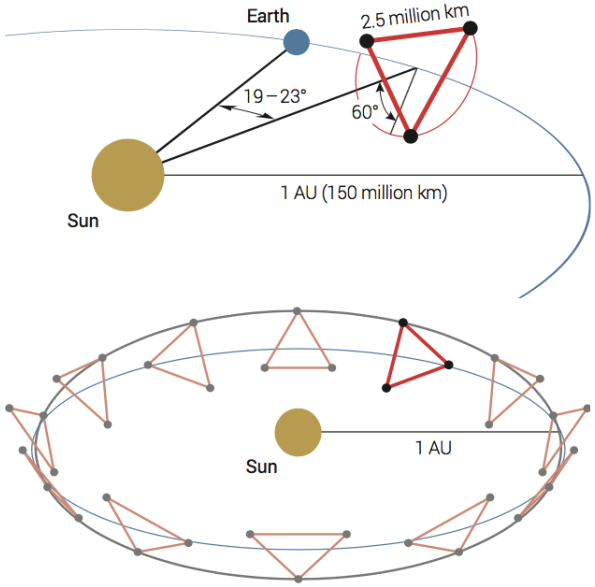


Figure 9.6: Depiction of the LISA orbits: The three space-craft trail Earth and are on orbits that are slightly inclined and slightly eccentric such that the triangular configuration rotates once around its axis as per orbit around the sun. (Figure from `https://www.elisascience.org/files/publications/LISA_L3_20170120.pdf`)

Figure 9.6 depicts the orbital configuration of the three satellites. The three satellites will allow for three interferometers, utilizing the three different combinations of two sides. The arm-length $L = 2.5\times 10^6$km sets the light-travel time, and that in turn sets the typical frequency range for LISA:

$$f \sim \frac{c}{L} \sim 0.1\text{Hz}. \tag{9.11}$$

Gravitational waves above this frequency oscillate multiple times per light-travel time, and so their effect on the travel-time averages out. Therefore, the light-travel time indicates roughly the upper end of LISA's frequency range.

The complete proposed LISA senitivity curve is shown in Fig. 9.7. Also included in this figure are the major science-targets of LISA:

- Supermassive BH-BH binaries with total mass of $10^5 \cdots 10^7 M_\odot$. Depending on their total mass, such systems spend between days and a full year in the LISA frequency band. Supermassive BH-BH mergers happen when the supermassive black holes at the centres of galaxies merge, following the merger of their host galaxies. Tracing the supermassive BH-BH mergers, LISA explores the hierarchical formation of galaxies in the universe.

- Extreme mass ratio inspirals (EMRI), where a stellar mass object (neutron star of black hole) inspirals into a supermassive black hole near the center of a galaxy. Because of the very small symmetric mass-ratio ν, such systems spend $O(1/\nu) \sim 100'000$ GW cycles in the LISA band. The orbits of the small objects in the EMRI

are perturbed by galactic stars in the vicinity of the supermassive blak hole, and so EMRIs are generally on eccentric orbits leading to very complex waveforms which carry detailed information about the strong-field gravity around the supermassive black hole, cf. Fig. 2.5.

- Short period binary stars in the Milky Way, especially binary white dwarfs. Such systems have orbital periods of minutes. For such low-mass systems at such low frequencies, the GW amplitudes are quite small, so the visibility is only the Milky Way. Nevertheless, LISA will detect every such binary in the Milky Way.

- Finally, the breed of stellar mass BH-BH systems recently discovered by LIGO with total masses $\sim 60M_\odot$ will pass through the LISA band some time before they enter the LIGO band. From Eq. (5.89), we find

$$\frac{dx}{x^5} = \frac{64}{5}\frac{c^3\nu}{Gm}dt, \tag{9.12}$$

$$-\frac{1}{4}x^{-4} = \frac{64}{5}\frac{c^3\nu}{Gm}\left(t - t_{\rm merger}\right), \tag{9.13}$$

$$t - t_{\rm merger} = -\frac{5}{256}\frac{Gm}{c^3\nu}\left(\frac{\pi G m f_{\rm GW}}{c^3}\right)^{-8/3}. \tag{9.14}$$

In the second line, we chose the integration constant $t_{\rm merger}$ such that at merger where $x \to \infty$, the time $t \to t_{\rm merger}$. In the third line, we substituted in the definition $x = \left(Gm\Omega/c^3\right)^{2/3}$, and used the fact that the GW-frequency is twice the orbital frequency, i.e. $f_{\rm GW} = \pi\Omega$. Substituting in $m = 60M_\odot$, $\nu = 1/4$ (equal mass), and using a frequency $f_{\rm GW} = 0.02$Hz at the upper end of the LISA band-width, one finds

$$|t - t_{\rm merger}| \approx 2\text{years}. \tag{9.15}$$

That means, a BH-BH merger that is visble in LIGO will start out in the LISA band about two years earlier!

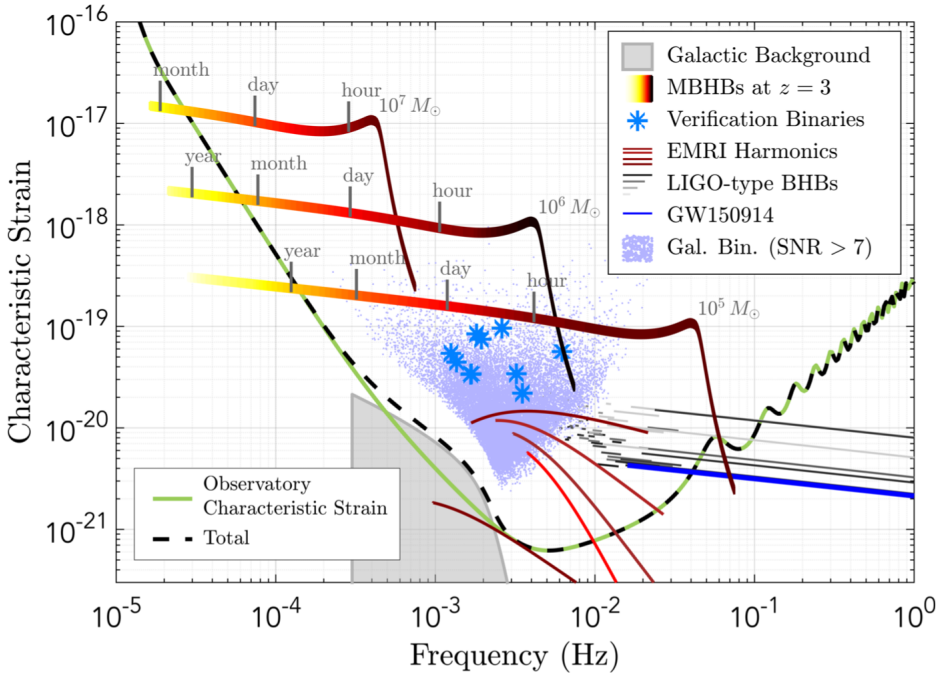


Figure 9.7: Sensitivity-curve of LISA, overlaid with the most important scientific targets. Shown are (1) Supermassive BH-BH binaries (labeled "MBHB") with total mass $10^7 M_\odot, 10^6 M_\odot, 10^5 M_\odot$. (2) The five most important harmonics of an extreme mass ratio inspiral (EMRI) (3) Galactic binaries (primarily WD-WD, but also NS-NS and BH-BH), and (4) LIGO-type BH-BH systems at larger separation than in LIGO. (Figure from `https://www.elisascience.org/files/publications/LISA_L3_20170120.pdf`)