PART 1

GRAVITATIONAL WAVES, SOURCES AND DETECTORS

Bernard F Schutz[1] *and Franco Ricci*[2]

[1] *Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Golm bei Potsdam, Germany and Department of Physics and Astronomy, Cardiff University, Wales*
[2] *University 'La Sapienza', Rome, Italy*

Synopsis

Gravitational waves and their detection are becoming increasingly important both for the theoretical physicist and the astrophysicist. In fact, technological developments have enabled the construction of such sensitive detectors (bars and interferometers) that the detection of gravitational radiation could become a reality during the next few years. In these lectures we give a brief overview of this interesting and challenging field of modern physics.

The topics covered are divided into six lectures. We begin (chapter 2) by describing gravitational waves in linearized general relativity, where one can examine most of the basic properties of gravitational radiation itself; propagation, gauge invariance and interactions with matter (and in particular with detectors).

The second lecture (chapter 3) deals with gravitational-wave detectors: how they operate, what their most important sources of noise are, and what mechanisms are used to overcome noise. We report here on the most important detectors planned or under construction (both ground-based and space-based ones), their likely sensitivity and their prospects for making detections. Other speakers will go into much more detail on specific detectors, such as LISA.

The third lecture (chapter 4) deals with the astrophysics of likely sources of gravitational waves: binary systems, neutron stars, pulsars, x-ray sources, supernovae/hypernovae, γ-ray bursts and the big bang. We estimate the expected wave amplitude h and the suitability of specific detectors for seeing waves from each source.

The fourth lecture (chapter 5) is much more theoretical. Here we develop the mathematical theory of gravitational waves in general, their effective stress-energy tensor, the energy carried by gravitational waves, and the energy in a random wave field (gravitational background generated by the big bang).

The fifth lecture (chapter 6) takes the theory further and examines the generation of gravitational radiation in linearized theory. We show in some detail how both mass-quadrupole and current-quadrupole radiation is generated, including how characteristics of the radiation such as its polarization are related to the motion of the source. Current-quadrupole radiation has become important very recently and may indeed be one of the first forms of gravitational radiation to be detected. We attempt to give a physical description of the way it is generated.

particular we choose units in which $c = G = 1$; Greek indices run from 0 to 3; Latin indices run from 1 to 3; repeated indices are summed; commas in subscripts or superscripts denote partial derivatives; and semicolons denote covariant derivatives. The metric has positive signature. These above two textbooks and others referred to at the end of these chapters give more details on the theory that we outline here. For an even simpler introduction, based on a scalar analogy to general relativity, see [1].

2.1 Mathematics of linearized theory

Consider a perturbed flat spacetime. Its metric tensor can be written as

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}, \quad |h_{\alpha\beta}| \ll 1, \quad \alpha, \beta = 0, \ldots, 3 \tag{2.1}$$

where $\eta_{\alpha\beta}$ is the Minkowski metric $(-1, 1, 1, 1)$ and $h_{\alpha\beta}$ is a very small perturbation of the flat spacetime metric. Linearized theory is an approximation to general relativity that is correct to first order in the size of this perturbation. Since the size of tensor components depends on coordinates, one must be careful with such a definition. What we require for linearized theory to be valid is that there should exist a coordinate system in which equation (2.1) holds in a suitably large region of spacetime. Even though $\eta_{\alpha\beta}$ is not the true metric tensor, we are free to *define* raising and lowering indices of the perturbation with $\eta_{\alpha\beta}$, as if it were a tensor on flat spacetime. We write

$$h^{\alpha\beta} := \eta^{\alpha\gamma}\eta^{\beta\delta}h_{\gamma\delta}.$$

This leads to the following equation for the inverse metric, correct to first order (all we want in linearized theory):

$$g^{\alpha\beta} = \eta^{\alpha\beta} - h^{\alpha\beta}. \tag{2.2}$$

The mathematics is simpler if we define the *trace-reversed* metric perturbation:

$$\bar{h}_{\alpha\beta} := h_{\alpha\beta} - \tfrac{1}{2}\eta_{\alpha\beta}h, \tag{2.3}$$

where $h := \eta_{\alpha\beta}h^{\alpha\beta}$. There is considerable coordinate freedom in the components $h_{\alpha\beta}$, since we can wiggle and stretch the coordinate system with a comparable amplitude and change the components. This coordinate freedom is called *gauge freedom*, by analogy with electromagnetism. We use this freedom to enforce the *Lorentz (or Hilbert)* gauge:

$$\bar{h}^{\alpha\beta}{}_{,\beta} = 0. \tag{2.4}$$

In this gauge the Einstein field equations (neglecting the quadratic and higher terms in $h^{\alpha\beta}$) are just a set of decoupled linear wave equations:

$$\left(-\frac{\partial^2}{\partial t^2} + \nabla^2\right)\bar{h}^{\alpha\beta} = -16\pi T^{\alpha\beta}. \tag{2.5}$$

To understand wave propagation we look for the easiest solution of the vacuum gravitational field equations:

$$\Box\bar{h}^{\alpha\beta} \equiv \left(-\frac{\partial^2}{\partial t^2} + \nabla^2\right)\bar{h}^{\alpha\beta} = 0. \tag{2.6}$$

Plane waves have the form:

$$\bar{h}_{\alpha\beta} = \mathcal{A}\boldsymbol{e}_{\alpha\beta}\exp(ik_\gamma x^\gamma) \tag{2.7}$$

where the amplitude $\mathcal{A}$, polarization tensor $\boldsymbol{e}^{\alpha\beta}$ and wavevector k^γ are all constants. (As usual one has to take the real part of this expression.)

The Einstein equations imply that the wavevector is 'light-like', $k^\gamma k_\gamma = 0$, and the gauge condition implies that the amplitude and the wavevector are orthogonal: $\boldsymbol{e}^{\alpha\beta}k_\beta = 0$.

Linearized theory describes a classical gravitational field whose quantum description would be a massless spin 2 field that propagates at the speed of light. We expect from this that such a field will have only two independent degrees of freedom (helicities in quantum language, polarizations in classical terms). To show this classically we remember that $h_{\alpha\beta}$ is symmetric, so it has ten independent components, and that the Lorentz gauge applies four independent conditions to these, reducing the freedom to six. However, the Lorentz gauge does not fully fix the coordinates. In fact if we perform another infinitesimal coordinate transformation ($x^\mu \to x^\mu + \xi^\mu$ with $\xi^\mu{}_{,\nu} = O(h)$) and impose $\Box\xi^\mu = 0$, we remain in Lorentz gauge. We can use this freedom to demand:

$$\boldsymbol{e}^{0\alpha} = 0 \Longrightarrow \boldsymbol{e}^{ij}k_j = 0 \quad \text{(transverse wave)}, \tag{2.8}$$

$$\boldsymbol{e}^{i}{}_{i} = 0 \quad \text{(traceless wave)}. \tag{2.9}$$

These conditions can only be applied outside a sphere surrounding the source. Together they put the metric into the *transverse-traceless* (TT) gauge. We will explicitly construct this gauge in chapter 5.

2.2 Using the TT gauge to understand gravitational waves

The TT gauge leaves only *two independent polarizations* out of the original ten, and it ensures that $\bar{h}_{\alpha\beta} = h_{\alpha\beta}$. In order to understand the polarization degrees of freedom, let us take the wave to move in the z-direction, so that $k_z = \omega$, $k^0 = \omega$, $k_x = 0$, $k_y = 0$; the TT gauge conditions in equations (2.8) and (2.9) lead to $\boldsymbol{e}^{0\alpha} = \boldsymbol{e}^{z\alpha} = 0$ and $\boldsymbol{e}^{xx} = -\boldsymbol{e}^{yy}$. This leaves only two independent components of the polarization tensor, say $\boldsymbol{e}^{xx}$ and $\boldsymbol{e}^{xy}$ (which we denote by the symbols $\oplus$, $\otimes$).

A wave for which $\boldsymbol{e}^{xy} = 0$ (pure $\oplus$ polarization) produces a metric of the form:

$$ds^2 = -dt^2 + (1 + h_+)\,dx^2 + (1 - h_+)\,dy^2 + dz^2. \tag{2.10}$$

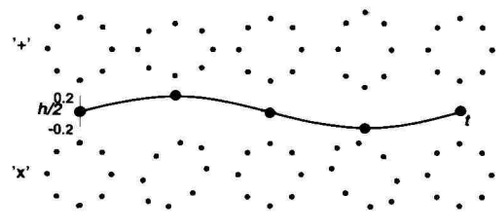

Figure 2.1. Illustration of two linear polarizations and the associated wave amplitude.

where $h_+ = \mathcal{A}e^{xx}\exp[-i\omega(t-z)]$. Such a metric produces opposite effects on proper distance at the two transverse axes, contracting one while expanding the other.

If $e^{xx} = 0$ we have pure $\otimes$ polarization $h_\times$ which can be obtained from the previous case by a simple 45° rotation, as in figure 2.1. Since the wave equation and TT conditions are linear, a general wave will be a linear combination of these two polarization tensors. A circular polarization basis would be:

$$\boldsymbol{e}_{\mathrm{R}} = \frac{1}{\sqrt{2}}(\boldsymbol{e}_+ + i\boldsymbol{e}_\times), \quad \boldsymbol{e}_{\mathrm{L}} = \frac{1}{\sqrt{2}}(\boldsymbol{e}_+ - i\boldsymbol{e}_\times), \tag{2.11}$$

where $\boldsymbol{e}_+$, $\boldsymbol{e}_\times$ are the two linear polarization tensors and $\boldsymbol{e}_{\mathrm{R}}$ and $\boldsymbol{e}_{\mathrm{L}}$ are polarizations that rotate in the right-handed and left-handed directions, respectively. It is important to understand that, for circular polarization, the polarization pattern rotates around the central position, but test particles themselves rotate only in small circles relative to the central position.

Now we compute the effects of a wave in the TT gauge on a particle at rest in the flat background metric $\eta_{\alpha\beta}$ before the passage of the gravitational wave. The geodesic equation

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu{}_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} = 0$$

implies in this case:

$$\frac{\mathrm{d}^2x^i}{\mathrm{d}\tau^2} = -\Gamma^i{}_{00} = -\frac{1}{2}(2h_{i0,0} - h_{00,i}) = 0, \tag{2.12}$$

so that the particle *does not move*. The TT gauge, to first order in $h_{\alpha\beta}$, represents a coordinate system that is comoving with freely-falling particles. Because $h_{0\alpha} = 0$, TT time is proper time on the clock of freely-falling particles at rest.

Tidal forces show the action of the wave independently of the coordinates. Let us consider the equation of geodesic deviation, which governs the separation of two neighbouring freely-falling test particles A and B. If the particles are initially at rest, then as the wave passes it produces an oscillating curvature tensor, and the separation ξ of the two particles is:

$$\frac{\mathrm{d}^2\xi^i}{\mathrm{d}t^2} = R^i{}_{0j0}\xi^j. \tag{2.13}$$

To calculate the component $R^i{}_{0j0}$ of the Riemann tensor in equation (2.13), we can use the metric in the TT gauge, because the Riemann tensor is gauge-invariant at linear order (see exercise (d) at the end of this chapter). Therefore, we can replace $R^i{}_{0j0}$ by $R^i{}_{0j0} = \frac{1}{2}h^{\mathrm{TT}i}{}_{j,00}$ and write:

$$\frac{\mathrm{d}^2\xi^i}{\mathrm{d}t^2} = \frac{1}{2}h^{\mathrm{TT}i}{}_{j,00}\xi^j. \tag{2.14}$$

This equation, with an initial condition $\xi^j_{(0)} = \text{constant}$, describes the oscillations of Bs location as measured in the proper reference frame of A. The validity of equation (2.14) is the same as that of the geodesic deviation equation: geodesics have to be close to one another, in a neighbourhood where the change in curvature is small. In this approximation a gravitational wave is like an extra force, called a *tidal force*, perturbing the proper distance between two test particles. If there are other forces on the particles, so that they are not free, then as long as the gravitational field is weak, one can just add the tidal forces to the other forces and work as if the particle were in special relativity.

2.3 Interaction of gravitational waves with detectors

We have shown above that the TT gauge is a particular coordinate system in which the polarization tensor of a plane gravitational wave assumes a very simple form. This gauge is comoving for freely-falling particles and so it is not the locally Minkowskian coordinate system that would be used by an experimenter to analyse an experiment. In general relativity one must always be aware of how one's coordinate system is defined.

We shall analyse two typical situations:

- the detector is small compared to the wavelength of the gravitational waves it is measuring; and
- the detector is comparable to or larger than that wavelength.

In the first case we can use the geodesic deviation equation above to represent the wave as a simple extra force on the equipment. Bars detectors can always be analysed in this way. Laser interferometers on the Earth can be treated this way too. In these cases a gravitational wave simply produces a force to be measured. There is no more to say from the relativity point of view. The rest of the detection story is the physics of the detectors. Sadly, this is not as simple as gravitational wave physics!

In the second case, the geodesic deviation equation is not useful because we have to abandon the 'local mathematics' of geodesic deviation and return to the 'global mathematics' of the TT gauge and metric components $h^{\mathrm{TT}}{}_{\alpha\beta}$. Space-based interferometers like LISA, accurate ranging to solar-system spacecraft and pulsar timing are all in this class. Together with ground interferometers, these are *beam detectors*: they use light (or radio waves) to register the waves.

To study these detectors, it is easiest to remain in the TT gauge and to calculate the effect of the waves on the (coordinate) speed of light. Let us consider, for example, the $\oplus$ metric from equation (2.10) and examine a null geodesic moving in the x-direction. The speed along this curve is:

$$\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 = \frac{1}{1+h_+}. \tag{2.15}$$

This is only a *coordinate speed*, not a contradiction to special relativity.

To analyse the way in which detectors work, suppose one arm of an interferometer lies along the x-direction and the wave, for simplicity, is moving in the z-direction with a $\oplus$ polarization of *any* waveform $h_+(t)$ along this axis (it is a plane wave, so its waveform does not depend on x). Then a photon emitted at time t from the origin reaches the other end, at a fixed coordinate position $x = L$, at the coordinate time

$$t_{\mathrm{far}} = t + \int_0^L \sqrt{1+h_+(t(x))}\,\mathrm{d}x, \tag{2.16}$$

where the argument $t(x)$ denotes the fact that one must know the time to reach position x in order to calculate the wave field. This implicit equation can be solved in linearized theory by using the fact that h_+ is small, so we can use the first-order solution of equation (2.15) to calculate $h_+(t)$ to sufficient accuracy.

To do this we expand the square root in powers of h_+, and consider as a zero-order solution a photon travelling at the speed of light in the x-direction of a flat spacetime. We can set $t(x) = t + x$. The result is:

$$t_{\mathrm{out}} = t + L + \tfrac{1}{2}\int_0^L h_+(t+x)\,\mathrm{d}x. \tag{2.17}$$

In an interferometer, the light is reflected back, so the return trip takes

$$t_{\mathrm{return}} = t + L + \tfrac{1}{2}\left[\int_0^L h_+(t+x)\,\mathrm{d}x + \int_0^L h_+(t+x+L)\,\mathrm{d}x\right]. \tag{2.18}$$

What one monitors is changes in the time taken by a return trip as a function of time at the origin. If there were no gravitational waves t_{return} would be constant because L is fixed, so changes indicate a gravitational wave.

The rate of variation of the return time as a function of the start time t is

$$\frac{\mathrm{d}t_{\mathrm{return}}}{\mathrm{d}t} = 1 + \frac{1}{2}[h_+(t+2L) - h_+(t)]. \tag{2.19}$$

This depends only on the wave amplitude when the beam leaves and when it returns.

Let us consider now a more realistic geometry than the previous one, and in particular suppose that the wave travels at an angle θ to the z-axis in the x–z plane. If we redo this calculation, allowing the phase of the wave to depend on x in an appropriate way, and taking into account the fact that $h_+^{\mathrm{TT}xx}$ is reduced if the wave is not moving in a direction perpendicular to x, we find (see exercise (a) at the end of this chapter for the details of the calculation)

$$\begin{aligned}\frac{\mathrm{d}t_{\mathrm{return}}}{\mathrm{d}t} = \frac{1}{2}\{&(1-\sin\theta)h_+^{xx}(t+2L) - (1+\sin\theta)h_+^{xx}(t)\\ &+ 2\sin\theta h_+^{xx}[t + L(1-\sin\theta)]\}.\end{aligned} \tag{2.20}$$

This three-term relation is the starting point for analysing the response of all beam detectors. This is directly what happens in radar ranging or in transponding to spacecraft, where a beam in only one direction is used. In long-baseline interferometry, one must analyse the second beam as well. We shall discuss these cases in turn.

2.4 Analysis of beam detectors

2.4.1 Ranging to spacecraft

Both NASA and ESA perform experiments in which they monitor the return time of communication signals with interplanetary spacecraft for the characteristic effect of gravitational waves. For missions to Jupiter and Saturn, the return times are of the order 2–4 $\times$ 10^3 s. Any gravitational wave event shorter than this will leave an imprint on the delay time three times: once when the wave passes the Earth-based transmitter, once when it passes the spacecraft, and once when it passes the Earth-based receiver. Searches use a form of pattern matching to look for this characteristic imprint. There are two dominant sources of noise: propagation-time irregularities caused by fluctuations in the solar wind plasma, and timing noise in the clocks used to measure the signals. The plasma delays depend on the radio-wave frequency, so by using two transmission frequencies one can model and subtract the plasma noise. Then if one uses the most stable atomic clocks, it is possible to achieve sensitivities for h of the order 10^{-13}. In the future, using higher radio frequencies, such experiments may reach 10^{-15}. No positive detections have yet been made, but the chances are not zero. For example, if a small black hole fell into a massive black hole in the centre of the Galaxy, it would produce a signal with a frequency of about 10 mHz and an amplitude significantly bigger than 10^{-15}. Rare as this might be, it would be a dramatic event to observe.

2.4.2 Pulsar timing

Many pulsars, in particular old millisecond pulsars, are extraordinarily regular clocks, whose random timing irregularities are too small for even the best atomic clocks to measure. Other pulsars have weak but observable irregularities. Measurements of or even upper limits on any of these timing irregularities for single pulsars can be used to set *upper limits* on any background gravitational wave field with periods comparable to or shorter than the observing time. Here the three-term formula is replaced by a simpler two-term expression (see exercise (b) at the end of this chapter), because we only have a one-way transmission from the pulsar to Earth. Moreover, the transit time of a signal to Earth from the pulsar may be thousands of years, so we cannot look for correlations between the two terms in a given signal. Instead, the delay time is a combination of the effects of uncorrelated waves at the pulsar when the signal was emitted and at the Earth when it is received.

If one simultaneously observes two or more pulsars, the Earth-based part of the delay is correlated between them, and this offers a means of actually detecting long-period gravitational waves. Observations require a timescale of several years in order to achieve the long-period stability of pulse arrival times, so this method is suited to looking for strong gravitational waves with periods of several years.

2.4.3 Interferometry

An interferometer essentially measures changes in the difference in the return times along two different arms. It does this by looking for changes in the interference pattern formed when the returning light beams are superimposed on one another. The response of each arm will follow the three-term formula in equation (2.20), but with a different value of θ for each arm, depending in a complicated way on the orientation of the arms relative to the direction of travel and the polarization of the wave. Ground-based interferometers are small enough to use the small-L formulae we derived earlier. However, LISA, the space-based interferometer that is described by Bender in this book, is larger than a wavelength of gravitational waves for frequencies above 10 mHz, so a detailed analysis of its sensitivity requires the full three-term formula.

2.5 Exercises for chapter 2

Suggested solutions for these exercises are at the end of chapter 7.

(a) 1. *Derive the full three-term return equation, reproduced here:*

$$\frac{\mathrm{d}t_{\mathrm{return}}}{\mathrm{d}t} = \frac{1}{2}\{(1-\sin\theta)h_{+}^{TT}(t+2L) - (1+\sin\theta)h_{+}^{TT}(t) + 2\sin\theta h_{+}^{TT}[t+L(1-\sin\theta)]\}. \tag{2.21}$$

2. *Show that, in the limit where L is small compared to the wavelength of the gravitational wave, the derivative of the return time is the derivative of the excess proper distance $\delta L = Lh_{+}^{xx}(t)\cos^2\theta$ for small L. Make sure you know how to interpret the factor of $\cos^2\theta$.*
3. *Examine the limit of the three-term formula when the gravitational wave is travelling along the x-axis too ($\theta = \pm\frac{\pi}{2}$): what happens to light going parallel to a gravitational wave?*

(b) *Derive the two-term formula governing the delays induced by gravitational waves on a signal transmitted only one-way, for example from a pulsar to Earth.*

(c) *A frequently asked question is: if gravitational waves alter the speed of light, as we seem to have used here, and if they move the ends of an interferometer closer and further apart, might these effects not cancel, so that there would be no measurable effects on light? Answer this question. You may want to examine the calculation above: did we make use of the changing distance between the ends, and why or why not?*

(d) *Show that the Riemann tensor is gauge-invariant in linearized theory.*

Chapter 3

Gravitational-wave detectors

Gravitational radiation is a central prediction of general relativity and its detection is a key test of the integrity of the theoretical structure of Einstein's work. However, in the long run, its importance as a tool for observational astronomy is likely to be even more important. We have excellent observational evidence from the Hulse–Taylor binary pulsar system (described in chapter 4) that the predictions of general relativity concerning gravitational radiation are quantitatively correct. However, we have incomplete information from astronomy today about the likely sources of detectable radiation.

The gravitational wave spectrum is completely unexplored, and whenever a new electromagnetic waveband has been opened to astronomy, astronomers have discovered completely unexpected phenomena. This seems to me just as likely to happen again with gravitational waves, especially because gravitational waves carry some kinds of information that electromagnetic radiation cannot convey. Gravitational waves are generated by bulk motions of masses, and they encode the mass distributions and speeds. They are coherent and their low frequencies reflect the dynamical timescales of their sources.

In contrast, electromagnetic waves come from individual electrons executing complex and partly random motions inside their sources. They are incoherent, and individual photons must be interpreted as samples of the large statistical ensemble of photons being emitted. Their frequencies are determined by microphysics on length scales much smaller than the structure of the astronomical system emitting them. From electromagnetic observations we can make inferences about this structure only through careful modelling of the source. Gravitational waves, by contrast, carry information whose connection to the source structure and motion is fairly direct.

A good example is that of massive black holes in galactic nuclei. From observations that span the electromagnetic spectrum from radio waves to x-rays, astrophysicists have inferred that black holes of masses up to $10^9 M_{\odot}$ are responsible for quasar emissions and control the jets that power the giant radio emission regions. The evidence for the black hole is very strong but

indirect: no other known object can contain so much mass in such a small volume. Gravitational wave observations will tell us about the dynamics of the holes themselves, providing unique signatures from which they can be identified, measuring their masses and spins directly from their vibrational frequencies. The interplay of electromagnetic and gravitational observations will enrich many branches of astronomy.

The history of gravitational-wave detection started in the 1960s with J Weber at the University of Maryland. He built the first *bar detector*: it was a massive cylinder of aluminium ($\sim 2 \times 10^3$ kg) operating at room temperature (300 K) with a resonant frequency of about 1600 Hz. This early prototype had a modest sensitivity, around 10^{-13} or 10^{-14}.

Despite this poor sensitivity, in the late 1960s Weber announced the detection of a population of coincident events between two similar bars at a rate far higher than expected from instrumental noise. This news stimulated a number of other groups (at Glasgow, Munich, Paris, Rome, Bell Laboratories, Stanford, Rochester, LSU, MIT, Beijing, Tokyo) to build and develop bar detectors to check Weber's results. Unfortunately for Weber and for the idea that gravitational waves were easy to detect, none of these other detectors found anything, even at times when Weber continued to find coincidences. Weber's observations remain unexplained even today. However, the failure to confirm Weber was in a real sense a confirmation of general relativity, because theoretical calculations had never predicted that reasonable signals would be strong enough to be seen by Weber's bars.

Weber's announcements have had a mixed effect on gravitational-wave research. On the one hand, they have created a cloud under which the field has laboured hard to re-establish its respectability in the eyes of many physicists. Even today the legacy of this is an extreme cautiousness among the major projects, a conservatism that will ensure that the next claim of a detection will be ironclad. On the other hand, the stimulus that Weber gave to other groups to build detectors has directly led to the present advanced state of detector development.

From 1980 to 1994 groups developed detectors in two different directions:

- *Cryogenic bar detectors*, developed primarily at Rome/Frascati, Stanford, LSU and Perth (Australia). The best of these detectors reach below 10^{-19}. They are the only detectors operating continuously today and they have performed a number of joint coincidence searches, leading to upper limits but no detections.
- *Interferometers*, developed at MIT, Garching (where the Munich group moved), Glasgow, Caltech and Tokyo. The typical sensitivity of these prototypes was 10^{-18}. The first long coincidence observation with interferometers was the Glasgow/Garching 100 hr experiment in 1989 [2].

In fact, interferometers had apparently been considered by Weber, but at that time the technology was not good enough for this kind of detector. Only 10–15 years later, technology had progressed. Lasers, mirror coating and polishing

techniques and materials science had advanced far enough to allow the first practical interferometers, and it was clear that further progress would continue unabated. Soon afterwards several major collaborations were formed to build large-scale interferometric detectors:

- LIGO: Caltech and MIT (NSF) LIGO;
- VIRGO: France (CNRS) and Italy (INFN)
- GEO600: Germany (Max Planck) and UK (PPARC).

Later, other collaborations were formed in Australia (AIGO) and Japan (TAMA and JGWO). At present there is still considerable effort in building successors to Weber's original resonant-mass detector: ultra-cryogenic bars are in operation in Frascati and Padua, and they are expected to reach below 10^{-20}. Further, there are proposals for a new generation of spherical or icosahedral solid-mass detectors from the USA (LSU), Brazil, the Netherlands and Italy. Arrays of smaller bars have been proposed for observing the highest frequencies, where neutron star normal modes lie.

However, the real goal for the near future is to break through the 10^{-21} level, which is where theory predicts that it is not unreasonable to expect gravitational waves of the order of once per year (see the discussion in chapter 4 later). The first detectors to reach this level will be the large-scale interferometers that are now under construction. They have very long arms: LIGO, Hanford (WA) and Livingstone (LA), 4 km; VIRGO: Pisa, 3 km; GEO600: Hannover, 600 m; TAMA300: Tokyo, 300 m.

The most spectacular detector in the near future is the space-based detector LISA, which has been adopted by ESA (European Space Agency) as a Cornerstone mission for the twenty-first century. The project is now gaining a considerable amount of momentum in the USA, and a collaboration between ESA and NASA seems likely. This mission could be launched around 2010.

3.1 Gravitational-wave observables

We have described earlier how different gravitational-wave observables are from electromagnetic observables. Here are the things that we want to measure when we detect gravitational waves:

- $h_+(t)$, $h_\times(t)$, $phase(t)$: the amplitude and polarization of the wave, and the phase of polarization, as functions of time. These contain most of the information about gravitational waves.
- θ, ϕ: the direction on the sky of the source (except for observations of a stochastic background).

From this it is clear that gravitational-wave detection is not the same as electromagnetic-radiation detection. In electromagnetic astronomy one almost always rectifies the electromagnetic wave, while we can follow the oscillations of

the gravitational wave. Essentially in electromagnetism one detects the power in the radiation, while for gravitational radiation, as we have said before, one detects the wave coherently.

Let us consider now what we can infer from a detection. If the gravitational wave has a short duration, of the order of the sampling time of the signal stream, then each detector will usually give just a single number, which is the amplitude of the wave projected on the detector (a projection of the two polarizations h_+ and $h_\times$). If the wave lasts more than one sampling time, then this information is a function of time.

If the signal lasts for a sufficiently long time, then both the amplitude and the phase of the wave can be affected by the motion of the detector, which moves and turns with the motion of the Earth. This produces an amplitude and phase modulation which is not intrinsic to the signal. If the signal's intrinsic form is understood, then this modulation can be used to determine the location of the source. We distinguish three distinct kinds of signals, from the point of view of observations.

Bursts have a duration so short that modulation due to detector motion is not observable. During the detection, the detector is effectively stationary. In this case we need at least three, and preferably four, interferometers to triangulate the positions of bursts on the sky and to find the two polarizations h_+ and $h_\times$. (See discussions in Schutz 1989.) A network of detectors is essential to extract all the information in this case.

Continuous waves by definition last long enough for the motion of the detector to induce amplitude and phase modulation. In this case, assuming a simple model for the intrinsic signal, we can use the information imprinted on the signal (the amplitude modulation and phase modulation) to infer the position and polarization amplitude of the source on the sky. A single detector, effectively, performs aperture synthesis, finding the position of the source and the amplitude of the wave entirely by itself. However, in order to be sure that the signal is not an artefact, it will be important that the signal is seen by a second or third detector.

Stochastic backgrounds can be detected just like noise in a single detector. If the detector noise is well understood, this excess noise may be detected as a stochastic background. This is closely analogous to the way the original microwave background detection was discovered.

A more reliable method for detecting stochastic radiation is the cross-correlation between two detectors, which experience the same cosmological noise but have a different intrinsic noise. Coherent cross-correlation between two detectors eliminates much detector noise and works best when detectors are closer than a wavelength.

In general, detection of gravitational waves requires joint observing by a network of detectors, both to increase the confidence of the detection and to provide accurate information on other physical observables (direction, amplitude and so on). Networks can be assembled from interferometers, bars, or both.

3.2 The physics of interferometers

Interferometric gravitational-wave detectors are the most sensitive instruments, and among the most complex, that have ever been constructed. They are remarkable for the range of physics that is important for their construction. Interferometer groups work at the forefront of the development in lasers, mirror polishing and coating, quantum measurement, materials science, mechanical isolation, optical system design and thermal science. In this section we shall only be able to take a fairly superficial look at one of the most fascinating instrumentation stories of our age. A good introduction to interferometer design is Saulson (1994).

Interferometers use laser light to compare the lengths of two perpendicular arms. The simplest design, originated by Michelson for his famous experiment on the velocity of light, uses light that passes up and down each arm once, as in the first panel in figure 3.1. Imagine such an instrument with identical arms defined by mirrors that hang from supports, so they are free to move horizontally in response to a gravitational wave. If there is no wave, the arms have the same length, and the light from one arm returns exactly in phase with that from the other. When the wave arrives, the two arms typically respond differently. The arms are no longer the same length, and so the light that arrives back at the centre from one arm will no longer be in phase with that arriving back from the other arm. This will produce a shift in the interference fringes between the two beams. This is the principle of detection.

Real detectors are designed to store the light in each arm for longer than just one reflection (see figure 3.1(*b*)). It is optimum to store the light for half of the period of the gravitational wave, so that on each reflection the light gains an added phase shift. Michelson-type *delay-line* interferometers store the light by arranging multiple reflections. *Fabry–Perot* interferometers store the light in cavities in each arm, allowing only a small fraction to escape for the interference measurement (figure 3.1(*e*)).

An advantage of interferometers as detectors is that the gravitational-wave-induced phase shift of the light can be made larger simply by making the arm length larger, since gravitational waves act by tidal forces. A detector with an arm length $l = 4$ km responds to a gravitational wave with an amplitude of 10^{-21} with

$$\delta l_{\rm gw} \sim \tfrac{1}{2} h l \sim 2 \times 10^{-18} \ \text{m} \qquad (3.1)$$

where $\delta l_{\rm gw}$ is the change in the length of one arm. If the orientation of the interferometer is optimum, then the other arm will change by the same amount in the opposite direction, so that the interference fringe will shift by twice this length.

If the light path is folded or resonated, as in figure 3.1(*b*) and (*d*), then the effective number of bounces can be traded off against overall length to achieve a given desired total path length, or storage time. Shorter interferometers with many bounces have a disadvantage, however: even though they can achieve the same response as a longer interferometer, the extra bounces introduce noise from the mirrors, as discussed below. There is, therefore, a big advantage to long-arm interferometers.

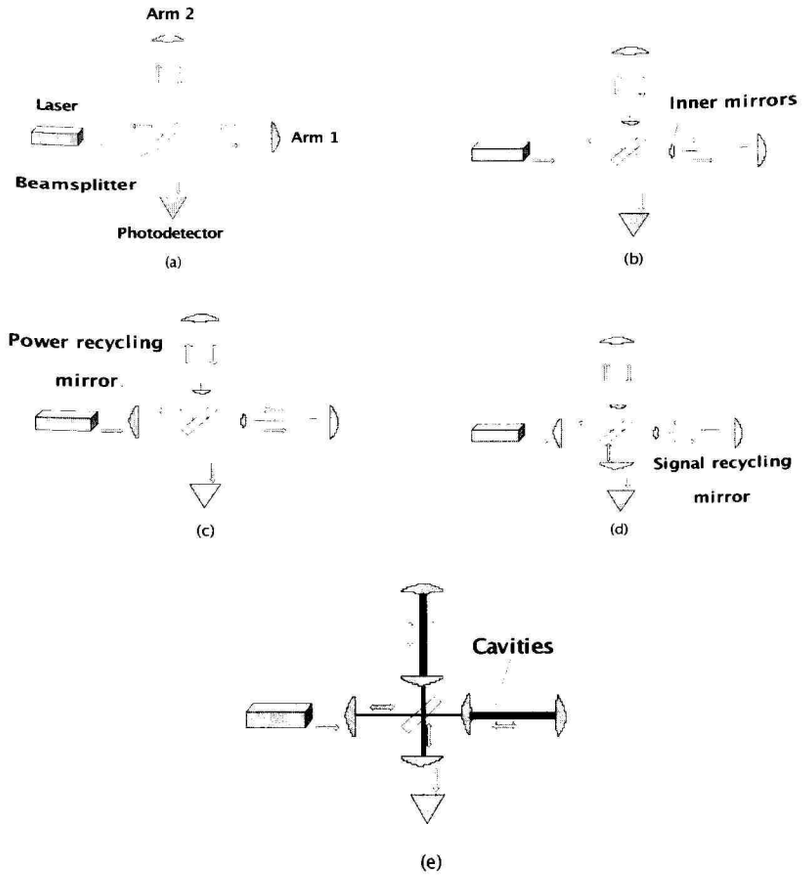

Figure 3.1. Five steps to a gravitational-wave interferometer. (*a*) The simple Michelson. Notice that there are two return beams: one goes toward the photodetector and the other toward the laser. (*b*) Delay line: a Michelson with multiple bounces in each arm to enhance the signal. (*c*) Power recycling. The extra mirror recycles the light that goes towards the laser, which would otherwise be wasted. (*d*) Signal recycling. The mirror in front of the photodetector recycles only the signal sidebands, provided that in the absence of a signal no light goes to the photodetector. (*e*) Fabry–Perot interferometer. The delay lines are converted to cavities with partially silvered interior mirrors.

There are three main sources of noise in interferometers: thermal, shot and vibrational. To understand the way they are controlled, it is important to think in frequency space. Observations with ground-based detectors will be made in a range from perhaps 10 Hz up to 10 kHz, and initial detectors will have a much smaller observing bandwidth within this. Disturbances by noise that occur at frequencies outside the observation band can simply be filtered out. The goal of noise control is to reduce disturbances in the observation band.

- *Thermal noise*. Interferometers work at room temperature, and vibrations of the mirrors and of the suspending pendulum can mask gravitational waves. To control this noise, scientists take advantage of the fact that thermal noise has its maximum amplitude at the frequency of the vibrational mode, and if the resonance of the mode is narrow (a high quality factor Q) then the amplitude at other frequencies is small. Therefore, pendulum suspensions are designed with the pendulum frequency at about 1 Hz, well below the observing window, and mirror masses are designed to have principal vibration modes above 1 kHz, well above the optimum observing frequency for initial interferometers. These systems are constructed with high values of Q (10^6 or more) to reduce the noise in the observing band. Even so, thermal noise is typically a dominant noise below 100 or 200 Hz.
- *Shot noise*. This is the principal limitation to sensitivity at higher frequencies, above 200–300 Hz. It arises from the quantization of photons. When photons form interference fringes, they arrive at random times and make random fluctuations in the light intensity that can look like a gravitational wave signal; the more photons one uses, the smoother will be the interference fringe. We can easily calculate this intrinsic noise. If N is the number of photons emitted by the laser during our measurement, then as a random process the fluctuation number δN is proportional to the square root of N. If we are using light with a wavelength λ (for example infrared light with $\lambda \sim 1\ \mu$m) one can expect to measure lengths to an accuracy of

$$\delta l_{\text{shot}} \sim \frac{\lambda}{2\pi\sqrt{N}}.$$

To measure a gravitational wave at a frequency f, one has to make at least $2f$ measurements per second, so one can accumulate photons for a time $1/2f$. If P is the light power, one has

$$N = \frac{P}{\frac{hc}{\lambda} \cdot \frac{1}{2f}}.$$

It is easy to work out from this that, for δl_{shot} to be equal to δl_{gw} in equation (3.1), one needs light power of about 600 kW. No continuous laser could provide this much light to an interferometer.

The key to reaching such power levels inside the arms of a detector is a technique called power recycling (see Saulson 1994) first proposed by Drever and independently by Schilling. Normally, interferometers work on a 'dark fringe', that is they are arranged so that the light reaching the photodetector is zero if there is no gravitational wave. Then, as shown in figure 3.1(a), the whole of the input light must emerge from the interferometer travelling towards the laser. If one places another mirror, correctly positioned, between the laser and the beam splitter (figure 3.1(c)), it will reflect this wasted light back into the interferometer in such a way that it adds coherently in phase with light emerging from the laser. In this way, light can be recycled and the required power levels in the arms achieved.

Of course, there will be a maximum recycling gain, which is set by mirror losses. Light power builds up until the laser merely re-supplies the losses at the mirrors, due to scattering and absorption. The maximum power gain is

$$P = \frac{1}{1 - R^2}$$

where $1 - R^2$ is the total loss summed over all the optical surfaces. For the very high-quality mirrors used in these projects, $1 - R^2 \sim 10^{-5}$. This reduces the power requirement for the laser by the same factor, down to about 6 W. This is attainable with modern laser technology.

- *Ground vibration* and *mechanical vibrations* are another source of noise that must be screened out. Typical seismic vibration spectra fall sharply with frequency, so this is a problem primarily below 100 Hz. Pendulum suspensions are excellent mechanical filters above the pendulum frequency: it is a familiar elementary-physics demonstration that one can wiggle the suspension point of a pendulum vigorously at a high frequency and the pendulum itself remains undisturbed. Suspension designs typically involve multiple pendula, each with a frequency around 1 Hz. These provide very fat roll-off of the noise above 1 Hz. Interferometer spectra normally show a steep low-frequency noise 'wall': this is the expected vibrational noise amplitude.

In addition, there are noise sources that are not dominant in the present interferometers but will become important as sensitivity increases.

- *Quantum effects: uncertainty principle noise*. Shot noise is a quantum noise, but in addition there are other effects similar to those that bar detectors face, as described below: zero-point vibrations of suspensions and mirror surfaces, and back-action of light pressure fluctuations on the mirrors. These are small compared to present operating limits of detectors, but they may become important in five years or so. Practical schemes to reduce this noise have already been demonstrated in principle, but they need to be improved considerably. This is the subject of considerable theoretical work at the moment.
- *Gravity gradient noise*. Gravitational-wave detectors respond to any changes in the gradients (tidal forces) of the local gravitational field, not just

those carried by waves. The environment always contains changes in the Newtonian fields of nearby objects. Besides obvious ones, like people, there are changes caused by density waves in ground vibrations, atmospheric pressure changes, and many other disturbances. Below about 1 Hz, these gravity gradient changes will be stronger than waves expected from astronomical objects, and they make it impossible to do observing at low frequencies from Earth. This is the reason that scientists have proposed the LISA mission, discussed later. Above 1 Hz, this noise does not affect the sensitivity of present detectors, but in ten years this could become a limiting factor.

Besides these noise sources, which are predictable and therefore can be controlled by detector design, it is possible that there will be unexpected or unpredicted noise sources. Interferometers will be instrumented with many kinds of environmental monitors, but there may occasionally be noise that is impossible to identify. For this reason, short bursts of gravitational radiation must be identified at two or more separated facilities. Even if detector noise is not at all understood, it is relatively easy to estimate from the observed noise profile of the individual detectors what the chances are of a coincident noise event between two detectors.

3.2.1 New interferometers and their capabilities

Interferometers work over a broad bandwidth and they do not have any natural resonance in their observing band. They are ideal for detecting *bursts*, since one can perform pattern-matching over the whole bandwidth and detect such signals optimally. They are also ideal for searching for unknown *continuous signals*, such as surveying the sky for neutron stars. And in observations of *stochastic signals* by cross-correlating two detectors, they can give information about the spectrum of the signal.

If an interferometer wants to study a signal with a known frequency, such as known pulsars, then there is another optical technique available to enhance its sensitivity in a narrow bandwidth, at the expense of sensitivity outside that band. This is called *signal recycling* [3]. In this technique, a further mirror is placed in front of the photodetector, where the signal emerges from the interferometer (see figure 3.1(*d*)). If the mirror is chosen correctly, it will build up the signal, but only in a certain bandwidth. This modifies the shot noise in the detector, but not other noise sources. Therefore, it can improve sensitivity only at the higher frequencies where shot noise is the limiting factor.

Four major interferometer projects are now under construction, and they could begin acquiring good data in the period between 2000–2003. They will all operate initially with a sensitivity approaching 10^{-21} over a bandwidth between 50–1000 Hz. Early detections are by no means certain, but recent work has made prospects look better for an early detection than when these detectors were funded.

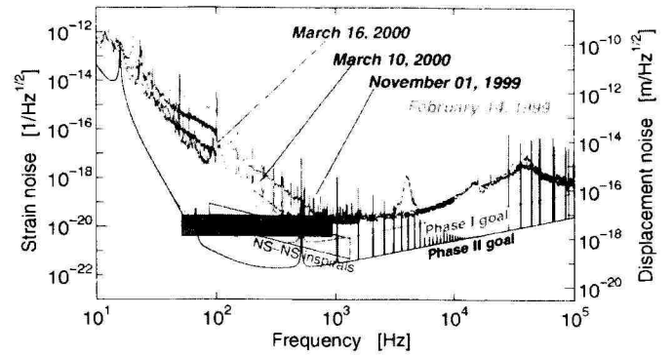

Figure 3.2. TAMA300 sensitivity as a function of frequency. The vertical axis is the 1σ noise level, measured in strain per root Hz. To get a limit on the gravitational wave amplitude h, one must multiply the height of the curve by the square root of the bandwidth of the signal. This takes into account the fact that the noise power at different frequencies is independent, so the power is proportional to bandwidth. The noise amplitude is therefore proportional to the square root of the bandwidth.

TAMA300 [4] (Japan) is located in Tokyo, and its arm length is 300 m. It began taking data without power recycling in 1999, but its sensitivity is not yet near 10^{-21}. Following improvements, especially power recycling, it should get to within a factor of ten of this goal. However, it is not planned as an observing instrument: it is a prototype for a kilometre-scale interferometer in Japan, currently called JGWO. By 2005 this may be operating, possibly with cryogenically cooled mirrors.

GEO600 [5] (Germany and Britain) is located near Hannover (Germany). Its arm length is 600 m and the target date for first good data is now the end of 2001. Unlike TAMA, GEO600 is designed as a leading-edge-technology detector, where high-performance suspensions and optical tricks like signal recycling can be developed and applied. Although it has a short baseline, it will have a similar sensitivity to the larger LIGO and VIRGO detectors at first. At a later stage, LIGO and VIRGO will incorporate the advanced methods developed in GEO, and at that point they will advance in sensitivity, leaving GEO behind.

As we can see from figure 3.3 the sensitivity of GEO600 depends on its bandwidth, which in its turn depends on the signal recycling factor. GEO600 can change its observing bandwidth in response to observing goals. By choosing low or high reflectivity for the signal recycling mirror, scientists can make GEO600 wide-band or narrow-band, respectively. The centre frequency of the observing band (in the right-hand panel of figure 3.3 it is ~600 Hz) can be tuned to any desired frequency by shifting the position of the signal recycling mirror, thus

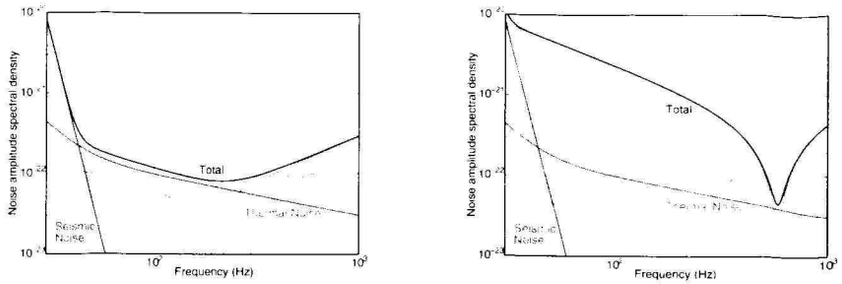

Figure 3.3. GEO600 noise curves. As for the TAMA curve, these are calibrated in strain per root Hz. The figure on the left-hand side shows GEO's wideband configuration; that on the right-hand side shows a possible narrowband operating mode.

changing the resonance frequency of the signal recycling cavity. This feature could be useful when interferometers work with bars or when performing wideband surveys.

LIGO [6] (USA) is building two detectors of arm length 4 km. One is located in Hanford WA and the other in Livingstone LA. The target date for observing is mid-2002. The two detectors are placed so that their antenna patterns overlap as much as possible and yet they are far enough apart that there will be a measurable time delay in most coincident bursts of gravitational radiation. This delay will give some directional information. The Hanford detector also contains a half-length interferometer to assist in coincidence searches. The two LIGO detectors are the best placed for doing cross-correlation for a random background of gravitational waves. LIGO's expected initial noise curve is shown in figure 3.4. These detectors have been constructed to have a long lifetime. With such long arms they can benefit from upgrades in laser power and mirror quality. LIGO has defined an upgrade goal called LIGO II, which it hopes to reach by 2007, which will observe at 10^{-22} or better over a bandwidth from 10 Hz up to 1 kHz.

VIRGO [7] (Italy and France) is building a 3 km detector near Pisa. Its target date for good data is 2003. Its expected initial noise curve is shown in figure 3.4. Like LIGO, it can eventually be pushed to much higher sensitivities with more powerful lasers and other optical enhancements. VIRGO specializes in sophisticated suspensions, and the control of vibrational noise. Its goal is to observe at the lowest possible frequencies from the ground, at least partly to be able to examine as many pulsars and other neutron stars as possible.

3.3 The physics of resonant mass detectors

The principle of operation of bar detectors is to use the gravitational tidal force of the wave to stretch a massive cylinder along its axis, and then to measure the elastic vibrations of the cylinder.

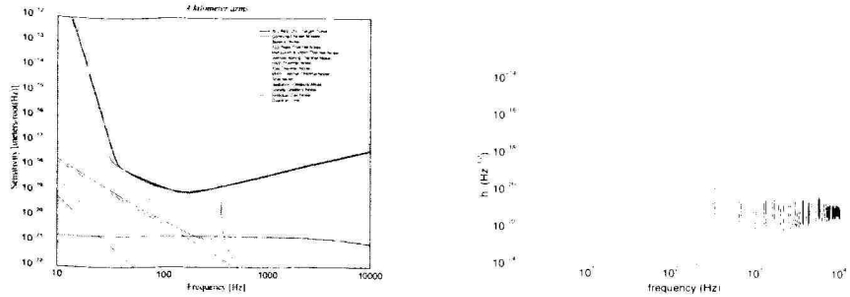

Figure 3.4. Noise curves of the initial LIGO (left) and VIRGO (right) detectors. The VIRGO curve is in strain per root Hz, as the GEO curves earlier. The LIGO curve is calibrated in metres per root Hz, so to convert to a limit on h one multiplies by the square root of the bandwidth and divides by the length of the detector arm, 4000 m.

Let us suppose we have a typical bar with length $L \sim 1$ m. (In the future, spheres may go up to 3 m.) Depending on the length of the bar and its material, the resonant frequency will be $f \sim 500$ Hz to 1.5 kHz and mass $M \sim 1000$ kg. A short burst gravitational wave h will make the bar vibrate with an amplitude

$$\delta l_{\rm gw} \sim hl \sim 10^{-21}\ {\rm m}.$$

Unlike the interferometers, whose response is simply given by this equation, the bars respond in a complicated way depending on all their internal forces. However, if the duration of the wave is short, the amplitude will be of the same order as that given here. If the wave has long duration and is near the bars resonant frequency, then the signal can build up to much larger amplitudes. Normally, bar detector searches have been targeted at short-duration signals.

The main sources of noise that compete with this very small amplitude are:

- *Thermal noise*. This is the most serious source of noise. Interferometers can live with room-temperature thermal noise because their larger size makes their response to a gravitational wave larger, and because they observe at frequencies far from the resonant frequency, where the noise amplitude is largest. However, bars observe at the resonant frequency and have a very short length, so they must reduce thermal noise by going to low temperatures. The best ultra-cryogenic bars today operate at about $T = 100$ mK, where the rms amplitude of vibration is found by setting the kinetic energy of the normal mode, $M(\delta\dot{l})^2/2$, equal to $kT/2$, the equipartition thermal energy of a single degree of freedom. This gives then

$$\langle \delta l^2 \rangle_{\rm th}^{\frac{1}{2}} = \left(\frac{kT}{4\pi^2 M f^2} \right)^{\frac{1}{2}} \sim 6 \times 10^{-18}\ {\rm m}.$$

This is far larger than the gravitational wave amplitude. In order to detect gravitational waves against this noise, bars are constructed to have a very high Q, of order 10^6 or better.

The reason that bars need a high Q is different from the reason that interferometers also strive for high-Q systems. To see how Q helps bars, we recall that Q is defined as $Q = f \cdot \tau$ where f is the resonant frequency of the mode and τ is the decay time of the oscillations. If Q is large, then the decay time is long. If the decay time is long, then the amplitude of oscillation changes very slowly in thermal equilibrium. Essentially, the bar's mode of vibration changes its amplitude by a random walk with very small steps, taking time $Q/f \sim 1000$ s to change by the full amount. On the other hand, a gravitational wave burst will cause an amplitude change in time of the order 1 ms, during which the thermal noise will have random walked to an expected amplitude change that is $Q^{\frac{1}{2}} = (\frac{1000\text{ s}}{1\text{ ms}})^{\frac{1}{2}}$ times smaller. In this case

$$\langle \delta l^2 \rangle^{\frac{1}{2}}_{\text{th: 1 ms}} = \left(\frac{kT}{4\pi^2 M f^2 Q} \right)^{\frac{1}{2}} \sim 6 \times 10^{-21} \text{ m}.$$

Thus, thermal noise only affects a measurement to the extent that it changes the amplitude of vibration during the time of the gravitational-wave burst. This change is similar to that produced by a gravitational wave of amplitude 6×10^{-21}. It follows that, if thermal noise were the only noise source, bars would be operating at around 10^{-20} today. Bar groups expect in fact to reach this level during the next few years, as they reduce the other competing sources of noise. Notice that the effect of thermal noise has nothing to do with the frequency of the disturbance, so it is not the reason that bars observe near their resonant frequency. In fact, both thermal impulses and gravitational-wave forces are mechanical forces on the bar, and the ratio of their induced vibrations is the same at all frequencies for a given applied impulsive force.

- *Sensor noise*. Because the oscillations of the bar are very small, bars require a *transducer* to convert the mechanical energy of vibration into electrical energy, and an *amplifier* that increases the electrical signal to a level where it can be recorded. If the amplifier were perfect, then the detector would in fact be broadband: it would amplify the smaller off-resonant responses just as well as the on-resonance ones. Conversely, real bars are narrow-band because of sensor noise, not because of their mechanical resonance.

  Unfortunately sensing is not perfect: amplifiers introduce noise and this makes small amplitudes harder to measure. The amplitudes of vibration are largest in the resonance band near the resonant frequency f_0, so amplifier noise limits the detector sensitivity to frequencies near f_0. Now, the signal (a typical gravitational-wave burst) has a duration time $\tau_w \sim 1$ ms, so the amplifier's bandwidth should be at least $1/\tau_w$ in order for it to be able to record a signal every τ_w. In other words, bars require amplifiers with very small noise in a large bandwidth ($\sim$1000 Hz) near f_0 (note that this band is much larger than f/Q). Today typical bandwidths of realizable amplifiers are 1 Hz, but in the very near future it is hoped to extend these to 10 Hz, and eventually to 100 Hz.
- *Quantum limit*. According to the Heisenberg uncertainty principle, the zero-point vibrations of a bar with a frequency of 1 kHz have rms amplitude

$$\langle \delta l^2 \rangle^{\frac{1}{2}}_{\text{quant}} = \left(\frac{\hbar}{2\pi M f} \right)^{\frac{1}{2}} \sim 4 \times 10^{-21} \text{ m}.$$

  This is bigger than the expected signal, and comparable to the thermal limit over 1 ms. It represents the accuracy with which one can measure the amplitude of vibration of the bar. So as soon as current detectors improve their thermal limits, they will run into the quantum limit, which must be overcome before a signal at 10^{-21} can be seen with such a detector. One way to overcome this limit is by increasing the size of the detector and even by making it spherical. This increases its mass dramatically, pushing the quantum limit down below 10^{-21}.

  Another way around the quantum limit is to avoid measuring δl, but instead to measure other observables. After all, the goal is to infer the gravitational-wave amplitude, not to measure the state of vibration of the bar. It is possible to define a pair of conjugate observables that have the property that one of them can be measured arbitrarily accurately repeatedly, so that the resulting inaccuracy of knowing the conjugate variable's value does not disturb the first variable's value. Then, if the first variable responds to the gravitational wave, the gravitational wave may be measured accurately, even though the full state of the bar is poorly known. This method is called *'back reaction evasion'*. The theory was developed in a classic paper by Caves *et al* [8]. However, no viable schemes to do this have been demonstrated for bar detectors so far.

3.3.1 New bar detectors and their capabilities

Resonant-mass detectors are limited by properties of materials and, as we have just explained, they have their best sensitivity in a narrow band around their resonant frequency. However, they can usefully explore higher frequencies (above 500 Hz), where the interferometer noise curves are rising (see earlier figures).

From the beginning, bars were designed to detect *bursts*. If the burst radiation carries significant energy in the bar's bandwidth, then the bar can do well. Standard assumptions about gravitational collapse suggest a signal with a broad spectrum to 1 kHz or more, so that most of the sensitive bars today would be suited to observe such a signal. Binary coalescence has a spectrum that peaks at low frequencies, so bars are not partiularly well suited for such signals. On the other hand, neutron-star and stellar-mass black-hole normal modes range in

frequency from about 1 kHz up to 10 kHz, so suitably designed bars could, in principle, go after these interesting signals.

A bar gets all of its sensitivity in a relatively narrow bandwidth, so if a bar and an interferometer can both barely detect a burst of amplitude 10^{-20}, then the bar has much greater sensitivity than the interferometer in its narrow band, and much worse at other frequencies. This has led recently to interest in bars as detectors of *continuous signals*. If the signal frequency is in the observing band of the bar, it can do very well compared to interferometers. Signals from millisecond pulsars and possible signals from x-ray binaries are suitable if they have the right frequency. However, most known pulsars will radiate at frequencies rather low compared to the operating frequencies of present-day bars.

The excellent sensitivity of bars in their narrow bandwidth also suits them to detecting *stochastic signals*. Cross-correlations of two bars or of bars with interferometers can be better than searches with first-generation interferometers [9]. One gets no spectral information, of course, and in the long run expected improvements in interferometers will overtake bars in this regard.

Today's best bar detectors are orders of magnitude more sensitive than the original Weber bar. Two *ultra-cryogenic* bars have been built and are operating at thermodynamic temperatures below 100 mK: *NAUTILUS* [10] at Frascati, near Rome, and [11] in Legnaro. With a mass of several tons, these may be the coldest massive objects ever seen anywhere in the universe. These are expected soon to reach a sensitivity of 10^{-20} near 1 kHz. Already they are performing coincidence experiments with bars at around 4 K at Perth, Australia, and at LSU.

Proposals exist in the Netherlands, Brazil, Italy, and the USA for *spherical* or *icosahedral detectors* (see links from [10]). These detectors have more mass, so they could reach 10^{-21} near 1 kHz. Because of their shape, they have omnidirectional antenna patterns; if they are instrumented so that all five independent fundamental quadrupolar modes of vibration can be monitored, they can do all-sky observing and determine directions as well as verify detections using coincidences between modes of the same antenna.

3.4 A detector in space

As we have noted earlier, gravitational waves from astronomical objects at frequencies below 1 Hz are obscured by Earth-based gravity-gradient noise. Detectors must go into space to observe in this very interesting frequency range.

The *LISA* [12] mission is likely to be the first such mission to fly. LISA will be a triangular array of spacecraft, with arm lengths of 5×10^6 km, orbiting the Sun in the Earth's orbit, about 20° behind the Earth. The spacecraft will be in a plane inclined to the ecliptic by 60°. The three arms can be combined in different ways to form two independent interferometers. During the mission the configuration of spacecraft rotates in its plane, and the plane rotates as well, so that LISA's antenna pattern sweeps the sky.

LISA has been named a cornerstone mission of the European Space Agency (ESA), and NASA has recently formed its own team to study the same mission, with a view toward a collaboration with ESA. LISA will be sensitive in a range from 0.3 mHz to about 0.1 Hz, and it will be able to detect known binary star systems in the Galaxy and binary coalescences of supermassive black holes anywhere they occur in the universe. A joint ESA–NASA project looks very likely, aiming at a launch around 2010. A technology demonstration mission might be launched in 2005 or 2006.

LISA's technology is fascinating. We can only allude to the most interesting parts of the mission here. A full description can be found in the pre-Phase A study document [13]. The most innovative aspect of the mission is *drag-free control*. In order to guarantee that the interferometry is not disturbed by external forces, such as fluctuations in solar radiation pressure, the mirror that is the reference point for the interferometry is on a free mass inside the spacecraft. The spacecraft acts as an active shield, sensing the position of the free mass, firing jets to counteract external forces on itself and ensure that it does not disturb the free mass. The jets themselves are remarkable, in that they must be very weak compared to most spacecraft's control jets, and they must be capable of very precise control. They will work by expelling streams of ions, accelerated and controlled by a high-voltage electric field. Fuel for these jets is not a problem: 1 g will be enough for a mission lifetime of ten years!

LISA interferometry is not done with reflection from mirrors. When a laser beam reaches one spacecraft from the other, it is too weak to reflect: the sending spacecraft would only get the occasional photon! Instead, the incoming light is sensed, and an on-board laser is slaved to it, returning an amplified beam with the same phase and frequency as the incoming one. No space mission has yet implemented this kind of laser-transponding. The LISA team had to ensure that there was enough information in all the signals to compensate for inevitable frequency fluctuations among all six on-board lasers.

A further serious problem that the LISA team had to solve was how to compensate for the relative motions of the spacecraft. The laser signals converging on a single spacecraft from the other two corners will be Doppler shifted so that their fringes change at MHz frequencies. This has to be sensed on board and removed from the signal that is sent back to Earth, which can only be sampled a few tens of times per second.

When LISA flies it will, on a technical as well as a scientific level, be a worthy counterpart to its Earth-based interferometer cousins!

3.4.1 LISA's capabilities

In the low-frequency LISA window, most sources will be relatively long lived, at least a few months. During an observation, LISA will rotate and change its velocity by a significant amount. This will induce Doppler shifts into the signals, and modulate their amplitudes, so that LISA should be able to infer the position,

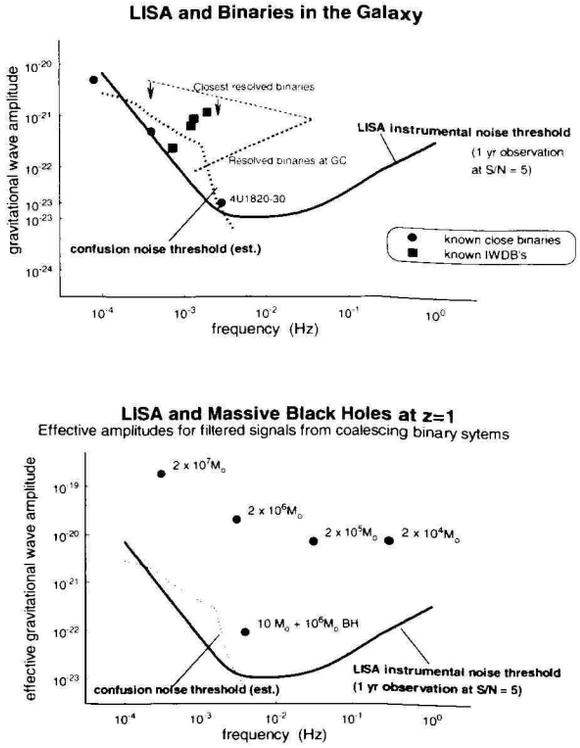

Figure 3.5. LISA sensitivity to binary systems in the Galaxy (top) and to massive black hole coalescences (bottom). The top figure is calibrated in the intrinsic amplitude of the signal, and the noise curve shows the detection threshold (5σ) for a one-year observation. It also shows the confusion limit due to unresolved binary systems. The bottom panel shows the effective amplitude of signals from coalescences of massive black holes. Since some such events last less than one year, what is shown is the expected signal-to-noise ratio of the observation.

polarization and amplitude of sources entirely from its own observations. Below about 1 mHz, this information weakens, because the wavelength of the radiation becomes comparable to or greater than the radius of LISA's orbit. The amplitude modulation is the only directional information in this frequency range.

3.5 Gravitational and electromagnetic waves compared and contrasted

To conclude this lecture it is useful to discuss the most important differences and similarities between gravitational waves and electromagnetic ones. We do this in the form of a table.

Table 3.1.

Electromagnetism	General relativity
Two signs of charges—large bodies usually neutral—waves usually emitted by single particles, often incoherently—waves carry 'local' information.	One sign of mass—gravity accumulates—waves emitted more strongly by larger body—waves carry 'global' information.
A genuine physical force, acting differently on different bodies. Detected by measuring accelerations.	Equivalence principle: gravity affects all bodies in the same way. Represented as a spacetime curvature rather than a force. Detected only by tidal forces—differential accelerations.
Maxwell's equations are *linear*. Physical field is $F_{\mu\nu}$ ($\boldsymbol{E}$ and $\boldsymbol{B}$). Gauge field is vector potential A.	Einstein's equations are *nonlinear*. Physical field is Riemann curvature tensor $R_{\mu\nu\alpha\beta}$. Gauge fields are metric $g_{\mu\nu}$ and connection $\Gamma^{\alpha}_{\mu\nu}$. Gauge transformations are coordinate transformations.
Source is charge-current density J_μ. Charge creates electric field, current magnetic field.	Source is stress-energy tensor $T_{\mu\nu}$. Mass creates a Newtonian-like field, momentum as gravito-magnetic effects. Stress creates field too.
Moderately strong force on the atomic scale: $\frac{e^2/4\pi\epsilon_0}{Gm_p^2} = 10^{39}$.	Weaker than 'weak' interaction.
Wave generation for A_μ: $\partial^\beta\partial_\beta A_\mu = 4\pi\epsilon_0 J_\mu$ in a convenient gauge (Lorentz gauge).	Wave generation for $h_{\mu\nu} = g_{\mu\nu} - \eta_{\mu\nu}$: $\partial^\beta\partial_\beta(h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h^\alpha{}_\alpha) = 8\pi T_{\mu\nu}$ in a convenient gauge.
Propagate at the speed of light, amplitude falls as $1/r$.	Propagate at the speed of light, amplitude falls as $1/r$.
Conservation of charge $\Rightarrow$ radiation by low-velocity charges is dominated by dipole component.	Conservation of mass and momentum $\Rightarrow$ radiation by low-velocity masses is dominated by quadrupole component.

Table 3.1. (Continued)

Electromagnetism	General relativity
Simple detector: oscillating charge. Action is along a line, transverse to the directions of propagation. Spin $s = 1$ and two states of linear polarization that are inclined to each other at an angle of 90°.	Simple detector: distorted ring of masses. Action is elliptic in a plane transverse to the direction of propagation. Spin $s = 2$ and two states of linear polarization that are inclined to each other at an angle of 45°. Equivalence principle ⇒ action depends only on $h_{\mu\nu}$, which is dimensionless.
Strength of force ⇒ waves scatter and refract easily.	Weakness of gravity ⇒ waves propagate almost undisturbed and transfer energy very weakly. Dimensionless amplitude h is small.
Local energy and flux well defined: Poynting vector etc.	Equivalence principle ⇒ local energy density cannot be defined exactly. Only *global* energy balance is exact.
Multipole expansion in slow-motion limit is straightforward, radiation reaction well defined.	Multipole expansion different if fields are weak or strong. For quasi-Newtonian case fields are weak, and the resulting post-Newtonian expansion is delicate. Radiation reaction is still not fully understood.
Exact solutions, containing waves, are available and can guide the construction of approximation methods for more complicated situations.	Fully realistic exact solutions for dynamical situations of physical interest are not available. Extensive reliance on approximation methods.

Chapter 4

Astrophysics of gravitational-wave sources

There are a large number of possible gravitational-wave sources in the observable waveband, which spans eight orders of magnitude in frequency: from 10^{-4} Hz (lower bound of current space-based detector designs) to 10^{4} Hz (frequency limit of likely ground-based detectors). Some of these sources are highly relativistic and not too massive, especially above 10 Hz: a black hole of mass $1000M_{\odot}$ has a characteristic frequency of 10 Hz, and larger holes have lower frequencies in inverse proportion to the mass. Neutron stars have even higher characteristic frequencies. Other systems are well described by Newtonian dynamics, such as binary orbits.

For nearly-Newtonian sources the post-Newtonian approximation (see chapter 6) provides a good framework for calculating gravitational waves. More relativistic systems, and unusual sources like the early universe, require more sophisticated approaches (see chapter 7).

4.1 Sources detectable from ground and from space

4.1.1 Supernovae and gravitational collapse

The longest expected and still probably the least understood source, gravitational collapse is one of the most violent events known to astronomy. Yet, because we have little direct information about the deep interior, we cannot make reliable predictions about the gravitational radiation from it.

Supernovae are triggered by the gravitational collapse of the interior degenerate core of an evolved star. According to current theory the result should be a neutron star or black hole. The collapse releases an enormous amount of energy, about $0.15M_{\odot}c^2$, most of which is carried away by neutrinos; an uncertain fraction is converted into gravitational waves. One mechanism for producing this radiation could be dynamical instabilities in the rapidly rotating core before it becomes a neutron star. Another likely source of radiation is the r-mode instability (see chapter 7). This could release $\sim 0.1M_{\odot}c^2$ in radiation every time a neutron star is formed.

However, both kinds of mechanisms are difficult to model. The problem with gravitational collapse is that perfectly spherical motions do not emit gravitational waves, and it is still not possible to estimate in a reliable way the amount of asymmetry in gravitational collapse. Even modern computers are not able to perform realistic simulations of gravitational collapse in three dimensions, including all the important nuclear reactions and neutrino- and photon-transport. Similarly, it is hard to model the r-mode instability because its evolution depends on nonlinear hydrodynamics and on poorly known physics, such as the cooling and viscosity of neutron stars.

An alternative approach is to use general energy considerations. If, for example, we assume that 1% of the available energy is converted into gravitational radiation, then, from formulae we will derive in the next chapter, the amplitude h would be large enough to be detected by the first ground-based interferometers (LIGO/GEO600/VIRGO) at the distance of Virgo Cluster (18 Mpc) if the emission centres at 300 Hz. Moreover, bar and spherical-mass detectors with an effective sensitivity of 10^{-21} and the right resonant frequency could see these signals as well.

The uncertainties in our predictions have a positive aspect: it is clear that if we can detect radiation from supernovae, we will learn much that we do not know about the end stages of stellar evolution and about neutron-star physics.

4.1.2 Binary stars

Binary systems have given us our best proof of the reliability of general relativity for gravitational waves. The most famous example of such systems is the binary pulsar PSR1916+16, discovered by Hulse and Taylor in 1974; they were awarded the Nobel Prize for this discovery in 1993. From the observations of the modulation of the pulse period as the stars move in their orbits, one knows many important parameters of this system (orbital period, eccentricity, masses of the two stars, etc), and the data also show directly the decrease of the orbital period due to the emission of gravitational radiation. The observed value is 2.4×10^{-12} s/s. Post-Newtonian theory allows one to predict this from the other measured parameters of the system, without any free parameters (see chapter 7); the prediction is 2.38×10^{-12}, in agreement within the measurement errors.

Unfortunately the radiation from the Hulse–Taylor system will be too weak and of too low a frequency to be detectable by LISA.

4.1.3 Chirping binary systems

If a binary gives off enough energy for its orbit to shrink by an observable amount during an observation, it is said to *chirp*: as the orbit shrinks, the frequency and amplitude go up. LISA will see a few chirping binaries. If a binary system is compact enough to radiate above 10^{-3} Hz, it will always chirp within one year, provided its components have a mass above about $1M_{\odot}$. If they are above

Table 4.1. The range for detecting a $2 \times 1.4M_{\odot}$ NS binary coalescence. The threshold for detection is taken to be 5σ. The binary and detector orientations are assumed optimum. The average S/N ratio for randomly oriented systems is reduced from the optimum by $1/\sqrt{5}$.

Detector:	TAMA300	GEO600	LIGO I	VIRGO	LIGO II
Range (S/N = 5)	3 Mpc	14 Mpc	30 Mpc	36 Mpc	500 Mpc

about $10^3 M_{\odot}$, the binary will go all the way to coalescence within the one-year observation.

Chirping binary systems are more easily detectable than gravitational collapse events because one can model with great accuracy the gravitational waveform during the inspiral phase. There will be radiation, possibly with considerable energy, during the poorly understood *plunge* phase (when the objects reach the last stable orbit and fall rapidly towards one another) and during the merger event, but the detectability of such systems rests on tracking their orbital emissions.

The major uncertainty about this kind of source is the event rate. Current pulsar observations suggest that there will be $\sim$1 coalescence per year of a Hulse-Taylor binary out to about 200 Mpc. This is a *lower limit* on the event rate, since it comes from systems we actually observe. It is possible that there are other kinds of binaries that we have no direct knowledge of, which will boost the event rate.

Theoretical modelling of binary populations gives a wide spectrum of mutually inconsistent predictions. Some authors [14] suggest that there may be a large population that escapes pulsar surveys but brings the nearest neutron star coalescence in one year as far as 30 Mpc, only slightly farther than the Virgo cluster; but other models [15] put the rate near to the observational limit.

The most exciting motivation for detecting coalescing binaries is that they could be associated with gamma-ray bursts. The event rates are consistent, and neutron stars are able to provide the required energy. If gamma-bursts are associated with neutron-star coalescence, then observations of coalescence radiation should be followed within a second or so by a strong gamma-ray burst.

LISA will see a few chirping binaries in the Galaxy, but the sensitivity of the first generation of ground-based detectors is likely to be *too poor* to see many such events (see table 4.1).

A certain fraction of such systems could contain black holes instead of neutron stars. In fact black holes should be overrepresented in binary systems (relative to their birth rate) because their formation is much less likely to disrupt a binary system (there is much less mass lost) than the formation of a neutron star would be. Pulsar observations have not yet turned up a black-hole/neutron-star system, and of course one does not expect to see binary black holes

Table 4.2. The range for detecting a $10M_\odot$ black-hole binary. Conventions as in table 4.1.

Detector:	GEO600	LIGO I	VIRGO	LIGO II
Range (S/N = 5)	75 Mpc	160 Mpc	190 Mpc	2.6 Gpc

electromagnetically. So we can only make theoretical estimates, and there are big uncertainties.

Some evolution calculations [14] suggest that the coalescence rate of BH–BH systems may be of the same order as the NS–NS rate. Other models [15] suggest it could even be zero, because stellar-wind mass loss (significant in very massive stars) could drive the stars far apart before the second BH forms, leading to coalescence times longer than the age of the universe. A recent proposal identifies globular clusters as 'factories' for binary black holes, forming binaries by three-body collisions and then expelling them [16]. Gamma-ray bursts may also come from black-hole/neutron-star coalescences. If the more optimistic event rates are correct, then black-hole coalescences may be among the first sources detected by ground-based detectors (table 4.2).

4.1.4 Pulsars and other spinning neutron stars

There are a number of ways in which a spinning neutron star may give off a continuous stream of gravitational waves. They will be weak, so they will require long continuous observation times, up to many months. Here are some possible emission mechanisms for neutron stars.

The r-modes. Neutron stars are born hot and probably rapidly rotating. Before they cool (during their first year) they have a family of *unstable normal modes, the r-modes.* These modes are excited to instability by the emission of gravitational radiation, as predicted originally by Andersson [17]. They are particularly interesting theoretically because the radiation is gravitomagnetic, generated by mass currents rather than mass asymmetries. We will study the theory of this radiation in chapter 6. In chapter 7 we will discuss how the emission of this radiation excites the instability (the CFS instability mechanism).

Being unstable, young neutron stars will presumably radiate away enough angular momentum to reduce their spin and become stable. This could lower the spin of a neutron star to ~100 Hz within one year after its formation [18]. The energy emitted in this way should be a good fraction of the star's binding energy, so in principle this radiation could be detected from the Virgo Cluster by LIGO II, provided matched filtering can be used effectively.

We discuss a possible stochastic background of gravitational waves from the r-modes below.

Accreting neutron stars (figure 4.1) are the central objects of most of the

Figure 4.1. Accreting neutron star in a low-mass x-ray binary system.

binary x-ray sources in the Galaxy. Astronomers divide them into two distinct groups: the low-mass and high-mass binaries, according to the mass of the companion star. In these systems mass is pulled from the low- or high-mass giant by the tidal forces exerted by its neutron star companion. In low-mass x-ray binaries (LMXBs) the accretion lasts long enough to spin the neutron star up to the rotation rates of millisecond pulsars. Astronomers have therefore supposed for some time that the neutron stars in LMXBs would have a range of spins, from near zero (young systems) to near 500 or 600 Hz (at the end of the accretion phase). Until the launch of the Rossi X-ray Timing Explorer (RXTE), there was no observational evidence for the neutron star spins. However, in the last two years there has been an accumulation of evidence that most, if not all, of these stars have angular velocities in a narrow range around 300 Hz [19]. It is not known yet what mechanism regulates this spin, but a strong candidate is the emission of gravitational radiation.

A novel proposal by Bildsten [20] suggests that the temperature gradient across a neutron star that is accreting preferentially at its magnetic poles should lead to a composition and hence a density gradient in the deep crust. Spinning at 300 Hz, such a star could radiate as much as it accretes. It would then be a steady source for as long as accretion lasts, which could be millions of years.

In this model the gravitational-wave energy flux is proportional to the observed x-ray energy flux. The strongest source in this model is Sco X-1, which could be detected by GEO600 in a two-year-long narrow-band mode if the appropriate matched filtering can be done. LIGO II would have no difficulty in detecting this source.

Older stars may also be *lumpy*. For known pulsars, this is constrained by the rate of spindown: the energy radiated in gravitational waves cannot exceed the

total energy loss. In most cases, this limit is rather weak, and stars would have to sustain strains in their crust of order 10^{-3} or more. It is unlikely that crusts could sustain this kind of strain, so the observational limits are probably significant overestimates for most pulsars. However, millisecond pulsars have much slower spindown rates, and it would be easier to account for the strain in their crusts, for example as a remnant Bildsten asymmetry. Such stars could, in principle, be radiating more energy in gravitational waves than electromagnetically.

Observations of individual neutron stars would be rich with information about astrophysics and fundamental nuclear physics. So little is known about the physics of these complex objects that the incentive to observe their radiation is great.

However, making such observations presents challenges for data analysis, since the motion of the Earth puts a strong phase modulation on the signal, which means that even if its rest-frame frequency is constant it cannot be found by simple Fourier analysis. More sophisticated pattern-matching (matched-filtering) techniques are needed, which track and match the signal's phase to within one cycle over the entire period of measurement. This is not difficult if the source's location and frequency are known, but the problem of doing a wide-area search for unknown objects is very challenging [21]. Moreover, if the physics of the source is poorly known, such as for LMXBs or r-mode spindown, the job of building an accurate family of templates is a difficult one. These questions are the subject of much research today, but they will need much more in the future.

4.1.5 Random backgrounds

The big bang was the most violent event of all, and it may have created a significant amount of gravitational radiation. Other events in the early universe may also have created radiation, and there may be backgrounds from more recent epochs. We have seen earlier, for example, that compact binary systems in the Galaxy will merge into a confusion-limited noise background in LISA observations below about 1 mHz.

Let us consider the r-modes as another important example. This process may have occurred in a good fraction of all neutron stars formed since the beginning of star formation. The sum of all of their r-mode radiation will be a stochastic background, with a spectrum that extends from a lower cut-off of about 200 Hz in the rest frame of the emitter to an upper limit that depends on the initial angular velocity of stars. If significant star formation started at, say, a redshift of five, then this background should extend down to about 25 Hz. If 10^{-3} of the mass of the Galaxy is in neutron stars, and each of them radiates 10% of its mass in this radiation, then the gravitational-wave background should have a density equal to 10^{-4} of the mean cosmological density of visible stars. Expressed as a fraction $\Omega_{\rm gw}$ of the closure density of the universe, per logarithmic frequency interval, this converts to

$$\Omega_{\rm gw}^{r\text{-modes}}(25\text{–}1000\ {\rm Hz}) \approx 10^{-8}\text{–}10^{-7}.$$

This background would be easily detectable by LIGO II.

There may also be a cosmological background from either topological defects (e.g. cosmic strings) or from inflation (which amplifies initial quantum gravitational fluctuations as it does the scalar ones that lead to galaxy formation). Limits from COBE observations suggest that standard inflation could not produce a background stronger than $\Omega_{\rm gw}^{\rm inflation} \sim 10^{-14}$ today. This is too weak for any of the planned detectors to reach, but it remains an important long-range goal for the field. However, there could also be a component of background radiation that depends on what happened *before* inflation: string cosmological models, for example, predict spectra growing with frequency [22].

First-generation interferometers are not likely to detect these backgrounds: they may not be able to go below the upper limit set by the requirement that gravitational waves should not disturb cosmological nucleosynthesis, which is $\Omega_{\rm gw} = 10^{-5}$. (This limit does not apply to backgrounds generated after nucleosynthesis, like the r-mode background.) Bar detectors may do as well or better than the first generation of interferometers for a broad-spectrum primordial background: as we have noted earlier, their noise levels within their resonance bands are very low. However, their frequencies are not right for the r-mode background.

Second-generation interferometers may be able to reach to 10^{-11} of closure or even lower, by cross-correlation of the output of the two detectors. However, they are unlikely to get to the inflation target of 10^{-14}. LISA may be able to go as low as 10^{-10} (if we have a confident understanding of the instrumental noise), but it is likely to detect only the confusion background of binaries, which is expected to be much stronger than a cosmological background in the LISA band.

4.1.6 The unexpected

At some level, we are bound to see things we did not expect. LISA, with its high signal-to-noise ratios for predicted sources, is particularly well placed to do this. Most of the universe is composed of dark matter whose existence we can infer only from its gravitational effects. It would not be particularly surprising if a component of this dark matter produced gravitational radiation in unexpected ways, such as from binaries of small exotic compact objects of stellar mass. We will have to wait to see!

Chapter 5

Waves and energy

Here we discuss wave-like perturbations $h_{\mu\nu}$ of a general background metric $g_{\mu\nu}$. The mathematics is similar to that of linearized theory: $h_{\mu\nu}$ is a tensor with respect to background coordinate transformations (as it was for Lorentz transformations in linearized theory) and it undergoes a gauge transformation when one makes an infinitesimal coordinate transformation. As in linearized theory, we will assume that the amplitude of the waves is small. Moreover, the waves must have a wavelength that is short compared to the radius of curvature of the background metric. These two assumptions allow us to visualize the waves as small ripples running through a curved and slowly changing spacetime.

5.1 Variational principle for general relativity

We start our analysis of the small perturbation $h_{\mu\nu}$ by introducing the standard Hilbert variational principle for Einstein's equations. The field equations of general relativity can be derived from an action principle using the Ricci scalar curvature as the Lagrangian density. The Ricci scalar (second contraction of the Riemann tensor) is an invariant quantity which contains in addition to $g_{\mu\nu}$ and its first derivatives also the second derivatives of $g_{\mu\nu}$, so our action can be written symbolically as:

$$I[g_{\mu\nu}] = \frac{1}{16\pi}\int R(g_{\mu\nu}, g_{\mu\nu,\alpha}, g_{\mu\nu,\alpha\beta})\sqrt{-g}\,\mathrm{d}^4x \tag{5.1}$$

where $\sqrt{-g}$ is the square root of the determinant of the metric tensor. As usual in variational principles, the metric tensor components are varied $g_{\mu\nu} \to g_{\mu\nu} + h_{\mu\nu}$, and one demands that the resulting change in the action should vanish to first order in any small variation $h_{\mu\nu}$ of compact support:

$$\begin{aligned}\delta I &= I[g_{\mu\nu} + h_{\mu\nu}] - I[g_{\mu\nu}] \\ &= \frac{1}{16\pi}\int \frac{\delta(R\sqrt{-g})}{\delta g_{\mu\nu}} h_{\mu\nu}\,\mathrm{d}^4x + O(2)\end{aligned} \tag{5.2}$$

$$= -\frac{1}{16\pi}\int G^{\mu\nu}h_{\mu\nu}\sqrt{-g}\,\mathrm{d}^4x + O(2) \tag{5.3}$$

where '$O(2)$' denotes terms quadratic and higher in $h_{\mu\nu}$. All the divergences obtained in the intermediate steps of this calculation integrate to zero since $h_{\mu\nu}$ is of compact support. This variational principle therefore yields the vacuum Einstein equations: $G^{\mu\nu} = 0$.

Let us consider how this changes if we include matter. This will help us to see how we can treat gravitational waves as a new kind of 'matter' field on spacetime.

Suppose we have a matter field, described by a variable Φ (which may represent a vector, a tensor or a set of tensors). It will have a Lagrangian density $L_{\mathrm{m}} = L_{\mathrm{m}}(\Phi, \Phi_{,\alpha}, \ldots, g_{\mu\nu})$ that depends on the field and also on the metric. Normally derivatives of the metric tensor do not appear in L_{m}, since by the equivalence principle, matter fields should behave locally as if they were in flat spacetime, where of course there are no metric derivatives. Variations of L_{m} with respect to Φ will produce the field equation(s) for the matter system, but here we are more interested in variations with respect to $g_{\mu\nu}$, which is how we will find the matter field contribution to the gravitational field equations. The total action has the form:

$$I = \int (R + 16\pi L_{\mathrm{m}})\sqrt{-g}\,\mathrm{d}^4x, \tag{5.4}$$

whose variation is

$$\delta I = \int \frac{\delta(R\sqrt{-g})}{\delta g_{\mu\nu}} h_{\mu\nu}\,\mathrm{d}^4x + \int 16\pi \frac{\partial(L_{\mathrm{m}}\sqrt{-g})}{\partial g_{\mu\nu}} h_{\mu\nu}\,\mathrm{d}^4x. \tag{5.5}$$

This variation must yield full Einstein equations, so we must have the following result for the stress-energy tensor of matter:

$$T^{\mu\nu}\sqrt{-g} = 2\frac{\partial L_{\mathrm{m}}\sqrt{-g}}{\partial g_{\mu\nu}}. \tag{5.6}$$

leading to

$$\boxed{G^{\mu\nu} = 8\pi T^{\mu\nu}}. \tag{5.7}$$

This way of deriving the stress-energy tensor of the matter field has deep connections to the conservation laws of general relativity, to the way of constructing conserved quantities when the metric has symmetries and to the so-called pseudotensorial definitions of gravitational wave energy (see Landau and Lifshitz 1962) [23]. We shall use it in the latter sense.

5.2 Variational principles and the energy in gravitational waves

Before we introduce the mathematics of gravitational waves, it is important to understand which geometries we are going to examine. We have said that these

geometries consist of a slowly and smoothly changing *background metric* which is altered by *perturbations* of small amplitude and high frequency. If L and λ are the characteristic lengths over which the background and 'ripple' metrics change significantly, we assume that the ratio λ/L will be very much smaller than unity and that $|h_{\mu\nu}|$ is of the same order of smallness as λ/L. In this way the total metric remains slowly changing on a macroscopic scale, while the high-frequency wave, when averaged over several wavelengths, will be the principal source of the curvature of the background metric. This is the 'short-wave' approximation [24]. Obviously this is a direct generalization of the treatment in chapter 2.

5.2.1 Gauge transformation and invariance

Consider an infinitesimal coordinate transformation generated by a vector field ξ^α,

$$x^\alpha \to x^\alpha + \xi^\alpha. \tag{5.8}$$

In the new coordinate system, neglecting quadratic and higher terms in $h^{\alpha\beta}$, it is not hard to show that the general gauge transformation of the metric is

$$h_{\mu\nu} \to h_{\mu\nu} - \xi_{\mu;\nu} - \xi_{\nu;\mu}, \tag{5.9}$$

where a semicolon denotes the covariant derivative. We assume that the derivatives of the coordinate displacement field are of the same order as the metric perturbation: $|\xi^{\alpha,\beta}| \sim |h^{\alpha\beta}|$.

Isaacson [24] showed that the gauge transformation of the Ricci and Riemann curvature tensors has the property

$$\bar{R}^{(1)}_{\mu\nu} - R^{(1)}_{\mu\nu} \approx \left(\frac{\lambda}{L}\right)^2 \tag{5.10}$$
$$\bar{R}^{(1)}_{\alpha\mu\beta\nu} - R^{(1)}_{\alpha\mu\beta\nu} \approx \left(\frac{\lambda}{L}\right)^2$$

where $R^{(1)}_{\mu\nu}$ and $R^{(1)}_{\alpha\mu\beta\nu}$ are the first order of Ricci and Riemann tensors (in powers of perturbation $h_{\mu\nu}$) and an overbar denotes their values after the gauge transformation. In our high-frequency limit, therefore, these tensors are gauge-invariant to linear order, just as in linearized theory.

5.2.2 Gravitational-wave action

Let us suppose that we are in vacuum so we have only the metric, no matter fields, but we work in the high-frequency approximation. The full metric is $g_{\mu\nu}$ (smooth *background* metric) $+h_{\mu\nu}$ (high-frequency perturbation). Our purpose is to show that the wave field can be treated as a 'matter' field, with a Lagrangian and its

own stress-energy tensor. To do this we have to expand the action out to second order in the metric perturbation,

$$\begin{aligned} I[g_{\mu\mu} + h_{\mu\nu}] &= \int R(g_{\mu\nu} + h_{\mu\nu}, g_{\mu\nu,\alpha} + h_{\mu\nu,\alpha}, \ldots)\sqrt{-g[g_{\mu\nu} + h_{\mu\nu}]}\,\mathrm{d}^4x \\ &= \int R(g_{\mu\nu}, \ldots)\sqrt{-g}\,\mathrm{d}^4x + \int \frac{\delta(R\sqrt{-g})}{\delta g_{\mu\nu}} h_{\mu\nu}\,\mathrm{d}^4x \\ &\quad + \frac{1}{2}\int \left(\frac{\partial^2(R\sqrt{-g})}{\partial g_{\mu\nu}\partial g_{\alpha\beta}} h_{\mu\nu}h_{\alpha\beta} + 2\frac{\partial^2(R\sqrt{-g})}{\partial g_{\mu\nu}\partial g_{\alpha\beta,\gamma}} h_{\mu\nu}h_{\alpha\beta,\gamma} \right. \\ &\quad \left. + \frac{\partial^2(R\sqrt{-g})}{\partial g_{\mu\nu,\tau}\partial g_{\alpha\beta,\gamma}} h_{\mu\nu,\tau}h_{\alpha\beta,\gamma} + 2\frac{\partial^2(R\sqrt{-g})}{\partial g_{\mu\nu}\partial g_{\alpha\beta,\gamma\tau}} h_{\mu\nu}h_{\alpha\beta,\gamma\tau} \right) \mathrm{d}^4x \\ &\quad + O(3). \end{aligned}$$

The first term is the action for the background metric $g_{\mu\nu}$. The second term vanishes (see equation (5.2)), since we assume that the background metric is a solution of the Einstein vacuum equation itself, at least to lowest order.

If we compare the above equation with equation (5.4), we can see that the third term, complicated as it seems, is an effective 'matter' Lagrangian for the gravitational field. Indeed, if one varies it with respect to $h_{\mu\nu}$ holding $g_{\mu\nu}$ fixed (as we would do for a physical matter field on the background), then the complicated coefficients are fixed and one can straightforwardly show that one gets exactly *the linear perturbation of the Einstein tensor itself*. Its vanishing is the equation for the gravitational-wave perturbation $h_{\mu\nu}$. In this way we have shown that, for a *small amplitude* perturbation, the gravitational wave can be treated as a 'matter' field with its own Lagrangian and field equations.

Given this Lagrangian, we should be able to calculate the effective stress-energy tensor of the wave field by taking the variations of the effective Lagrangian with respect to $g_{\mu\nu}$, holding the 'matter' field $h_{\mu\nu}$ fixed:

$$T^{(\mathrm{GW})\alpha\beta}\sqrt{-g} = 2\frac{\partial L^{(\mathrm{GW})}[g_{\mu\nu}, h_{\mu\nu}]\sqrt{-g}}{\partial g_{\alpha\beta}} \tag{5.11}$$

with

$$\begin{aligned} L^{(\mathrm{GW})}\sqrt{-g} = \frac{1}{32\pi}\left(\frac{\partial^2(R\sqrt{-g})}{\partial g_{\mu\nu}\partial g_{\alpha\beta}} h_{\mu\nu}h_{\alpha\beta} + 2\frac{\partial^2(R\sqrt{-g})}{\partial g_{\mu\nu}\partial g_{\alpha\beta,\gamma}} h_{\mu\nu}h_{\alpha\beta,\gamma} \right. \\ \left. + \frac{\partial^2(R\sqrt{-g})}{\partial g_{\mu\nu,\tau}\partial g_{\alpha\beta,\gamma}} h_{\mu\nu,\tau}h_{\alpha\beta,\gamma} + 2\frac{\partial^2(R\sqrt{-g})}{\partial g_{\mu\nu}\partial g_{\alpha\beta,\gamma\tau}} h_{\mu\nu}h_{\alpha\beta,\gamma\tau} \right). \end{aligned} \tag{5.12}$$

This quantity is quadratic in the wave amplitude $h_{\mu\nu}$. It could be simplified further by integrations by parts, such as by taking a derivative off $h_{\alpha\beta,\gamma\tau}$. This would change the coefficients of the other terms. We will not need to worry about finding the 'best' form for the expression (4.12), as we now show.

As in linearized theory, so also in the general case, the quantity $h_{\mu\nu}$ behaves like a tensor with respect to background coordinate transformations, and so does $T^{(\mathrm{GW})}_{\mu\nu}$. However, it is not gauge-invariant and so it is not physically observable. Since the integral of the action *is* independent of coordinate transformations that have compact support, so too is the integral of the effective stress-energy tensor. In practical terms, this makes it possible to localize the energy of a wave to within a region of about one wavelength in size where the background curvature does not change significantly. In fact, if we restrict our gauge transformations to have a length scale of a wavelength, and if we average (integrate) the stress-energy tensor of the waves over such a region, then any gauge changes will be small surface terms.

By evaluating the effective stress-energy tensor on a smooth background metric in a Lorentz gauge, and performing the averaging (denoted by symbol $\langle\cdots\rangle$), one arrives at the *Isaacson tensor*:

$$T^{(\mathrm{GW})}_{\alpha\beta} = \frac{1}{32\pi}\langle h_{\mu\nu;\alpha}h^{\mu\nu}{}_{;\beta}\rangle. \tag{5.13}$$

This is a convenient and compact form for the gravitational stress-energy tensor. It localizes energy in short-wavelength gravitational waves to regions of the order of a wavelength. It is interesting to remind ourselves that our only experimental evidence of gravitational waves today is the observation of the effect on a binary orbit of the loss of energy to the gravitational waves emitted by the system. So this energy formula, or equivalent ones, is central to our understanding of gravitational waves.

5.3 Practical applications of the Isaacson energy

If we are far from a source of gravitational waves, we can treat the waves by linearized theory. Then if we adopt the TT gauge and specialize the stress-energy tensor of the radiation to a flat background, we get

$$T^{(\mathrm{GW})}_{\alpha\beta} = \frac{1}{32\pi}\langle h^{\mathrm{TT}}_{ij,\alpha}h^{\mathrm{TT}ij}{}_{,\beta}\rangle. \tag{5.14}$$

Since there are only two components, a wave travelling with frequency f (wavenumber $k = 2\pi f$) and with a typical amplitude h in both polarizations carries an energy F_{gw} equal to (see exercise (f) at the end of this lecture)

$$F_{\mathrm{gw}} = \frac{\pi}{4}f^2h^2. \tag{5.15}$$

Putting in the factors of c and G and scaling to reasonable values gives

$$F_{\mathrm{gw}} = 3\ \mathrm{mW\ m^{-2}}\left[\frac{h}{1\times10^{-22}}\right]^2\left[\frac{f}{1\ \mathrm{kHz}}\right]^2, \tag{5.16}$$

which is a very large energy flux even for this weak a wave. It is twice the energy flux of a full moon! Integrating over a sphere of radius r, assuming a total duration of the event τ, and solving for h, again with appropriate normalizations, gives

$$h = 10^{-21}\left[\frac{E_{\mathrm{gw}}}{0.01M_\odot c^2}\right]^{\frac{1}{2}}\left[\frac{r}{20\ \mathrm{Mpc}}\right]^{-1}\left[\frac{f}{1\ \mathrm{kHz}}\right]^{-1}\left[\frac{\tau}{1\ \mathrm{ms}}\right]^{-\frac{1}{2}}. \tag{5.17}$$

This is the formula for the 'burst energy', normalized to numbers appropriate to a gravitational collapse occurring in the Virgo cluster. It explains why physicists and astronomers regard the 10^{-21} threshold as so important. However, this formula could also be applied to a binary system radiating away its orbital gravitational binding energy over a long period of time τ, for example.

5.3.1 Curvature produced by waves

We have assumed that the background metric satisfies the vacuum Einstein equations to linear order, but now it is possible to view the full action principle as a principle for the background with a wave field $h_{\mu\nu}$ on it, and to let the wave energy affect the background curvature [24]. This means that the background will actually solve, in a self-consistent way, the equation

$$G_{\alpha\beta}[g_{\mu\nu}] = 8\pi T^{\mathrm{GW}}_{\alpha\beta}[g_{\mu\nu} + h_{\mu\nu}]. \tag{5.18}$$

This does not contradict the vanishing of the first variation of the action, which we needed to use above, because now we have an Einstein tensor that is of quadratic order in $h_{\mu\nu}$, contributing a term of cubic order to the first-variation of the action, which is of the same order as other terms we have neglected.

5.3.2 Cosmological background of radiation

This self-consistent picture allows us to talk about, for example, a cosmological gravitational wave background that contributes to the curvature of the universe. Since the energy density is the same as the flux (when $c = 1$), we have

$$\varrho_{\mathrm{gw}} = \frac{\pi}{4}f^2h^2, \tag{5.19}$$

but now we must interpret h in a statistical way. This will be treated in the contribution by Babusci *et al*, but basically it is done by replacing h^2 by a statistical mean square amplitude per unit frequency (Fourier transform power), so that the energy density per *unit frequency* is proportional to $f^2|\bar{h}|^2$. It is then conventional to talk about the energy density per unit logarithm of frequency, which means multiplying by f. The result, after being careful about averaging over all directions of the waves and all independent polarization components, is

$$\frac{\mathrm{d}\varrho_{\mathrm{gw}}}{\mathrm{d}\ln f} = 4\pi^2f^3|\bar{h}(f)|^2. \tag{5.20}$$

Finally, what is of the most interest is the energy density as a fraction of the closure or critical cosmological density, given by the Hubble constant H_0 as $\varrho_c = 3H_0^2/8\pi$. The resulting ratio is the symbol $\Omega_{\rm gw}(f)$ that we met in the previous lecture:

$$\Omega_{\rm gw}(f) = \frac{32\pi^3}{3H_0^2} f^3 |\bar{h}(f)|^2. \tag{5.21}$$

5.3.3 Other approaches

We finish this lecture by observing that there is no unique approach to defining energy for gravitational radiation or indeed for any solution of Einstein's equations. Historically this has been one of the most difficult areas for physicists to get to grips with. In the textbooks you will find discussions of pseudotensors, of energy measured at null infinity and at spacelike infinity, of Noether theorems and formulae for energy, and so on. None of these are worse than we have presented here, and in fact all of them are now known to be consistent with one another, if one does not ask them to do too much. In particular, if one wants only to localize the energy of a gravitational wave to a region of the size of a wavelength, and if the waves have short wavelength compared to the background curvature scale, then pseudotensors will give the same energy as the one we have defined here. Similarly, if one takes the energy flux defined here and evaluates it at null infinity, one gets the so-called Bondi flux, which was derived by H Bondi in one of the pioneering steps in the understanding of gravitational radiation. Many of these issues are discussed in the Schutz–Sorkin paper referred to earlier [23].

5.4 Exercises for chapter 5

(e) *In the notes above we give the general gauge transformation*

$$h_{\mu\nu} \to h_{\mu\nu} - \xi_{\mu;\nu} - \xi_{\nu;\mu}.$$

Use the formula for the derivation of Einstein's equations from an action principle,

$$\delta I = \frac{1}{16\pi} \int \frac{\delta(R\sqrt{-g})}{\delta g_{\mu\nu}} h_{\mu\nu}\, \mathrm{d}^4x$$

with

$$\frac{\delta(R\sqrt{-g})}{\delta g_{\mu\nu}} = -G^{\mu\nu}\sqrt{-g},$$

but insert a pure gauge $h_{\mu\nu}$. Argue that since this is merely a coordinate transformation, the action should be invariant. Integrate the variation of the action to prove the contacted Bianchi identity

$$G^{\mu\nu}{}_{;\nu} = 0.$$

This shows that the divergence-free property of $G^{\mu\nu}$ is closely related to the coordinate invariance of Einstein's theory.

(f) *Suppose a plane wave, travelling in the z-direction in linearized theory, has both polarization components h_+ and $h_\times$. Show that its energy flux in the z-direction, $T^{(\rm GW)0z}$, is*

$$\langle T^{(\rm GW)0z} \rangle = \frac{k^2}{32\pi}(A_+^2 + A_\times^2),$$

where the angle brackets denote an average over one period of the wave.

Chapter 6

Mass- and current-quadrupole radiation

In this lecture we focus on the wave amplitude itself, and how it and the polarization depend on the motions in the source. Consider an isolated source with a stress-energy tensor $T^{\alpha\beta}$. As in chapter 2, the Einstein equation is

$$\left(-\frac{\partial^2}{\partial t^2}+\nabla^2\right)\bar{h}^{\alpha\beta}=-16\pi T^{\alpha\beta} \tag{6.1}$$

($\bar{h}^{\alpha\beta}=h^{\alpha\beta}-\frac{1}{2}\eta^{\alpha\beta}h$ and $\bar{h}^{\alpha\beta}{}_{,\beta}=0$). Its general solution is the following retarded integral for the field at a position x^i and time t in terms of the source at a position y^i and the retarded time $t-R$:

$$\bar{h}^{\alpha\beta}(x^i,t)=4\int\frac{1}{R}T^{\alpha\beta}(t-R,y^i)\,\mathrm{d}^3y. \tag{6.2}$$

where we define

$$R^2=(x^i-y^i)(x_i-y_i). \tag{6.3}$$

6.1 Expansion for the far field of a slow-motion source

Let us suppose that the origin of coordinates is in or near the source, and the field point x^i is far away. Then we define $r^2=x^ix_i$ and we have $r^2\gg y^iy_i$. We can, therefore, expand the term R in the dominator in terms of y^i. The lowest order is r, and all higher-order terms are smaller than this by powers of r^{-1}. Therefore, they contribute terms to the field that fall off faster than r^{-1}, and they are negligible in the far zone. Therefore, we can simply replace R by r in the dominator, and take it out of the integral.

The R inside the time-argument of the source term is not so simple. If we suppose that $T^{\alpha\beta}$ does not change very fast we can substitute $t-R$ by $t-r$ (the retarded time to the origin of coordinates) and expand

$$t-R=t-r+n^iy_i+O\left(\frac{1}{r}\right),\quad\text{with } n^i=\frac{x^i}{r},\ n^in_i=1. \tag{6.4}$$

The two conditions $r\gg y^iy_i$ and the slow-motion source, can be expressed quantitatively as:

$$r\gg\bar{\lambda}$$
$$R\ll\bar{\lambda}$$

where $\bar{\lambda}$ is the reduced wavelength $\bar{\lambda}=\lambda/2\pi$ and R is the size of source.

The terms of order r^{-1} are negligible for the same reason as above, but the first term in this expansion must be taken into account. It depends on the direction to the field point, given by the unit vector n^i. We use this by making a Taylor expansion in time on the time-argument of the source. The combined effect of these approximations is

$$\bar{h}^{\alpha\beta}=\frac{4}{r}\int[T^{\alpha\beta}(t',y^i)+T^{\alpha\beta}{}_{,0}(t',y^i)n^jy_j+\tfrac{1}{2}T^{\alpha\beta}{}_{,00}(t',y^i)n^jn^ky_jy_k+\tfrac{1}{6}T^{\alpha\beta}{}_{,000}(t',y^i)n^jn^kn^ly_jy_ky_l+\cdots]\,\mathrm{d}^3y. \tag{6.5}$$

We will need all the terms of this Taylor expansion out to this order.

The integrals in expression (5.5) contain moments of the components of the stress-energy. It is useful to give these names. Use M for moments of the density T^{00}, P for moments of the momentum T^{0i} and S for the moments of the stress T^{ij}. Here is our notation:

$$M(t')=\int T^{00}(t',y^i)\,\mathrm{d}^3y,\quad M^j(t')=\int T^{00}(t',y^i)y^j\,\mathrm{d}^3y,$$
$$M^{jk}(t')=\int T^{00}(t',y^i)y^jy^k\,\mathrm{d}^3y,\quad M^{jkl}(t')=\int T^{00}(t',y^i)y^jy^ky^l\,\mathrm{d}^3y,$$
$$P^l(t')=\int T^{0l}(t',y^i)\,\mathrm{d}^3y,\quad P^{lj}(t')=\int T^{0l}(t',y^i)y^j\,\mathrm{d}^3y,$$
$$P^{ljk}(t')=\int T^{0l}(t',y^i)y^jy^k\,\mathrm{d}^3y,$$
$$S^{lm}(t')=\int T^{lm}(t',y^i)\,\mathrm{d}^3y,\quad S^{lmj}(t')=\int T^{lm}(t',y^i)y^j\,\mathrm{d}^3y.$$

These are the moments we will need.

Among these moments there are some identities that follow from the conservation law in linearized theory, $T^{\alpha\beta}{}_{,\beta}=0$, which we use to replace time derivatives of components of T by divergences of other components and then integrate by parts. The identities we will need are

$$\dot{M}=0,\quad \dot{M}^k=P^k,\quad \dot{M}^{jk}=P^{jk}+P^{kj},\quad \dot{M}^{jkl}=P^{jkl}+P^{klj}+P^{ljk}. \tag{6.6}$$

$$\dot{P}^j=0,\quad \dot{P}^{jk}=S^{jk},\quad \dot{P}^{jkl}=S^{jkl}+S^{jlk}. \tag{6.7}$$

These can be applied recursively to show, for example, two further very useful relations

$$\frac{d^2M^{jk}}{dt^2} = 2S^{jk}, \quad \frac{d^3M^{jkl}}{dt^3} = 6\dot{S}^{(jkl)} \tag{6.8}$$

where the round brackets on indices indicate full symmetrization.

Using these relations and notations it is not hard to show that

$$\bar{h}^{00}(t,x^i) = \frac{4}{r}M + \frac{4}{r}P^j n_j + \frac{4}{r}S^{jk}(t')n_j n_k + \frac{4}{r}\dot{S}^{jkl}(t')n_j n_k n_l + \cdots \tag{6.9}$$

$$\bar{h}^{0j}(t,x^i) = \frac{4}{r}P^j + \frac{4}{r}S^{jk}(t')n_k + \frac{4}{r}\dot{S}^{jkl}(t')n_k n_l + \cdots \tag{6.10}$$

$$\bar{h}^{jk}(t,x^i) = \frac{4}{r}S^{jk}(t') + \frac{4}{r}\dot{S}^{jkl}(t')n_l + \cdots. \tag{6.11}$$

In these three formulae there are different orders of time-derivatives, but in fact they are evaluated to the same final order in the slow-motion approximation. One can see that from the gauge condition $\bar{h}^{\alpha\beta}{}_{,\beta} = 0$, which relates time-derivatives of some components to space-derivatives of others.

In these expressions, one must remember that the moments are evaluated at the retarded time $t' = t - r$ (except for those moments that are constant in time), and they are multiplied by components of the unit vector to the field point $n^j = x^j/r$.

6.2 Application of the TT gauge to the mass quadrupole field

In the expression for the amplitude that we derived so far, the final terms are those that represent the current-quadrupole and mass-octupole radiation. The terms before them represent the static parts of the field and the mass-quadrupole radiation. In this section we treat just these terms, placing them into the TT gauge. This will be simpler than treating it all at once, and the procedure for the next terms will be a straightforward generalization.

6.2.1 The TT gauge transformations

We are already in Lorentz gauge, and this can be checked by taking derivatives of the expressions for the field that we have derived above. However, we are manifestly not in the TT gauge. Making a gauge transformation consists of choosing a vector field ξ^α and modifying the metric by

$$h_{\alpha\beta} \to h_{\alpha\beta} - \xi_{\alpha,\beta} - \xi_{\beta,\alpha}. \tag{6.12}$$

The corresponding expression for the potential $\bar{h}^{\alpha\beta}$ is

$$\bar{h}^{\alpha\beta} \to \bar{h}^{\alpha\beta} + \xi^{\alpha,\beta} + \xi^{\alpha,\beta} - \eta^{\alpha\beta}\xi^{\mu}{}_{,\mu}. \tag{6.13}$$

For the different components this implies changes

$$\delta\bar{h}^{00} = \xi^{0,0} + \xi^{j}{}_{,j} \tag{6.14}$$

$$\delta\bar{h}^{0j} = \xi^{0,j} + \xi^{j,0} \tag{6.15}$$

$$\delta\bar{h}^{jk} = \xi^{j,k} + \xi^{k,j} - \delta^{jk}\xi^{\mu}{}_{,\mu} \tag{6.16}$$

where δ^{jk} is the Kronecker delta (unit matrix). In practice, when taking derivatives, the algebra is vastly simplified by the fact that we are keeping only r^{-1} terms in the potentials. This means that spatial derivatives do not act on $1/r$ but only on $t' = t - r$. It follows that $\partial t'/\partial x^j = -n_j$, and $\partial h(t')/\partial x^j = -\dot{h}(t')n_j$.

It is not hard to show that the following vector field puts the metric into the TT gauge to the order we are working:

$$\xi^0 = \frac{1}{r}P^k{}_k + \frac{1}{r}P^{jk}n_j n_k + \frac{1}{r}S^l{}_{lk}n^k + \frac{1}{r}S^{ijk}n_i n_j n_k, \tag{6.17}$$

$$\xi^i = \frac{4}{r}M^i + \frac{4}{r}P^{ij}n_j - \frac{1}{r}P^k{}_k n^i - \frac{1}{r}P^{jk}n_j n_k n^i + \frac{4}{r}S^{ijk}n_j n_k - \frac{1}{r}S^l{}_{lk}n^k n^i - \frac{1}{r}S^{jlk}n_j n_l n_k n^i. \tag{6.18}$$

6.2.2 Quadrupole field in the TT gauge

The result of applying this gauge transformation to the original amplitudes is

$$\bar{h}^{TT00} = \frac{4M}{r}, \tag{6.19}$$

$$\bar{h}^{TT0i} = 0, \tag{6.20}$$

$$\bar{h}^{TTij} = \frac{4}{r}\left[\perp^{ik}\perp^{jl}S_{lk} + \frac{1}{2}\perp^{ij}(S_{kl}n^k n^l - S^k{}_k)\right]. \tag{6.21}$$

Remember that here we are not including $\dot{S}^{jkl}$, because it is a third-order effect.

The notation $\perp^{ik}$ represents the projection operator perpendicular to the direction n^i to the field point.

$$\perp^{jk} = \delta^{jk} - n^j n^k. \tag{6.22}$$

It can be verified that this tensor is transverse to the direction n^i and is a projection, in the sense that it projects to itself

$$\perp^{jk} n_k = 0, \quad \perp^{jk}\perp_k{}^l = \perp^{jl}. \tag{6.23}$$

The spherical component of the field is not totally eliminated in this gauge transformation: the time–time component of the metric must contain the constant Newtonian field of the source. (In fact we have succeeded in eliminating the

dipole, or momentum part of the field, which is also part of the non-wave solution. Our gauge transformation has incorporated a Lorentz transformation that has put us into the rest frame of the source.) The time-dependent part of the field is now purely spatial, transverse (because everything is multiplied by $\perp$), and traceless (as can be verified by explicit calculation).

The expression for the spatial part of the field actually does not depend on the trace of S_{jk}, as can be seen by constructing the trace-free part of the tensor, defined as:

$$\tilde{S}^{jk} = S^{jk} - \tfrac{1}{3}\delta^{jk}S^{l}{}_{l}. \tag{6.24}$$

In fact, it is more conventional to use the mass moment here instead of the stress, so we also define

$$\tilde{M}^{jk} = M^{jk} - \frac{1}{3}\delta^{jk}M^{l}_{l}, \quad \tilde{S}^{jk} = \frac{1}{2}\frac{\mathrm{d}^2\tilde{M}^{jk}}{\mathrm{d}t^2}. \tag{6.25}$$

In terms of $\tilde{M}$ the far field is

$$\bar{h}^{\mathrm{TT}ij} = \frac{2}{r}\left(\perp^{ik}\perp^{jl}\ddot{\tilde{M}}_{kl} + \frac{1}{2}\perp^{ij}\ddot{\tilde{M}}_{kl}n^{l}n^{k}\right). \tag{6.26}$$

This is the usual formula for the mass-quadrupole field. In textbooks the notation is somewhat different than we have adopted here. In particular, our tensor $\tilde{M}$ is what is called I in Misner *et al* (1973) and Schutz (1985). It is the basis of most gravitational-wave source estimates. We have derived it only in the context of linearized theory, but remarkably its form is identical if we go to the post-Newtonian approximation, where the gravitational waves are a perturbation of the Newtonian spacetime rather than of flat spacetime.

Given this powerful formula, it is important to try to interpret it and understand it as fully as possible. One obvious conclusion is that the dominant source of radiation, at least in the slow-motion limit, is the second time-derivative of the second moment of the mass density T^{00} (the mass-quadrupole moment). This is a very important difference between gravitational waves and electromagnetism, in which the most important source is the electric-dipole. In our case the mass-dipole term is not able to radiate because it is constant, reflecting conservation of the linear momentum of the source. In electromagnetism, however, if the dipole term is absent for some reason (all charges positive, for example) then the quadrupole term dominates and it looks very similar to equation (6.26).

6.2.3 Radiation patterns related to the motion of sources

The projection operators in equation (6.26) show that the radiative field is transverse, as we expect. However, the form of equation (6.26) hides two equally important messages:

- the only motions that produce the radiation are the ones transverse to the line of sight; and
- the induced motions in a detector mirror the motions of the source projected onto the plane of the sky.

To see why these are true, we define the *transverse traceless quadrupole tensor*

$$M^{\mathrm{TT}}_{ij} = \perp^{k}{}_{i}\perp^{l}{}_{j}M_{kl} - \tfrac{1}{2}\perp_{ij}\perp^{kl}M_{kl}. \tag{6.27}$$

(Notice that some of our definitions of tracelessness involve subtracting $\frac{1}{3}$ of the trace, as in equation (6.24), and sometimes $\frac{1}{2}$ of the trace, as in equation (6.27). The appropriate factor is determined by the effective dimensionality (rank) of the tensor. Although we have three spatial dimensions, the projection tensor $\perp$ projects the mass-quadrupole tensor onto a two-dimensional plane, where the trace involves only two components, not three.)

Now, if in equation (6.26) we replace $\tilde{M}_{ij}$ by its definition in terms of M_{ij}, and then collect terms appropriately, it is not hard to show that the equation simplifies to its most natural form:

$$\bar{h}^{\mathrm{TT}ij} = \frac{2}{r}\ddot{M}^{\mathrm{TT}ij}. \tag{6.28}$$

This could of course have been derived directly by applying the TT operation to equations (6.9)–(6.11). Since this equation involves only the TT part of M, our first assertion above is proved. According to this equation, in order to calculate the quadrupole radiation that a particular observer will receive, one need only compute the mass-quadrupole tensor's second time-derivative, project it onto the plane of the sky as seen by the observer looking toward the source, take away its trace, and rescale it by a factor $2/r$. In particular, the TT tensor that describes the action of the wave (as in the polarization diagram in figure 2.1) is a copy of the TT tensor of the mass distribution. This proves our second assertion above.

Looking again at figure 2.1 we imagine a detector consisting of two free masses whose separation is being monitored. If the wave causes them to oscillate relative to one another along the x-axis (the $\oplus$ polarization), this means that the source motion contained a component that did the same thing. If the source is a binary, then the binary orbit projected onto the sky must involve motion of the stars back and forth along either the x- or the y-axis.

It is possible from this to understand many aspects of quadrupole radiation in a simple way. Consider a binary star system with a circular orbit. Seen by a distant observer in the orbital plane, the projected source motion is linear, back and forth. The received polarization will be linear, the polarization ellipse aligned with the orbit. Seen by a distant observer along the axis of the orbit of the binary, the projected motion is circular, which is a superposition of two linear motions separated in phase by 90°. The received radiation will also have circular polarization. Because both linear polarizations are present, the amplitude of the

wave emitted up the axis is twice that emitted in the plane. In this way we can completely determine the radiation pattern of a binary system.

Notice that, when viewed at an arbitrary angle to the axis, the radiation will be elliptically polarized, and the degree of ellipticity will directly measure the inclination of the orbital plane to the line of sight. This is a very special kind of information, which one cannot normally obtain from electromagnetic observations of binaries. It illustrates the complementarity of the two kinds of observing.

6.3 Application of the TT gauge to the current-quadrupole field

Now we turn to the problem of placing next-order terms of the wave field, the current quadrupole and mass octupole, into the TT gauge. Our interest here is to understand current-quadrupole radiation in the same physical way as we have just done for mass-quadrupole radiation. So we shall put the field into the TT gauge and then see how to separate the current-quadrupole part from the mass-octupole, which we will discard from the present discussion.

6.3.1 The field at third order in slow-motion

The next order terms in the non-TT metric bear a simple relationship to the mass-quadrupole terms (see equations (6.9)–(6.11)). In each of the metric components, just replace S^{jk} by $\dot{S}^{jkl}n_l$ to go from one order to the next.

This means that we can just skip to the end of the application of the gauge transformations in equations (6.17) and (6.18) and write the next order of the final field, only using S again, not M:

$$\bar{h}^{\mathrm{TT}ij} = \frac{4}{r}\left[\perp^{ik}\perp^{jl}\dot{S}_{lkm}n^m + \frac{1}{2}\perp^{ij}(\dot{S}_{klm}n^k n^l n^m - \dot{S}^k{}_{kl}n^l)\right], \tag{6.29}$$

or more compactly

$$\bar{h}^{\mathrm{TT}ij} = \frac{4}{r}\left(\perp^{ik}\perp^{jl}\dot{\tilde{S}}_{klm}n^m + \frac{1}{2}\perp^{ij}\dot{\tilde{S}}_{klm}n^l n^k n^m\right). \tag{6.30}$$

The tilde on S represents a trace-free operation *on the first two indices*.

$$\dot{\tilde{S}}_{klm} = \dot{S}_{lkm} - \tfrac{1}{3}\delta_{kl}\dot{S}^i{}_{im}.$$

These are the indices that come from the indices of T^{jk}, so the tensor is symmetric on these. By analogy with the quadrupole calculation, we can also define the TT part of S_{ijk} by doing the TT projection on the first two indices,

$$S^{\mathrm{TT}}_{ijm} = \perp^k{}_i\perp^l{}_j S_{klm} - \tfrac{1}{2}\perp_{ij}\perp^{kl}S_{klm}. \tag{6.31}$$

The TT projection of the equation for the metric is

$$h^{\mathrm{TT}ij} = \frac{4}{r}\dot{S}^{\mathrm{TT}ijk}n_k. \tag{6.32}$$

6.3.2 Separating the current quadrupole from the mass octupole

The last equation is compact, but it does not have the ready interpretation that we have at quadrupole order. This is because the moment of the stress, S_{ijk}, does not have such a clear physical interpretation. We see from equations (6.6)–(6.8) that S_{ijk} is a complicated mixture of moments of momentum and density. To gain more physical insight into radiation at this order, we need to separate these different contributions. It is straightforward algebra to see that the following identity follows from the earlier ones:

$$\dot{S}^{ijk} = \tfrac{1}{6}\dddot{M}^{ijk} + \tfrac{2}{3}\ddot{P}^{[jk]i} + \tfrac{2}{3}\ddot{P}^{[ik]j}, \tag{6.33}$$

where square brackets around indices mean antisymmetrization:

$$A^{[ik]} := \tfrac{1}{2}(A^{ik} - A^{ki}).$$

This is a complete separation of the mass terms (in M) from the momentum terms (in P) because the only identities relating the momentum moments to the mass moments involve the symmetric part of P^{ijk} on its first two indices, and this is absent from equation (6.33).

The first term in equation (6.33) is the third moment of the density, and this is the source of the *mass-octupole* field. It produces radiation through the third time-derivative. Since we are in a slow-motion approximation, this is smaller than the mass-quadrupole radiation by typically a factor v/c. Unless there were some very special symmetry conditions, one would not expect the mass octupole to be anything more than a small correction to the mass quadrupole. For this reason we will not treat it here.

The second and third terms in equation (6.33) involve the second moment of the momentum, and together they are the source of the *current-quadrupole* field. It involves two time-derivatives, just as the mass quadrupole does, but these are time-derivatives of the momentum moment, not the mass moment, so these terms produce a field that is also v/c smaller than the typical mass-quadrupole field. However, it requires less of an accident for the mass quadrupole to be absent and the current quadrupole present. It just requires motions that leave the density unchanged to lowest order. This happens in the r-modes. Therefore, the current quadrupole deserves more attention, and we will work exclusively with these terms from now on.

The terms in equation (6.33) that we need are the ones involving $\ddot{P}^{ijk}$. These are antisymmetrized on the first two indices, which involves effectively a vector product between the momentum density (first index) and one of the moment indices. This is essentially the angular momentum density. To make

the angular momentum explicit and to simplify the expression, we introduce the angular momentum and the first moment of the angular momentum density

$$J^i := \epsilon^{ijk} P_{jk}, \tag{6.34}$$

$$J^{il} := \epsilon^{ijk} P_{jk}{}^{l}, \tag{6.35}$$

where ϵ^{ijk} is the fully antisymmetric (Levi-Civita) symbol in three dimensions. It follows from this that

$$P^{[jk]l} = \tfrac{1}{2}\epsilon^{jki} J_i{}^{l}.$$

These terms enter the TT projection of the field (6.32) with the last index of S always contracted with the direction n^i to the observer from the source. According to equation (6.33), this contraction always occurs on one of the antisymmetrized indices, or if we use the form in the previous equation then we will always have a contraction of n^i with ϵ^{ijk}. This is a simple object, which we call

$$_{\perp}\epsilon^{jk} := n_i \epsilon^{ijk}. \tag{6.36}$$

This is just the two-dimensional Levi-Civita object in the plane perpendicular to n^i, which is the plane of the sky as seen by the observer. These quantities will be used in the current-quadrupole field, which contains projections on all the indices. Therefore, the only components of J^{jk} that enter are those projected onto the sky, and so it will simplify formulae to define the sky-projected moment of the angular momentum $_{\perp}J$

$$_{\perp}J^{ij} := {}_{\perp}^{\ i}{}_{l}\, {}_{\perp}^{\ j}{}_{m} J^{lm}. \tag{6.37}$$

Using this assembled notation, the current-quadrupole field is

$$h^{\mathrm{TT}ij} = \frac{4}{3r}({}_{\perp}\epsilon^{ik}\, {}_{\perp}\ddot{J}_k{}^{j} + {}_{\perp}\epsilon^{jk}\, {}_{\perp}\ddot{J}_k{}^{i} + {}_{\perp}^{\ ij}\, {}_{\perp}\epsilon^{km}\, {}_{\perp}\ddot{J}_{km}). \tag{6.38}$$

This is similar in form and complexity to the mass-quadrupole field expression. The interpretation of the contributions is direct. Only components of the angular momentum in the plane of the sky contribute to the field. Similarly only moments of this angular momentum transverse to the line of sight contribute. If one wants, say, the xx component of the field, then the $_{\perp}\epsilon$ factor tells us it is determined by the y-component of momentum, i.e. the component perpendicular to the x-direction in the sky. In fact, it is much simpler just to write out the actual components, assuming that the wave travels toward the observer along the z-direction. Then we have

$$h^{\mathrm{TT}xx} = \frac{4}{3r}(\ddot{J}^{xy} + \ddot{J}^{yx}), \tag{6.39}$$

$$h^{\mathrm{TT}xy} = \frac{4}{3r}(\ddot{J}^{yy} - \ddot{J}^{xx}), \tag{6.40}$$

and the remaining components can be found from the usual symmetries of the TT-metric. I have dropped the prefix $_{\perp}$ on J because in this coordinate system the given components are already transverse.

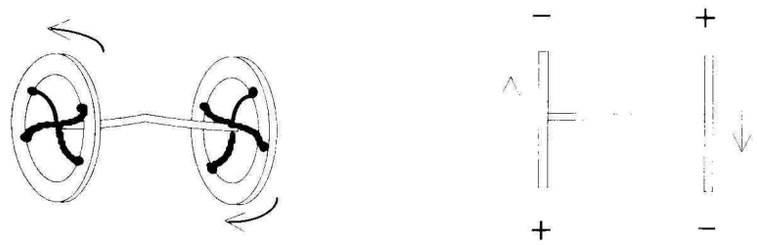

Figure 6.1. A simple current-quadrupole radiator. The left-hand panel shows how the two wheels are connected with blade springs to a central axis. The wheels turn in opposite directions, each oscillating back and forth about its rest position. The right-hand panel shows the side view of the system, and the arrows indicate the motion of the near side of the wheels at the time of viewing. The + signs indicate where the momentum of the mass of the wheel is toward the viewer and the − signs indicate where it is away from the viewer.

The simplicity of these expressions is striking. There are two basic cases where one gets current-quadrupole radiation.

- If there is an oscillating angular momentum distribution with a dipole moment along the angular momentum axis, as projected onto the sky, then in an appropriate coordinate system $\ddot{J}^{xx}$ will be nonzero and we will have $\otimes$ radiation. To have a non-vanishing dipole moment, the angular momentum density could, for example, be symmetrical under reflection through the origin along its axis, so that it points in opposite directions on opposite sides.
- If there is an oscillating angular momentum distribution with a dipole moment along an axis perpendicular to the angular momentum axis, as projected onto the sky, then in an appropriate coordiate system $\ddot{J}^{xy}$ will be nonzero and we will have $\oplus$ radiation.

6.3.3 A model system radiating current-quadrupole radiation

To see that the first of these two leads to physically sensible results, let us consider a simple model system that actually bears a close resemblance to the r-mode system. Imagine, as in the left panel of figure 6.1, two wheels connected by an axis, and the wheels are sprung on the axis in such a way that if a wheel is turned by some angle and then released, it will oscillate back and forth about the axis. Set the two wheels into oscillation with opposite phases, so that when one wheel rotates clockwise, the other rotates anticlockwise, as seen along the axis.

Then when viewed along the axis, the angular momentum has no component transverse to the line of sight, so there is no radiation along the axis. This

is sensible, because when projected onto the plane of the sky the two wheels are performing exactly opposite motions, so the net effect is that there is zero projected momentum density.

When viewed from a direction perpendicular to the axis, with the axis along the x-direction, then the angular momentum is transverse, and it has opposite direction for the two wheels. There is therefore an x-moment of the x-component of angular momentum, and the radiation field will have the $\otimes$ orientation.

To see that this has a physically sensible interpretation, look back again at the polarization diagram, figure 2.1, and look at the bottom row of figures illustrating the $\otimes$ polarization. See what the particles on the x-axis are doing. They are moving up and down in the y-direction. What motions in the source could be producing this?

At first one might guess that it is the up-and-down motion of the mass in the wheels as they oscillate, because in fact the near side of each wheel does exactly what the test particles at the observer are doing. However, this cannot be the explanation, because the far side of each wheel is doing the opposite, and when they both project onto the sky they cancel. What in fact gives the effect is that at the *top* of the wheel the momentum density is first positive (towards the observer) and then negative, while at the *bottom* of the wheel it is first negative and then positive. On the other wheel, the signs are reversed.

Current-quadrupole radiation is produced, at least in simple situations like the one we illustrate here, by (the second time-derivative derivative of) the component of source momentum along the line of sight. If this is positive in the sense that it is towards the observer, then the momentum density acts as a positive gravitational 'charge'. If negative, then it is a negative 'charge'. The wheels have an array of positive and negative spots that oscillates with time, and the test particles in the polarization diagram are drawn toward the positive ones and pushed away from the negative ones. Interestingly, in electromagnetism, magnetic dipole and magnetic quadrupole radiation are also generated by the component of the electric current along the line of sight.

This is a rather simple physical interpretation of some rather more complex equations. It is possible to re-write equation (6.38) to show explicitly the contribution of the line-of-sight momentum, but the expressions become even more complicated. Instead of dwelling on this, I will turn to the question of calculating the total energy radiated by the source.

6.4 Energy radiated in gravitational waves

We have calculated the energy flux in equation (5.14), and we now have the TT wave amplitudes. We need only integrate the flux over a distant sphere to get the total luminosity. We do this for the mass and current quadrupoles in separate sections.

6.4.1 Mass-quadrupole radiation

The mass-quadrupole radiation field in equation (6.26) must be put into the energy flux formula, and the dependence on the direction n^i can then be integrated over a sphere. It is not a difficult calculation, but it does require some angular integrals over multiple products of the vector n^i, which depends on the angular direction on the sphere. By symmetry, integrals of odd numbers of factors of n^i vanish. For even numbers of factors, the result is essentially determined by the requirement that after integration the result must be fully symmetric under interchange of any two indices and it cannot have any special directions (so it must depend only on the Kronecker delta $\delta^i{}_j$). The identities we need are

$$\int n^i n^j \, \mathrm{d}\Omega = \frac{4\pi}{3}\delta^{ij}, \tag{6.41}$$

$$\int n^i n^j n^k n^l \, \mathrm{d}\Omega = \frac{4\pi}{15}(\delta^{ij}\delta^{kl} + \delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}). \tag{6.42}$$

Using these, one gets the following simple formula for the total luminosity of mass-quadrupole radiation

$$\boxed{L_{\mathrm{gw}}^{\mathrm{mass}} = \tfrac{1}{5}\langle \dddot{\tilde{M}}^{jk}\, \dddot{\tilde{M}}_{jk}\rangle}. \tag{6.43}$$

Here we still preserve the angle brackets of equation (5.14), because this formula only makes sense in general if we average in time over one cycle of the radiation.

6.4.2 Current-quadrupole radiation

The energy radiated in the current quadrupole is nearly as simple to obtain as the mass-quadrupole formula. The extra factor of n^i in the radiation field makes the angular integrals longer, and requires two further identities:

$$\int n^i n^j n^k n^l n^p n^q \, \mathrm{d}\Omega = \frac{4\pi}{7}\delta^{(ij}\delta^{kl}\delta^{pq)}, \tag{6.44}$$

$$\begin{aligned}\epsilon^{ijk}\epsilon^{i'j'k'} &= \delta^{ii'}\delta^{jj'}\delta^{kk'} + \delta^{ij'}\delta^{jk'}\delta^{ki'} + \delta^{ik'}\delta^{ji'}\delta^{kj'} \\ &\quad - \delta^{ii'}\delta^{jk'}\delta^{kj'} - \delta^{ij'}\delta^{ji'}\delta^{kk'} - \delta^{ik'}\delta^{jj'}\delta^{ki'},\end{aligned} \tag{6.45}$$

where the round brackets indicate full symmetrization on all indices. The expression is simplest if we define

$$\tilde{J}^{jk} := \epsilon^{jlm}\tilde{P}_{lm}{}^{k} + \epsilon^{klm}\tilde{P}_{lm}{}^{j},$$

where

$$\tilde{P}^{kij} := P^{kij} - \tfrac{1}{3}\delta^{ij}P^{kl}{}_{l}.$$

The result of the integration of the flux formula over a distant sphere is [18, 25], in our notation,

$$L_{\mathrm{gw}}^{\mathrm{current}} = \tfrac{4}{5}\langle \dddot{\tilde{J}}^{jk}\, \dddot{\tilde{J}}_{jk}\rangle. \tag{6.46}$$

6.5 Radiation in the Newtonian limit

The calculation so far has been within the assumptions of linearized theory. Real sources are likely to have significant self-gravity. This means, in particular, that there will be a significant component of the source energy in gravitational potential energy, and this must be taken into account.

In fact a more realistic equation than equation (6.1) would be

$$\Box\bar{h}^{\alpha\beta} = -16\pi(T^{\alpha\beta} + t^{\alpha\beta}) \tag{6.47}$$

where $t^{\alpha\beta}$ is the stress-energy *pseudotensor* of gravitational waves. This is hard to work with: equation (6.47) is an implicit equation because $t^{\alpha\beta}$ depends on $\bar{h}^{\alpha\beta}$.

Fortunately, the formulae that we have derived are more robust than they seem. It turns out that the *leading order* radiation field from a Newtonian source has the same formula as in linearized theory. By leading order we mean the dominant radiation. If there is mass-quadrupole radiation, then the mass-octupole radiation from a Newtonian source will not be given by the formulae of the linearized theory. On the other hand, current-quadrupole and mass-quadrupole radiation can coexist, because they have different symmetries, so the work we have done here is generally applicable.

More details on how one calculates radiation to higher order in the Newtonian limit will be given in Blanchet's contribution in this book. This is particularly important for computing the radiation to be expected from coalescing binary systems, whose orbits become highly relativistic just before coalescence and which are, therefore, not well described by linearized theory.

Chapter 7

Source calculations

Now that we have the formulae for the radiation from a system, we can use them for some simple examples.

7.1 Radiation from a binary system

The most numerous sources of gravitational waves are binary stars systems. In just half an orbital period, the non-spherical part of the mass distribution returns to its original configuration, so the angular frequency of the emitted gravitational waves is twice the orbital angular frequency.

We shall calculate here the mass-quadrupole moment for two stars of masses m_1 and m_2, orbiting in the x–y plane in a circular orbit with angular velocity Ω, governed by Newtonian dynamics. We take their total separation to be R, which means that the orbital radius of mass m_1 is $m_2R/(m_1+m_2)$ while that of mass m_2 is $m_1R/(m_1+m_2)$. We place the origin of coordinates at the centre of mass of the system. Then, for example, the xx-component of M^{ij} is

$$\begin{aligned} M_{xx} &= m_1\left(\frac{m_2R\cos(\Omega t)}{m_1+m_2}\right)^2 + m_2\left(\frac{m_1R\cos(\Omega t)}{m_1+m_2}\right)^2 \\ &= \mu R^2\cos^2(\Omega t), \end{aligned} \tag{7.1}$$

where $\mu := m_1m_2/(m_1+m_2)$ is the reduced mass. By using a trigonometric identity and throwing away the part that does not depend on time (since we will use only time-derivatives of this expression) we have

$$M_{xx} = \tfrac{1}{2}\mu R^2\cos(2\Omega t). \tag{7.2}$$

By the same methods, the other nonzero components are

$$M_{yy} = -\tfrac{1}{2}\mu R^2\cos(2\Omega t), \quad M_{xy} = \tfrac{1}{2}\mu R^2\sin(2\Omega t).$$

This shows that the radiation will come out at twice the orbital frequency.

In this case the trace-free moment $\tilde{M}^{ij}$ differs from M^{ij} only by a constant, so we can use these values for M^{ij} to calculate the field and luminosity.

As an example of calculating the field, let us compute $\bar{h}^{\mathrm{TT}xx}$ as seen by an observer at a distance r from the system along the y-axis, i.e. lying in the plane of the orbit. We first need the TT part of the mass-quadrupole moment, from equation (6.27):

$$M^{\mathrm{TT}xx} = M^{xx} - \tfrac{1}{2}(M^{xx} + M^{zz}).$$

However, since $M^{zz} = 0$, this is just $M^{xx}/2$. Then from equation (6.28) we find

$$\bar{h}^{\mathrm{TT}xx} = -2\frac{\mu}{r}(R\Omega)^2 \cos[2\Omega(t-r)]. \tag{7.3}$$

Similarly, the result for the luminosity is

$$L_{\mathrm{gw}} = \tfrac{32}{5}\mu^2 R^4 \Omega^6. \tag{7.4}$$

The various factors in these two equations are not independent, because the angular velocity is determined by the masses and separations of the stars. When observing such a system, we cannot usually measure R directly, but we can infer Ω from the observed gravitational-wave frequency, and we may often be able to make a guess at the masses (we will see below that we can actually measure an important quantity about the masses). So we eliminate R using the Newtonian orbit equation

$$R^3 = \frac{m_1 + m_2}{\Omega^2}. \tag{7.5}$$

If in addition we use the gravitational-wave frequency $\Omega_{\mathrm{gw}} = 2\Omega$, we get

$$\bar{h}^{\mathrm{TT}xx} = -2^{1/3}\frac{\mathcal{M}^{5/3}\Omega_{\mathrm{gw}}^{2/3}}{r}\cos[\Omega_{\mathrm{gw}}(t-r)], \tag{7.6}$$

$$L_{\mathrm{gw}} = \frac{4}{5\times 2^{1/3}}(\mathcal{M}\Omega_{\mathrm{gw}})^{\frac{10}{3}}, \tag{7.7}$$

where we have introduced the symbol for the *chirp mass* of the binary system:

$$\mathcal{M} := \mu^{3/5}(m_1 + m_2)^{2/5}.$$

Notice that both the field and the luminosity depend only on $\mathcal{M}$, not on the individual masses in any other combination.

The power represented by L_{gw} must be supplied by the orbital energy, $E = -m_1 m_2/2R$. By eliminating R as before we find the equation

$$E = -\frac{1}{2^{5/3}}\mathcal{M}^{5/3}\Omega_{\mathrm{gw}}^{2/3}.$$

This is remarkable because it too involves only the chirp mass $\mathcal{M}$. By setting the rate of change of E equal to the (negative of the) luminosity, we find an equation for the rate of change of the gravitational-wave frequency

$$\dot{\Omega}_{\mathrm{gw}} = \frac{12\times 2^{1/3}}{5}\mathcal{M}^{5/3}\Omega_{\mathrm{gw}}^{11/3}. \tag{7.8}$$

As we mentioned in chapter 4, since the frequency increases, the signal is said to 'chirp'.

These results show that the chirp mass is the only mass associated with the binary that can be deduced from observations of its gravitational radiation, at least if only the Newtonian orbit is important. Moreover, if one can measure the field amplitude (e.g. $h^{\mathrm{TT}xx}$) plus Ω_{gw} and $\dot{\Omega}_{\mathrm{gw}}$, one can deduce from these the value of $\mathcal{M}$ *and* the distance r to the system! *A chirping binary with a circular orbit, observed in gravitational waves,* is a standard candle: *one can infer its distance purely from the gravitational-wave observations.* To do this one needs the full amplitude, not just its projection on a single detector, so one generally needs a network of detectors or a long-duration observation with a single detector to get enough information.

It is very unusual in astronomy to have standard candles, and they are highly prized. For example, one can, in principle, use this information to measure Hubble's constant [26].

7.1.1 Corrections

In the calculation above we made several simplifying assumptions. For example, how good is the assumption that the orbit is circular? The Hulse–Taylor binary is in a highly eccentric orbit, and this turns out to enhance its gravitational-wave luminosity by more than a factor of ten, since the elliptical orbit brings the two stars much nearer to one another for a period of time than a circular orbit with the same period would do. So there are big corrections for this system.

However, systems emitting at frequencies observable from ground-based interferometers are probably well approximated by circular orbits, because they have arrived at their very close separation by gravitational-wave-driven in-spiral. This process removes eccentricity from the orbit faster than it shrinks the orbital radius, so by the time they are observed they have insignificant eccentricity.

Another assumption is that the orbit is well described by Newtonian theory. This is not a good assumption in most cases. Post-Newtonian orbit corrections will be very important in observations. This is not because the stars eventually approach each other closely. It is because they spend a long time at wide separations where the small post-Newtonian corrections accumulate systematically, eventually changing the phase of the orbit by an observable amount. So it is very important for observations that we match signals with a template containing high-order post-Newtonian corrections, as described in Blanchet's contribution. But even so, the information contained in the Newtonian part of the radiation is still there, so all our conclusions above remain important.

7.2 The *r*-modes

We consider rotating stars in Newtonian gravity and look at the effect that the emission of gravitational radiation has on their oscillations. One might expect

that the loss of energy to gravitational waves would damp out any perturbations, and indeed this is normally the case. However, it was a remarkable discovery of Chandrasekhar [27] that the opposite sometimes happens.

A rotating star is idealized as an axially symmetric perfect-fluid system. In the Newtonian theory the pulsations of a perturbed fluid can be described by normal modes which are the solutions of perturbed Euler and gravitational field equations. If the star is stable, the eigenfrequencies σ of the normal modes are real; if the star is unstable, there is at least one pair of complex-conjugate frequencies, one of which represents an exponentially growing mode and the other a decaying mode. (We take the convention that the time-dependence of a mode is $\exp(i\sigma t)$.)

In general relativity, the situation is, in principle, the same, except that there is a boundary condition on the perturbation equations that insists that gravitational waves far away be outgoing, i.e. that the star loses energy to gravitational waves. This condition forces all eigenfrequencies to be complex. The sign of the imaginary part of the frequency determines stability or instability.

The loss of energy to gravitional radiation can destabilize a star that would otherwise (i.e. in Newtonian theory) be stable. This is because it opens a pathway to lower-energy configurations that might not be accessible to the Newtonian star. This normally happens because gravitational radiation also carries away angular momentum, a quantity that is conserved in the Newtonian evolution of a perturbation.

The sign of the angular momentum lost by the star is a critical diagnostic for the instability. A wave that moves in the positive angular direction around a star will radiate positive angular momentum to infinity. A wave that moves in the opposite direction, as seen by an observer at rest far away, will radiate negative angular momentum. In a spherical star, both actions result in the damping of the perturbation because, for example, the positive-going wave has intrinsically positive angular momentum, so when it radiates its angular momentum decreases and so its amplitude decreases. Similarly, the negative-going wave has negative angular momentum, so when it radiates negative angular momentum its amplitude decreases.

The situation can be different in a rotating star, as first pointed out by Friedman and Schutz [28]. The angular momentum carried by a wave depends on its pattern angular velocity *relative to the star's angular velocity*, not relative to an observer far away. If a wave pattern travels backwards relative to the star, it represents a small effective slowing down of the star and therefore carries negative angular momentum. This can lead to an anomalous situation: if a wave travels backwards relative to the star, but forwards relative to an inertial observer (because its angular velocity relative to the star is smaller than the star's angular velocity), then it will have negative angular momentum but it will radiate positive angular momentum. The result will be that its intrinsic angular momentum will get more negative, and its amplitude will grow.

This is the mechanism of the Chandrasekar–Friedman–Schutz (CFS) instability. In an ideal star, it is always possible to find pressure-driven waves of short enough wavelength around the axis of symmetry (high enough angular eigenvalue m) that satisfy this condition. However, it turns out that even a small amount of viscosity can damp out the instability in such waves , so it is not clear that pressure-driven waves will ever be significantly unstable in realistic stars.

However, in 1997 Andersson [17] pointed out that there was a class of modes called r-modes (Rossby modes) that no-one had previously investigated, and that were formally unstable in all rotating stars. Rossby waves are well known in oceanography, where they play an important role in energy transport around the Earth's oceans. They are hard to detect, having long wavelengths and very low-density perturbations. They are mainly *velocity perturbations* of the oceans, whose restoring force is the Coriolis effect, and that is their character in neutron stars too. Because they have very small density perturbation, the gravitational radiation they emit is dominated by the current-quadrupole radiation.

For a slowly-rotating, nearly-spherical Newtonian star, the following velocity perturbation is characteristic of r-modes:

$$\delta v^a = \varsigma(r)\epsilon^{abc}\nabla_b r \nabla_c Y_{lm}, \tag{7.9}$$

where $\varsigma(r)$ is some function of r determined by the mode equations. This velocity is a curl, so it is divergence-free; since it has no radial component, it does not change the density. If the star is perfectly spherical, these perturbations are simply a small rotation of some of the fluid, and it continues to rotate. They have no oscillation, and have zero frequency.

If we consider a star with a small rotational angular velocity Ω, then the frequency σ is no longer exactly zero and a Newtonian calculation to first order in Ω shows that there is a mode with *pattern speed* $\omega_{\rm p} = -\sigma/m$ equal to

$$\omega_{\rm p} = \Omega\left[1 - \frac{2}{l\,(l+1)}\right]. \tag{7.10}$$

These modes are now oscillating currents that move (approximately) along the equipotential surfaces of the rotating star.

For $l \geqslant 2$, $\omega_{\rm p}$ is positive but slower than the speed of the star, so by the CFS mechanism these modes are unstable to the emission of gravitational radiation for an arbitrarily slowly rotating star.

The velocity pattern given in equation (7.9) for ($l = 2$, $m = 2$) is closely related to the wheel model we described for current-quadrupole radiation in figure 6.1. Take two such wheels and orient their axes along the x- and y-axes, with the star rotating about the z-axis. Choose the sense of rotation so that the wheels at positive-x and positive-y are spinning in the opposite sense at any time, i.e. so that their adjacent edges are always moving in the *same* direction. Then this relationship will be reproduced for all other adjacent pairs of wheels: adjacent edges move together.

When seen from above the equatorial plane, the line-of-sight momenta of the wheels reinforce each other, and we get the same kind of pattern that we saw

when looking at one wheel from the side. However, in this case the pattern rotates with the angular velocity $2\Omega/3$ of equation (7.10). Since the pattern of line-of-sight momenta repeats itself every half rotation period, the gravitational waves are circularly polarized with frequency $4\Omega/3$. Seen along the x-axis, the wheel along the x-axis contributes nothing, but the other wheel contributes fully, so the radiation amplitude in this direction is half that going out the rotation axis. Seen along a line at 45° to the x-axis, the line-of-sight momenta of the wheels on the front part of the star cancel those at the back, so there is no radiation. Thus, along the equator there is a characteristic series of maxima and zeros, leading to a standard $m = 2$ radiation pattern. This pattern also rotates around the star, but the radiation in the equator remains linearly polarized because there is only the $\otimes$ component, not the $\oplus$. Again, the radiation frequency is twice the pattern speed because the radiation goes through a complete cycle in half a wave rotation period.

This discussion cannot go into the depth required to understand the r-modes fully. There are many issues of principle: what happens beyond linear order in Ω; what happens if the star is described in relativity and not Newtonian gravity; what is the relation between r-modes and the so-called g-modes that can have similar frequencies; what happens when the amplitude grows large enough that the evolution is nonlinear; what is the effect of magnetic fields on the evolution of the instability? The literature on r-modes is developing rapidly. We have included references where some of the most basic issues are discussed [17, 18, 29–31], but the interested student should consult the current literature carefully.

7.2.1 Linear growth of the r-modes

We have seen how the r-mode becomes unstable when coupled to gravitational radiation, and now we turn to the practical question: is it important? This will depend on the balance between the growth rate of the mode due to relativistic effects and the damping due to viscosity.

When coupled to gravitational radiation and viscosity, the mode has a complex frequency. If we define $\Im(\sigma) := 1/\tau$, then τ is the characteristic damping time. When radiation and viscosity are treated as small effects, their contributions to the eigenfrequencies add, so we have that the total damping is given by

$$\frac{1}{\tau(\Omega)} = \frac{1}{\tau_{\rm GR}} + \frac{1}{\tau_{\rm v}}, \qquad \frac{1}{\tau_{\rm v}} = \frac{1}{\tau_{\rm s}} + \frac{1}{\tau_{\rm b}}, \tag{7.11}$$

where $1/\tau_{\rm GR}$, $1/\tau_{\rm v}$ are the contributions due to gravitational radiation emission and viscosity, and where the latter has been further divided between shear viscosity ($\frac{1}{\tau_{\rm s}}$) and bulk viscosity ($\frac{1}{\tau_{\rm b}}$).

If we consider a 'typical' neutron star with a polytropic equation of state $p = k\rho^2$ (for which k has been chosen so that a $1.5M_\odot$ model has a radius $R = 12.47$ km), and if we express the angular velocity in terms of the scale for

Table 7.1. Gravitational radiation and viscous timescales, in seconds. Negative values indicate instability, i.e. a growing rather than damping mode.

l	m	$\tau_{\rm gw}$ (s)	$p_{\rm gw}$	$\tau_{\rm bv}$ (s)	$p_{\rm bv}$	$\tau_{\rm sv}$ (s)
2	2	−20.83	5.93	9.3×10^{10}	1.77	2.25×10^{8}
3	3	−316.1	7.98	1.89×10^{10}	1.83	3.53×10^{7}

the approximate maximum speed $\sqrt{\pi G\bar{\rho}}$ and the temperature in terms of 10^9 K, then it can be shown that [30]

$$\frac{1}{\tau} = \frac{1}{\tau_{\rm gw}\left(\frac{1\ {\rm ms}}{P}\right)^{p_{\rm gw}} + \frac{1}{\tau_{\rm bv}}\left(\frac{1\ {\rm ms}}{P}\right)^{p_{\rm bv}}}\left(\frac{10^9\ {\rm K}}{T}\right)^6 + \frac{1}{\tau_{\rm sv}}\left(\frac{T}{10^9\ {\rm K}}\right)^2, \tag{7.12}$$

where the scaling parameters $\tilde{\tau}_{\rm sv}$, $\tilde{\tau}_{\rm bv}$, $\tilde{\tau}_{\rm gw}$ and the exponents $p_{\rm gw}$ and $p_{\rm bv}$ have to be calculated numerically. Some representative values relevant to the r-modes with $2 \leqslant l \leqslant 6$ are in table 7.1 [30].

The physics of the viscosity is interesting. It is clear from equation (7.12) that gravitational radiation becomes a stronger and stronger destabilizing influence as the angular velocity of a star increases, but the viscosity is much more complicated. There are two contributions: shear and bulk. Shear viscosity comes mainly from electrons scattering off protons and other electrons. This effect falls with increasing temperature, just as does viscosity of everyday materials. So a cold, slowly rotating star will not have the instability, where a hotter star might. However, at high temperatures, bulk viscosity becomes dominant. This effect arises in neutron stars from nuclear physics. Neutron-star matter always contains some protons and electrons. When it is compressed, some of these react to form neutrons, emitting a neutrino. When it is expanded, some of the neutrons beta-decay to protons and electrons, again emitting a neutrino. The emitted neutrino is not trapped in the star; within a short time, of the order of a second or less, it escapes. This irreversible loss of energy each time the star is compressed creates a bulk viscosity. Now, bulk viscosity acts only due to the density perturbation, which is small in r-modes. So the effect of bulk viscosity only dominates at very high temperatures.

The balance of the viscous and gravitational effects is illustrated in figure 7.1 [30]. This is indicative, but not definitive: much more work is needed on the physics of viscosity and the structure of the modes at large values of Ω (small P).

7.2.2 Nonlinear evolution of the star

Our description so far is only a linear approximation. To understand the full evolution of the r-modes we have to treat the nonlinear hydrodynamical effects

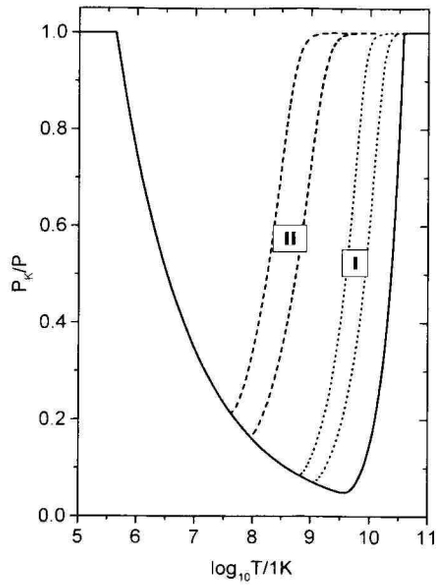

Figure 7.1. The balance of viscous and gravitational radiation effects in the r-modes is illustrated in a diagram of rotation speed, showing the ratio of the maximum period P_K to the rotation period P versus the temperature of the star. The solid curve indicates the boundary between viscosity-dominated and radiation-dominated behaviour: stars above the line are unstable. The dashed curves illustrate possible nonlinear evolution histories as a young neutron star cools.

that become important as the modes grow. This could only be done with a numerical simulation, which some groups are now working on. However, it is possible to make simple estimates analytically.

We characterize the initial configuration with just two paramters: the uniform angular velocity Ω, and the amplitude α of the r-modes perturbation. The star is assumed to cool at the accepted cooling rate for neutron stars, independently of whether it is affected by the r-mode instability or not. The star is assumed to lose angular momentum to gravitational radiation at a rate given by the linear radiation field, with its large amplitude α. This loss is taken to drive the star through a sequence of equilibrium states of lower and lower angular momentum. Details of this approximation are in [31], here we report only the results. The

evolution turns out to have three phases.

(a) Initially the angular velocity Ω of the hot rapidly rotating neutron star is nearly constant, evolving on the viscous timescale $1/\tau_v$, while the amplitude α grows exponentially on the gravitational radiation timescale $1/\tau_{GR}$.

(b) After a short time nonlinear effects become important and stop the growth of the amplitude α. Most of the initial angular momentum of the star is radiated away by gravitational radiation. The star spins down and evolves to a point where the angular velocity Ω and the temperature is sufficiently low that the r-mode is stable.

(c) Finally gravitational radiation and viscosity damp out the r-mode and drive the star into its final equilibrium configuration.

This may take about a year, a timescale governed by the cooling time of the star. During this year, the star would radiate away most of its angular momentum and rotational kinetic energy. This could be a substantial fraction of a solar mass in energy.

7.2.3 Detection of r-mode radiation

The large amount of energy radiated into the r-modes makes them attractive for detection, but detection will not be trivial. The r-mode event occurs at the rate of supernovae: some fraction (hopefully large) of all supernovae leave behind a rapidly spinning neutron star that spins down over a one-year period. This means we should have sufficient sensitivity to reach the Virgo Cluster (20 Mpc distance). Estimates [31] suggest that a neutron star in the Virgo Cluster could be detected by second generation of LIGO and VIRGO gravitational-wave detectors with an amplitude signal-to-noise of about eight, provided one can use matched filtering (exact template matching).

It will not be easy to use matched filtering, since one must follow all cycles of the signal as the star spins down, and we will not know this well because of many uncertainties: initial temperature, initial spin distribution, detailed physics of viscosity, and so on. However, it would be helpful to have a parametrized model to take account of the uncertainties, so that we could look for a significant fit to one or more of the parameters.

In addition, it is likely that, if a significant proportion of all neutron stars went through the r-mode instability, then the universe has been filled by their radiation. There should be a background with an energy density Ω_{gw} that is a good fraction of the closure density. Its lower frequency limit should be around 200 Hz in the rest frame of the star. When we see radiation cosmologically, its lower frquency limit will indicate the epoch at which star formation began.

It is clear that the discovery of this new source of gravitational waves will open several prospects for astronomy. Observations could be used as supernovae detectors, revealing supernovae hidden in clouds of dust, identifying them about a year after they are formed. The existence of the radiation raises several prospects

and questions about the physics of neutron stars, not least the interaction of magnetic fields with the instability.

7.3 Conclusion

These lectures (chapters 2 to 7) have taken us through the basic theory of gravitational radiation and its applications in astrophysics, so far as we can understand and predict them now. In a few years, perhaps as little as two, perhaps as many as eight, we will start to make observations of gravitational radiation from astrophysical sources. If gravitational-wave astronomy follows other branches of observational astronomy, it will not be long before completely unexpected signals are seen, or unexpected features in long-predicted signals. To interpret these will require joining a physical understanding of the relationship between gravitational radiation and its source to a wide knowledge of astronomical phenomena. We encourage the students who have attended these lectures, and others who may study them, to get themselves ready to contribute to this activity. It will be an exciting time!

References

We have divided the references into two sections. The first gives generally useful references—books, conference summaries, etc—that interested students should go to for background and a more complete discussion of the theory. The second section contains specific references to the research literature. General references are indicated in the text by the author name plus year, as Misner *et al* (1973). The specific references are indicated by numbers, for example [1,3].

General references

General relativity

There are a number of good text books on GR. The following cover, at different levels of difficulty and completeness, linearized theory, gauges and the definition of energy:

Ciufolini I and Wheeler J L 1995 *Gravitation and Inertia* (Princeton, NJ: Princeton University Press)

Landau L and Lifshitz E M 1962 *The Classical Theory of Fields* (New York: Pergamon)

Misner C W, Thorne K S and Wheeler J L 1973 *Gravitation* (San Francisco, CA: Freeman)

Ruffini R and Ohanian H C 1997 *Gravitazione e Spazio-Tempo* (Bologna: Zanichelli)

Schutz B F 1995 *A First Course in General Relativity* (Cambridge: Cambridge University Press)

Wald R M 1994 *General Relativity* (Chicago, IL: Chicago University Press)

Weinberg S 1972 *Gravitation and Cosmology* (New York: Wiley)

Gravitational-wave detectors

Conference volumes on detector progress appear more than once per year these days. You can find progress reports on detectors on the web sites of the different groups, which you will find in the list of literature references below. The two references below are more tutorial, aimed at introducing the subject.

Blair D G 1991 *The Detection of Gravitational Waves* (Cambridge: Cambridge University Press)

Saulson P R 1994 *Fundamentals of Interferometric Gravitational Wave Detectors* (Singapore: World Scientific)

Sources of gravitational waves

Again, there have been a number of conference publications on this subject. The third and fourth references are recent reviews of sources. The first two survey the problem of data analysis.

Schutz B F (ed) 1989 *Gravitational Wave Data Analysis* (Dordrecht: Kluwer)

——1997 The detection of gravitational waves *Relativistic Gravitation and Gravitational Radiation* ed J A Marck and J P Lasota (Cambridge: Cambridge University Press) pp 447–75

Thorne K S 1987 Gravitational radiation *300 Years of Gravitation* ed S W Hawking and W Israel (Cambridge: Cambridge University Press) pp 330–458

——1995 *Proc. 1994 Summer Study on Particle and Nuclear Astrophysics and Cosmology* ed E W Kolb and R Peccei (Singapore: World Scientific)

Text references

[1] Schutz B F 1984 Gravitational waves on the back of an envelope *Am. J. Phys.* **52** 412–19

[2] Nicholson D *et al* 1996 Results of the first coincident observations by 2 laser-interferometric gravitational-wave detectors *Phys. Lett.* A **218** 175–80

[3] Meers B J 1988 Recycling in laser-interferometric gravitational wave detectors *Phys. Rev.* D **38** 2317–26

[4] TAMA300 project website, http://tamago.mtk.nao.ac.jp/

[5] GEO600 project website, http://www.geo600.uni-hannover.de/

[6] LIGO project website, http://ligo.caltech.edu/

[7] VIRGO project website, http://virgo4p.pg.infn.it/virgo/

[8] Caves C M *et al* 1980 *Rev. Mod. Phys.* **52** 341–92

[9] Astone P, Lobo J A and Schutz B F 1994 Coincidence experiments between interferometric and resonant bar detectors of gravitational waves *Class. Quantum Grav.* **11** 2093–112

[10] Rome gravitational wave group website, http://www.roma1.infn.it/rog/rogmain.html

[11] Auriga detector website, http://axln01.lnl.infn.it/

[12] LISA project websites, http://www.lisa.uni-hannover.de/ and http://lisa.jpl.nasa.gov/

[13] Bender P *et al LISA: Pre-Phase A Report* MPQ 208 (Max-Planck-Institut für Quantenoptik, Garching, Germany) (also see the 2nd edn, July 1998)

[14] Lipunov V M, Postnov K A and Prokhorov M E 1997 Formation and coalescence of relativistic binary stars: the effect of kick velocity *Mon. Not. R. Astron. Soc.* **288** 245–59

[15] Yungelson L R and Swart S F P 1998 Formation and evolution of binary neutron stars *Astron. Astrophys.* **332** 173–88

[16] Zwart S F P and McMillain S L W 2000 Black hole mergers in the universe *Astrophys. J.* **528** L17–20 (gr-qc/9910061)

[17] Andersson N 1998 A new class of unstable modes of rotating relativistic stars *Astrophys. J.* **502** 708–13 (gr-qc/9706075)

[18] Lindblom L, Owen B J and Morsink S M 1998 Gravitational radiation instability in hot young neutron stars *Phys. Rev. Lett.* **80** 4843–6 (gr-qc 9803053)

[19] van der Klis M 1998 *The Many Faces of Neutron Stars* ed R Buccheri, J van Paradijs and M A Alpar (Dordrecht: Kluwer Academic)

[20] Bildsten L 1998 Gravitational radiation and rotation of accreting neutron stars *Astrophys. J.* **501** L89–93 (astro-ph/9804325)
See also the recent paper Ushomirsky G, Cutler C and Bildsten L 1999 Deformations of accreting neutron star crusts and gravitational wave emission *Preprint* (astro-ph/0001136)

[21] Brady P R, Creighton T, Cutler C and Schutz B F 1998 Searching for periodic sources with LIGO *Phys. Rev.* D **57** 2101–16

[22] Brustein R, Gasperini M, Giovannini M and Veneziano G 1996 Gravitational radiation from string cosmology *Proc. Int. Europhysics Conf. (HEP 95)* ed C Van der Velde, F Verbeure and J Lemonne (Singapore: World Scientific)

[23] Schutz B F and Sorkin R 1977 Variational aspects of relativistic field theories, with application to perfect fluids *Ann. Phys., NY* **107** 1–43

[24] Isaacson R *Phys. Rev.* **166** 1263–71
Isaacson R 1968 *Phys. Rev.* **166** 1272–80

[25] Thorne K S 1980 Multipole expansions of gravitational radiation *Rev. Mod. Phys.* **52** 299–340

[26] Schutz B F 1986 Determining the Hubble constant from gravitational wave observations *Nature* **323** 310–11

[27] Chandrasekar S 1970 *Phys. Rev. Lett.* **24** 611

[28] Friedman J L and Schutz B F 1978 *Ap. J.* **222** 281

[29] Friedman J L and Morsink S M 1998 Axial instability of rotating relativistic stars *Astrophys. J.* **502** 714–20 (gr-qc/9706073)

[30] Andersson N, Kokkotas K and Schutz B F 1999 Gravitational radiation limit on the spin of young neutron stars *Astrophys. J.* **510** 846–8539 (astro-ph/9805225)

[31] Owen B, Lindblom L, Cutler C, Schutz B F, Vecchio A and Andersson N 1998 Gravitational waves from hot young rapidly rotating neutron stars *Phys. Rev.* D **5808** 4020 (gr-qc/9804044)

Solutions to exercises

Chapter 2

Exercise (a)

1. Let us take the form of the wave to be

$$h^{\mathrm{TT}jk} = e^{jk}_{\oplus} h_+(t - \hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{x}})$$

where $e^{jk}_{\oplus}$ is the polarization tensor for the $\oplus$ polarization, and where $\hat{\boldsymbol{n}}$ is the unit vector in the direction of travel of the wave. We will let h_+ be an arbitrary function of its phase argument.

If the wave travels in the x–z plane at an angle θ to the z-direction, then the unit vector in our coordinates is

$$\hat{n}^i = (\sin\theta, 0, \cos\theta).$$

We need to calculate the polarization tensor's components in x, y, z coordinates. We do this by rotating the $\oplus$ polarization tensor from its TT form in coordinates parallel to the wavefront to its form in our coordinates. This requires a simple rotation around the y-axis. The transformation matrix is:

$$\Lambda^{j'}{}_{k} = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}.$$

The polarization tensor in our coordinates (primed indices) becomes:

$$\begin{aligned} e^{j'k'} &= \Lambda^{j'}{}_{l}\Lambda^{k'}{}_{m} e^{lm} \\ &= \begin{pmatrix} \cos^2\theta & 0 & -\sin\theta\cos\theta \\ 0 & -1 & 0 \\ -\sin\theta\cos\theta & 0 & \sin^2\theta \end{pmatrix}. \end{aligned}$$

Notice that the new polarization tensor is again traceless.

The gravitational wave will be, at an arbitrary time t and position (x, z) in our (x, z)-plane,

$$h^{\mathrm{TT}j'k'} = e^{j'k'} h_+(t - x\sin\theta - z\cos\theta).$$

For this problem we need the $x'x'$-component because the photon is propagating along this direction, and we will always stay at $z = 0$, so we have

$$h^{\mathrm{TT}x'x'} = \cos^2\theta h_+(t - x\sin\theta).$$

We see that for this geometry the wave amplitude is reduced by a factor of $\cos^2\theta$.

Generalizing the argument in the text, the relation between time and position for the photon on its trip outwards along the x-direction is $t = t_0 + x$, where t_0 is the starting time. The analogous relation after the photon is reflected is $t = t_0 + L + (L - x)$, since in this case x decreases in time from L to 0. If we put these into the equation for the linearized corrections to the return time, we get

$$\begin{aligned} t_{\mathrm{return}} = t_0 + 2L + \tfrac{1}{2}\cos^2\theta \Big\{ &\int_0^L h_+[t_0 + (1 - \sin\theta)x]\,\mathrm{d}x \\ &\times \int_0^L h_+[t_0 + 2L - (1 + \sin\theta)x]\,\mathrm{d}x \Big\}. \end{aligned}$$

This expression must be differentiated with respect to t_0 to find the variation of the return time as a function of the start time. The key point is how to handle differentiation within the integrals. Consider, for example, the function $h_+[t_0 + (1 - \sin\theta)x]$. It is a function of a single argument,

$$\xi := t_0 + (1 - \sin\theta)x$$

so derivatives with respect to t_0 can be converted to derivatives with respect to x as follows

$$\begin{aligned} \frac{\mathrm{d}h_+}{\mathrm{d}t_0} &= \frac{\mathrm{d}h_+}{\mathrm{d}\xi}\frac{\mathrm{d}\xi}{\mathrm{d}t_0} = \frac{\mathrm{d}h_+}{\mathrm{d}\xi}; \\ \frac{\mathrm{d}h_+}{\mathrm{d}x} &= \frac{\mathrm{d}h_+}{\mathrm{d}\xi}\frac{\mathrm{d}\xi}{\mathrm{d}x} = (1 - \sin\theta)\frac{\mathrm{d}h_+}{\mathrm{d}\xi}. \end{aligned}$$

It follows that

$$\frac{\mathrm{d}h_+}{\mathrm{d}t_0} = \frac{\mathrm{d}h_+}{\mathrm{d}x}/(1 - \sin\theta).$$

On the return trip the factor will be $-(1 + \sin\theta)$. So when we differentiate we can convert the derivatives with respect to t_0 inside the integrals into derivatives with respect to x. Taking account of the factor $\cos^2\theta = (1 - \sin\theta)(1 + \sin\theta)$ in front of the integrals, the result is

$$\begin{aligned} \frac{\mathrm{d}t_{\mathrm{return}}}{\mathrm{d}t_0} = 1 &+ \frac{1}{2}(1 + \sin\theta)\int_0^L \frac{\mathrm{d}h_+}{\mathrm{d}x}[t_0 + (1 - \sin\theta)x]\,\mathrm{d}x \\ &+ \frac{1}{2}(1 - \sin\theta)\int_0^L \frac{\mathrm{d}h_+}{\mathrm{d}x}[t_0 + 2L - (1 + \sin\theta)x]\,\mathrm{d}x. \end{aligned}$$

The integrals can now be done, since they simply invert the differentiation by x. Evaluating the integrands at the end points of the integrals gives equation (2.20):

$$\frac{\mathrm{d}t_{\mathrm{return}}}{\mathrm{d}t_0} = 1 - \tfrac{1}{2}(1+\sin\theta)h_+(t_0) + \sin\theta h_+[t_0 + (1-\sin\theta)L] + \tfrac{1}{2}(1-\sin\theta)h_+(t_0+2L).$$

2. If we Taylor-expand this equation in powers of L about $L = 0$, the leading term vanishes, and the first-order term is:

$$\frac{\mathrm{d}t_{\mathrm{return}}}{\mathrm{d}t_0} = L\sin\theta(1-\sin\theta)\dot{h}_+(t_0) + L(1-\sin\theta)\dot{h}_+(t_0),$$
$$= L\cos^2\theta\dot{h}_+(t_0).$$

This is just what was required. The factor of $\cos^2\theta$ comes, as we saw above, from the projection of the TT field on the x-coordinate direction.

3. All the terms cancel and there is no effect on the return time.

Exercise (b)

This is part of the calculation in the previous example. All we need is the segment where the light travels from the distant end to the centre:

$$t_{\mathrm{out}} = t_0 + \tfrac{1}{2}\cos^2\theta\int_0^L h_+[t_0 + L - (1+\sin\theta)x]\,\mathrm{d}x$$

and so $\mathrm{d}t_{\mathrm{out}}/\mathrm{d}t_0$ is:

$$\frac{\mathrm{d}t_{\mathrm{out}}}{\mathrm{d}t_0} = 1 + \tfrac{1}{2}(1-\sin\theta)[h_+(t_0 - \sin\theta L) - h_+(t_0+L)].$$

Exercise (c)

This question is frequently asked, but not by people who have done the calculation. The answer is that the two effects occur in different gauges, not in the same one. So they cannot cancel. The apparent speed of light changes in the TT gauge, but then the positions of the ends remain fixed, so that the effect is all in the coordinate speed. In a local Lorentz frame tied to one mass, the ends do move back and forth, but then the speed of light is invariant.

Exercise (d)

To first order we have

$$R^{\mu}_{\alpha\beta\nu} = \Gamma^{\mu}_{\alpha\nu,\beta} - \Gamma^{\mu}_{\alpha\beta,\nu}.$$
$$\Gamma^{\mu}_{\alpha\beta,\nu} = \tfrac{1}{2}\eta^{\mu\sigma}(h_{\sigma\beta,\alpha\nu} + h_{\sigma\alpha,\beta\nu} - h_{\alpha\beta,\sigma\nu}). \quad \text{(i)} \qquad (7.13)$$
$$\Gamma^{\mu}_{\alpha\nu,\beta} = \tfrac{1}{2}\eta^{\mu\sigma}(h_{\sigma\nu,\alpha\beta} + h_{\alpha\sigma,\nu\beta} - h_{\alpha\nu,\sigma\beta}). \quad \text{(ii)} \qquad (7.14)$$

The gauge transformation for a perturbation in linearized theory is

$$h'_{\alpha\beta} = h_{\alpha\beta} - \xi_{\alpha,\beta} - \xi_{\beta,\alpha}. \quad \text{(iii)} \qquad (7.15)$$

Substituting (iii) into (i) and (ii), we obtain

$$\text{(i)} = \tfrac{1}{2}\eta^{\mu\sigma}(h'_{\sigma\beta,\alpha\nu} + \xi_{\sigma,\beta\alpha\nu} + h'_{\sigma\alpha,\beta\nu} + \xi_{\sigma,\alpha\beta\nu} - h'_{\alpha\beta,\sigma\nu})$$
$$\text{(ii)} = \tfrac{1}{2}\eta^{\mu\sigma}(h'_{\sigma\nu,\alpha\beta} + \xi_{\sigma,\nu\alpha\beta} + h'_{\sigma\alpha,\beta\nu} + \xi_{\sigma,\alpha\beta\nu} - h'_{\alpha\nu,\sigma\beta}).$$

If we find the difference between the two formulae above we get

$$\underline{R^{\mu}_{\alpha\beta\nu}} = \text{(ii)} - \text{(i)} = \Gamma'^{\mu}_{\alpha\nu,\beta} - \Gamma'^{\mu}_{\alpha\beta,\nu} = \underline{R'^{\mu}_{\alpha\beta\nu}}.$$

Chapter 5

Exercise (e)

The action principle is:

$$\delta I = \int \frac{\delta(R\sqrt{-g})}{\delta g_{\mu\nu}}h_{\mu\nu}\,\mathrm{d}^4x = -\int G^{\mu\nu}\sqrt{-g}h_{\mu\nu}\,\mathrm{d}^4x = 0. \quad \text{(i)} \qquad (7.16)$$

If we perform an infinitesimal coordinate transformation $x^{\mu} \to x^{\mu} + \xi^{\mu}$ without otherwise varying the metric, then the action I must not change:

$$0 = \delta I = \int G^{\mu\nu}(\xi_{\mu;\nu} + \xi_{\nu;\mu})\sqrt{-g}\,\mathrm{d}^4x$$
$$= 2\int G^{\mu\nu}\xi_{\mu;\nu}\,\mathrm{d}^4x.$$

This can be transformed in the following way:

$$\delta I = \int (G^{\mu\nu}\xi_{\mu})_{;\nu}\sqrt{-g}\,\mathrm{d}^4x - \int (G^{\mu\nu}{}_{;\nu}\xi_{\mu})\sqrt{-g}\,\mathrm{d}^4x = 0.$$

The first integral is a divergence and vanishes. The second, because of the arbitrariness of ξ_{μ}, gives the Bianchi's identities:

$$G^{\mu\nu}{}_{;\nu} = 0.$$

Exercise (f)

The two polarization components are $h^{xx}_{+} = -h^{yy}_{+} = A_+\mathrm{e}^{-\mathrm{i}k(t-z)}$ and $h^{xy}_{\times} = A_{\times}\mathrm{e}^{-\mathrm{i}k(t-z)}$. The energy flux is the negative of

$$\langle T^{(\mathrm{GW})}_{0z}\rangle = \frac{1}{32\pi}\langle h^{ij}{}_{,0}h_{ij,z}\rangle = -\frac{k^2}{16\pi}(A_+^2 + A_{\times}^2)\langle\sin^2 k(t-z)\rangle$$
$$= -\frac{k^2}{32\pi}(A_+^2 + A_{\times}^2).$$