# Gravitational Waves

*Volume 1*

## Theory and Experiments

Michele Maggiore
*Département de Physique Théorique*
*Université de Genève*

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York



First published 2008



British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

Printed in Great Britain
on acid-free paper by
Anthony Rowe, Cippenham

ISBN 978–0–19–857074–5 (Hbk)

10 9 8 7 6 5 4 3 2 1

*A Maura, Sara, Ilaria e Lorenzo*

# Contents

# 7 Data analysis techniques



In this chapter we begin our study of GW experiments. The functioning principles and the sensitivities of existing or planned detectors will be examined in great detail in the subsequent chapters. In this chapter we rather introduce a number of general concepts which characterize any GW detector, and we discuss the crucial problem of how to extract a GW signal from the (typically much larger) detector noise.

In Section 7.1 we will see how the various noise generated inside a the detector can be conveniently treated referring them to the detector input, and are characterized by a *spectral strain sensitivity*, which has dimensions $1/\sqrt{\text{Hz}}$. In Section 7.2 we introduce the pattern functions that encode the detector angular sensitivity. We will then discuss in Section 7.3 the optimum filtering techniques that must be applied to the detector output. The importance of this procedure stems from the fact that, with existing detectors and with reasonable estimates of the GW signal, we expect that the GW signal will be buried into a much larger noise. The fact that we try to extract a small signal from noisy detectors is certainly not a new situation in physics. Rather on the contrary, it is a typical problem in many fields, e.g. in radio engineering where it has been much studied in connections with radars, or in radio astronomy for application to pulsar searches, and standard filtering techniques have been developed. We will see how these techniques are adapted to the problem of GW detection. The proper interpretation of the results obtained with matched filtering relies on notions of probability and statistics, that we discuss in Section 7.4. Here, after an introduction to the frequentist and the Bayesian frameworks, we discuss how to reconstruct the parameters of the source and how to examine the statistical significance of the observation of an event with a given signal-to-noise ratio. Then, in Sections 7.5–7.8, we will examine the application of these concepts to various classes of GW signals, i.e. bursts, periodic signals, coalescing binaries and stochastic backgrounds.

## 7.1 The noise spectral density

The output of any GW detector is a time series, which describes for instance the oscillation state of a resonant mass, or the phase shift of the light recombined after traveling in the two arms of an interferometer. This output will be a combination of a true GW signal (hopefully) and of noise. To understand how signal and noise combine, it is useful to think of a GW detector as a linear system. At its input there is the GW signal that we want to detect. More precisely, the input and output of the detector are scalar quantities, while the GW is described by a tensor $h_{ij}$. So, in general, the input of the detector will have the form

$$h(t) = D^{ij} h_{ij}(t)\,, \tag{7.1}$$

where $D^{ij}$ is a constant tensor which depends on the detector geometry, and is known as the *detector tensor*. For example, for a detector which is driven only by the $(x, x)$ component of $h_{ij}$ (which, as we will see, is the case for a resonant bar oriented along the $x$ axis), $D^{ij} = 1$ if $i = j = 1$ and $D^{ij} = 0$ otherwise. We will later compute the explicit form of $D^{ij}$ for interferometers and for resonant masses.

For a linear system, the output of the detector is a linear function, in frequency space, of the input $h(t)$, that is, the output $h_{\rm out}(t)$ of the detector (in the absence of noise) is related to the input $h(t)$ by

$$\tilde{h}_{\rm out}(f) = T(f)\tilde{h}(f)\,, \tag{7.2}$$

where $T(f)$ is known as the *transfer function* of the system. However, in the output of any real detector there will also be noise, so the output $s_{\rm out}(t)$ will be rather given by

$$s_{\rm out}(t) = h_{\rm out}(t) + n_{\rm out}(t)\,. \tag{7.3}$$

More precisely, a detector can be modeled as a linear system with many stages, labeled by $i = 1, \ldots, N$, each one with its own transfer function $T_i(f)$, so the total transfer function is $T(f) = \prod_i T_i(f)$. For example, we will see in Chapter 8 that resonant-bar detectors are composed of a heavy aluminum cylinder which is set into oscillation by an incoming GW; its energy is then transferred to a lighter mechanical oscillator, coupled to the heavy bar, which works as a mechanical amplifier, then it is transformed into an electric signal by an LC circuit coupled to the light oscillator, and then this electric signal is further amplified by one or more SQUIDs, and recorded. Clearly, noise can be generated at each of these stages. Each noise will propagate to the output with a transfer function which depends on the point of the linear system at which it first appeared, see Fig. 7.1, and will contribute to total noise $n_{\rm out}(t)$ at the output. It is convenient to refer each noise to the detector input, defining the quantity $n(t)$ from

$$\tilde{n}(f) = T^{-1}(f)\tilde{n}_{\rm out}(f)\,, \tag{7.4}$$

where $n_{\rm out}(t)$ is the total noise measured at the output. That is, $n(t)$ is a fictitious noise that, if it were injected at the detector input, and if there were no other noise inside the detector, would produce at the output the noise $n_{\rm out}(t)$ that is actually observed. It is therefore the quantity that we can compare directly with $h(t)$, i.e. to the effect due to the GW. We then define

$$s(t) = h(t) + n(t)\,, \tag{7.5}$$

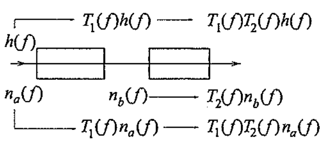


**Fig. 7.1** A schematic representation of a detector as a linear system. The full transfer function $T(f)$ is the product of the separate transfer function. Here $T(f) = T_1(f)T_2(f)$, and $\tilde{n}_{\rm out}(f) = T_1(f)T_2(f)n_a(f) + T_2(f)n_b(f)$.



and we can simply think of the detector as if $s(t)$ were its output, composed of a noise $n(t)$ and a GW signal $h(t)$,[1] and the detection problem is how to distinguish $h(t)$ from $n(t)$. In the following, when we speak of the detector output, we will always refer to $s(t)$.[2] If one has a theoretical model for a given source of noise $n_i(t)$, which appears at a given stage of the linear system, we can compare it with $h(t)$ simply multiplying it by the inverse of the appropriate transfer function, in order to refer this noise to the detector input. Equivalently, of course, one could refer both the noise and the signal to the true detector output, and compare $n_{\rm out}(t)$ to the quantity $h_{\rm out}(t)$ whose Fourier transform is given by eq. (7.2). However, the great advantage of referring everything to the input is that $n(t)$ gives a measure of the minimum value of $h(t)$ that can be detected and $h(t)$, apart from the geometrical factor $D^{ij}$ which is always of order one, depends only on the incoming GW. In contrast, $h_{\rm out}(t)$ depends on the transfer function of the system, and different detectors can have transfer functions which differ by many orders of magnitude. Thus, the use of $n_{\rm out}(t)$ and $h_{\rm out}(t)$ would be very unpractical when we want to compare the performances of different detectors.

So, in the above sense, we take $n(t)$ to be the detector's noise. If the noise is stationary, as we assume for the moment, the different Fourier components are uncorrelated, and therefore the ensemble average[3] of the Fourier components of the noise is of the form

$$\langle \tilde{n}^*(f)\tilde{n}(f')\rangle = \delta(f-f')\frac{1}{2}S_n(f)\,. \tag{7.6}$$

The above equation defines the function $S_n(f)$. Since $n(t)$ is real, $\tilde{n}(-f) = \tilde{n}^*(f)$ and therefore $S_n(-f) = S_n(f)$. If $n(t)$ is dimensionless, as we will assume, $S_n(f)$ has dimensions $\mathrm{Hz}^{-1}$. Without loss of generality, we can also assume that

$$\langle n(t)\rangle = 0\,. \tag{7.7}$$

Observe that, for $f = f'$, the right-hand side of eq. (7.6) diverges. However, in any real experiment we have a finite value of the time $T$ used to measure $\tilde{n}(f)$, see Note 3. Restricting the time interval to $-T/2 < t < T/2$ we have

$$\delta(f=0) \to \left[\int_{-T/2}^{T/2} dt\, e^{i2\pi ft}\right]\Bigg|_{f=0} = T\,. \tag{7.8}$$

Then, from eq. (7.6) with $f = f'$, we get

$$\langle\, |\tilde{n}(f)|^2\, \rangle = \frac{1}{2}S_n(f)T\,. \tag{7.9}$$

For a function defined on the interval $[-T/2, T/2]$, the Fourier modes have discrete frequencies $f_n = n/T$, so the resolution in frequency is given by

$$\Delta f = \frac{1}{T}\,. \tag{7.10}$$

[1] One often multiplies the detector output by $T^{-1}(f)$ already at the level of data acquisition, so in this sense $s(t)$ is really the output of the data acquisition system.

[2] Some more nomenclature: we will always use the word "event" to indicate that in the detector happened something, which deserves further scrutiny. At this stage, it could be due to a GW or (much more likely) to noise. An event which is already assumed to have been generated by a GW will be called a "GW signal". The letter $s$ conventionally used to denote the detector output $s(t) = h(t) + n(t)$ does not stand for "signal" (the signal in this nomenclature is $h(t)$). It can rather be taken to denote the "strain amplitude" of the detector.

[3] The ensemble average is the average over many possible "realizations" of the system. In practice we have only one physical system, our detector, but we can follow it in time, so the ensemble average is replaced by a time average (this implicitly assumes that the system is ergodic). Then the ensemble average is computed measuring the noise $n(t)$ over a given time interval $T$, and considering this as a "realization" of the system. From this we obtain $\tilde{n}(f)$ (with a resolution in frequency $\Delta f = 1/T$). We then repeat the procedures over a subsequent time stretch, again of duration $T$ and separated by a sufficient time shift from the first realization, so that the correlation between the noise $n(t)$ in the two stretches can be neglected, and we define this as a second independent realization of the system. Finally, we average $\tilde{n}(f)$ over many independent realizations. It is useful to keep in mind that a time-scale $T$ is implicit in this procedure, and will indeed appear in the equations below.

We can then write eq. (7.9) also in the form

$$\frac{1}{2}S_n(f) = \langle\, |\tilde{n}(f)|^2 \,\rangle\, \Delta f\,. \quad (7.11)$$

The factor 1/2 is conventionally inserted in the definition (7.6) of $S_n(f)$, so that $\langle n^2(t)\rangle$ is obtained integrating $S_n(f)$ over the physical range $0 \leqslant f < \infty$, rather than from $-\infty$ to $\infty$,

$$\begin{aligned}\langle n^2(t)\rangle &= \langle n^2(t=0)\rangle \\ &= \int_{-\infty}^{\infty} df df' \langle n^*(f) n(f')\rangle \\ &= \frac{1}{2}\int_{-\infty}^{\infty} df\, S_n(f) \\ &= \int_0^{\infty} df\, S_n(f)\,. \end{aligned} \quad (7.12)$$

The function $S_n(f)$ is known as the *noise spectral density* (or the noise spectral sensitivity, or the noise power spectrum). More precisely, it is called a ***single-sided spectral density***, to emphasize that $\langle n^2(t)\rangle$ is obtained from it integrating only over the physical range of frequencies $f > 0$. Alternatively, we can write

$$\langle n^2(t)\rangle = \int_{-\infty}^{\infty} df\, S_n^{\text{double sided}}(f)\,, \quad (7.13)$$

with $S_n^{\text{double sided}}(f) = (1/2)S_n(f)$. Throughout this book, when we will use the term spectral density or power spectrum, we will always refer to the single-sided quantity.

Equivalently, the noise of a detector can be characterized by $\sqrt{S_n(f)}$, which is called the ***spectral strain sensitivity***, or *spectral amplitude*, and has dimensions $\text{Hz}^{-1/2}$. Note that, if the noise increases by a factor $\lambda$, $n(t) \to \lambda n(t)$, then $S_n(f) \to \lambda^2 S_n(f)$ while the strain sensitivity scales linearly.

Actually the definition (7.6), even if rather intuitive, is not mathematically rigorous, because the function $n(t)$ in general does not satisfy the conditions necessary for having a well-defined Fourier transform; for instance, on the interval $-\infty < t < \infty$, $n(t)$ does not necessarily go to zero at $t \to \pm\infty$, so $\tilde{n}(f)$ in general does not exist. A more precise definition of the spectral density is obtained considering the auto-correlation function of the noise,

$$R(\tau) \equiv \langle n(t+\tau) n(t)\rangle \quad (7.14)$$

A Gaussian stochastic process $n(t)$ is characterized uniquely by its average value $\langle n(t)\rangle$, that for a stationary noise is a constant and can be set to zero with a constant shift of $n(t)$, and by its auto-correlation function. Typically, the knowledge of the noise at time $t$ gives us very little information on the value of the noise at a subsequent time $t+\tau$ with $\tau$ sufficiently large, that is, for $|\tau| \to \infty$, $R(\tau)$ goes to zero quite fast, e.g. exponentially, $R(\tau) \sim \exp\{-|\tau|/\tau_c\}$. The limiting case is white noise, in which the noise at time $t$ and at any subsequent time $t+\tau$ are totally uncorrelated, so for $\tau \neq 0$ we have $\langle n(t+\tau)n(t)\rangle = \langle n(t+\tau)\rangle\langle n(t)\rangle = 0$, and $R(\tau) \sim \delta(\tau)$.

The auto-correlation function therefore goes to zero very fast as $\tau \to \pm\infty$, and it satisfies the requirements for performing the Fourier transform. We can then define the (one-sided) noise spectral density $S_n(f)$ by

$$\frac{1}{2}S_n(f) \equiv \int_{-\infty}^{\infty} d\tau\, R(\tau)\, e^{i2\pi f\tau}\,. \quad (7.15)$$

The reality of $R(\tau)$ implies $S_n(-f) = S_n^*(f)$, while invariance under time translations gives $R(-\tau) = \langle n(t-\tau)n(t)\rangle = \langle n(t)n(t+\tau)\rangle = R(\tau)$, which implies $S_n(-f) = S_n(f)$. Inverting eq. (7.15),

$$\begin{aligned} R(\tau) &\equiv \langle n(t+\tau)n(t)\rangle \\ &= \frac{1}{2}\int_{-\infty}^{\infty} df\, S_n(f) e^{-i2\pi f\tau}\,, \end{aligned} \quad (7.16)$$

and in particular

$$\begin{aligned} R(0) &= \langle n^2(t)\rangle \\ &= \frac{1}{2}\int_{-\infty}^{\infty} df\, S_n(f) \\ &= \int_0^{\infty} df\, S_n(f)\,. \end{aligned} \quad (7.17)$$

Comparing this result with eq. (7.12) we see that, when $\tilde{n}(f)$ exists, eqs. (7.6) and (7.15) are equivalent definitions of $S_n$. Otherwise, only eq. (7.15) applies. Equation (7.15) is known as the Wiener–Khintchin relation.

If $R(\tau) \sim \delta(\tau)$, we see from eq. (7.15) that $S_n(f)$ is independent of frequency and therefore we have white noise. If instead $S_n(f)$ depends on $f$, one speaks generically of colored noise. A typical example is $1/f$ noise, which is a generic denomination for a noise where $S_n(f)$ has a power-law behavior, $S_n(f) \sim 1/f^{\gamma}$, over many decades in frequency.



## 7.2 Pattern functions and angular sensitivity

From eq. (1.58), we know that a GW with a given propagation direction $\hat{\mathbf{n}}$ can be written as

$$h_{ij}(t,\mathbf{x}) = \sum_{A=+,\times} e_{ij}^A(\hat{\mathbf{n}}) \int_{-\infty}^{\infty} df\, \tilde{h}_A(f)\, e^{-2\pi i f(t-\hat{\mathbf{n}}\cdot\mathbf{x}/c)}\,, \quad (7.18)$$

where $e_{ij}^A$ are the polarization tensors given in eq. (1.54). We take $\mathbf{x}=0$ as the location of the detector. For a detector which is sensitive only to GWs with a reduced wavelength much larger than its size, such as resonant masses and ground-based interferometers, we have $2\pi f \hat{\mathbf{n}} \cdot \mathbf{x} = \hat{\mathbf{n}} \cdot \mathbf{x}/\lambda\!\!\!^{-} \ll 1$ over the whole detector, and we can neglect the spatial dependence of $h_{ab}(t, \mathbf{x})$. So, to study the interaction of GWs with such detectors we can write simply

$$h_{ij}(t) = \sum_{A=+,\times} e^A_{ij}(\hat{\mathbf{n}}) \int_{-\infty}^{\infty} df\, \tilde{h}_A(f)\, e^{-2\pi i f t} = \sum_{A=+,\times} e^A_{ij}(\hat{\mathbf{n}}) h_A(t) . \tag{7.19}$$

Combining this with eq. (7.1) we see that the contribution of GWs to the scalar output of the detector can be written as

$$h(t) = \sum_{A=+,\times} D^{ij} e^A_{ij}(\hat{\mathbf{n}}) h_A(t) . \tag{7.20}$$

It is then convenient to define the *detector pattern functions* $F_A(\hat{\mathbf{n}})$,

$$\boxed{F_A(\hat{\mathbf{n}}) = D^{ij} e^A_{ij}(\hat{\mathbf{n}}) .} \tag{7.21}$$

The pattern functions depend on the direction $\hat{\mathbf{n}} = (\theta, \phi)$ of propagation of the wave, and in terms of them eq. (7.20) becomes

$$\boxed{h(t) = h_+(t) F_+(\theta, \phi) + h_\times(t) F_\times(\theta, \phi) .} \tag{7.22}$$

The above equations assume that we have chosen a system of axes $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$, in the plane orthogonal to the propagation direction $\hat{\mathbf{n}}$ of the wave, with respect to which the polarizations $h_+$ and $h_\times$ are defined. It is interesting to see what happens if we change this system of axes, performing a rotation by an angle $\psi$ in the transverse plane. Then the axes $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ are rotated to new axes $(\hat{\mathbf{u}}', \hat{\mathbf{v}}')$ given by

$$\hat{\mathbf{u}}' = \hat{\mathbf{u}} \cos\psi - \hat{\mathbf{v}} \sin\psi , \qquad \hat{\mathbf{v}}' = \hat{\mathbf{u}} \sin\psi + \hat{\mathbf{v}} \cos\psi , \tag{7.23}$$

where we used the same conventions on the sign of $\psi$ as in eqs. (2.188) and (2.194). With respect to the $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ axes, the amplitudes of the plus and cross polarizations have values $h_+$ and $h_\times$, while with respect to the $(\hat{\mathbf{u}}', \hat{\mathbf{v}}')$ axes, they have the values $h'_+$ and $h'_\times$. Equations (1.49) and (1.50) show that $h'_+$ and $h'_\times$ are related to $h_+$ and $h_\times$ by

$$h'_+ = h_+ \cos 2\psi - h_\times \sin 2\psi , \tag{7.24}$$

$$h'_\times = h_+ \sin 2\psi + h_\times \cos 2\psi . \tag{7.25}$$

In the new frame, the definition (1.54) states that the polarization tensors are given by

$$(e^+_{ij})'(\hat{\mathbf{n}}) = \hat{\mathbf{u}}'_i \hat{\mathbf{u}}'_j - \hat{\mathbf{v}}'_i \hat{\mathbf{v}}'_j , \qquad (e^\times_{ij})'(\hat{\mathbf{n}}) = \hat{\mathbf{u}}'_i \hat{\mathbf{v}}'_j + \hat{\mathbf{v}}'_i \hat{\mathbf{u}}'_j . \tag{7.26}$$



Then, using eq. (7.23), we find

$$(e^+_{ij})'(\hat{\mathbf{n}}) = e^+_{ij}(\hat{\mathbf{n}}) \cos 2\psi - e^\times_{ij} \sin 2\psi , \tag{7.27}$$

$$(e^\times_{ij})'(\hat{\mathbf{n}}) = e^+_{ij}(\hat{\mathbf{n}}) \sin 2\psi + e^\times_{ij} \cos 2\psi . \tag{7.28}$$

The pattern functions $F_A$ depends on the polarization tensors $e^A_{ij}$ through eq. (7.21). Since the detector tensor is a fixed quantity, independent of $\psi$, we find that in the new frame

$$F'_+(\hat{\mathbf{n}}) = F_+(\hat{\mathbf{n}}) \cos 2\psi - F_\times(\hat{\mathbf{n}}) \sin 2\psi , \tag{7.29}$$

$$F'_\times(\hat{\mathbf{n}}) = F_+(\hat{\mathbf{n}}) \sin 2\psi + F_\times(\hat{\mathbf{n}}) \cos 2\psi . \tag{7.30}$$

Combining this transformation of the pattern functions with the transformation of $h_+, h_\times$ given in eqs. (7.24) and (7.25), we see that $h(t)$ in eq. (7.22) is independent of $\psi$.

Of course, once a choice of the axes $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ used to define the polarization is made, then the pattern functions $F_A$ depends on $\theta$ and $\phi$ only. However, it is sometime useful to keep generic the definition of the $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ axes in the transverse plane, and to parametrize the possible choices by the angle $\psi$. In this case, the pattern functions depend also on $\psi$, and

$$F_+(\hat{\mathbf{n}}; \psi) = F_+(\hat{\mathbf{n}}; 0) \cos 2\psi - F_\times(\hat{\mathbf{n}}; 0) \sin 2\psi , \tag{7.31}$$

$$F_\times(\hat{\mathbf{n}}; \psi) = F_+(\hat{\mathbf{n}}; 0) \sin 2\psi + F_\times(\hat{\mathbf{n}}; 0) \cos 2\psi . \tag{7.32}$$

A useful identity satisfied by the pattern functions, independently of the specific form of the detector tensor $D_{ij}$, is[4]

$$\int \frac{d^2\hat{\mathbf{n}}}{4\pi} F_+(\hat{\mathbf{n}}) F_\times(\hat{\mathbf{n}}) = 0 , \tag{7.33}$$

where as usual $d^2\hat{\mathbf{n}} = d\cos\theta d\phi$ is the integral over the solid angle. As for the integral over $d^2\hat{\mathbf{n}}$ of $F_+^2$ and of $F_\times^2$, with a generic choice of the angle $\psi$ they are different. We will see for instance that one can choose $\psi$ so that $F_\times$ vanishes while $F_+$ is non-zero, or viceversa. However, if we average over the angle $\psi$, we find

$$\int_0^{2\pi} \frac{d\psi}{2\pi} F_+^2(\hat{\mathbf{n}}; \psi) = \int_0^{2\pi} \frac{d\psi}{2\pi} F_\times^2(\hat{\mathbf{n}}; \psi) . \tag{7.34}$$

In fact, inserting eqs. (7.31) and (7.32) into eq. (7.34), the equality follows from $\int d\psi \sin 2\psi \cos 2\psi = 0$ and $\int d\psi \sin^2\psi = \int d\psi \cos^2\psi$. From this, it also trivially follows that

$$\langle F_+^2(\hat{\mathbf{n}}; \psi) \rangle = \langle F_\times^2(\hat{\mathbf{n}}; \psi) \rangle , \tag{7.35}$$

where

$$\langle \ldots \rangle \equiv \int_0^{2\pi} \frac{d\psi}{2\pi} \int \frac{d^2\hat{\mathbf{n}}}{4\pi} (\ldots) . \tag{7.36}$$

For later use we also define the angular efficiency factor

$$F = \langle F_+^2 \rangle + \langle F_\times^2 \rangle = 2 \langle F_+^2 \rangle . \tag{7.37}$$

[4]Equation (7.33) can be shown in full generality writing

$$\int d^2\hat{\mathbf{n}}\, F_+(\hat{\mathbf{n}}) F_\times(\hat{\mathbf{n}}) = D_{ab} D_{cd} \int d^2\hat{\mathbf{n}}\, e^+_{ab}(\hat{\mathbf{n}}) e^\times_{cd}(\hat{\mathbf{n}}) ,$$

and using eq. (1.54), which shows that $e^+_{ab}(\hat{\mathbf{n}}) e^\times_{cd}(\hat{\mathbf{n}})$ is a sum of terms such as $\hat{\mathbf{u}}_a \hat{\mathbf{u}}_b \hat{\mathbf{u}}_c \hat{\mathbf{v}}_d$, which has three factors $\hat{\mathbf{u}}$ and one factor $\hat{\mathbf{v}}$, and of similar terms with $\hat{\mathbf{u}} \leftrightarrow \hat{\mathbf{v}}$. A simple way to see that the integral over $d^2\hat{\mathbf{n}}$ vanishes is then to observe that, when we integrate over all possible values of $\hat{\mathbf{n}}$, for each term $\hat{\mathbf{u}}_a \hat{\mathbf{u}}_b \hat{\mathbf{u}}_c \hat{\mathbf{v}}_d$ there is also a corresponding term obtained with $\hat{\mathbf{u}} \to -\hat{\mathbf{u}}$ and $\hat{\mathbf{v}} \to +\hat{\mathbf{v}}$, which cancels it.

**Table 7.1** The pattern functions $F(\theta,\phi;\psi=0)$ for various detectors. For interferometers, the arms are perpendicular and along the $(x,y)$ axis, $(\theta,\phi)$ are the usual polar angles defined using the $z$ axis as polar axis and, for a wave propagating along the $z$ axis, $\psi$ is the angle in the $(x,y)$ plane measured from the $x$ axis, just as $\phi$. For cylindrical bars, $\theta$ is measured from the longitudinal axis of the bar and, if we denote by $x$ the longitudinal axis, for a wave propagating along the $z$ axis, again $\psi$ is the angle in the $(x,y)$ plane measured from the $x$ axis. For resonant spheres, the modes $m=0,1c,1s,2c,2s$ are combinations of the five quadrupolar modes with $m=-2,\dots,2$, defined in Zhou and Michelson (1995). The angular efficiency factor $F$ is defined in eqs. (7.36) and (7.37). Observe that the mode $m=0$ of a sphere has the same pattern functions as a cylindrical bar (apart from a constant), while the mode $m=2c$ has the same pattern functions as an interferometer.

| Detector | $F_+(\theta,\phi;\psi=0)$ | $F_\times(\theta,\phi;\psi=0)$ | $F$ |
|---|---|---|---|
| interferometers | $\frac{1}{2}(1+\cos^2\theta)\cos 2\phi$ | $\cos\theta\sin 2\phi$ | 2/5 |
| cylindrical bars | $\sin^2\theta$ | 0 | 8/15 |
| resonant spheres | | | |
| $m=0$ | $(\sqrt{3}/2)\sin^2\theta$ | 0 | 2/5 |
| $m=1s$ | $-\sin\theta\cos\theta\sin\phi$ | $\sin\theta\cos\phi$ | 2/5 |
| $m=1c$ | $\sin\theta\cos\theta\cos\phi$ | $\sin\theta\sin\phi$ | 2/5 |
| $m=2s$ | $-\frac{1}{2}(1+\cos^2\theta)\sin 2\phi$ | $\cos\theta\cos 2\phi$ | 2/5 |
| $m=2c$ | $\frac{1}{2}(1+\cos^2\theta)\cos 2\phi$ | $\cos\theta\sin 2\phi$ | 2/5 |

We will compute the explicit forms of $F_{+,\times}(\theta,\phi;\psi)$ for bars and interferometers, in their respective chapters. We find useful to collect here the result that we will find for interferometers, cylindrical bars and resonant spheres; in Table 7.1 we give the value of $F(\theta,\phi;\psi=0)$ (with appropriate definitions of the angles, discussed in the table caption and, in more detail, in their respective chapters), and the values of the angular efficiency factor $F$.

As we see from the above table, the pattern functions are relatively smooth functions of the position of the source in the sky. On the one hand, this has the positive consequence that GW detectors have a large sky coverage, of almost $4\pi$, except for some blind directions. This is very different from conventional astronomy, where a telescope must point the source very precisely to detect it. The reverse of the coin, however, is that with a single GW detector we cannot determine the position of the source in the sky. A single detector has an output $h(t)$ that, according to eq. (7.22), depends on four unknown: the two functions $h_{+,\times}(t)$ and the angles $(\theta,\phi)$ that give the source position. To disentangle these quantities we need a coincident observation by a network of detectors. With two detectors we have at our disposal their two outputs $h_1(t)$ and $h_2(t)$ and the delay time $\tau_{12}$ between these two signals. These three quantities are not yet sufficient to solve for the four unknown $h_+(t), h_\times(t), \theta$ and $\phi$. However, with three interferometers we have five measured quantities, the three functions $h_i(t)$, $i=1,2,3$, and two independent delay times, so we can solve for $h_+(t), h_\times(t), \theta$ and $\phi$. The actual accuracy of the reconstruction depends on the signal-to-noise ratio. For typical expected signals, at first-generation interferometers the angular resolution could be of order one square degree.



## 7.3 Matched filtering

We have seen above that the detector output will be of the general form $s(t)=h(t)+n(t)$. Naively, one might then think that we can detect a GW signal only when $|h(t)|$ is larger than $|n(t)|$. This would be very unfortunate since we will see that, with plausible estimates of the expected GW signals bathing the Earth, and with the sensitivity of the present generation of detectors, we will rather be in the situation $|h(t)| \ll |n(t)|$.

The fundamental question that we ask in this section is then how can we dig out the GW signal from a much larger noise. This is a classical problem in many fields of physics or in radio engineering, and the answer is that we can detect values of $h(t)$ much smaller than the floor of the noise *if* we know, at least to some level of accuracy, the form of $h(t)$.[5] To understand the basic idea, we can first illustrate a simple version of this "filtering" procedure, before moving to optimal filtering. Suppose that $s(t)=h(t)+n(t)$, and that we know the form of the GW signal $h(t)$ that we are hunting for. Then we can multiply the output $s(t)$ by $h(t)$, integrate over an observation time $T$, and divide by $T$,

$$\frac{1}{T}\int_0^T dt\, s(t)h(t) = \frac{1}{T}\int_0^T dt\, h^2(t) + \frac{1}{T}\int_0^T dt\, n(t)h(t)\,. \tag{7.38}$$

The crucial point now is that $h(t)$ and $n(t)$, separately, are oscillating functions. However, the integrand of the first integral on the right-hand side is definite positive; it might be for instance the integral of something like $\cos^2\omega t$, times a slowly varying function of time; this integral then grows, for large $T$, as $T$. Its value averaged over a time $T$ is therefore of order one in $T$,

$$\frac{1}{T}\int_0^T dt\, h^2(t) \sim h_0^2\,, \tag{7.39}$$

where $h_0$ is the characteristic amplitude of the oscillating function $h(t)$. In contrast, since the noise $n(t)$ and our chosen function $h(t)$ are uncorrelated, the quantity $n(t)h(t)$ is oscillating, and its integral will grow only as $T^{1/2}$ for large $T$ (as is typical of systems performing a random walk), so

$$\frac{1}{T}\int_0^T dt\, n(t)h(t) \sim \left(\frac{\tau_0}{T}\right)^{1/2} n_0 h_0\,, \tag{7.40}$$

where $n_0$ is the characteristic amplitude of the oscillating function $n(t)$, and $\tau_0$ a typical characteristic time, e.g. the period of the oscillating

[5] More precisely, we must know $h(t)$ and have an idea of the typical scales of variations of the noise, in order to exploit their different behaviors.

function $h(t)$. Thus, in the limit $T \to \infty$, the second term on the right-hand side of eq. (7.38) averages to zero, and we have "filtered out" the contribution of the noise from the output. Of course, in practice we cannot sent $T$ to infinity, either because the signal $h(t)$ itself has a limited temporal duration or because we are limited by the total available observation time. Still we see that, to detect the signal given in eq. (7.39) against the background of eq. (7.40), it is not necessary to have $h_0 > n_0$ but it suffices to have $h_0 > (\tau_0/T)^{1/2} n_0$. For example, for a periodic signal with a period $\tau_0 \sim 1$ ms, such as a millisecond pulsar, observed for $T = 1$ yr, we have $(\tau_0/T)^{1/2} \sim 10^{-5}$. We can therefore dig very deeply into the noise floor.

After having discussed the intuitive idea, let us see how the above procedure can be made more precise mathematically, and optimized in order to obtain the highest possible value of the signal-to-noise ratio. We define

$$\hat{s} = \int_{-\infty}^{\infty} dt\, s(t) K(t)\,, \tag{7.41}$$

where $K(t)$ is called the *filter* function. We assume that we know what GW signal we are looking for, i.e. we know the form of $h(t)$. We then ask what is the filter function that maximizes the signal-to-noise ratio for such a signal. Since the filter function is chosen so to "match" the signal that we are looking for, the technique is called matched filtering.[6]

The signal-to-noise ratio (in amplitude) is defined as $S/N$, where $S$ is the expected value of $\hat{s}$ when the signal is present, and $N$ is the rms value of $\hat{s}$ when the signal is absent. Since $\langle n(t)\rangle = 0$, we have

$$\begin{aligned} S &= \int_{-\infty}^{\infty} dt\, \langle s(t)\rangle K(t) \\ &= \int_{-\infty}^{\infty} dt\, h(t) K(t) \\ &= \int_{-\infty}^{\infty} df\, \tilde{h}(f)\tilde{K}^*(f)\,, \end{aligned} \tag{7.42}$$

while

$$\begin{aligned} N^2 &= \left[\langle \hat{s}^2(t)\rangle - \langle \hat{s}(t)\rangle^2\right]_{h=0} \\ &= \langle \hat{s}^2(t)\rangle_{h=0} \\ &= \int_{-\infty}^{\infty} dt dt'\, K(t)K(t')\, \langle n(t) n(t')\rangle \\ &= \int_{-\infty}^{\infty} dt dt' K(t)K(t') \int_{-\infty}^{\infty} df df'\, e^{2\pi i f t - 2\pi i f' t'}\, \langle \tilde{n}^*(f)\tilde{n}(f')\rangle\,. \end{aligned} \tag{7.43}$$

Using eq. (7.6) we obtain

$$N^2 = \int_{-\infty}^{\infty} df\, \frac{1}{2} S_n(f) |\tilde{K}(f)|^2\,, \tag{7.44}$$

and therefore

$$\frac{S}{N} = \frac{\int_{-\infty}^{\infty} df\, \tilde{h}(f)\tilde{K}^*(f)}{\left[\int_{-\infty}^{\infty} df\, (1/2) S_n(f) |\tilde{K}(f)|^2\right]^{1/2}}\,. \tag{7.45}$$

[6] We limit ourselves to linear filters, i.e. filters in which $\hat{s}$ is linear in $s(t)$, as in eq. (7.41).



We now ask what is the filter $K(t)$ that maximizes $S/N$, for a given $h(t)$. This variational problem is elegantly solved by defining the scalar product between two real functions $A(t)$ and $B(t)$, by

$$\begin{aligned} (A|B) &= \mathrm{Re} \int_{-\infty}^{\infty} df\, \frac{\tilde{A}^*(f)\tilde{B}(f)}{(1/2)S_n(f)} \\ &= 4\,\mathrm{Re} \int_{0}^{\infty} df\, \frac{\tilde{A}^*(f)\tilde{B}(f)}{S_n(f)}\,, \end{aligned} \tag{7.46}$$

where Re denotes the real part, and the second line holds because we take $A(t)$ and $B(t)$ to be real functions, so that $\tilde{A}(-f) = \tilde{A}^*(f)$ (recall also that $S_n(-f) = S_n(f)$). Since $S_n(f) > 0$, this scalar product is positive definite. Then eq. (7.45) can be written as

$$\frac{S}{N} = \frac{(u|h)}{(u|u)^{1/2}}\,. \tag{7.47}$$

where $u(t)$ is the function whose Fourier transform is

$$\tilde{u}(f) = \frac{1}{2} S_n(f) \tilde{K}(f)\,. \tag{7.48}$$

In this form, the solution is clear. We are searching for the "vector" of unit norm $\hat{n} = u/(u|u)^{1/2}$, such that its scalar product with the "vector" $h$ is maximum. This is obtained choosing $\hat{n}$ and $h$ parallel, i.e. $\tilde{u}(f)$ proportional to $\tilde{h}(f)$, so we get

$$\boxed{\tilde{K}(f) = \text{const.}\, \frac{\tilde{h}(f)}{S_n(f)}\,.} \tag{7.49}$$

The constant is arbitrary, since rescaling $\hat{s}$ by an overall factor does not change its signal-to-noise ratio. Equation (7.49) defines the matched filter (or Wiener filter).[7] In particular, if we are looking for a signal $h(t)$ embedded into white noise, so that $\tilde{S}_n(f)$ is a constant, then the best filter is provided by the signal itself, which is the filtering discussed in eq. (7.38). However, when $\tilde{S}_n(f)$ is not flat, eq. (7.49) tells us that we must weight less the frequency region where the detector is more noisy, a very natural result.

Inserting the solution (7.49) into eq. (7.48) we get $\tilde{u} = \text{const.} \times \tilde{h}$. Plugging this into eq. (7.47), the overall constant cancels and we get the optimal value of $S/N$,

$$\left(\frac{S}{N}\right) = (h|h)^{1/2}\,, \tag{7.50}$$

that is

$$\boxed{\left(\frac{S}{N}\right)^2 = 4\int_0^{\infty} df\, \frac{|\tilde{h}(f)|^2}{S_n(f)}\,,} \tag{7.51}$$

which is the optimal value of the signal-to-noise ratio.[8] The above equations are completely general, and independent of the form of $\tilde{h}(f)$. In Sections 7.5–7.8 we will apply them to some specific signals.

[7] It is also common in the literature to write eq. (7.41) in the form $\hat{s} = \int_{-\infty}^{\infty} dt\, s(t) G(-t)$, and to call $G(t)$ the filter function. So $G(t) = K(-t)$ and $\tilde{G}(f) \sim \tilde{h}^*(f)/S_n(f)$.

[8] Recall from Section 7.1 that our $S_n(f)$ is single-sided. In terms of the double-sided spectral density, defined after eq. (7.13), we have $(S/N)^2 = \int_{-\infty}^{\infty} df\, |\tilde{h}(f)|^2 / S_n^{\text{double sided}}(f)$.

## 7.4 Probability and statistics

The matched filtering technique discussed above (as well as other techniques that we will meet later in this chapter) provide us with a way to optimize the signal-to-noise ratio, assuming that a given signal is indeed present in our data stream. The issue that the experimenter normally faces (especially in the field of GW experiments) is however different. We do not know a priori whether a GW signal is present or not in a given stream of data, and we know even less its waveform. We can apply the matched filtering technique repeating it with many possible different filters, e.g. many possible starting times for the putative signal, many possible parameters describing a family of waveforms, etc. and we will correspondingly extract from our data stream a number of "events",[9] with various values of the signal-to-noise ratio $S/N$. What can we conclude from this? When can we claim detection of GWs? And, if we can claim detection, what can we learn from it, in particular how can we reconstruct the properties of the source (such as, for an astrophysical source, its direction, its distance, its mass, etc.), and with what accuracy?

To address these questions we need to use statistical reasoning. Before looking into the technical aspects, it is however useful to discuss more generally the statistical frameworks that one can use, as we do in the next subsection.

[9] See Note 2 on page 337 for the distinction between events, GW signals, and detector output.

### 7.4.1 Frequentist and Bayesian approaches

An abstract definition of probability can be obtained by considering a set $S$ with subsets $A, B, \ldots$, whose interpretation for the moment is left open, and defining the probability $P$ as a real-valued function that satisfies the Kolmogorov axioms: 1. For every $A$ in $S$, $P(A) \geqslant 0$. 2. For disjoint subsets (i.e. $A \cap B = 0$), $P(A \cup B) = P(A) + P(B)$, and 3. $P(S) = 1$. Furthermore, one defines the conditional probability $P(A|B)$ (i.e. the probability of $A$ given $B$) as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{7.52}$$

There exist two main approaches to probability, frequentist (also called classical) and Bayesian, depending on the interpretations of the subsets $A, B, \ldots$.

In the frequentist interpretation, $A, B, \ldots$ are the outcome of a repeatable experiment, and the probability $P(A)$ is defined as the frequency of occurrence of $A$, in the limit of an infinite number of repetitions. In this interpretation, the probabilities of obtaining some data are of course well-defined, and it also makes sense to consider the conditional probability of obtaining some data, given some hypothesis (or given a theory, or given the value of the parameters in a theory).[10] Therefore, quantities such as $P(\text{data}|\text{hypothesis})$ or $P(\text{data}|\text{parameters})$ make sense. However, one is never allowed to speak of the probability that the parameters take a given value, nor of the probability that a hypothesis, or a theory, is correct. Hypotheses, or theories, are not the outcome of a repeatable experiment. Rather, they are correct or they are wrong, and similarly the true value of a parameter in a theory is what it is, and these are facts that are not subject to probabilistic analysis.

[10] The kind of example that appears in all textbooks: we toss a coin five times. What is the probability of getting all five times head (data), given that the coin has 50% probability of heads and tails (hypothesis)?



In the Bayesian approach, instead, one is allowed to consider the probability of a hypothesis, or of a theory, or the probability that a parameter within a theory takes a given value. To define these probabilities, one starts from the identities $P(A \cap B) = P(A|B)P(B)$ and $P(B \cap A) = P(B|A)P(A)$, which follow from the definition (7.52) of conditional probability. On the other hand, $A \cap B = B \cap A$ and therefore

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \tag{7.53}$$

which is Bayes' theorem. Observe also that, from the axioms of probability given above, it follows that

$$P(B) = \sum_i P(B|A_i)P(A_i), \tag{7.54}$$

for any $B$ and for $A_i$ disjoints and such that $\cup_i A_i = S$. Therefore eq. (7.53) can be rewritten as

$$P(A|B) = \frac{P(B|A)P(A)}{\sum_i P(B|A_i)P(A_i)}, \tag{7.55}$$

so the denominator is just a normalization factor. As long as $A$ and $B$ are the outcome of a repeatable experiment, eq. (7.55) would be accepted also by frequentists. In the Bayesian approach, however, one applies this to $A$ = hypothesis (or parameters, or theory) and $B$ = data. Then one finds that

$$P(\text{hypothesis}|\text{data}) \propto P(\text{data}|\text{hypothesis})\, P(\text{hypothesis}). \tag{7.56}$$

The probability of the hypothesis given the data is called the *posterior probability*, and eq. (7.56) states that it is proportional to the product of two factors. The first is the probability of the data given the hypothesis (a "honest" frequentist probability), which is called the *likelihood function*. The second is the probability of the hypothesis, and is called the *prior probability* (or, simply, the prior). The latter cannot be determined just by performing identical trials (so it makes no sense to a frequentist) and, in the Bayesian approach, one must make assumptions to determine it. In fact, this prior probability in general can even depend on subjective factors, and on the state of knowledge of the person that makes the analysis. In the Bayesian interpretation, $P(\text{hypothesis})$ can be seen as the "degree of belief" that the hypothesis is true, and eq. (7.56) describes the evolution of this degree of belief due to the fact that we have performed the measurement. The prior probability describes the degree of belief in the hypothesis before the measurement was made, and the posterior probability describes the degree of belief after.[11]

[11] Observe also that eq. (7.56) is stated as a proportionality, so $P(\text{hypothesis}|\text{data})$ must be normalized summing over all possible hypothesis (or theories) that we want to compare, or integrating over a given domain of values for the continuous parameters.

This difference in approach implies also an important difference among the frequentist and the Bayesian notions of confidence interval and of confidence level (C.L.). The expression "confidence interval", without further qualifications, refers to the frequentist definition, and has the following meaning. Suppose that we are performing repeated identical measurements of a physical quantity $x$. We want to express our result saying that, at a given confidence level, say 90%, $x_1 < x < x_2$. What is meant by this is the following. The true (unknown) value of $x$ is a fixed number $x_t$, which is always the same in all repetitions of the experiment; each repetition provides a different interval $[x_1, x_2]$, that we want to construct in such a way that $x_t$ will be contained inside this interval in 90% (or whatever the specified C.L.) of the repetitions, no matter what the true value $x_t$ is. This is the frequentist concept of *coverage*. There is a general construction, given by Neyman in a famous 1937 paper, that allows us to construct the frequentist confidence intervals. We illustrate it in the simple case in which we know that the experimental apparatus provides values distributed as a Gaussian around the true value $x_t$, with a standard deviation $\sigma$,

$$P(x|x_t) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{(x-x_t)^2}{2\sigma^2}\right\}. \qquad (7.57)$$

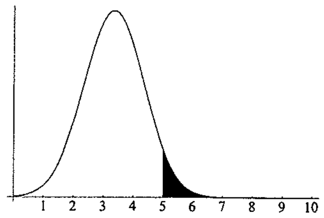

**Fig. 7.2** The Neyman construction for the lower limit $x_1$ of the confidence interval. Here the measured value was $x_0 = 5$ and, to get the interval at 90% C.L., we look for a Gaussian distribution such that its area at $x \geqslant 5$ (shaded region) is 5% of the total area. This is a Gaussian centered in $x_1 \simeq x_0 - 1.64485\sigma$ (here we used $\sigma = 1$).

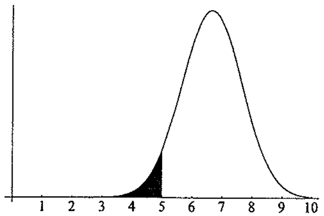

**Fig. 7.3** The same as the previous figure, for the upper limit $x_2$ of the confidence interval. The Gaussian is now centered in $x_2 \simeq x_0 + 1.64485\sigma$.

Suppose that a given repetition of the experiment yields the value $x_0$. The Neyman's construction (using for definiteness 90% C.L.) proceeds by finding a value $x_1 < x_0$ such that 5% of the area under $P(x|x_1)$ is at $x > x_0$. That is, we fix $x_1$ by requiring that a Gaussian distribution centered on $x_1$, only in 5% of the cases produces values of $x$ higher than $x_0$, see Fig. 7.2. If the true value $x_t$ were smaller than such $x_1$, then the value $x_0$ that we observed was due to a statistical fluctuation that takes place in less than 5% of the repetitions, so choosing in this way the lower limit of the interval, we are wrong at most in 5% of the cases. The upper limit of the confidence interval is obtained similarly, by finding a value $x_2 > x_0$ such that 5% of the area under $P(x|x_2)$ is at $x < x_0$, see Fig. 7.3. Observe that the probabilistic variables in this construction are $x_1$ and $x_2$, while the true value $x_t$ is fixed (and unknown).

In contrast, the Bayesian approach constructs a probability distribution for the true value $x_t$. This is obtained from the likelihood function $P$(data|hypothesis) in eq. (7.56), where the hypothesis is that the true value of $x$ is $x_t$ and the data is the observed value $x_0$. We denote the likelihood function as $\Lambda(x_0|x_t)$. In our case, this is the same as the Gaussian given in eq. (7.57), so $\Lambda(x_0|x_t) = P(x_0|x_t)$. As long as we interpret it as the probability of obtaining the value $x_0$, given that the true value is $x_t$, the likelihood function is a legitimate frequentist concept. However, in the Bayesian approach, it is inserted into eq. (7.50) together in this case with a flat prior in $x_t$, to get a probability density function (p.d.f.) in the variable $x_t$, given the observed value $x_0$,

$$P(x_t|x_0) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{(x_t-x_0)^2}{2\sigma^2}\right\}. \qquad (7.58)$$



So, in this case we have a Gaussian distribution centered on $x_0$ (rather than on $x_1$ or on $x_2$ as in the Neyman construction), and we use it as a p.d.f. for $x_t$. The most probable value of $x_t$ is found by maximizing this p.d.f., which of course gives $x_t = x_0$, and the Bayesian 90% confidence interval is defined as the interval which subtends an area equal to 90% of the total area of the p.d.f.[12] In the case of a Gaussian distribution, the Bayesian and frequentist definitions give the same result for $x_1$ and $x_2$, even if the interpretation is different. However, in a general situation, the two definitions do not agree. The frequentist confidence interval, by construction, always has the prescribed coverage, i.e. we are sure that in the limit of a large number of repetitions, 90% (or whatever the chosen C.L.) of the confidence intervals obtained by the different repetitions of the experiment will include ("cover") the true value $x_t$, no matter what $x_t$ is. This covering properties is not necessarily true in the Bayesian procedure, which in certain cases yields intervals whose frequentist coverage is less than the stated C.L. (i.e. they undercover). This can happen in particular for event-counting experiments, that obey Poisson statistics, when the data sample is small.

Beside the situation when we have small numbers, the other typical situation where the Bayesian and frequentist approaches can give sensibly different answers is when the variable $x$, for physical reasons, has a bounded domain, and the measured values are close to the boundaries of the domain. An instructive example, that nicely illustrates the different results that can be obtained with the Bayesian and the frequentist approaches in such situations, is the following.[13] Nowadays, we know from oscillation experiments that the three neutrinos have a small mass, with squared masses (more precisely, squared mass differences) between $10^{-5}$ and $10^{-4}\,\mathrm{eV}^2$. Before these results, a number of other experiments attempted a direct measure of the mass $m_\nu$ of the electron neutrino (or more precisely, of $m_\nu^2$) from tritium beta decay. In the early 1990s the experimental situation was that various experiments reported *negative* values for their best estimate of $m_\nu^2$. This is not surprising in principle since, if $m_\nu^2$ were really zero, or anyway much smaller than the experimental accuracy (as indeed it was), and if the distribution of the data is an unbiased Gaussian, on average half of the ensemble of the experiments should report negative values, and statistical fluctuations can drive the average over the experiments in the unphysical region $m_\nu^2 < 0$. However, these negative fluctuations happened to be so large that even the frequentist upper limit at 90% C.L. was negative, and was $m_\nu^2 < -16\,\mathrm{eV}^2$.[14] To say the least, it is quite disturbing to set up a complicated experiment to come out with the conclusion that $m_\nu^2$ is smaller than a negative value. The point is that this statement holds at 90% C.L., so it should be false in 10% of the cases, and here we know for sure that we are in this false 10%.[15]

A possible alternative in this case is to include our prior information that $m_\nu^2 \geqslant 0$. This suggests to take a Bayesian approach with a prior p.d.f. $P(m_\nu^2)$ which is zero when $m_\nu^2 < 0$, and uniform for $m_\nu^2 \geqslant 0$, and to use the resulting posterior p.d.f. to set the bound on $m_\nu^2$. Here however

[12] Such an interval is selected uniquely by imposing an extra requirement, typically that it is symmetric around the maximum, or that it is the minimum length interval. For a Gaussian distribution, these two conditions give of course the same result.

[13] We follow the paper by Cousins (1995), "Why isn't every physicist a Bayesian?", where the reader can find a very clear exposition of the difference between the Bayesian and frequentist approaches.

[14] Since the early 1990s, direct experiments (i.e. experiments not based on oscillations) on the electron neutrino mass squared have improved, but still their world average is negative, see Yao *et al.* [Particle Data Group] (2006).

[15] It should be mentioned that a strict application of the frequentist Neyman construction can never produces an upper limit in the unphysical region, but rather an empty confidence interval (which is equally disturbing). There is however a generalization of the Neyman construction that produces non-empty intervals in the physical region, see Feldman and Cousins (1998).

the problem arises as to whether, in the region $m_\nu^2 \geqslant 0$, the prior should be uniform in the variable $m_\nu^2$, or in $m_\nu$, or in $\log m_\nu$, etc. Of course a distribution $P(m_\nu^2)dm_\nu^2$ with $P(m_\nu^2) = \text{const.}$ is flat with respect to the variable $m_\nu^2$ but, since $dm_\nu^2 = 2m_\nu dm_\nu$, it is linearly raising with respect to $m_\nu$. The issue is significant since the resulting upper bound depends on the choice. In this specific problem the consensus finally settled on using a flat prior in $m_\nu^2$, which gave an upper bound, at 90% C.L., $m_\nu^2 < 26.6\,\text{eV}^2$.

A physicist that is not too much interested in the philosophical aspects of the debate, can take a pragmatic attitude and use a frequentist or a Bayesian approach, depending on the kind of experiment to be analyzed. In particular, elementary particle physics is very well suited for the frequentist approach. This basically stems from the fact that in this case it is the physicist that controls the parameters of the experiment (e.g. the kind of particle used in the beams, the beam energy, etc.) and can reproduce them accurately many times. We are therefore in the situation where the frequentist notion of repeated trials fits very well.[16] The advantage is that this allows us to report objectively the outcome of the experiment, without the need of incorporating prior (and possibly subjective) beliefs.

[16]In fact, in the standard compilation of experimental high-energy physics data, the Particle Data Group (PDG) "Reviews of Particle Properties", essentially all measurements and their statistical uncertainties are reported within the frequentist framework.

On the other hand in astrophysics, and even more in GW astrophysics, the sources can be rare, they are not under the control of the experimenter, and each one is very interesting individually. If a single BH-BH binary coalesces, and we detect its signal in a GW experiment, we would obviously be very interested in questions such as in which direction the binary system was, at what distance from us, what were the masses of the two black holes, their spins, etc. A strict frequentist approach is inapplicable here. We do not have at our disposal an ensemble of identical BH-BH binaries located in that position, with the same values of the masses, etc. We just have that unique event, and we want to get the maximum out of it. In this case, a Bayesian approach can be more appropriate, since it allows us to ask questions such as "What was the most likely value of the position, masses, spin, etc. of the BHs?". For this reason, while negative results, giving upper limits on the rate of GW signals, should normally be expressed in frequentist terms, the discussion of parameter estimation from a given positive detection, to which we turn next, should rather be performed within the Bayesian framework.

### 7.4.2 Parameters estimation

In Section 7.3, when we introduced the matched filtering technique, we assumed that the form of $h(t)$ is known. In practice, however, $h(t)$ will necessarily depend on a number of free parameters. For instance, if $h(t)$ is a short burst of GWs, among its parameters we will certainly have the time of arrival $t_0$. When searching for very short bursts we might simply use a Dirac delta, so $h(t) = h_0\delta(t - t_0)$, but more generally we might also wish to include its temporal width $\Delta t$ and possibly more parameters describing the shape of the pulse. For a coalescing binary, among the parameters we will have the time of entry in the interferometer bandwidth, the distance to the source, the star masses, etc.



Therefore, we must consider a family of possible waveforms, or *templates*, that we denote generically as $h(t;\theta)$, where $\theta = \{\theta_1, \ldots, \theta_N\}$ is a collection of parameters. Correspondingly, we have a family of optimal filters $K(t;\theta)$, determined through eq. (7.49), $\tilde{K}(f;\theta) \sim \tilde{h}(f;\theta)/S_n(f)$. In practice, this means that we must discretize the $\theta$-space, and repeat the filtering procedure many times, once for each point of this discretized parameter space (except that for some parameters the maximization procedure can be done analytically, see below).

The problem that we address in this section is the following. Suppose that a GW signal has indeed been detected, which means that for some template $h(t;\theta)$ the value of $S/N$, determined by the optimal Wiener filtering (or by any other procedure that we specified in advance) has exceeded a predetermined threshold, and the signal satisfies further criteria that we might have set for claiming detection, such as coincidences between different detectors (we will see in more detail in Sections 7.4.3 and 7.5.3 some possible criteria that could allow us to claim a detection, at a given confidence level). How do we reconstruct the most probable value of the parameters of the source, and how we compute the error on these parameters?

This question is Bayesian in nature, so its answer is contained in the posterior probability. To compute the likelihood function, and hence the posterior probability, we assume for simplicity that the noise $n(t)$ is stationary and Gaussian. From eq. (7.6) we see that the variance of the Fourier mode of the noise with frequency $f$ is proportional to $(1/2)S_n(f)$, so the corresponding Gaussian probability distribution for the noise is

$$p(n_0) = \mathcal{N} \exp\left\{-\frac{1}{2}\int_{-\infty}^{\infty} df\, \frac{|\tilde{n}_0(f)|^2}{(1/2)S_n(f)}\right\}, \tag{7.59}$$

where $\mathcal{N}$ is a normalization constant. This is the probability that the noise $n(t)$, which is a random variable with zero mean, has a given realization $n_0(t)$. The above result can be rewritten very simply in terms of the scalar product (7.46) as[17]

$$p(n_0) = \mathcal{N} \exp\{-(n_0|n_0)/2\}. \tag{7.62}$$

We are assuming that the output of the detector satisfies the condition for claiming detection, i.e. it is of the form $s(t) = h(t;\theta_t) + n_0(t)$, where $n_0(t)$ is the specific realization of the noise in correspondence to this event, and $\theta_t$ is the (unknown) true value of the parameters $\theta$. The likelihood function for the observed output $s(t)$, given the hypothesis that there is a GW signal corresponding to the parameters $\theta_t$, is obtained plugging $n_0 = s - h(\theta_t)$ into eq. (7.62),

$$\Lambda(s|\theta_t) = \mathcal{N} \exp\left\{-\frac{1}{2}(s - h(\theta_t)|s - h(\theta_t))\right\}, \tag{7.63}$$

[17]For simplicity, we limit ourselves to the case of a single detector. The formalism can however be extended straightforwardly to multiple detectors. In this case the definition of the noise spectral density, eq. (7.6), is replaced by

$$\langle \tilde{n}_a^*(f)\tilde{n}_b(f')\rangle = \delta(f - f')\frac{1}{2}[S_n(f)]_{ab}, \tag{7.60}$$

where the indices $a, b$ label the detectors. This definition takes into account the possibility of correlated noise. Let $\mathbf{A}(t)$ and $\mathbf{B}(t)$ be vectors whose components $A_a(t)$ and $B_a(t)$ are output of the single detectors, and let $[S_n^{-1}]^{ab}$ denote the inverse matrix. The equations of this section can then be generalized to multipole detectors, using the scalar product

$$(\mathbf{A}|\mathbf{B}) = 4\,\text{Re} \int_0^\infty df\, \tilde{A}_a^*(f)[S_n^{-1}(f)]^{ab}B_b(f), \tag{7.61}$$

which generalizes eq. (7.46). See the Further Reading for details.

or, introducing the short-hand notation $h_t \equiv h(\theta_t)$,

$$\Lambda(s|\theta_t) = \mathcal{N} \exp\left\{(h_t|s) - \frac{1}{2}(h_t|h_t) - \frac{1}{2}(s|s)\right\}. \qquad (7.64)$$

In the Bayesian approach, according to eq. (7.56), we also introduce a prior probability $p^{(0)}(\theta_t)$.[18] Then, the posterior probability distribution for the true value $\theta_t$, given the observed output $s$,

$$\boxed{p(\theta_t|s) = \mathcal{N} p^{(0)}(\theta_t) \exp\left\{(h_t|s) - \frac{1}{2}(h_t|h_t)\right\},} \qquad (7.65)$$

where, since we are considering $p(\theta_t|s)$ as a distribution in $\theta_t$ for a fixed output $s$, we have reabsorbed into the normalization factor $\mathcal{N}$ the term $(s|s)/2$ which appears in the exponential in eq. (7.64).

Once the prior distribution is given, eq. (7.65) gives the p.d.f. in the parameter space, so in principle it contains all the information that we need. However, in this form the information might not be very manageable. The $\theta$-space will in general be a multi-dimensional space of large dimension. For example, for a binary coalescence the parameters $\theta^i$ that determine the waveform, at the post-Newtonian level, are the distance, the source's location (two angles), the orientation of the normal to the orbit (two more angles), the time at which the signal enters in the interferometer's bandwidth, the orbital phase at that moment, the two masses of the stars, and their spins, so 15 parameters in all.[19] From the probability distribution function (7.65) in such a complicated space we would like to extract some more approximate, but also more manageable information; essentially, we want the most probable value of the variables $\theta_t$, that we denote by $\hat{\theta}$, and also their corresponding errors.

There is no unique way of defining what is the most probable value of $\theta_t$. A rule for assigning the most probable value is called an estimator. The most important properties that an estimator must have are:[20] (a) Consistency: the estimator must converge to the true value as the amount of data increases. This property is so important that it is possessed by all commonly used estimators. (b) The bias $b$ is defined as the difference between the expectation value of the estimator, $E(\hat{\theta})$ (taken over a hypothetical set of similar experiments in which $\hat{\theta}$ is constructed in the same way), and the true value $\theta_t$, $b \equiv E(\hat{\theta}) - \theta_t$. When $b = 0$ the estimator is said to be unbiased. (c) Efficiency: we want the smallest possible value for the variance of $\hat{\theta}$, and (d) Robustness, i.e. the property of being relatively insensitive to small departure in the assumed p.d.f. due to factors such as noise.

Two choice of estimators seems especially reasonable. The first is to define $\hat{\theta}$ as the value which maximizes the probability distribution function (7.65). Another natural option is to define it as the average of $\theta_t$, over the distribution (7.65). We discuss these options below.

### Maximum likelihood estimator

Let us consider first the situation in which the prior probability is flat.[21]

[18] As an example of prior information, one of the typical parameters entering in the waveform is the distance $r$ to the source, and we might be searching for signals from a population of stars with a known distribution in space, e.g. a distribution $p^{(0)}(r)dr \sim r^2 dr$ for isotropic sources, or $p^{(0)}(r)dr \sim r dr$ for sources within a few kpc from us, in the Galactic disk. Another typical parameter is the mass of the star and, for neutron stars, we know from astrophysical observations that their mass distribution is strongly concentrated around $1.35 M_\odot$.

[19] Assuming that the eccentricity can be neglected, since the orbit should be highly circular by the time the signal enters in the bandwidth of a ground-based detector, as we saw in Section 4.1.3.

[20] For details see, e.g. the statistics section of Yao *et al.* [Particle Data Group] (2006).

[21] As we already mentioned when discussing the example of the neutrino mass on page 350, a distribution which is flat with respect to the variables $\theta_t$ is no longer flat if we make a non-linear transformation of the parameters. Therefore this prior distribution assumes a definite choice of coordinates in the parameter space.



Then, maximization of the posterior probability becomes the same as maximization of the likelihood $\Lambda(s|\theta_t)$. The value of $\theta_t$ that maximizes $\Lambda(s|\theta_t)$ defines the maximum likelihood estimator, and we denote it by $\hat{\theta}_{\rm ML}(s)$. It is the most widely used estimator in general situations.[22] It is usually simpler to maximize $\log\Lambda$. From eq. (7.64),

$$\log\Lambda(s|\theta_t) = (h_t|s) - \frac{1}{2}(h_t|h_t). \qquad (7.66)$$

Since we are working at fixed $s$, the term $(-1/2)(s|s)$ in eq. (7.64) is an irrelevant constant, and we omitted it. Denoting $\partial/\partial\theta_t^i$ simply by $\partial_i$, the value of $\hat{\theta}_{\rm ML}$ is found by solving the equations

$$(\partial_i h_t|s) - (\partial_i h_t|h_t) = 0. \qquad (7.67)$$

The errors $\Delta\theta^i$ can then be defined in terms of the width of the probability distribution function (7.65) at the peak.

Typically, (7.67) is a set of equations that must be solved numerically (except for some parameter such as the overall amplitude that can be eliminated analytically, see below). However, they have a rather simple geometric interpretation. The set of all possible waveforms $h(t;\theta)$ defines a manifold, called the manifold of the signals, parametrized by the coordinates $\theta^i$. This is a subset of zero measure in the space of all possible functions, so the addition of generic noise $n(t)$ to a function $h(t;\theta)$ will necessarily bring us out of this manifold. In Fig. 7.4 we illustrate the situation with a two-dimensional manifolds of the signals. The point labeled $\theta_t$ represents the true value of the signal, and therefore lies on the manifold. The addition of noise carries us outside this manifold. Since we are minimizing $(s-h|s-h)$, see eq. (7.63), the maximum likelihood estimator actually searches the point on the signal manifold which is closest to the output $s$, where distances are defined with respect to the scalar product $(\ |\ )$.

To summarize, in the Bayesian framework $\hat{\theta}_{\rm ML}$ is determined assuming a flat prior distribution and requiring the maximization of the posterior probability (7.65), i.e. maximizing our "degree of belief" in the hypothesis that there is a GW signal. A natural question, at this point, is what is the relation between $\hat{\theta}_{\rm ML}$ and the value of $\theta$ that provides the highest signal-to-noise ratio in the matched filtering. We now prove that in fact they are the same. To show it, we write the generic template as $h(t;\theta) = a h_a(t;\xi)$, where $a$ is an amplitude, and is a free parameter, while the normalization of $h_a$ has been fixed imposing some condition. We have separated the parameters $\theta$ into $a$ and the remaining parameters, that we call $\xi$. The maximization with respect to $a$ of $\log\Lambda$ can be performed analytically since, from eq. (7.66),[23]

$$\log\Lambda(s|a,\xi) = a(h_a|s) - \frac{a^2}{2}(h_a|h_a). \qquad (7.68)$$

Requiring $\partial\log\Lambda/\partial a = 0$, we get the maximum likelihood estimate for $a$,

$$\hat{a}_{\rm ML}(s) = \frac{(h_a|s)}{(h_a|h_a)}. \qquad (7.69)$$

[22] See any textbook on statistics, e.g. Lyons (1986), Section 4.4, for an introduction to the maximum likelihood method and its virtues. Observe that the likelihood is a legitimate concept also in the frequentist approach. The most probable frequentist value is again identified with the maximum of the likelihood, and the confidence interval is usually defined in terms of the point where $2\log\Lambda$ decreases by one unit with respect to its value at the maximum. In the frequentist approach, however, we cannot use the likelihood as a p.d.f. for the true values of the parameters, i.e. we cannot consider areas under the curve, and of course we cannot include priors.

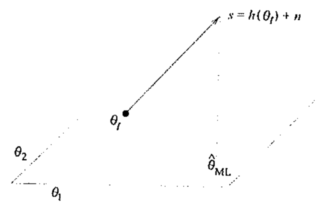


**Fig. 7.4** The manifolds of the signals, parametrized by the coordinates $(\theta_1,\theta_2)$. The point $\theta_t$ is the true value of the signal. The addition of noise to $h(\theta_t)$ brings us outside this manifold, and the maximum likelihood estimator searches the point on the signal manifold which is closest to $s$.

[23] To keep the notation lighter, we omit the subscript t (which stands for "true") from $a$ and $\xi$. We are anyway considering p.d.f. for the true values of the parameters.

The maximization with respect to the remaining variables $\xi$ can be performed substituting this expression for $a$ into $\log\Lambda$, obtaining

$$\log\Lambda(s|\xi) = \frac{1}{2}\frac{(h_a|s)^2}{(h_a|h_a)}\,. \tag{7.70}$$

The maximization of this quantity amounts to maximizing the overlap of the output $s$ with the normalized template $h_a/(h_a|h_a)^{1/2}$, where the overlap is measured using the scalar product ( | ), defined in terms of the noise spectral density $S_n(f)$. This is just the matched filtering procedure discussed in the Section 7.3. Thus, the maximum likelihood method provides a way of estimating the overall amplitude $a$ (which cannot be fixed just searching for the filter that maximizes the signal-to-noise ratio, since eq. (7.47) is unchanged by a multiplicative rescaling of the filter $u$) while, for the remaining parameters, it returns the values that maximize the signal-to-noise ratio according to matched filtering.[24]

[24] We also mention that another way of understanding the meaning of the maximum likelihood procedure is in term of the Neyman–Pearson criterium, which consists in maximizing the probability of detection, subject to a given false alarm probability, and leads again to the condition that $\log\Lambda$ be maximum.

### Maximum posterior probability

In various situations we do have important prior information, and we might want to include it in the analysis, see Note 18 for examples. In this case, rather then maximizing the likelihood function, we must determine the estimator by maximizing the full posterior probability $p(\theta_t|s)$ given in eq. (7.65), which takes into account the prior probability distribution.[25] For a generic prior, of course, the maximum of the posterior distribution will change, so it will no longer coincide with the value that gives the highest signal-to-noise ratio in the matched filtering. What happens is that the value suggested by matched filtering is weighted against our prior expectations (in a real sense, our "prejudices" ), to provide a new estimate of the most likely value for the true parameters.

[25] In the GW literature, the log of the prior $p^0(\theta)$ is sometime added to the exponential in eq. (7.65), and the resulting exponential is called again the log-likelihood function, $\log\Lambda$, so the corresponding estimator is called again the maximum likelihood estimator. This notation is however potentially confusing. For instance, one might be tempted to make a frequentist use of such a "$\log\Lambda$", which is obviously incorrect, since it involves a prior probability.

When we want to include non-trivial prior information, some conceptual complication may appear (apart from the issue of how to choose the appropriate prior). Suppose, for definiteness, that we have a two-dimensional parameter space $(\theta_1,\theta_2)$, as in Fig. 7.4, and that we are not interested in the variable $\theta_2$. Then, we can integrate the p.d.f. given in eq. (7.65) over $\theta_2$, to obtain a reduced p.d.f. in the variable $\theta_1$. From the geometric interpretation given in Fig. 7.4 it is clear that, as far as the likelihood function is concerned, the maximum in the variable $\theta_1$ is the same, independently of whether we integrated or not over $\theta_2$. However, once we include a generic non-flat prior probability $p^{(0)}(\theta)$, this nice geometric interpretation is lost and, in general, if $(\bar\theta_1,\bar\theta_2)$ is the maximum of the distribution function $p(\theta_1,\theta_2|s)$, it is no longer true that the $\bar\theta_1$ is the maximum of the reduced distribution function $\tilde p(\theta_1|s) = \int d\theta_2\, p(\theta_1,\theta_2|s)$, obtained integrating out $\theta_2$. Thus, there is an ambiguity on the value of the most probable value of $\theta_1$, which depends on whether we are interested or not in $\theta_2$. Another possible drawback, this one common to both the maximum likelihood and the maximum posterior methods, is that we might want an estimator that minimizes the error on the parameter determination, and this in general is not the case for the maximum likelihood or maximum posterior probability estimators. These issues motivate the consideration of the Bayes estimator, in the next subsection.



### Bayes estimator

In this case the most probable values of the parameters is defined by

$$\hat\theta^i_{\rm B}(s) \equiv \int d\theta\,\theta^i p(\theta|s)\,, \tag{7.71}$$

i.e. is the average with respect to the posterior probability distribution. The errors on the parameters is defined by the matrix

$$\Sigma^{ij}_B = \int d\theta[\theta^i - \hat\theta^i_{\rm B}(s)][\theta^j - \hat\theta^j_{\rm B}(s)]\,p(\theta|s)\,, \tag{7.72}$$

that is, in terms of the mean square deviations from $\hat\theta^i_{\rm B}(s)$, where the average is taken again with respect to $p(\theta|s)$. Even when there is a non-trivial prior probability function, the Bayes estimator is clearly independent on whether we integrate out some variable from $p(\theta|s)$, since we anyhow integrate over all the $\theta^i$ when computing $\hat\theta^i_{\rm B}$ and $\Sigma^{ij}_B$. Furthermore, it can also be shown that, if one wants to minimize the error on the parameters, averaged over the whole parameter space, the Bayes estimator is the optimal one.

The "operational" meaning of the Bayes estimator is the following. Suppose that, after a sufficiently long run, we end up with a large ensemble of detected signals, which correspond to actual GWs, and that among them there is still a large subensemble of GW signals that produced a given output $s(t)$. Each of these waves will be characterized by different values of the true parameters $\theta_t$ and therefore by a different $h(t;\theta_t)$ that, by combining each time with a different realization of the noise $n(t)$, has produced the same output $s(t)$. Then $\hat\theta^i_{\rm B}(s)$ is the value of $\theta^i$, averaged over this ensemble of signals, and $\Sigma^{ij}_B$ is the corresponding rms error.

Thus, the Bayes estimator has a well-defined operational meaning, and welcome mathematical properties, such as the independence on whether we integrate out some variable and the fact that it minimizes the error on parameter estimation. Its main drawback is its computational cost, since the evaluation of eq. (7.71) or of eq. (7.72) involves a multi-dimensional integral over the space of $\theta$ variables which, as we have seen, could have a dimensionality of order 15 or larger, and furthermore at each point of this parameter space we must compute the function $p(\theta|s)$, given in eq. (7.65), which requires the numerical computation of the integral over frequencies that defines the scalar product ( | ). The choice of the best estimator is therefore subject to various considerations, including computational cost, and depends on the specific situation. The use of the Bayes estimator goes also under the name of non-linear filtering.

Of course, in the limit of large signal-to-noise ratio (which unfortunately is not expected to be the appropriate one for GW detectors, at least in the near future) these issues becomes irrelevant, and all consistent estimators give the same answer. In this limit, there is also a very simple expression for the error on the parameters. If the SNR is large, the error that we make on the parameter estimation is small. For simplicity we assume that the prior $p^{(0)}(\theta)$ is nearly uniform near $\theta = \hat{\theta}$, where $\hat{\theta}$ is the value produced by (any) consistent estimator, say for definiteness the maximum likelihood estimator $\hat{\theta}_{\rm ML}$. That is, we assume that the prior information is irrelevant for reconstructing the parameters. In eq. (7.65) we can then write $\theta^i = \hat{\theta}^i_{\rm ML} + \Delta\theta^i$, and, since $\Delta\theta^i$ is small, we can expand the exponential in eq. (7.65) in powers of $\Delta\theta^i$. The linear term of the expansion vanishes because $\hat{\theta}^i_{\rm ML}$ is, by definition, the maximum of the distribution, and to quadratic order in $\Delta\theta$ we get

$$p(\theta|s) = \mathcal{N} \exp\left\{-\frac{1}{2}\Gamma_{ij}\Delta\theta^i\Delta\theta^j\right\}, \qquad (7.73)$$

where $\Gamma_{ij} = (\partial_i\partial_j h|h-s) + (\partial_i h|\partial_j h)$. Observe that, in the first term, we have $h - s = -n$ and, in the limit of large signal-to-noise ratio, $|n|$ is much smaller than $|h|$. So in this limit the first term can be neglected, and we get

$$\Gamma_{ij} = (\partial_i h|\partial_j h), \qquad (7.74)$$

evaluated at $\theta = \hat{\theta}_{\rm ML}$. This is called the *Fisher information matrix*. Then the expectation value of the errors $\Delta\theta^i$ are given by

$$\langle\Delta\theta^i\Delta\theta^j\rangle = (\Gamma^{-1})^{ij}. \qquad (7.75)$$

### 7.4.3 Matched filtering statistics

As we have discussed in the previous sections, a general data analysis strategy consists in performing matched filtering, applying many different templates $h(t;\theta)$ to the data. This will result in the generation of a list of "events" (in the sense of Note 2 on page 337), defined by the fact that the signal-to-noise ratio, in correspondence with some template, raises over a predetermined threshold. Applying the maximum likelihood criterium (or the maximization of the posterior probability, if we want to include prior information), we can then get the most probable value of the parameters $\theta$, *under the hypothesis that a GW signal $h(t;\theta)$ was present.* The issue that we want to address now is the following. How well such hypothesis performed? In other words, what is the statistical significance of the fact that we found events at a given level of signal-to-noise ratio?

The answer to this question depends crucially on the statistical properties of the noise so, first of all, it is important to realize that in any detector we can distinguish between two kinds of noise: "well-behaved" Gaussian noise, whose probability distribution is a Gaussian, and non-Gaussian noise, which is a generic denomination for anything else. A Gaussian distribution $\sim e^{-x^2/2}$ drops very fast for large values of its argument $x$. The intuitive idea, that we will formalize below, is therefore to eliminate Gaussian noise by setting a sufficiently large threshold for the signal-to-noise ratio. Non-Gaussian disturbances, however, have in general a totally different statistical distribution, characterized by long tails at large values of $S/N$, which decay only as a power law.[26]

These noises cannot be eliminated just by setting a high threshold, since they can produce events with values of $S/N$ that, in Gaussian distribution, would be inconceivably large. As a limiting case, any detector shaken by an earthquake will produce "events" with arbitrarily high values of $S/N$. Of course, these events cannot be eliminated just by setting a high threshold in $S/N$. Rather, they should be identified and vetoed. All detectors are equipped with sensors which monitor various aspects of the detector performance as well as environmental conditions (e.g. seismometers), so that non-Gaussian disturbances are vetoed as much as possible. However, it is practically impossible to be sure that one has identified and vetoed all possible non-Gaussian disturbances. So, while in principle one can study experimentally the noise distribution and then set a threshold so high that even non-Gaussian fluctuations would be very rare, in practice this is not possible because the resulting threshold would be much too high, and therefore would considerably degrade the sensitivity of the detector. Rather, the best way of eliminating non-Gaussian noise is to perform coincidences between two or more detectors. This is among the reasons why various different detectors have been built, and they are operated as a network.

In the following, we first discuss the statistical significance of obtaining a given value of the signal-to-noise ratio $S/N$, assuming that only Gaussian noise is present. This will tell us how to fix the threshold in $S/N$ so that, at some confidence level, we know that higher values of $S/N$ have not been produced by Gaussian noise alone, and allows us to generate, from the data stream of the detector, a list of "events". These events will then be subject to further scrutiny, using for instance coincidences between detectors whenever possible, with the aim of eliminating those which are due to non-Gaussian noise, and retaining the GW signals, if any. For the rest of this section we will be concerned only with Gaussian noise, while coincidences and other techniques will be discussed when we examine the various type of signals, in Sections 7.5–7.8.

In eqs. (7.42)–(7.45) we defined the signal-to-noise ratio in terms of the expectation value of the signal. Here however we want to study the full statistical distribution, rather than just its expectation value, so we define

$$\rho = \frac{\hat{s}}{N}, \qquad (7.76)$$

where $\hat{s}$ is the filtered output defined in eq. (7.41) and $N$ is given in eq. (7.43), that is $N$ is the root-mean-square (rms) of $\hat{s}$ when the signal is absent. The definition of $\rho$ is therefore analogous to the definition of the signal-to-noise ratio $S/N$, see eqs. (7.42)–(7.45), except that in the numerator we have $\hat{s}$ rather than its expectation value $\langle\hat{s}\rangle$. As a result,

[26] For instance, a large class of phenomena, characterized by what is called self-organized criticality, are such that the number $N$ of events that release an energy $E$ is distributed as $dN = E^{-\gamma}dE$ where, quite remarkably, the exponent $\gamma$ has approximately the same universal value, $\gamma \simeq 1.6$, in phenomena apparently very different. Such a law, together with the value $\gamma \simeq 1.6$, is in fact observed in earthquakes from different seismic faults (in which case it is called the Gutenberg–Richter law), in soft $\gamma$-ray bursts from highly magnetized neutron stars, as well as in numerical simulations of fractures in solids. The same distribution is experimentally observed when searching for short bursts in resonant-bar GW detectors, where they are likely due to microfractures inside the bar, and give an example of the non-Gaussian noise that we will have to fight. See Dubath, Foffa, Gasparini, Maggiore and Sturani (2005), and references therein.

the relation between $\rho$ and $S/N$ is $S/N = \langle\rho\rangle$. From

$$\hat{s} = \int_{-\infty}^{\infty} dt\,[h(t) + n(t)]\,K(t) \qquad (7.77)$$

we see that, when $h$ is absent, $\rho$ is a random variable with zero average and, since it has been normalized to its own rms, with variance equal to one. Thus, in the absence of a GW signal, the probability distribution of $\rho$ is

$$p(\rho|h=0)d\rho = \frac{1}{\sqrt{2\pi}}\,e^{-\rho^2/2}\,d\rho\,. \qquad (7.78)$$

In contrast, if in the output there is a GW signal $h$ with a signal-to-noise ratio $\bar{\rho}$, eqs. (7.76) and (7.77) give $\rho = \bar{\rho} + \hat{n}/N$, where $\hat{n} = \int dt\, n(t)K(t)$. Since $N$ is just the rms of $\hat{n}$, in this case $\rho - \bar{\rho}$ is a Gaussian variable with zero mean and unit variance, so

$$p(\rho|\bar{\rho})d\rho = \frac{1}{\sqrt{2\pi}}\,e^{-(\rho-\bar{\rho})^2/2}\,d\rho\,. \qquad (7.79)$$

The variable $\rho$ is the signal-to-noise ratio in amplitude. It is useful to introduce also $R \equiv \rho^2$, which is the signal-to-noise ratio in energy, since the energy of GWs is quadratic in the GW amplitude. Observe that $\rho$, being proportional to $h(t)$, is not positive definite, and runs between $-\infty$ and $+\infty$, while of course $0 \leqslant R < \infty$. The probability distribution for $R$, when there is in the output a GW signal with a signal-to-noise ratio in energy $\bar{R} = \bar{\rho}^2$, follows from eq. (7.79) observing that a single value $R$ is obtained from two values of the amplitude, $\rho = \pm\sqrt{R}$, so the probability of detecting an event with SNR in energy between $R$ and $R + dR$, when the SNR of the GW signal is $\bar{R}$, is given by

$$P(R|\bar{R})dR = p(\rho|\bar{\rho})d\rho + p(-\rho|\bar{\rho})d\rho\,, \qquad (7.80)$$

evaluated at $\rho = R^{1/2}$. Writing $d\rho = dR/(2R^{1/2})$, we get

$$P(R|\bar{R})dR = \frac{dR}{2\sqrt{R}}\,\frac{1}{\sqrt{2\pi}}\left[e^{-(\rho-\bar{\rho})^2/2} + e^{-(\rho+\bar{\rho})^2/2}\right]$$
$$= \frac{1}{\sqrt{2\pi R}}\,e^{-(\bar{R}+R)/2}\cosh\left[\sqrt{R\bar{R}}\right]\,dR\,. \qquad (7.81)$$

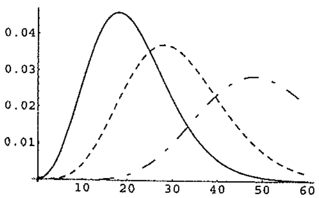

**Fig. 7.5** The probability distribution $P(R|\bar{R})$, as a function of $R$, for $\bar{R} = 20$ (solid line), $\bar{R} = 30$ (dotted line) and $\bar{R} = 50$ (dot-dashed line).

From this we can compute the average value of $R$ for a given $\bar{R}$,

$$\langle R\rangle = \int_0^{\infty} dR\, R\, P(R|\bar{R}) = 1 + \bar{R}\,. \qquad (7.82)$$

If we write $R = E/kT_n$, where $T_n$ has the physical meaning of an effective temperature of the noise after matched filtering, we can also rewrite eq. (7.82) as

$$\langle E\rangle = kT_n + \bar{E}\,. \qquad (7.83)$$

Therefore the average value of the detected energy is the sum of the energy $\bar{E}$ deposited in the detector by the GW, plus the energy $kT_n$ associated to the detector noise, a very natural result. In Fig. 7.5 we show the form of the probability distribution $P(R|\bar{R})$, as a function of $R$, for different values of $\bar{R}$. Observe that, while the average value is at $R = 1 + \bar{R}$, the maximum of the distribution is at a somewhat lower value. The corresponding distribution for $R$ in the absence of signal is obtained setting $\bar{R} = 0$ in eq. (7.81). In Fig. 7.6 we compare the probability distribution $P(R|\bar{R})$ when $\bar{R} = 10$ with the probability distribution in the absence of signal, $P(R|\bar{R} = 0)$.



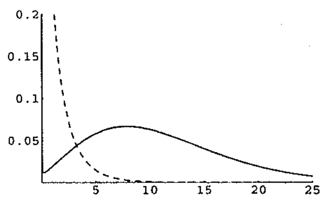

**Fig. 7.6** The probability distribution $P(R|\bar{R})$, as a function of $R$, for $\bar{R} = 10$ (solid line), compared to the probability distribution in the absence of signal, $P(R|\bar{R} = 0)$ (dashed).

The different behavior of the two distributions suggest that, when searching for a signal with a signal-to-noise ratio $\bar{R}$ in energy, we can discriminate a true GW signal from a fluctuation due to Gaussian noise setting a threshold in $R$, at a value $R_t$ that eliminates most of the noise, while retaining a large fraction of the signal distribution. Observe that anyway there will always be a *false alarm probability*, given by

$$p_{\rm FA} = \int_{R_t}^{\infty} dR\, P(R|\bar{R} = 0)$$
$$= 2\int_{\rho_t}^{\infty} d\rho\, e^{-\rho^2/2}$$
$$= 2\,\mathrm{erfc}\,(\rho_t/\sqrt{2})\,, \qquad (7.84)$$

where erfc $(z)$ is the complementary error function. Furthermore, there is a *false dismissal probability*, i.e. a probability of losing a real GW signal, given by[27]

$$p_{\rm FD} = \int_0^{R_t} dR\, P(R|\bar{R})\,. \qquad (7.85)$$

The threshold $R_t$ can be fixed deciding what is the maximum false alarm level that we are willing to tolerate. This depends also crucially on the number of trials that we do with different templates. For example, for a coalescing binary, one can estimate that of order $10^5$ templates might be needed to cover with good accuracy the possible range of values of masses and spins. Furthermore, to match the template to the signal one can estimate in about 3 ms the maximum temporal mismatch between the two. In one year of data ($\sim 3\times 10^7$ s), one must therefore try $\sim 10^{10}$ starting values of time, and for each value of time we have $10^5$ templates to cover the masses and spin parameters, so overall one might have to try $10^{15}$ templates.[28] Often the false alarm level is fixed so that the expected number of false alarms in a run will be of order a few. With a lower threshold one would be flooded by spurious events, while higher threshold have of course the effect of increasing the false dismissal probability. The few events obtained will then be subject to further scrutiny. Thus, if we search for a coalescence in a single detector, with one year of data and $10^{15}$ templates, we could chose a threshold in amplitude $\rho_t \simeq 8$, since this gives $p_{FA} \sim 2.5\times 10^{-15}$. However, performing coincidences between two detectors the probability of obtaining a false alarm simultaneously in the two detectors is the square of the single-detector probability, if the noise in the two detectors are uncorrelated, so in this case we might want $[2\,\mathrm{erfc}\,(\rho_t/\sqrt{2})]^2 \simeq 10^{-15}$, which gives $\rho_t \simeq 5.5$.

[27] In other words, whatever the value of $\bar{R}$, the distribution $P(R|\bar{R})$ is such that there is always some probability that $R$ be smaller, and even much smaller, than $\bar{R}$, and therefore the GW can go undetected even when the threshold $R_t$ was smaller than the value $\bar{R}$ due to the GW alone. If one thinks in terms of energies, it might be counterintuitive that the energy released inside the detector can be smaller than the value that would have been released by the GW alone, in the absence of noise. Recall however that GW detectors really measure an oscillation amplitude, and the amplitude induced by the GW combines with the amplitude induced by noise with a relative phase, so noise and signal can interfere constructively or destructively. In the latter case the overall output has a smaller energy than that due to the GW signal alone.

[28] We will see however in Section 7.7.1 that all these time shifts can be taken into account simultaneously performing a single Fast Fourier Transform, which makes the problem computationally feasible.

In the above discussion, we assumed that the output of the detector is a single quantity $\rho$ which, in the absence of noise, has a Gaussian distribution. Actually, we will meet below examples in which we have two outputs $x, y$, each one with its Gaussian noise, which are combined in quadrature, so that $\rho^2 = x^2 + y^2$. In this case the corresponding distribution function can be computed as follows. For the distribution $p_2(\rho|h=0)$ in the absence on signal (where the label 2 reminds us that we have two degrees of freedom $x, y$),[29] we simply have

$$p_2(x,y|h=0)dxdy = \frac{dx}{(2\pi)^{1/2}}\frac{dy}{(2\pi)^{1/2}}e^{-(x^2+y^2)/2} = \rho d\rho \frac{d\theta}{2\pi}e^{-\rho^2/2}. \tag{7.86}$$

If we are not interested in the phase $\theta$ we simply integrate over it, and we get

$$p_2(\rho|h=0) = \rho\, e^{-\rho^2/2}, \tag{7.87}$$

which is called a Rayleigh distribution, or a $\chi^2$ distribution with two degrees of freedom. To compute the distribution in presence of signal, we start from the probability distribution of $x, y$, given that the true GW signal has the values $\bar{x}, \bar{y}$

$$p_2(x,y|\bar{x},\bar{y})dxdy = \frac{1}{2\pi}e^{-\frac{1}{2}[(x-\bar{x})^2+(y-\bar{y})^2]}. \tag{7.88}$$

We pass to polar coordinates, $x = \rho\cos\theta$, $y = \rho\sin\theta$, with $\rho^2 = R$, so $dxdy = \rho\, d\rho d\theta = (1/2)dRd\theta$. To obtain the probability distribution $P_2(R|\bar{R})$ we integrate over the phase $\theta$, and we also integrate over all the values of $\bar{x}, \bar{y}$ with the constraint $\bar{x}^2 + \bar{y}^2 = \bar{R}$, that is,

$$P_2(R|\bar{R})dR = c\frac{dR}{2}\int_0^{2\pi} d\theta \int_{-\infty}^{\infty} d\bar{x}d\bar{y}\,\delta(\bar{x}^2+\bar{y}^2-\bar{R}) \times \frac{1}{2\pi}\exp\{-\frac{1}{2}[(x-\bar{x})^2+(y-\bar{y})^2]\}, \tag{7.89}$$

where $c$ is a normalization constant. The integrals are easily performed expressing also $\bar{x}, \bar{y}$ in polar coordinates, $\bar{x} = r\cos\theta'$, $\bar{y} = r\sin\theta'$, so

$$P_2(R|\bar{R}) = \text{const.}\int_0^{2\pi} d\theta \int_0^{2\pi} d\theta' \int_0^{\infty} d(r^2)\delta(r^2-\bar{R}) \times \exp\{-\frac{1}{2}(R+\bar{R}) + \sqrt{R\bar{R}}\cos(\theta-\theta')\} = \text{const}'.\, e^{-(R+\bar{R})/2}\int_0^{2\pi} d\alpha\, e^{\sqrt{R\bar{R}}\cos\alpha}, \tag{7.90}$$

where $\alpha = \theta - \theta'$. The integral over $\alpha$ gives a modified Bessel function $I_0$. We fix the normalization constant requiring that $\int_0^\infty P_2(R|\bar{R})dR = 1$, and we get

$$P_2(R|\bar{R}) = \frac{1}{2}e^{-(R+\bar{R})/2}I_0\left(\sqrt{R\bar{R}}\right). \tag{7.91}$$

[29] We reserve the capital letter $P$ for the distribution in energy, i.e. in $R$, and use $p$ for the distribution in amplitude, i.e. in $\rho$.



More generally, if $\rho^2 = x_1^2 + \ldots x_n^2$, performing a computation similar to that presented above one finds[30]

$$P_n(R|\bar{R}) = \frac{1}{2}\left(\frac{R}{\bar{R}}\right)^{(n-2)/4} e^{-(R+\bar{R})/2}I_{\frac{n}{2}-1}\left(\sqrt{R\bar{R}}\right). \tag{7.92}$$

In Fig. 7.7 we show the function $P(R|\bar{R})$ given in eq. (7.81), which is appropriate for the case of a single degree of freedom, together with the functions $P_n(R|\bar{R})$ for $n = 2$ and $n = 10$ degrees of freedom, as obtained from eq. (7.92). These distribution functions are known as the non-central chi-squared densities with $n$-degrees of freedom. The average value of $R$ with $n$ degrees of freedom is

$$\langle R\rangle = \int_0^\infty dR\, R\, P_n(R|\bar{R}) = n + \bar{R}, \tag{7.93}$$

and therefore

$$\langle E\rangle = n(kT_n) + \bar{E}, \tag{7.94}$$

while the variance is given by

$$\langle R^2\rangle - \langle R\rangle^2 = 2n + 4\bar{R}. \tag{7.95}$$

## 7.5 Bursts

We now begin to apply the general theory that we have developed, to specific classes of GW signals. We begin with GW bursts. A number of astrophysical phenomena, like supernova explosions or the final merging of a neutron star–neutron star binary system, can liberate a large amount of energy in GWs in a very short time, typically less than a second, and sometimes as small as few milliseconds. We will refer to such signals as GW bursts, and we denote their duration by $\tau_g$. In Fourier space, a GW burst therefore has a continuum spectrum of frequency over a broad range, up to a maximum frequency $f_{\max} \sim 1/\tau_g$.

### 7.5.1 Optimal signal-to-noise ratio

In principle, if we know the form of $\tilde{h}(f)$, we can just plug it into eq. (7.51) to obtain the $S/N$ for a given noise spectral density of the detector. However, bursts come from explosive and complicated phenomena, and it is very difficult to predict accurately their waveform. We can first of all make some simple order-of-magnitude estimates, distinguishing two cases.

#### Narrow-band detectors

In this case the detector is sensitive only to frequencies in a bandwidth $\Delta f$, centered around a frequency $f_0$, and we assume that $\Delta f$ is small with respect to the typical variation scale of the signal in frequency space.

[30] See, e.g. McDonough and Whalen (1995), Sections 4.8 and 4.9.

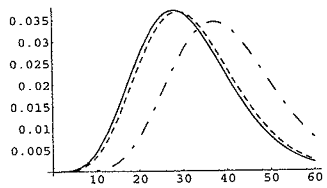


**Fig. 7.7** The probability distribution $P(R|\bar{R})$ given in eq. (7.81) (solid line) compared to $P_n(R|\bar{R})$ with $n = 2$ (dashed line) and with $n = 10$ (dot-dashed), as a function of $R$, for $\bar{R} = 30$.

Outside this interval, the detector is blind and $1/S_n(f)$ in eq. (7.51) becomes practically zero. Inside this small bandwidth $\tilde{h}(f)$ cannot change much, so our ignorance of the precise waveform becomes irrelevant, and in the integrand in eq. (7.51) we can approximate $\tilde{h}(f)$ with $\tilde{h}(f_0)$. Then eq. (7.51) becomes

$$\left(\frac{S}{N}\right)^2 \simeq 4|\tilde{h}(f_0)|^2 \frac{\Delta f}{S_n}\,, \tag{7.96}$$

where $1/S_n$ is an average value of $1/S_n(f)$ in a bandwidth $\Delta f$ centered on $f_0$. This was the typical situation of resonant mass detectors until the 1990s, when the bandwidth $\Delta f$ was only of order a few Hz, around a frequency $f_0 \sim 1$ kHz.[31]

[31] As we will see in Chapter 8, the bandwidth of resonant-mass detectors has subsequently evolved, reaching values of order $\Delta f/f_0 \sim 0.1$, but still, in a first approximation, eq. (7.96) applies. We will also see that, for resonant-mass detectors, $S_n(f)$ is not at all a slowly varying function of $f$ around the resonance frequency $f_0$, so in the estimate (7.96) we must really take an average of $1/S_n(f)$ over the whole useful bandwidth $\Delta f$, and we cannot simply use $1/S_n(f_0)$.

## Broad-band detectors

In this case we get the signal in a bandwidth $(f_{\min}, f_{\max})$ where $f_{\max}$ is the maximum frequency contained in the burst, if the detector is sensitive up to $f_{\max}$, or otherwise is the maximum frequency to which the detector is sensitive. The detailed form of the signal is therefore important, but a first order-of-magnitude estimate can still be obtained writing eq. (7.51) as

$$\left(\frac{S}{N}\right)^2 \sim 4|\tilde{h}|^2 \frac{f_{\max}}{S_n}\,, \tag{7.97}$$

where $\tilde{h}$ is a characteristic value of $\tilde{h}(f)$ over the detector bandwidth and $S_n$ is a characteristic value of $S_n(f)$.

We can translate these order-of-magnitude estimates into limits on the value of the dimensionless GW amplitude $h(t)$ that can be measured. For this we assume for definiteness that the wave comes from a direction such that $F_+ = 1$ and $F_\times = 0$, so that $h(t)$ is the same as the amplitude $h_+(t)$ of the + polarization. In the most general situation, we will also have a factor which depends on $F_+$ and $F_\times$ and reflects the sensitivity of the detector to the given direction and polarization of the wave. To express eq. (7.51) in terms of $h(t)$ we need a model for the signal. For a GW burst of amplitude $h_0$ and duration $\tau_g$, a crude choice could be

$$h(t) = h_0 \quad \text{if } |t| < \tau_g/2 \tag{7.98}$$

and $h(t) = 0$ if $|t| > \tau_g/2$. We can write it more compactly as

$$h(t) = h_0 \tau_g\, \delta_{\rm reg}(t)\,, \tag{7.99}$$

where $\delta_{\rm reg}(t)$ has a rectangular shape of unit area, $\delta_{\rm reg}(t) = 1/\tau_g$ for $|t| < \tau_g/2$ and $\delta_{\rm reg}(t) = 0$ for $|t| > \tau_g/2$. For $\tau_g \to 0$, $\delta_{\rm reg}(t)$ becomes a Dirac delta. More generally, for a burst we can model $h(t)$ as in eq. (7.99), choosing for $\delta_{\rm reg}(t)$ a smooth function of unit area which goes to zero rather fast for $|t| \gtrsim \tau_g$. Performing the Fourier transform this gives

$$|\tilde{h}(f)| \sim h_0 \tau_g\,, \tag{7.100}$$



times a dimensionless function of the frequency, numerically of order one, and whose details depend on the precise waveform $\delta_{\rm reg}(t)$ chosen.

Actually, rather than using a function $\delta_{\rm reg}(t)$ with a unit area, it can be more convenient to write $h(t) = h_0 g(t)$, with $g(t)$ some function peaked at $t = 0$ and with $g(0) = O(1)$, so that the value of $h(t)$ near the peak is of order $h_0$ (rather than $h_0 \delta_{\rm reg}(0)$ as in eq. (7.99)). A simple waveform of this type is a Gaussian,

$$h(t) = h_0\, e^{-t^2/\tau_g^2}\,, \tag{7.101}$$

whose Fourier transform is

$$\tilde{h}(f) = h_0 \tau_g \sqrt{\pi}\, e^{-(\pi f \tau_g)^2}\,. \tag{7.102}$$

A waveform with a somewhat more realistic shape is a sine-Gaussian, i.e. a Gaussian modulated by a frequency $f_0$,

$$h(t) = h_0 \sin(2\pi f_0 t)\, e^{-t^2/\tau_g^2}\,, \tag{7.103}$$

shown in Fig. 7.8. Its Fourier transform is

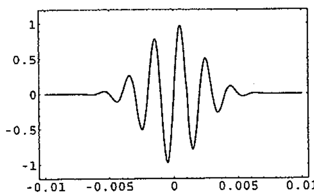


**Fig. 7.8** The sine-Gaussian function $\sin(2\pi f_0 t)e^{-t^2/\tau_g^2}$, for $\tau_g = 3$ ms and $f_0 = 500$ Hz, as a function of $t$ (in seconds).

$$\tilde{h}(f) = h_0 \tau_g\, i \frac{\sqrt{\pi}}{2} \left[e^{-\pi^2 (f-f_0)^2 \tau_g^2} - e^{-\pi^2 (f+f_0)^2 \tau_g^2}\right], \tag{7.104}$$

and is shown in Fig. 7.9. If $4\pi^2 f_0^2 \tau_g^2 \gg 1$, near $f = f_0$ the second term in brackets is negligible with respect to the first (while close to $f = 0$ it cancels the first term so that $\tilde{h}(0) = 0$), and we basically have a Gaussian in frequency space, centered at $f = f_0$, and with a value at the maximum

$$|\tilde{h}(f_0)| \simeq h_0 \tau_g \frac{\sqrt{\pi}}{2}\,. \tag{7.105}$$

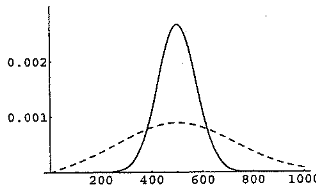


**Fig. 7.9** $|\tilde{h}(f)|$ (in units of $h_0$) for a sine-Gaussian function with $f_0 = 500$ Hz, $\tau_g = 3$ ms (solid line) and for a sine-Gaussian function with $f_0 = 500$ Hz, $\tau_g = 1$ ms (dashed line).

Writing $f = f_0 + \Delta f$ we see that the width of the maximum $\Delta f$ is of order $1/(\pi\tau_g)$, so $\Delta f/f_0 \sim 1/(\pi f_0 \tau_g)$. For $\pi f_0 \tau_g \ll 1$, $\tilde{h}(f)$ becomes relatively flat while for $\pi f_0 \tau_g \gg 1$ it is sharply peaked around $f_0$. Using eqs. (7.96) and (7.97) we can estimate the minimum value of the dimensionless GW amplitude $h_0$ that can be detected at a given level of the signal-to-noise ratio $S/N$. For narrow-band detectors eq. (7.96) gives, using for definiteness the value $|\tilde{h}(f_0)| \simeq h_0 \tau_g (\sqrt{\pi}/2)$ appropriate for a sine-Gaussian waveform,

$$(h_0)_{\min} \sim \frac{1}{\tau_g} \left(\frac{S_n}{\pi \Delta f}\right)^{1/2} (S/N)\,, \tag{7.106}$$

while for broad-band detectors eq. (7.97) gives

$$(h_0)_{\min} \sim \frac{1}{\tau_g} \left(\frac{S_n}{\pi f_{\max}}\right)^{1/2} (S/N)\,. \tag{7.107}$$

The precise numerical factors, of course, depend on the choice of the waveform, so to fix the numerical coefficients in eqs. (7.106) and (7.107) we must know the shape of the signal $\tilde{h}(f)$, use the exact form of the noise $S_n(f)$, and perform the integral in eq. (7.51). We see from eqs. (7.106) and (7.107) that in a narrow-band detector the minimum detectable amplitude is higher by a factor $(f_{\max}/\Delta f)^{1/2}$, compared to a detector which is able to maintain the same typical sensitivity $S_n$ over a broad bandwidth. This reflects the fact that the narrow-band detector has access only to a portion of the Fourier modes of the burst.

Rather than expressing the result in terms of $h_0$, it is also common to use the so-called root-sum-square (rss) amplitude $h_{\rm rss}$, defined by

$$h_{\rm rss}^2 = \int_{-\infty}^{\infty} dt\, h^2(t) = \int_{-\infty}^{\infty} df\, |\tilde{h}(f)|^2 . \tag{7.108}$$

For the Gaussian (7.101) we have

$$h_{\rm rss}^2 = h_0^2 \tau_g \sqrt{\frac{\pi}{2}}, \qquad \text{(Gaussian)}, \tag{7.109}$$

while, for the sine-Gaussian (7.103),

$$h_{\rm rss}^2 = h_0^2 \tau_g \sqrt{\frac{\pi}{8}}\, (1 - e^{-2\pi^2 f_0^2 \tau_g^2}), \qquad \text{(sine-Gaussian)}. \tag{7.110}$$

Observe that, dimensionally, $h_{\rm rss} \sim ({\rm time})^{1/2}$, so $h_{\rm rss}$ is conventionally quoted in ${\rm Hz}^{-1/2}$, as the strain sensitivity.

To have an idea of the numerical values of $h_{\rm rss}$ (or, equivalently, of $h_0$) that could be obtained from astrophysical phenomena, we can compute the energy released in GWs by an event which produced, at the detector, a given value of $h_{\rm rss}$. This can be obtained from the expression for $dE/dAdf$ given in eq. (1.159). Observe however that, for a wave coming from an arbitrary direction and with arbitrary polarization, a detector does not measure directly $\tilde{h}_+(f)$ and $\tilde{h}_\times(f)$ but rather the combination $\tilde{h}(f) = F_+\tilde{h}_+(f) + F_\times \tilde{h}_\times(f)$, where $F_{+,\times}$ are the detector pattern functions. For definiteness, we consider a GW coming from the optimal direction for the + polarization, so we take $F_+ = 1$ and $F_\times = 0$, and for $\tilde{h}_+(f)$ we take the sine-Gaussian waveform (7.104). We substitute this into eq. (1.159) and we get the total energy $\Delta E_{\rm rad}$ radiated by the source in GWs,[32]

$$\Delta E_{\rm rad} \simeq \left(\frac{\pi}{2}\right)^{3/2} \frac{\pi r^2 c^3}{G}\, h_0^2 \tau_g f_0^2 = \frac{\pi^2 r^2 c^3}{G}\, h_{\rm rss}^2 f_0^2 . \tag{7.111}$$

Inserting the numerical values,

$$\boxed{\Delta E_{\rm rad} \simeq 1\times 10^{-2} M_\odot c^2 \left(\frac{r}{8\,{\rm kpc}}\right)^2 \left(\frac{h_{\rm rss}}{10^{-19}\,{\rm Hz}^{-1/2}}\right)^2 \left(\frac{f_0}{1\,{\rm kHz}}\right)^2 ,} \tag{7.112}$$

[32] In the computation, we neglect the term $\exp\{-\pi^2(f+f_0)^2\tau_g^2\}$ in eq. (7.104), which is small with respect to $\exp\{-\pi^2(f-f_0)^2\tau_g^2\}$ and, when we integrate eq. (1.159) over $df$, we replace $f^2\exp\{-\pi^2(f-f_0)^2\tau_g^2\}$ with $f_0^2\exp\{-\pi^2(f-f_0)^2\tau_g^2\}$ and we extend the resulting integral from $-\infty$ to $+\infty$.

where in the second line we normalized $r$ to a value of order of the distance to the galactic center. Recall that in the above we assumed a wave coming from optimal direction. For an ensemble of waves with arbitrary direction and polarization, we must also take into account the average over the pattern functions of the detector. For an interferometer, this is a factor 2/5 (see Table 7.1), so on average a burst coming from arbitrary direction, in order to produce a given signal $h_{\rm rss}$ in the detector, had an energy larger by a factor 5/2, compared to eq. (7.112). We see that a burst at the kHz, with $h_{\rm rss} = 10^{-19}\,{\rm Hz}^{-1/2}$, carries away about $10^{-2}$ solar masses in GWs, if it comes from a source located at typical galactic distances.

Taking $10^{-2} M_\odot c^2$ as a reference value for $\Delta E_{\rm rad}$ (which, as we will see in Vol. 2, is the maximum value that can be reasonably expected in cataclysmic events involving solar mass objects. Even larger energies can be released in the merging of very massive black holes), we see that a detector must reach at least a sensitivity to $h_{\rm rss}$ of order $10^{-19}\,{\rm Hz}^{-1/2}$ to have some chance of detecting GW bursts from the galactic center. To be able to see a burst which releases $10^{-2}$ solar masses in the Virgo cluster of galaxies, which is at $r \sim 14$ Mpc, one rather needs to be able to reach $h_{\rm rss} \simeq 6\times 10^{-23}\,{\rm Hz}^{-1/2}$ or, from eq. (7.110) with $\tau_g = 1$ ms, a value of the dimensionless amplitude $h_0 \simeq 2\times 10^{-21}$.

### 7.5.2 Time–frequency analysis

The matched filtering technique that we have discussed in Section 7.3 works well if we know the form of the signal, or if we can parametrize it with a limited number of free parameters, so that it becomes practically feasible to put a sufficiently fine grid in this parameter space, and repeat the search for each point of the grid. As we will discuss in the next sections, this can be the case for the inspiral of compact binaries and, partly, for the signals due to pulsars.

Concerning bursts, the situation is different. In general, bursts may come from complicated explosive phenomena, such as supernovae, or from processes such as the final merging of coalescing binaries, which are difficult to model. In a narrow-band detector, such as resonant bars, we only have access to a narrow range of Fourier components of the signal. Thus, in a first approximation it is reasonable to model the signal as flat in frequency, i.e. as a Dirac delta in time and, as a next step, we can use more realistic modelizations such as the Gaussian and the sine-Gaussian waveforms described above. However, in a broad-band detector, the difference between these simple modelizations and the real waveform will become important. Thus, to exploit optimally the capabilities of a broad-band detector, one is lead to consider also other methods, which are sub-optimal with respect to matched filtering when the waveform is know precisely (since we have seen that, if the waveform is known, matched filtering is the optimal strategy) but might be more robust in the absence of detailed knowledge of the signal.

Such search algorithms can be obtained working in the time–frequency plane. To understand the usefulness of the time–frequency representation, suppose at first that we have a function $s(t)$ defined on the whole real axis $-\infty < t < \infty$. We can take its Fourier transform $\tilde{s}(f)$ and compute from it the power spectrum $|\tilde{s}(f)|^2$. A plot of the power spectrum against the frequency will enable us to see immediately what are the dominant Fourier modes. However, this power spectrum knows nothing about *when* things happened.[33]

The simplest way to recover partly this information is to take the Fourier transform not on the whole real line, but on segments of length $\delta t$. When we Fourier transform the function $s(t)$ on the interval $0 < t < \delta t$ and we plot the resulting power spectrum, we find the Fourier modes that dominated the function, during this temporal span. We can then repeat it for $\delta t < t < 2\delta t$, etc. Of course, on a finite segment of length $\delta t$, the resolution in frequency is finite, and is $1/\delta t$, so we are giving up the fine details in frequency space, but we gain an understanding of when things happened. That is, rather than working in frequency space, with an arbitrarily good resolution, it can be convenient to work in the time–frequency plane, making a good compromise between the accuracy in frequency and the accuracy in time.

This is particularly important when we are looking for transient phenomena, such as GW bursts. Suppose that we are unable to compute the detailed waveform of a burst, as it is typically the case, but still we can give a reasonable estimate of its total duration $\delta t$, and of the frequency range $f_1 < f < f_2$ where most of its power should be concentrated. Then, a useful search strategy is as follows.

First of all, it is convenient to work in a discretize space. Recall that the output of a detector is sampled at some rate $1/\Delta t$. Then, we can split the output into time segments, and inside each segment the output $s(t)$ is given by the discrete set of values

$$s_j \equiv s(t_{\rm start} + j\Delta t)\,, \tag{7.113}$$

where $t_{\rm start}$ is the start time of the segment considered, $j = 0, \ldots, N$, and $\delta t = N\Delta t$ is its length. We can then perform a discrete Fourier transform over the segment $\delta t$ by writing

$$\tilde{s}_k = \sum_{j=0}^{N-1} n_j \exp\left\{\frac{2\pi i}{N} jk\right\}, \tag{7.114}$$

or

$$\tilde{s}_k = \sum_{j=0}^{N-1} n(t_j) \exp\{2\pi i (t_j - t_{\rm start}) f_k\}\,. \tag{7.115}$$

where $t_j = t_{\rm start} + j\Delta t$ and

$$f_k = \frac{k}{N\Delta t} = \frac{k}{\delta t}\,. \tag{7.116}$$

We see that frequencies are spaced by $1/\delta t$, up to a maximum frequency equal to $N/\delta t$ (since eq. (7.114) is periodic under $k \to k + N$), which of course is just the sampling frequency $1/\Delta t$.[34] We can write as usual $s_i = n_i + h_i$, where $n_i$ is the noise and $h_i$ a putative signal, and we define the Fourier transforms $\tilde{n}_k$ and $\tilde{h}_k$ as in eq. (7.114). The discrete version of eq. (7.6) is obtained replacing the Dirac delta by a Kronecker delta,

$$\langle \tilde{n}_k^* \tilde{n}_{k'} \rangle = \delta_{kk'} \frac{1}{2} S_k\,, \tag{7.117}$$

where we used the short-hand notation $S_k \equiv S_n(f_k)$.

If the only theoretical expectation that we have about a signal is that it should have a duration $\delta t$, and should have most of its power in a frequency band $f_1 < f < f_2$, with $f_1 = k_1/\delta t$, $f_2 = k_2/\delta t$, and $f_2 - f_1 \equiv \delta f$, we can form, for each possible start time $t_{\rm start}$, the quantity

$$\mathcal{E} = 4 \sum_{k=k_1}^{k_2} \frac{|\tilde{s}_k|^2}{S_k}\,, \tag{7.118}$$

which is called the *excess power statistic*. We collect the values of $\mathcal{E}$ for all possible start time and, if we find a value above some given threshold, we record it as an event.[35] To understand what is a statistically significant value of $\mathcal{E}$, observe that $\mathcal{E}$ is formed from $k_2 - k_1$ independent complex variables $s_k$. Since $(k_2 - k_1)/\delta t = f_2 - f_1 \equiv \delta f$, the number of independent real variables is

$$\mathcal{N} = 2\delta f \delta t\,, \tag{7.119}$$

i.e. twice the area of the time–frequency plane explored. Therefore, even in the absence of any GW signal in the data, the average value of $\mathcal{E}$ is of order $\mathcal{N}$.[36] This means that a real GW signal, in order to be visible in $\mathcal{E}$ against the noise with a signal-to-noise ratio of order one, must give a contribution to $\mathcal{E}$ of order $\mathcal{N}$. From eq. (7.119), $\mathcal{N} \geqslant 2$ (the uncertainty principle, in a quantum langauge) and, depending on the situation, one can have $\mathcal{N} \gg 1$. Comparing with eq. (7.51), we see that, if we knew the waveform and we could make a matched filtering, such a signal would produce a value of $S/N$ of order $\mathcal{N}^{1/2}$. In other words, using the excess power statistic, we can detect with a signal-to-noise ratio of order one, a signal that with matched filtering would emerge with signal-to-noise ratio of order $\mathcal{N}^{1/2}$.[37] This is not surprising since we know that, when we have the waveform, the matched filtering maximizes the signal-to-noise ratio. However, the excess power statistic only needs very crude information about the signal, namely its duration and its typical frequency range, and is therefore much more robust. Furthermore, it can be proved that, if the only information on the signal is its duration and its bandwith, the excess power method is the optimal one.

From the above discussion, it is clear that the method is viable only when $\mathcal{N}$ is not too large. For instance, for the inspiral phase of a coalescing NS-NS binary, as observed in a ground-based interferometer, we see from eq. (4.21) that the signal enters the bandwidth of the interferometer, say at 40 Hz, when the time to coalescence is $\tau = 25$ s, and sweeps in frequency up to the kHz. Taking $\delta f \sim 1000$ Hz and $\delta t \sim 25$ s, we get $\mathcal{N} \sim 5 \times 10^4$ and $\mathcal{N}^{1/2} \sim 200$, so the excess power method would allow

[33] The information about time localization, which is encoded in $s(t)$, is of course still encoded in $\tilde{s}(f)$, since from $\tilde{s}(f)$ we can get back $s(t)$ uniquely. However, it is lost in $|\tilde{s}(f)|^2$, since it was contained in the phase relation between the Fourier components, and this information is obliterated when taking the modulus. A nice example (taken from van den Berg (1999), a textbook on wavelets) is obtained if we take $s(t)$ to be a classical symphony. Then its power spectrum will immediately reveal the dominating keys: the groundtones and their harmonics. Suppose now that we play the parts in a different order, and we even interchange smaller parts within the parts. The power spectrum would not change at all, while to the ear, which actually makes a time–frequency analysis, the result is very different.

[34] To fix the ideas, one can consider that the sampling rate $1/\Delta t$ of an interferometer is typically of order 10 – 20 kHz. We can imagine that we are searching for bursts of duration $\delta t = 0.5$ s, so frequency space is split into bins of width 2 Hz. Restricting the search to the frequency range where interferometers are most sensitive, which corresponds to a bandwidth $\Delta f = O(200)$ Hz around peak sensitivity, we have a total of $O(100)$ bins in frequency, for each value of the start time of the segment.

[35] In the sense defined in Note 2 on page 337.

[36] More precisely, in the presence of Gaussian noise $\mathcal{E}$ follows a $\chi^2$ distribution with $\mathcal{N}$ degrees of freedom, while in the presence of signal it follows the corresponding non-central $\chi^2$ distribution, see eq. (7.92).

[37] For a more accurate estimate of the signal-to-noise ratio obtained restricting the frequency bandwidth, i.e. performing a band-pass filter, see Section II of Flanagan and Hughes (1998a).

us to detect signals only when their signal-to-noise ratio, with matched filtering, is of order several hundreds. Thus, for inspiraling binaries, the excess power method is not at all competitive. Furthermore it is not needed, since in this case we have precise calculations of the waveform in the inspiral phase, as we saw in Chapter 5.

The situation is different for the merging phase of a BH-BH coalescence. In this case the maximum value of $f$ can be estimated to be of order $f_{\rm qnr}$, where $f_{\rm qnr}$ is the ringing frequency of the fundamental quasi-normal mode of the black hole. To include the power radiated by the BH in its higher quasi-normal modes, an estimate of order $2f_{\rm qnr}$ could be more appropriate. Black hole normal modes will be discussed in Vol. 2, and we will see that $f_{\rm qnr}$ can reach a maximum value (for rapidly spinning BHs of mass $m$) $f_{\rm qnr} = c^3/(2\pi Gm)$. Observe that this is quite larger than the maximum frequency (4.39) at which the inspiral phase ends, so we finally take $\delta f \sim 2f_{\rm qnr} \sim c^3/(\pi Gm)$. As for the merging time, we can roughly estimate that it should not be much larger than $r_{\rm ISCO}/c$, where $r_{\rm ISCO} = 6Gm/c^2$ is the radius of the innermost stable circular orbit in a Schwarzschild geometry, see eq. (4.38). Taking for instance $\delta t \sim 2r_{\rm ISCO}/c = 12Gm/c^3$, we get the estimate $\mathcal{N}^{1/2} \sim 2$, so the loss in sensitivity with respect to optimal filtering is not large. This is quite important, considering that the merging phase is very difficult to model.[38]

The time–frequency method discussed here can be generalized in various directions. One possibility is to consider wavelets. These are generalizations of the Fourier transform, in which to a function $s(t)$ is associated a function $S(f, t_0)$ of two parameters, of which $f$ is the frequency and $t_0$ is the position in time of the signal,

$$S(f,t_0) = \int_{-\infty}^{\infty} \psi^*_{f,t_0}(t)s(t)\,. \tag{7.120}$$

The simplest example consists in taking

$$\psi_{f,t_0}(t) = e^{-ift}\psi(t-t_0)\,, \tag{7.121}$$

where $\psi(t-t_0)$ is a window function centered around $t_0$. This Windowed Fourier Transform, or Gabor transform, as it is called, is essentially what we have used above (more precisely, we used its discrete version), with a sharp window function. Other choices of window functions, such as a Gaussian, are more commonly used in signal analysis.

A possible drawback of a choice such as eq. (7.121) is that the temporal window has a fixed size, independently of the frequency. In most type of signals, however, there is a correlation between the characteristic frequency of a given segment and the time duration of the segment, such that low-frequency pieces tend to last longer. To take this into account, the wavelet transform is defined by choosing a window function of the form $\psi(f(t-t_0))$ (times a normalization constant $\sqrt{f}$), which depends explicitly on $f$.[39] In this way, at high frequencies the temporal window is shorter, so we have a better time resolution. In a sense, wavelets provide a "microscope" that, at each point in time of the signal, zooms in and out, depending on the frequency scale of the signal. The other crucial property is that it is possible to choose wavelets so that they form an orthonormal basis, and the signal can therefore be decomposed uniquely into its component with respect to this wavelet basis, just as in the Fourier transform. Wavelets are by now widely used in signal analysis in many branches of science, and many possible choices of wavelets are available, depending on the problem at hand, see the Further Reading.

Another generalization of the time–frequency analysis discussed here consists in marking as "black" the bins in the time–frequency plane where an indicator such as the excess power statistic goes above a threshold value, and searching for structures of black bins, such as clusters. This is basically a variant of the Hough transform that we will discuss in Section 7.6.3, in the context of periodic signals.

[38] For fast spinning BHs the coalescence time will be longer, since they must first shed some angular momentum before setting into their final state, in order that the angular momentum of the final BH does not exceed the maximum value allowed for rotating BHs. The estimate of $\mathcal{N}$ depends on the angular momentum of the BHs, but a typical value can be $\mathcal{N}^{1/2} \sim 5$. See Flanagan and Hughes (1998a), Section IIIE.

[39] In the literature on wavelets, this is actually written as $a^{-1/2}\psi((t-t_0)/a)$, where $a$ is a dimensionless quantity that rescales a characteristic frequency implicit in the function $\psi$.

### 7.5.3 Coincidences

Given that GW bursts can have a very short duration, even smaller than a millisecond, the output of ground-based detectors are sampled with a very high frequency, typically $O(10)$ kHz. In one year there are about $3\times10^{10}$ ms, so even a fluctuation with a probability $\sim 10^{-10}$ is bound to occur on average in one year worth of data. Then eq. (7.84) suggests, for bursts, a threshold on the amplitude signal-to-noise ratio of order $\rho_t \simeq 6$, in order to have just a few false alarms per year in a detector. However, this only eliminate Gaussian noise. GW bursts are particularly well simulated by non-Gaussian events such as microcreeps in the materials or sudden external mechanical or electromagnetic disturbances. In some cases the external disturbance can be identified, and the corresponding event is therefore vetoed, but in most cases this is impossible. To eliminate these non-Gaussian noise, the only possibility is to perform coincidences between different detectors.[40]

Using two or more detectors in coincidence is a standard practice in physics, at least since the early days of cosmic ray research. The idea is that, if two detectors are far apart, their noise are mostly uncorrelated,[41] and the probability of an accidental coincidence is small, while a GW should excite both detectors nearly simultaneously. Below we discuss some of the issues that must be addressed in order to apply this idea to GW detectors.

#### Relative orientation of the detectors

We have seen that the response of a detector to a GW depends on the relative orientation between the detector and the source. To perform coincidences between two or more detectors, it is therefore optimal to orient them, taking into account their difference in location, so that their response to an incoming GW signal is the same, or at least as similar as possible, for all of them. Otherwise, a real signal can be missed simply because, when one detector was oriented favorably with respect to the source, the other had a very poor sensitivity for the same direction.[42]

[40] Observe that the use of coincident detectors also allows us to lower the threshold necessary for eliminating Gaussian noise, since now the false alarm probability, for uncorrelated detector noise, is the square of (7.84). For instance, in the example above, the threshold $\rho_t \simeq 6$ valid for a single-detector search becomes $\rho_t \simeq 4.5$ in a two-detector correlation (even neglecting all consistency check discussed below).

[41] With some exceptions. For example, seismic or electromagnetic disturbances might propagate from one detector's site to the other.

[42] On the other hand, detectors with different orientation can perform independent measurement of the signal, allowing to disentangle the polarizations and the arrival direction of the wave, see the discussion on page 342.

### Coincidence window

Each GW detector has its temporal resolution, which might for instance be of the order of few ms.[43] Given two detectors, with variances $\sigma_1$ and $\sigma_2$ on the arrival time of their respective events, the corresponding variance in the coincidence search is $\sigma_{12} = \sqrt{\sigma_1^2 + \sigma_2^2}$, and therefore one can ask that the events be coincident within $k$ standard deviations $\sigma_{12}$ (e.g. $k = 3$ can be a typical choice). To this uncertainty one must add the light travel time $(\Delta t)_{\rm light}$ between the two detectors since, depending on the source location, either the first or the second detector will be hit by the wavefront a time up to $(\Delta t)_{\rm light}$ before the other. So finally one requires that the arrival times $t_1$ and $t_2$ in the two detectors are within a coincidence window

$$|t_1 - t_2| \leqslant (\Delta t)_{\rm light} + k \left(\sigma_1^2 + \sigma_2^2\right)^{1/2} . \qquad (7.122)$$

This typically results in a coincidence window of the order of a few tens of ms.

[43] This depends not only on the sampling time of the detector, but also on other factors, in particular on the signal-to-noise ratio of the event, since noise combines with the GW signal distorting and broadening its shape.

### Energy consistency

Another possible handle to discriminate between accidental coincidences and true GW signals is the compatibility of the signal in the two (or more) detectors. Ideally, if the GW signal is much larger than the noise, two detectors oriented in the same way should register the same energy flux, when a GW hits them. In contrast, two events due to noise which by chance happen simultaneously in the two detectors, should have uncorrelated energies. However, in practice, at moderate values of the signal-to-noise ratio the signal $h(t)$ induced by the GW combines with the noise $n(t)$ and, depending on the relative phase of these contributions, the output $h(t) + n(t)$ fluctuates and can be either larger or smaller than the value that would be induced by the GW. Therefore, as we computed in Section 7.4.3, one has a probability distribution for the amplitudes (or for the energies) measured in the two detectors, and the compatibility criterion must take into account this probability distribution. This procedure also requires that the two detectors have a sufficiently reliable calibration in energy.

### Waveform consistency

A broadband detector has rather detailed information on the waveform, and a consistency condition between the waveforms observed in the two detectors can be imposed. For instance, one of the algorithms used by LIGO for generating candidate events is based on the identification of connected regions ("clusters") in the time–frequency plane where the power is not consistent, statistically, with Gaussian noise, as discussed in Section 7.5.2. Then each event is characterized by its bandwidth $(f_{\rm min}, f_{\rm max})$, i.e. by the low and high frequency bounds of the cluster. One can then require, for instance, that the bandwidth of events in different detectors have an overlap, or at least that they are separated in frequency by no more than a fixed window $\Delta f$.



### Background estimation

After having applied all these cuts, we can still have accidental coincidences that, by chance, passed them. However, the residual number of accidental coincidences can be estimated very reliably. First of all, one can simply predict it from the observed event rate in a single detector, assuming that the noise is stationary. But in fact the most direct estimation of the background is obtained using a shifting algorithm which, together with many other techniques used in GW research, was introduced by Weber. The procedure consists simply in shifting the data stream of one detector with respect to the other by a time step significantly longer than the coincidence window, say 2 s, and counting the number of coincidences obtained after shifting (subject to the same requirements on the coincidence window and energy compatibility imposed on the coincidences at zero time shift). These coincidences, of course, are now all accidental, since the shift has been chosen much larger than the coincidence window and therefore of the uncertainties in the arrival times. We then repeat the procedure with a different shift, say 4 s, and we count again the number of accidentals. One can repeat the procedure for many different shifts (the overall time shift must however be short compared to the time-scale over which the event rate in a single detector changes substantially). We then average over these shifts, and we have a rather accurate estimate of the average number of accidental coincidences, its variance, and more generally their distribution (which is found experimentally to be a Poisson distribution, as expected whenever we count a number of discrete independent events), and we can also study how these quantities depend on the energy of the events.

## 7.6 Periodic sources

While a burst source is typically radiating only for a period of less than a second, a periodic source emits continuously an almost monochromatic signal, so the limit on its observation comes from the total available observation time, which can be of order of years. Our intuitive discussion of matched filtering showed that, if we can follow a signal for a time $T$, the minimum level of signal that we can extract from the noise scales as $1/T^{1/2}$, see eqs. (7.39) and (7.40). This means that, for periodic waves, we can extract from the noise a signal with an amplitude $h_0$ much smaller than the one that can be measured in the case of bursts. This opportunity, however, also comes at the expense of some complications, since we must able to track carefully the signal for a long period. We already met a similar situation in Chapter 6, where we studied the timing formula for the radio signals of pulsars, and we saw that there are two main issue to address: the intrinsic changes of the frequency of the source, and the modulation of the signal due to the motion of the

Earth.[44]

If, for a moment, we neglect these effects, a periodic source emitting GWs at a frequency $f_0$ produces in the detector a signal

$$h(t) = F_+(\theta,\phi)\,h_+(t) + F_\times(\theta,\phi)\,h_\times(t)\,, \tag{7.123}$$

where

$$h_+(t) = h_{0,+}\cos(2\pi f_0 t)\,, \tag{7.124}$$
$$h_\times(t) = h_{0,\times}\cos(2\pi f_0 t + \alpha)\,. \tag{7.125}$$

We take by definition $f_0 > 0$; $h_{0,A}$ are the real amplitudes for the two polarizations ($A = +, \times$), and $\alpha$ is their relative phase. We denote by $\theta, \phi$ the angles that define the propagation direction $\hat{\mathbf{n}}$ of the GW from the source to us, so the polar angles of the source, as seen from the Earth, are $\theta_s = \pi - \theta$ and $\phi_s = \phi + \pi$.

Assuming for the moment that the source is, intrinsically, perfectly periodic, still the motion of the Earth modifies eqs. (7.123)–(7.125) as follows.

- Because of the Earth's rotation, the apparent position of the source in the sky changes, so the angles $\theta$ and $\phi$ which appear in the pattern functions change with time, and are periodic functions of sidereal time, with period one sidereal day. If we are tracking a specific source in the sky, the time dependence of the pattern functions, $F_A(\theta(t), \phi(t))$, must therefore be taken into account, and this produces a modulation of the amplitude of the signal.
- Because of the Earth's rotation and of its revolution around the Sun (or, more precisely, because of its motion with respect to the Solar System Barycenter, as discussed in Chapter 6), the relative velocity of the Earth and the source changes with time, and this produces a time-varying Doppler shift in the frequency.

As a consequence, $h(t)$ is not a simple monochromatic signal. We will come back to these amplitude and phase modulations in Sections 7.6.1 and 7.6.2. For the moment, however, we restrict to an observation time $T$ sufficiently short, so that these amplitude and phase modulations can be neglected. For the amplitude modulation due to the Earth's rotation, this requires of course $T \ll 1$ day, while for the Doppler effect we will quantify this requirement in Section 7.6.2. In this limit $h(t)$ becomes monochromatic, with a frequency $f_0$.

In this simplified setting the form of the matched filter becomes obvious: we must limit ourselves to a bandwidth as small as possible around $f_0$, since enlarging the bandwidth we accept more noise but we add no further signal. If $T$ is the total observation time, our resolution in frequency is $1/T$, see eq. (7.10), and therefore a bandwidth as small as possible means $\Delta f \simeq 1/T$. Formally, we can obtain the same result using eq. (7.49). From eqs. (7.123)–(7.125) we have, for $f > 0$,

$$\tilde{h}(f) = \delta(f - f_0)\,\frac{1}{2}\left[F_+(\theta,\phi)h_{0,+} + F_\times(\theta,\phi)h_{0,\times}e^{-i\alpha}\right]\,, \tag{7.126}$$

[44] For GWs, propagation effects between the source and the Earth, such as dispersion in the interstellar medium, are totally irrelevant, given the smallness of gravitational cross-sections.



and therefore eq. (7.49) gives

$$\tilde{K}(f) = \delta(f - f_0)\,, \tag{7.127}$$

apart from an arbitrary constant, in which we also reabsorbed $1/S_n(f_0)$. Of course, the Dirac delta is a mathematical idealization, and if we measure for a total observation time $T$ we must replace it by a regularized Dirac delta,

$$\delta(f) = \int_{-\infty}^{\infty} dt\, e^{i2\pi f t} \rightarrow \int_{-T/2}^{T/2} dt\, e^{i2\pi f t}\,, \tag{7.128}$$

which has a support over a range $\Delta f \sim 1/T$ and satisfies $\delta(0) = T$. Then eq. (7.51) becomes

$$\left(\frac{S}{N}\right)^2 = \left|F_+(\theta,\phi)h_{0,+} + F_\times(\theta,\phi)h_{0,\times}e^{-i\alpha}\right|^2 \int_0^\infty df\, \frac{\delta(f - f_0)\delta(0)}{S_n(f)}$$
$$= \left|F_+(\theta,\phi)h_{0,+} + F_\times(\theta,\phi)h_{0,\times}e^{-i\alpha}\right|^2 \frac{T}{S_n(f_0)}\,. \tag{7.129}$$

Not surprisingly, the signal-to-noise ratio increases if we increase the observation time, and the dependence $S/N \sim \sqrt{T}$ is what we already found using heuristic arguments in eqs. (7.39) and (7.40).

In general, the frequency $f_0$ is not known in advance. However, for an exactly periodic signal, we do not need to repeat the matched filtering procedure separately for each value of the unknown parameter $f_0$. In fact, from eq. (7.42), when $\tilde{K}(f) = \delta(f - f_0)$ the signal is simply $S = \tilde{h}(f_0)$, and the values of $\tilde{h}(f)$ for all $f$ can be computed at once performing a single Fast Fourier Transform (FFT), which is a particularly efficient algorithm.

If this were the end of the story, the search for periodic signals would simply consist in performing a single FFT on a stretch of data of length $T$, and looking for lines in the power spectrum. The signal-to-noise ratio of these line should improve with the observation time as $\sqrt{T}$. We will see in Section 7.6.1 and especially in Section 7.6.2 that the full story is more complicated.

### 7.6.1 Amplitude modulation

As we pointed out above, the pattern functions depend on time because of the Earth's rotation, and are therefore periodic functions of sidereal time, with a period of one sidereal day. In the matched filtering, we must take this into account, and this results in a different amplitude modulation for each possible source position. We will discuss in the next sections how to efficiently scan the parameter space, in order to take this effect into account.

If we want to estimate the effect of this modulation on the sensitivity, we can however simply observe that, for integration times $T$ longer than one day, the effect of this amplitude modulation can be taken into account averaging eq. (7.129) over the apparent motion of the source in one sidereal day, i.e. averaging over all values of the right ascension of the source, and over a range of values of the declination which depend on the specific orbit of the source. In a first approximation, we can replace this average with an average over the solid angle and over the polarization angle $\psi$.[45] From eq. (7.129), using eqs. (7.33) and (7.35), we then find

$$\left(\frac{S}{N}\right)^2 = \langle F_+^2 \rangle \left(\frac{T}{S_n(f_0)}\right) h_0^2 \,, \tag{7.130}$$

where

$$h_0^2 = h_{0,+}^2 + h_{0,\times}^2 \,. \tag{7.131}$$

The values of $\langle F_+^2 \rangle$ for various detectors are given in Table 7.1, recalling that $\langle F_+^2 \rangle = F/2$. We can also rewrite eq. (7.130) as

$$\frac{S}{N} = \frac{h_0}{h_n} \,, \tag{7.132}$$

defining the dimensionless quantity $h_n$,

$$h_n = \frac{1}{\langle F_+^2 \rangle^{1/2}} \left(\frac{S_n(f_0)}{T}\right)^{1/2} . \tag{7.133}$$

Therefore $h_n$ is the GW amplitude that can be measured by the detector, for a periodic signal, at $S/N = 1$ (assuming that we have been able to correct for the phase modulation, see next section). More generally, the minimum amplitude that can be detected at a given value of $S/N$ is

$$(h_0)_{\rm min} = \frac{S/N}{\langle F_+^2 \rangle^{1/2}} \left(\frac{S_n(f_0)}{T}\right)^{1/2} . \tag{7.134}$$

It is instructive to compare this result with the minimum burst amplitude detectable at a broad-band detector, eq. (7.107). Recalling that $S_n(f)$ has dimensions 1/Hz, i.e. dimensions of time, we must divide it by a time in order to obtain a dimensionless quantity, such as a GW amplitude. For bursts, we see from eq. (7.107) that this time-scale is the duration $\tau_g = 1/f_{\rm max}$ of the burst, while for a periodic signal we see from eq. (7.133) that it is the observation time $T$. Since $T$ can be of the order of months or years, while $\tau_g$ is typically between the millisecond and a second, the minimum value of $h$ detectable for periodic signals is much smaller than for bursts. On the other hand, a periodic signal is intrinsically much weaker, since a burst emits a huge amount of energy in a very short time. We will estimate in Section 7.6.3 the maximum distances at which typical periodic signals can be seen.

For bursts, we assumed that the wave came from the optimal direction, and for this reason we wrote no angular factor in eq. (7.107). For periodic signals, an average over the source position is in any case necessary because of the apparent motion of the source in the sky, leading to the amplitude modulation, and produces the angular efficiency factor $\langle F_+^2 \rangle$ in eq. (7.134).

An alternative reference quantity which is often used is $h_{3/{\rm yr}}$, which is defined as the minimum value of $h_0$ that can be detected at a given value of $S/N$, integrating for $T = 10^7$ s (i.e. about 1/3 of a year),

$$h_{3/{\rm yr}} = \frac{S/N}{\langle F_+^2 \rangle^{1/2}} \sqrt{S_n(f_0) \times 10^{-7}\,{\rm Hz}} \,. \tag{7.135}$$

[45] If, rather than being interested in the sensitivity to a specific source, one wants to define an average sensitivity for an ensemble of sources, then one can improve this estimate taking care of the fact that there is a statistical preference for the angles and polarizations that give a larger $S/N$, since these can be seen to larger distances. This modifies $S/N$ by factors that can be approximately estimated to be of order $(3/2)^{1/2} \simeq 1.2$, see Thorne (1987).

### 7.6.2 Doppler shift and phase modulation

Even if an astrophysical source emitted exactly monochromatic GWs with a frequency $f_0$, for a detector on Earth the instantaneous value of the observed frequency $f$ would change with time because of the Doppler effect. Recall that, to first order in $v/c$, the frequency measured by an observer with a velocity $\mathbf{v}$ with respect to the source is

$$f = f_0 \left(1 + \frac{\mathbf{v} \cdot \hat{\mathbf{r}}}{c}\right) , \tag{7.136}$$

where $\hat{\mathbf{r}}$ is the unit vector in the direction of the source. If $\mathbf{v} \cdot \hat{\mathbf{r}}$ were a constant, this would cause little concern, since it would just amount to a constant offset in the frequency and, with a single FFT, monochromatic lines at all possible frequencies are searched simultaneously. However, the velocity of the detector with respect to the source changes in time because of the Earth's rotation and because of its revolution around the Sun and this induces a time-dependence in the observed frequency. We denote by $(\Delta v)_T$ the change of the component of the velocity in the direction of the source, in a time $T$. Then the frequency $f$ changes on the same time interval by an amount

$$(\Delta f)_{\rm Doppler} = f_0 \, \frac{(\Delta v)_T}{c} \,. \tag{7.137}$$

When we integrate the signal for a time $T$, the resolution in frequency is $\Delta f = 1/T$. As long as $(\Delta f)_{\rm Doppler}$ is smaller than this resolution, all the GW signal falls into a single frequency bin and the Doppler effect can be neglected. To estimate the maximum integration time for which the Doppler effect is negligible, we consider first the effect of the Earth rotation around its axis. At a latitude of 40 degrees, the rotational velocity of the Earth is $v_{\rm rot} = \omega_{\rm rot} R_{\oplus} \cos(40^\circ) \simeq 355$ m/s, where $\omega_{\rm rot} = (2\pi/24 \text{ hr})$ and $R_{\oplus} \simeq 6.38 \times 10^6$ m is the mean Earth equatorial radius. This gives $v_{\rm rot}/c \simeq 1.2 \times 10^{-6}$. During an integration time $T$, the Earth rotates by an angle $\Delta\theta = \omega_{\rm rot} T$ and, if $\Delta\theta \ll 1$, in order of magnitude the change of the component of the velocity in the direction of the source is given by $(\Delta v)_T / v_{\rm rot} \sim \Delta\theta$, i.e.

$$(\Delta v)_T \sim v_{\rm rot} \omega_{\rm rot} T \,. \tag{7.138}$$

(The precise numbers, of course, depends on the exact direction of the source with respect to the detector.) Then $(\Delta f)_{\rm Doppler}$ becomes of the order of the frequency resolution if

$$f_0\left(\frac{v_{\rm rot}}{c}\right)\omega_{\rm rot}T \sim \frac{1}{T}\,, \qquad (7.139)$$

which gives

$$T \sim 60\ {\rm min}\left(\frac{1\,{\rm kHz}}{f_0}\right)^{1/2}. \qquad (7.140)$$

Therefore, for waves with $f_0 \sim 1$ kHz, the Doppler effect due to the Earth's rotation around its axis becomes important after about one hour.[46] It reaches its maximum value after about 12 hr (the precise numbers, again, depend on the source position), when the detector has inverted its velocity with respect to the source, $\Delta v_{\rm rot} = 2v_{\rm rot}$, and in this time span the frequency has changed by a total amount

$$(\Delta f)^{\rm rot}_{\rm max} \sim 2f_0\,\frac{v_{\rm rot}}{c} \simeq 2.4\times10^{-3}\,{\rm Hz}\left(\frac{f_0}{1\,{\rm kHz}}\right). \qquad (7.141)$$

We can repeat the same reasoning for the orbital motion of the Earth around the Sun. For an order-of-magnitude estimate we can take the orbit as circular, with a radius $R = 1\,{\rm au} \simeq 1.5\times10^{11}$ m and $\omega_{\rm orb} = 2\pi/(365\,{\rm days})$, so $v_{\rm orb} \simeq 3\times10^4$ m/s and $v_{\rm orb}/c \simeq 10^{-4}$. The maximum frequency shift induced by the Earth revolution is then

$$(\Delta f)^{\rm orb}_{\rm max} \sim 2f_0\,\frac{v_{\rm orb}}{c} \simeq 0.2\,{\rm Hz}\left(\frac{f_0}{1\,{\rm kHz}}\right), \qquad (7.142)$$

and is much larger than that due to the Earth rotation around its axis, given in eq. (7.141), because $v_{\rm orb} \gg v_{\rm rot}$. However, the large drift (7.142) takes place over a six months period. In an integration time $T$ much shorter than six months, the orbital motion induces a variation $(\Delta v)_T \sim v_{\rm orb}\omega_{\rm orb}T$ and the corresponding frequency shift is $(\Delta f)_{\rm Doppler} \sim f_0\,(v_{\rm orb}/c)\,\omega_{\rm orb}T$. Similarly to eq. (7.139), the time after which the orbital Doppler shift becomes larger than the frequency resolution is given by

$$f_0\left(\frac{v_{\rm orb}}{c}\right)\omega_{\rm orb}T \sim \frac{1}{T}\,, \qquad (7.143)$$

i.e.

$$T \sim 120\ {\rm min}\left(\frac{1\,{\rm kHz}}{f_0}\right)^{1/2}. \qquad (7.144)$$

Therefore the Doppler shift due to the Earth rotation around its axis is the first to become important, when we increase the integration time (after about 1 hr if, for instance, $f_0 = 1$ kHz). The orbital Doppler shift becomes of the order of the frequency resolution shortly afterwards, after an integration times of about 2 hr for $f_0 = 1$ kHz, but then raises steadily; after less than one day it becomes more important than the contribution from the Earth's rotation around its axis, and it continues to raise for a six months period becoming, on the long term, the largely dominant effect.

[46] For frequencies $f_0 > O(40)$ Hz we have $T \ll 1$ day, so the approximation $\Delta\theta \ll 1$ used to write eq. (7.138) is consistent. Otherwise, a more accurate estimate is needed.



After an integration time of four months, i.e. $T \simeq 10^7$ s, the frequency resolution is $\Delta f = 10^{-7}$ Hz, which is many order of magnitudes smaller than the Doppler shifts (7.141) and (7.142). It is interesting to see what is the form of the frequency spectrum when we are sensitive enough to resolve the time-changing Doppler shift. To simplify the geometry, we assume at first that the detector performs a simple circular motion, with frequency $\omega_m$ and radius $R$, and that the source is in the plane of the orbital motion of the detector, as in Fig. 7.10. Since the source is at a very large distance, we have a plane wavefront propagating along the $y$ axis, and therefore proportional to $\cos[\omega_0(t+y/c)]$, where $\omega_0 = 2\pi f_0$ and $f_0$ is the GW frequency. The $y$ coordinate of the detector is a function of time; we choose for definiteness the origin of time so that $y(0) = 0$, and therefore $y(t) = R\sin(\omega_m t)$. Then the detector sees a signal proportional to

$$\cos\left[\omega_0\left(t+\frac{y(t)}{c}\right)\right] = \cos[\omega_0 t + \beta\sin(\omega_m t)]\,, \qquad (7.145)$$

where

$$\beta = \frac{\omega_0 R}{c} = \frac{\omega_0}{\omega_m}\,\frac{v}{c}\,, \qquad (7.146)$$

with $v = \omega_m R$. The parameter $\beta$ is called the modulation index, and $\omega_m = 2\pi f_m$, where $f_m$ is the modulation frequency. This signal can be written as a superposition of monochromatic waves using the identity

$$\cos[\omega_0 t + \beta\sin(\omega_m t)] = \sum_{k=-\infty}^{\infty} J_k(\beta)\cos[(\omega_0 + k\omega_m)t]\,, \qquad (7.147)$$

where $J_k(\beta)$ is the Bessel function.[47] The signal is therefore split into a carrier at the frequency $f_0$, plus an infinite number of sidebands at $f_0 \pm kf_m$, for all integer $k$, and the power in the $k$-th sideband is proportional to $J_k^2(\beta)$. The qualitative form of this spectrum depends strongly on the modulation index $\beta$. For $\beta \to 0$ and $k$ integer we have $J_k(\beta) \sim \beta^{|k|}$, so when $\beta \ll 1$ most of the power is in the carrier ($k = 0$), with smaller power in the sidebands $k = \pm1$, even smaller power at $k = \pm2$, etc. However, in our case $\beta$ is given by eq. (7.146) and it is large. In fact, for the rotation of the Earth around its axis, setting $\omega_m = 2\pi/(24\,{\rm hr})$ and $v/c \simeq 1.2\times10^{-6}$, eq. (7.146) gives $\beta \simeq 100\,(f_0/1\,{\rm kHz})$, while for the orbital motion $\beta \simeq 3\times10^6\,(f_0/1\,{\rm kHz})$. Therefore, in the range of frequencies relevant for ground-based interferometers ($f_0 > O(10)$ Hz), we are always in the regime $\beta \gg 1$.

The average number of sidebands into which the total power is distributed can be calculated using[48]

$$\langle k^2\rangle \equiv \frac{\sum_{k=-\infty}^{\infty} k^2 J_k^2(\beta)}{\sum_{k=-\infty}^{\infty} J_k^2(\beta)} = \frac{\beta^2}{2}\,, \qquad (7.148)$$

so the power is distributed in $O(\beta)$ sidebands, as shown in Fig. 7.11. Once the frequency resolution $1/T$ has become of the order of this

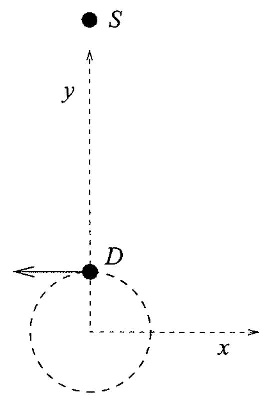


**Fig. 7.10** The simplified geometry discussed in the text. The detector $D$ performs a circular motion in the $(x,y)$ plane. The source $S$ is in the same plane, along the $y$ axis.

[47] This identity can be obtained writing $\cos[(\omega_0 + k\omega_m)t$, inside the sum, as $\cos(\omega_0 t)\cos(k\omega_m t) - \sin(\omega_0 t)\sin(k\omega_m t)$, and using Gradshteyn and Ryzhik (1980), 8.514.5 and 8.514.6, recalling that, for $k$ integer, $J_{-k}(z) = (-1)^k J_k(z)$.

[48] See Gradshteyn and Ryzhik (1980), 8.536.2.

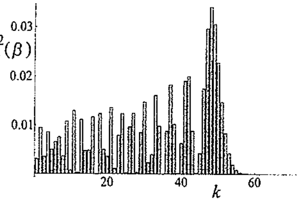


**Fig. 7.11** The quantity $J_k^2(\beta)$ for $\beta = 50$, as a function of $k$.

Doppler line broadening, increasing $T$ further does not improve substantially the signal-to-noise ratio. In fact, while a smaller frequency bin contains by definition less noise, it now also contains less signal, since the signal gets spread over many bins. However, if at this stage the signal already emerged from the noise, increasing $T$ further we improve the resolution by which we are able to reconstruct the line shape (and therefore the accuracy by which we can reconstruct the source position, see Note 49 below).

Luckily, there is in principle a very simple way (borrowed from pulsar radio-astronomy) to correct for the Doppler shift. In the simple geometrical situation illustrated in Fig. 7.10, we just need to define a new time variable $t' = t + y(t)/c$. In terms of this variable the signal (7.145) is simply proportional to $\cos(\omega_0 t')$ and, performing the Fourier transform with respect to $t'$, all sidebands collapse into a single frequency. For a generic source location, the redefinition of time that does the job is

$$t' = t + \frac{\mathbf{x}(t)\cdot\hat{\mathbf{r}}}{c}, \tag{7.149}$$

where $\mathbf{x}(t)$ is the position of the detector (measured for instance using the Solar System Barycenter (SSB) as a reference frame) and $\hat{\mathbf{r}}$ is the unit vector pointing toward the source. Observe that this is just the Roemer time delay that we already discussed in Section 6.2.2. We can therefore simply resample the output of the detector in terms of this new time, and we have corrected for the Doppler effect. The procedure has an added bonus: it is quite likely that, in the Fourier spectrum of the output, there will be monochromatic lines due to instrumental noise. If such a line has a frequency that is constant in time to a good accuracy, its signal-to-noise ratio will increase as $\sqrt{T}$, just as for a GW signal. However, when we apply the resampling procedure, a real GW signal, which was spread over many bins, is collapsed to a single frequency bin, an conversely an instrumental line which was monochromatic will be spread over many bins, and will finally be diluted into the noise. In other words, we are using the Doppler modulation as a powerful signature that discriminates a real GW signal from instrumental noise.

The simplicity of this solution comes however at a price: we need to know both $\mathbf{x}(t)$ and $\hat{\mathbf{r}}$ with great precision. We can assume that the motion of the Earth is known to a sufficient accuracy (although, if we want to integrate for a time $T \sim 1$ yr, we need to keep under control effects that can produce shifts $\Delta f \sim 10^{-7}$ Hz, and for this we must also include small effects like the oscillations of the Earth around the Earth–Moon barycenter, which however are precisely known), so the main error comes from the uncertainty on the angular position of the source. From eq. (7.136) we see that, in order to correct for the Doppler shift with an accuracy smaller than the experimental resolution $1/T$ on $f$, we need, in order of magnitude,

$$\frac{f_0}{c}\,(\Delta v)_T\,\Delta\theta < \frac{1}{T}, \tag{7.150}$$

where $(\Delta v)_T$ is the variation of the velocity of the Earth during the observation time $T$ (recall that only the temporal variation is relevant, otherwise the Doppler effect would give just a constant offset in $f_0$), while $\Delta\theta$ is the angular resolution (in radians) on the position of the source.

If we take $(\Delta v)_T \sim v_{\rm orb}\omega_{\rm orb}T$ we find that, to apply the Doppler correction, we need to know the source location to an accuracy

$$\Delta\theta < \frac{1}{f_0 (v_{\rm orb}/c)\,\omega_{\rm orb}T^2} \simeq 0.1\,\text{arcsec}\left(\frac{10^7\,\text{s}}{T}\right)^2\left(\frac{1\,\text{kHz}}{f_0}\right). \tag{7.151}$$

This expression is correct as long as the Doppler effect due to the orbital motion dominates that due to the Earth rotation around its axis, i.e. for $T \gtrsim 1$ day, and also as long as the angle $\omega_{\rm orb}T$ is small, since otherwise the approximation $(\Delta v)_T \sim v_{\rm orb}\omega_{\rm orb}T$ should rather be replaced by $(\Delta v)_T \sim v_{\rm orb}\sin\omega_{\rm orb}T$, so approximately eq. (7.151) is valid as long as 1 day $\lesssim T \lesssim$ 4 months.[49]

If we are targeting a specific source whose position is known to this accuracy or better, as is the case for many pulsars, this requirement does not pose special problems. However, as we will discuss in Vol. 2, there are many mechanisms that can produce periodic GWs, in particular in neutron stars, that are not necessarily associated to a strong electromagnetic emission or, as with pulsars, the electromagnetic emission could be beamed in a direction that does not intersect the Earth. It is in fact quite likely that most of the potentially interesting sources of periodic GWs have no detected optical counterpart. For example, the closest *observed* neutron star is at a distance $r \sim 100$ pc; however, population synthesis calculations indicate that the closest one should be at a distance $r \sim$ 5–10 pc, and then in a sphere of radius $r \sim 100$ pc there should be $O(10^3\text{–}10^4)$ neutron stars. It is therefore of the greatest interest to perform blind searches, i.e. searches for unknown sources over the whole sky. In principle, this means that we should partition the celestial sphere in pixels with a size given by eq. (7.151) (in fact even smaller, see Note 52 in the next section), and in each one we should apply a separate Doppler correction. As we will see in the next section, for integration times of months this is impossible, even with the maximum present or foreseeable computer power.

Furthermore, we have assumed until now that the intrinsic GW frequency $f_0$ of the source is stable, within the experimental resolution $\Delta f = 1/T$, and that the only modulation comes from the relative motion of the detector. This means that we are assuming a stability of the source frequency at the level

$$\frac{\Delta f_0}{f_0} = \frac{1}{Tf_0} \simeq 1\times 10^{-10}\left(\frac{10^7\,\text{s}}{T}\right)\left(\frac{1\,\text{kHz}}{f_0}\right). \tag{7.152}$$

Quite remarkably, rotating neutron stars can sometime have this sta-

[49] Actually, one could turn the argument around and observe that, if we are so lucky that there is a periodic signal so strong that can be extracted from the noise without correcting for the Doppler shift then, following the evolution of the frequency with time, we will be able to reconstruct the position of the source to the accuracy $\Delta\theta$ given by eq. (7.151). With present detector sensitivities, however, this possibility seems quite unlikely.

bility. The main mechanism that produces a drift in their frequency is the fact that a rotating NS radiates, both electromagnetically and gravitationally, and therefore loses energy. This energy is taken from the rotational energy of the NS, which therefore spins down. Pulsars are characterized by their spindown age $\tau$,

$$\tau = \frac{f}{|\dot{f}|}, \qquad (7.153)$$

where $f$ is their rotational frequency. As we saw in Section 4.2.1, for rotation around a principal axis and in the quadrupole approximation, the GWs emitted are monochromatic with a frequency $f_0 = 2f$, so $\tau = f_0/|\dot{f}_0|$.[50] During the observation time $T$, a pulsar with spindown age $\tau$ changes its GW frequency by an amount $\Delta f_0 = \dot{f}_0 T = -f_0 T/\tau$, i.e. by

$$\frac{\Delta f_0}{f_0} = -\frac{T}{\tau} \simeq -3.2\times 10^{-10}\left(\frac{10^9\,\text{yr}}{\tau}\right)\left(\frac{T}{10^7\,\text{s}}\right). \qquad (7.157)$$

Comparing with eq. (7.152) we see that, with an integration time $T = 10^7$ s, for a millisecond pulsar with $f_0 \sim 1$ kHz, the effect of the spindown is important if its spindown age is lower than $3\times 10^9$ yr while, if $f_0 = 10$ Hz, spindown is important, again over $T = 10^7$ s, if $\tau < 3\times 10^7$ yr. Therefore for many pulsars, and in particular for young pulsars, over such a long observation time the spindown must be taken into account. Actually in young pulsars the spindown rate can be so high that even the effect of the second derivative $\ddot{f}_0$ can become important.

For known pulsars the spindown can be measured and taken into account when we make the Doppler correction, while for blind searches it introduces new unknown parameters. Besides spindown, there are other reasons why the frequency of the GW emitted by a pulsar can change:

- Pulsars exhibit *glitches*, i.e. sudden jumps in the frequency related to rearrangements of their internal structure. These glitches can produce changes in the frequency as large as $\Delta f_0/f_0 \sim 10^{-6}$ and occur erratically, at a rate which depends strongly on the specific pulsar, but in general of the order one glitch every few years.
- A large fraction of the known millisecond pulsars are in binary systems. In this case, there will be an additional Doppler effect due to the motion of the source, as we saw in Section 6.2.
- Pulsar are the remnant of supernova explosions, and at birth they can receive a large kick; so their velocities can be larger than the typical velocities of the stars in their galactic neighborhood, and the pulsar proper motion can be important. Of course, if the motion is uniform, this only produces a constant shift in the frequency. However, accelerations due to gravitational fields can be important. In particular, many pulsars are found in globular clusters. In this case, the acceleration due to the Newtonian gravitational forces from all the other stars is known to produce frequency drifts comparable to the spindown rate.
- Even a uniform proper motion can be important if, during the observation time, it drives the NS out of the pixel in the sky where it was initially. For instance, a pulsar at a distance $r = 300$ pc, with a transverse velocity $v = 10^3$ km/s with respect to our line-of-sight, in a time $T = 10^7$ s moves by $\Delta\theta = vT/r \simeq 10^{-6}$ rad $\simeq 0.2$ arcsec which, according to eq. (7.151), is of order of the accuracy $\Delta\theta$ that we need, over such an integration time $T$, for a pulsar radiating GWs at $f_0 \sim 1$ kHz.

In the next section we will discuss how one can try to cope with these difficulties.

[50] The spindown age is of the order of the age of the pulsar if, throughout its lifetime, the pulsar frequency evolution can be described by the equation

$$\dot{f} = -af^n \qquad (7.154)$$

(where $a$ is a constant) and if the *braking index* $n > 1$. In fact, integrating the above equation we get

$$[f(t)]^{-(n-1)} - [f(0)]^{-(n-1)} = a\,(n-1)t, \qquad (7.155)$$

where $t = 0$ is the time at which the pulsar was born. If the frequency at birth, $f(0)$, was much bigger than the frequency today, and if $n > 1$, we can neglect the term $[f(0)]^{-(n-1)}$ and the age of a pulsar is related to its present values of $f$ and $\dot{f}$ by

$$t = \frac{1}{a(n-1)f^{n-1}} = \frac{f}{(n-1)|\dot{f}|} = \frac{\tau}{n-1}. \qquad (7.156)$$

Experimentally, the braking index $n$ typically has values $n \simeq 2$–$3$, depending on the specific pulsar.

### 7.6.3 Efficient search algorithms

#### Coherent searches

From the discussion of the previous section we know that, if we want to integrate the signal for a long time, we must resample the output of the detector in terms of the time $t'$ defined in eq. (7.149), plus further correction for the spindown or other effects that change the frequency. The GWs produced by a rotating NS, in the absence of spindown, has been computed in eq. (4.223). Including the Doppler effect of the detector and the spindown of the source we can write the signal received as

$$h(t) = F_+(\hat{\mathbf{n}}(t);\psi)h_0\,\frac{1+\cos^2\iota}{2}\,\cos\Phi(t) + F_\times(\hat{\mathbf{n}}(t);\psi)h_0\cos\iota\,\sin\Phi(t), \qquad (7.158)$$

where $h_0$ is given in eq. (4.224), and $\iota$ is the angle between the spin axis of the neutron star and the propagation direction $\hat{\mathbf{n}}$ of the GW; of course $\hat{\mathbf{n}} = -\hat{\mathbf{r}}$, where $\hat{\mathbf{r}}$ is the unit vector pointing toward the source, and depends on time because of the relative motion of the detector and source. The evolution of the accumulated phase $\Phi(t) = 2\pi\int dt f(t)$ observed by the detector can be described by a Taylor expansion, writing

$$f(t') = f_0 + \dot{f}_0(t'-t'_0) + \frac{1}{2}\ddot{f}_0(t'-t'_0)^2 + \ldots, \qquad (7.159)$$

where $t'$ is the resampled time given in eq. (7.149), i.e. the time of arrival of the signal in the Solar System Barycenter (SSB),[51] and $t'_0$ is a fiducial value, such that $\Phi(t'_0)$ has the value $\phi_0$. Then

$$\Phi(t) = \phi_0 + 2\pi\left[f_0(t'-t'_0) + \frac{1}{2}\dot{f}_0(t'-t'_0)^2 + \frac{1}{6}\ddot{f}_0(t'-t'_0)^3 + \ldots\right]. \qquad (7.160)$$

Of course, a truncated Taylor expansion is useful only if the higher order terms are small corrections during the whole observation time $T$. This is not the case for a neutron star in a binary system, which rather performs a circular motion around the center-of-mass of the system, so eq. (7.160) only applies to isolated neutron stars.

[51] Actually, the precise redefinition is

$$t' = t + \frac{\mathbf{x}(t)\cdot\hat{\mathbf{r}}}{c} + \Delta_{E\odot} - \Delta_{S\odot},$$

where $\Delta_{E\odot}$ and $\Delta_{S\odot}$ are the solar system Einstein and Shapiro time delays discussed in Section 6.2.2. However, given the detector and the source positions, the Einstein and Shapiro delays can be computed, as we did explicitly in Section 6.2.2, and introduce no new free parameter.

If our target is a given pulsar whose position, proper motion and spindown parameters are known to sufficient accuracy, the form of the signal (7.158) is fixed. Then we can simply demodulate the signal defining a new variable $t''$ as

$$t'' = (t' - t'_0) + \frac{\dot{f}_0}{2f_0}(t' - t'_0)^2 + \frac{\ddot{f}_0}{6f_0}(t' - t'_0)^3 + \dots, \qquad (7.161)$$

so that eq. (7.160) reads $\Phi = \phi_0 + 2\pi f_0 t''$. We resample the detector output with respect to this variable, and then all we need to do is to perform a single Fast Fourier Transform (FFT) on this resampled stretch of data, of length $T$. The number of spindown parameters $\dot{f}_0, \ddot{f}_0, \dots$ to be included to have sufficient accuracy depend on the source, and on the observation time $T$.

If however we want to perform a blind all-sky search, the problem becomes quickly intractable with increasing observation time $T$. In fact, our parameter space is given by the angles $(\theta_s, \phi_s)$ of the source and by the spindown parameters $\dot{f}_0/f_0, \ddot{f}_0/f_0$, etc. Observe that $f_0$ itself does not contribute to the dimension of the parameter space; the resampling of time (7.149) is independent of $f_0$, while eq. (7.161) depends only on the ratios $\dot{f}_0/f_0, \ddot{f}_0/f_0, \dots$, and not separately on $f_0, \dot{f}_0, \ddot{f}_0, \dots$. This is a crucial advantage of the resampling technique. If, rather than resampling the detector output, we directly used the Wiener filtering for the waveform given in eqs. (7.158) and (7.160), then $f_0$ would be an additional parameter to be searched, and the computational cost would increase dramatically.

Then, what we should do is to discretize this parameter space, and for each point of this parameter space we should perform the appropriate demodulation (7.161) and one FFT. This procedure is referred to as a *coherent* search. Its drawback is that, if we want to take advantage of the large integration time, the mesh in the discretized parameter space must become finer and finer when we increase $T$. For instance, even in the simplest case in which the spindown parameters are negligible, and therefore the parameter space is given only by the angles $(\theta_s, \phi_s)$, still the number of patches in the sky that we must consider is at least $N_{\rm patches} = 4\pi/(\Delta\theta)^2$ and scales at least as $T^4$, see eq. (7.151).[52]

More generally, the number of mesh points depends on the kind of search that we perform. For instance, old pulsars are less demanding than young pulsars of the same frequency, since their spindown rate is lower and therefore it can be taken into account using a larger mesh in the spindown parameter space. Similarly, we see from eq. (7.151) that slow pulsars (say, $f_0 < 200$ Hz) are easier to analyze that fast pulsars with $f_0 \sim 1$ kHz.

Since the time needed for data analysis grows with a large power of $T$, increasing $T$ we necessarily reach a point where the data analysis would take the same time as the observation time $T$ and beyond that point it would quickly becomes many orders of magnitude larger than the observation time. We can therefore take as a limit the condition that the time required by data analysis does not exceed the observation time used to take the same data. To have an idea of the computational requirements consider that, using $10^7$ s of data to search for periodic GWs with frequencies up to 500 Hz, requires the calculation of a FFT with $10^{10}$ points, which takes about 1 s on a teraflop computer (assuming that all $10^{10}$ points can be held simultaneously in fast memory), and we need one such FFT for each point of the parameter space. It can be estimated[53] that a coherent all-sky search of $T = 10^7$ s of data for slow, old pulsars ($\tau > 1000$ yr, $f_0 < 200$ Hz, i.e. the "easier" target) requires only one spindown parameter and $10^{10}$ independent points in parameter space, while for young, fast pulsar (frequencies up to $f_0 \simeq 1$ kHz, $\tau$ as low as 40 yr) three spindown parameters and $8 \times 10^{21}$ points in parameter space are required. Then, even in the "easy" case, the analysis of four months of data would require three centuries on a teraflop computer! Requiring that the data analysis does not last more than data taking, one finds that for slow, old pulsars the data stretch cannot be longer than $\sim 18$ days, while for young, fast pulsar the limit is less than one day. The disappointing conclusion is that, even if a detector can in principle take good data for months or years, a coherent blind all-sky search for pulsar using fully these data is impossible.

[52] A more careful argument shows that it even scales like $T^5$, because the approximation $(\Delta v)_T \sim v_{\rm orb}\omega_{\rm orb}T$ used to derive eq. (7.151) does not hold simultaneously for the right ascension and for the declination angles, see Brady, Creighton, Cutler and Schutz (1998).

[53] See Brady, Creighton, Cutler and Schutz (1998).

### Incoherent searches

A solution to the computational problem discussed above is to split the total observation time $T$ into $\mathcal{N}$ stacks of length $T_{\rm stack}$, with $T = \mathcal{N}T_{\rm stack}$. We choose $T_{\rm stack}$ so that a coherent search over such a time is computationally feasible. So the output of each coherent search over one stack is a collection of function $\tilde{h}(f)$, one for each value of the parameter space. For each point in parameter space we then add the quantity $|\tilde{h}(f)|^2$ over the $\mathcal{N}$ stacks. Since in this way the phase information between the different stacks gets lost, this is called an *incoherent search.* If we denote by $\tau_{\rm stack}$ the time needed to perform a coherent search on a stack of data of length $T_{\rm stack}$, the time needed for the full incoherent search is $\tau_{\rm inc} = \mathcal{N}\tau_{\rm stack}$, while the time needed for a full coherent search over the whole time $T$ is $\tau_{\rm coh} \simeq (T/T_{\rm stack})^n \tau_{\rm stack} = \mathcal{N}^n \tau_{\rm stack}$, so

$$\tau_{\rm inc} \simeq \frac{1}{\mathcal{N}^{n-1}}\, \tau_{\rm coh}\,, \qquad (7.162)$$

where the power $n$, as discussed above, is determined by the kind of pulsars that we are targeting. Since $n$ is large (at least $n = 5$, even when no spindown parameters are needed, see Note 52), it is clear that incoherent searches have a huge advantage in terms of computational cost and, for a given observation time $T$, taking $\mathcal{N}$ sufficiently large, i.e. $T_{\rm stack}$ sufficiently small, the computation becomes feasible.

From the point of view of sensitivity, the value of $(S/N)^2$ obtained from a single stack of length $T_{\rm stack}$ is given by eq. (7.130) replacing $T$ by $T_{\rm stack}$. Adding $\mathcal{N}$ of these spectra the variance is reduced by $1/\sqrt{\mathcal{N}}$ and therefore, for an incoherent search, eq. (7.130) becomes

$$\left(\frac{S}{N}\right)^2 = \langle F_+^2\rangle \left(\frac{\mathcal{N}^{1/2}T_{\rm stack}}{S_n(f_0)}\right) h_0^2$$

$$= \frac{1}{\sqrt{\mathcal{N}}} \langle F_+^2 \rangle \left( \frac{T}{S_n(f_0)} \right) h_0^2 \tag{7.163}$$

and the minimum amplitude detectable at a given $S/N$, eq. (7.134), becomes

$$(h_0)_{\min} = \eta \left( \frac{S_n(f_0)}{T} \right)^{1/2} . \tag{7.164}$$

where we have defined an efficiency factor

$$\eta = (S/N) \frac{\mathcal{N}^{1/4}}{\langle F_+^2 \rangle^{1/2}} , \tag{7.165}$$

which takes into account the desired level of the signal-to-noise ratio $S/N$, the average over the orbit of the source, which produces the factor $\langle F_+^2 \rangle^{1/2}$, and the need to separate the data into $\mathcal{N}$ stack for computational feasibility.

In practice, beside being forced by computational requirements, incoherent searches are also necessary because a detector never has months, and not even weeks, of continuous good data taking. There are always interruption due to maintenance, period of higher noise level that must be removed, etc. and the experimental precision that one has on the time at which data taking resumed is not sufficiently good to recombine coherently different stacks of data. The incoherent method, of course, can be applied even when the single stacks have not all the same duration, and when they are not consecutive.

When performing an incoherent search each stack is demodulated, as discussed in the previous section, using a mesh of points sufficient to confine the searched signal into a single bin. The individual power spectra, before being summed, must be corrected for their relative frequency drift using a finer parameter mesh suitable for removing the phase modulation over the whole observation period. The simplest implementation of this method consists in choosing stacks of about 30 min, so that the Doppler effect in each stack can simply be neglected, and within a single stack no demodulation is needed. In this case a period of $10^7$ sec of data is divided into $\mathcal{N} \simeq 5000$ stacks, and the minimum detectable value $(h_0)_{\min}$ in eq. (7.164) is larger than in a coherent search by a factor $\mathcal{N}^{1/4} \sim 8$. With the difference, of course, that a blind full-sky incoherent search of this type is computationally feasible while a blind full-sky coherent search is not.

Alternatively, one can choose longer stacks, say of the order of one day. These will need demodulation, but a relatively coarse mesh in parameter space will suffice to concentrate the whole signal into a single bin. Then we combine the separate stacks using a finer mesh. Of course, the longer the stack, the higher is the sensitivity, but the higher is also the computational cost. Incoherent searches can also be used as a first stage in a hierarchical search: an incoherent blind all-sky search can produce a number of interesting candidate signals, for certain values of the parameters. These points in parameter space can then be examined more thoroughly with a directed coherent search.

We can now compare the experimental sensitivity given by eq. (7.164) with the signal expected from a rotating NS, given in eq. (4.224). We then find that the maximum distance $r$ which a detector can reach in a blind full-sky search for periodic GWs from rotating neutron stars is

$$r = 5.8\,\mathrm{kpc} \left( \frac{10^{-23}\,\mathrm{Hz}^{-1/2}}{S_n^{1/2}(f_0)} \right) \left( \frac{T}{3\times 10^7\,\mathrm{s}} \right)^{1/2} \times \left( \frac{100}{\eta} \right) \left( \frac{\epsilon}{10^{-6}} \right) \left( \frac{I_{zz}}{10^{38}\,\mathrm{kg\,m^2}} \right) \left( \frac{f_0}{1\,\mathrm{kHz}} \right)^2 . \tag{7.166}$$

The reference value $\eta = 100$ corresponds to a search for a total time $T = 3\times 10^7$ s divided into stacks with $T_{\rm stack} \simeq 30$ min (so $\mathcal{N} \simeq 1.7\times 10^4$), a factor $1/\langle F_+^2 \rangle^{1/2} = \sqrt{5}$ as appropriate for interferometers, see Table 7.1, and a value $S/N \simeq 4$. The strain sensitivity $S_n^{1/2}$ has been normalized to the value expected for an advanced interferometer.

### The Hough transform

As we have seen above, in incoherent searches the observation time is divided into stacks, where the phase modulation due to Doppler effect and spindown is either negligible (if $T_{\rm stacks} \lesssim 30$ min) or anyway relatively easy to correct for, so that a GW signal, if present, falls into a single frequency bin. When we compare different stacks, the position in frequency of the bin that contains the signal changes, because of the Doppler effect and of the spindown. For each point in the parameter space $(\theta_s, \phi_s, \dot{f}_0/f_0, \ddot{f}_0/f_0, \ldots)$ we can compute how the position of the bin should change and we can correct for it, using the resampling technique discussed in the previous section. In this way, for each point of parameter spaces, the bins are "realigned", and the power in corresponding bins is summed.

An interesting variation on this scheme is given by the Hough transform, which is a technique used for pattern recognition in digital images.[54] In the Hough transform, as a first step, rather than summing up the power in the corresponding bins, we fix a threshold in each data stack. A bin is deemed "black" if the power in it exceeds the threshold, and "white" if it does not. In the time–frequency plane obtained aligning in frequency (with no correction) the various stacks, we therefore have a set of black pixels, as in Fig. 7.12.

In the case of Gaussian noise, where large fluctuations are unlikely, it would in principle be more convenient to sum up the power of the corresponding bins, rather than reducing all the information to a set of zeros (white) and ones (black). However, the Hough transform is more robust in the presence of non-Gaussian noise and large occasional external dis-

[54] It was developed in 1959 by Paul Hough at CERN, to analyze the tracks of particles in bubble chambers, and today is also used in astronomical data analysis.

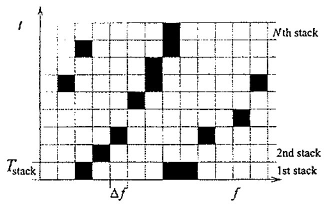


**Fig. 7.12** The time–frequency plane, with bins of length $\Delta t = T_{\rm stack}$ in time and $\Delta f = 1/T_{\rm stack}$ in frequency. Bins where the power exceeds a given threshold are marked in black.

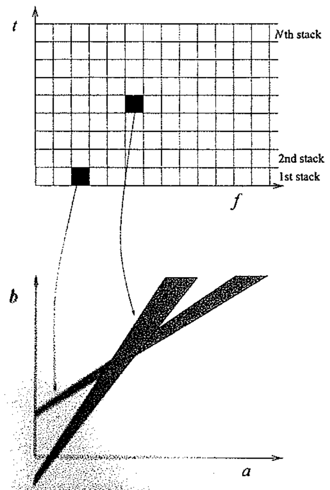


**Fig. 7.13** The map that to each black pixel associates a submanifold in the parameter space Σ.

turbances, which is the case in a real detector. Consider for instance the situation in which instrumental noise gives a very large spike in frequency, during a relatively short period, e.g. in only one stack. When summing the power, this single disturbance can give a large effect on the total sum, while collapsing all the information to black/white it contributes only to a single pixel. This method can therefore be appropriate when we search for a signal that is small, but is there during the whole observation time, and is embedded in a noise that occasionally can be much larger than the signal.

The next step is to perform a pattern recognition procedure in the set of black pixels, to see if some of the black pixels lie along a specified curve. To simplify the setting, suppose that we are searching for straight lines in the $(t, f)$ plane of Fig. 7.12. (The generalization to more complex curves is conceptually straightforward.) The set of all straight lines in this plane is parametrized by two parameters $(a, b)$, as $t = af + b$. We denote by Σ the manifold described by the parameters $(a, b)$; in our example of straight lines of course $\Sigma = \mathbb{R}^2$, but the notation is more general. Given a black pixel, we can find the set of points in the manifold Σ that are compatible with it; for instance, in our straight lines example, if a black pixel is centered at $(t_1, f_1)$, the straight lines consistent with it are those that satisfy $t_1 = af_1 + b$, and the corresponding submanifold of Σ is the curve $b = af_1 - t_1$ in the $(a, b)$ plane. More precisely, since the pixels in the $(t, f)$ plane have a finite resolution, we will rather get a bunch of straight lines in Σ. The transformation that, to each black pixel in the $(t, f)$ plane associates a submanifold in Σ, is illustrated graphically in Fig. 7.13.

In the absence of noise, the submanifolds in Σ obtained in this way from all the black pixels would have a non-empty intersection, which would define the point in parameter space compatible with the observations. Of course, in the presence of noise the intersection of all the curves will be empty. Still, we can try to recover the most probable value of the parameters in Σ as follows. First, we discretize the manifold Σ. Let us call $C_1$ the surface in Σ obtained from the first black pixel. We then assign +1 to all the bins in Σ that belong to $C_1$. We repeat the same for the second black pixel, adding +1 to the the bins in Σ that belong to $C_2$, and so on for all the $N$ black pixels. In conclusion, we have constructed a map that, to the set of black pixels, associates a histogram in the parameter space Σ.

In the GW detection problem, the manifold Σ becomes the parameter space $(\theta_s, \phi_s, \dot{f}_0/f_0, \ddot{f}_0/f_0, \ldots)$ and the straight lines of our example are replaced by the curves in the $(t, f)$ plane that describe how $f$ changes with time because of the Doppler effect and of the spindown. The points in parameter space whose number count is above a certain threshold are the candidates for a possible detection and can be further investigated, for instance with a coherent search.



## 7.7 Coalescence of compact binaries

The coalescence of compact binaries, such as BH-BH and NS-NS binaries, is a particularly interesting signal for broad-band GW detectors. This comes from a combination of two facts: first, we saw in eq. (4.44) that, in the last stages of the inspiral, a binary system can radiate away in GWs up to a few per cent of its total mass. This is a huge amount of energy, so the signal from an inspiral is quite strong, compared to most other GW sources. Second, the inspiral phase can be tracked for many cycles in a broad-band detector. We saw in eq. (4.23) that a ground-based interferometer can follow the inspiral phase of a compact binary system for $O(10^4)$ cycles. Thus, matched filtering can be very effective for extracting this signal from the noise. From eq. (7.40) and the discussion below it we see that, in order of magnitude, with matched filtering we can dig into the noise and catch the signal from a coalescence, even when the typical amplitude of the GW signal inside the interferometer bandwidth is smaller than the noise floor by a factor $\mathcal{N}_c^{1/2}$, where $\mathcal{N}_c$ is the number of cycles for which we are able to track carefully the signal with our template. Thus, we can gain a factor as large as $\mathcal{N}_c^{1/2} \sim 100$ in amplitude, if our template is so good that we can follow closely the signal from the time it enters in the interferometer bandwidth until the inspiral phase terminates and the two objects merge. Since the GW amplitude is proportional to $1/r$, a factor $O(100)$ in amplitude means that we gain a factor $O(100)$ in the maximum distance to which we can detect a source. For these reasons, we will see that interferometers have the potential of detecting coalescing binaries up to distances of order of hundreds of Mpc, and advanced ground-based interferometers could reach a few Gpc.

To exploit this opportunity we must however be able to follow closely the signal with a template. This means, first of all, that for a given value of the parameters of the binary system (time of coalescence, masses, spins, etc.), one must know the waveform accurately. We already quantified this requirement in Section 5.6, where we found that we need to compute the post-Newtonian corrections up to 3.5PN order. As we saw in Section 5.6, these remarkable computations have indeed been performed. The second aspect is that we do not know in advance the parameters of the system, and therefore we must scan a potentially large parameter space.

To leading Newtonian order we computed the waveform in eq. (4.29), and the corrections in the restricted post-Newtonian approximation were discussed in Section 5.6. Combining these results with the general expression $h(t) = F_+ h_+(t) + F_\times h_\times$, we see that the output $h(t)$ for a binary inspiral, in the restricted post-Newtonian approximation, is

$$h(t) = A_+ \left[\frac{\pi f_{\rm gw}(t)}{c}\right]^{2/3} \cos[\Phi(f_{\rm gw}(t)) + \Phi_0]$$
$$+ A_\times \left[\frac{\pi f_{\rm gw}(t)}{c}\right]^{2/3} \sin[\Phi(f_{\rm gw}(t)) + \Phi_0]\,, \qquad (7.167)$$

where, as discussed in Section 5.6.3, $\Phi(f_{gw})$ and $f_{gw}(t)$ are known up to 3.5PN order.[55] We have esplicitly displayed the arbitrary constant $\Phi_0$ in the phase, equivalent to the arbitrary constant $\omega_0$ in eq. (5.265), and we have defined

$$A_+ = \frac{4}{r}\left(\frac{GM_c}{c^2}\right)^{5/3} F_+(\theta,\phi)\frac{1+\cos^2\iota}{2} \tag{7.168}$$

$$A_\times = \frac{4}{r}\left(\frac{GM_c}{c^2}\right)^{5/3} F_\times(\theta,\phi)\cos\iota\,. \tag{7.169}$$

Writing $A_+ = A\cos\alpha$ and $A_\times = A\sin\alpha$, with $A = (A_+^2 + A_\times^2)^{1/2}$ and $\tan\alpha = A_\times/A_+$, we can rewrite this as

$$h(t) = A\left[\frac{\pi f_{gw}(t)}{c}\right]^{2/3} \cos[\Phi(f_{gw}(t)) + \varphi]\,, \tag{7.170}$$

with $\varphi = \Phi_0 - \alpha$.[56] Thus, in the waveform enter the distance $r$ to the source, its location, specified by the angles $(\theta,\phi)$ which appear in the pattern functions, the orientation of the orbit with respect to the line of sight (two angle, one of which is $\iota$, and the other identifies the axes with respect to which the plus and cross polarizations are defined), the reference time $t_*$ at which the signal enters in the detector bandwidth (which appears through $\Phi(t)$ and $f_{gw}(t)$), the constant phase $\varphi$, the masses of the two stars, and in principle also their spins (which we neglected in eq. (5.273)). So, in total, we have 15 parameters.[57] However, a number of simplifications are possible, as we discuss in the next subsection.

[55] Explicit expressions for $\Phi(f_{gw})$ and $f_{gw}(t)$ up to 2PN were given in eq. (5.273), and in eq. (5.270) or (5.272).

[56] Recall also from Section 4.1.4 that, for binaries at cosmological distances, i.e. at a non-negligible redshift $z$, the masses $m_1$ and $m_2$ must be multiplied by $(1+z)$, and the distance $r$ must be replaced by the luminosity distance $d_L(z)$.

[57] Furthermore, the angles $(\theta,\phi)$ change in time because of the Earth's motion. For a ground-based interferometer, which follows the coalescence only for 10–15 minutes, this dependence can be neglected. For a space-borne interferometer, instead, it must be taken into account.

### 7.7.1 Elimination of extrinsic variables

The variables that can be eliminated from the parameter space are generically called extrinsic. First, we observe that all possible shifts in time of the signal can be obtained at once with a single Fourier transform. Consider in fact the scalar product $(h(\theta,t_*)|s)$ between the output $s(t)$ of the detector and the template $h(t;\theta,t_*)$ where, from the parameters $\theta^i$, we singled out explicitly the arrival time $t_*$, defined as the time when the hypothetical signal enters into the interferometer bandwidth, say at $f_{gw} = 10$ Hz. The waveform $h(t;\theta,t_*)$ is obtained from $h(t;\theta,t_*=0)$ with a time translation, so if we denote by $\tilde{h}(f;\theta)$ the Fourier transform of $h(t;\theta,t_*)$ at $t_*=0$, the Fourier transform of $h(t;\theta,t_*)$ at $t_*$ generic is simply $\tilde{h}(f;\theta)e^{i2\pi f t_*}$. Thus, from the definition (7.46) of the scalar product, we have

$$(h(\theta,t_*)|s) = 4\,\mathrm{Re}\int_0^\infty df\, \frac{\tilde{h}^*(f;\theta)\tilde{s}(f)}{S_n(f)}\, e^{i2\pi f t_*}\,, \tag{7.171}$$

which is just the Fourier transform of $\tilde{h}^*(f;\theta)\tilde{s}(f)/S_n(f)$. Thus, performing a single FFT we can immediately locate the value of $t_*$ which gives the highest signal-to-noise ratio. This is of course of great practical importance. Typically we can expect that, to perform efficiently the



matched filtering, the maximum mismatch in arrival time that we can tolerate between the real signal and our template could be, say, of order 3 ms. If one should analyze one year of data ($3\times10^7$ s) computing a different scalar product every 3 ms, for each value of $\theta$ one should perform $10^{10}$ times the computation of the scalar product $h(t;\theta,t_*)$, while we see that just a single FFT does the job.[58] Thus, the arrival time $t_*$ is not really part of the parameter space that must be searched. Figure 7.14 shows the result of a simulation in which the signal corresponding to the coalescence of two BHs, each with a mass of $10M_\odot$, at a distance of 150 Mpc, is injected into the noise of the VIRGO detector. Performing the Fourier transform, we see that we have a spike in correspondence with the time at which this signal has been injected (in the figure, $t_*=1$, in arbitrary units).

[58] More precisely, if we have a time series with $N$ samples, computing the integral which defines the scalar product has a computational cost $O(N)$, and if one had to repeat it for all possible arrival times, the overall cost would be $O(N^2)$. With a single FFT, instead, the computational cost is $O(N\log N)$.

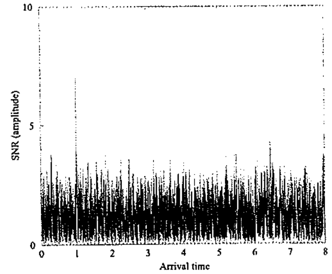


**Fig. 7.14** The result of a simulation in which the signal due to a BH-BH coalescence, each with a mass of $10M_\odot$, at a distance of 150 Mpc, is injected into the noise of the VIRGO detector. The arrival time is located from the position of the spike in the Fourier transform (7.171), which here is at $t_*=1$. (Courtesy of A. Viceré.)

Two more parameters that appear in eq. (7.167), which can be eliminated analytically from the matched filtering procedure, are the amplitude $A$ and the phase $\varphi$ of the signal. We already saw in Section 7.3 that the optimal filter is defined modulo an arbitrary constant, so the overall value of the amplitude $A$ does not enter when we search for the template that maximizes the signal-to-noise ratio. The maximization of the SNR with respect to $\varphi$ can be performed analytically, writing the template (7.170) in the form

$$h(t) = h_c(t)\cos\varphi + h_s(t)\sin\varphi\,. \tag{7.172}$$

If $s(t)$ is the detector output, after maximization of the log-likelihood function over the amplitude $A$, according to eq. (7.70) we want to further maximize

$$2\log\Lambda = \frac{(h|s)^2}{(h|h)} = \frac{[(h_c|s) + (h_s|s)\tan\varphi]^2}{(h_c|h_c) + (h_s|h_s)\tan^2\varphi + 2(h_c|h_s)\tan\varphi}\,. \tag{7.173}$$

This expression is easily maximized analytically with respect to $\tan\varphi$. The result is simpler if we introduce two new templates

$$h_p = h_c\cos\phi_p + h_s\sin\phi_p\,, \tag{7.174}$$

$$h_q = h_c\cos\phi_q + h_s\sin\phi_q\,, \tag{7.175}$$

where the angles $\phi_p$ and $\phi_q$ are chosen so that $h_p$ and $h_q$ satisfy $(h_p|h_q)=0$, i.e. they are orthogonal with respect to the scalar product $(\ |\ )$. In terms of these orthogonal templates the likelihood function, after performing the maximization over the amplitude $A$ and over the phase $\varphi$, takes the simple form

$$2\log\Lambda = \frac{(h_p|s)^2}{(h_p|h_p)} + \frac{(h_q|s)^2}{(h_q|h_q)}\,. \tag{7.176}$$

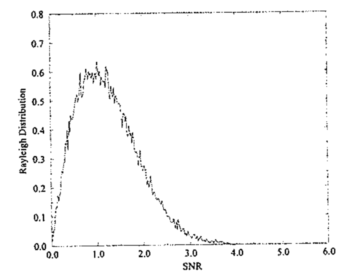


**Fig. 7.15** The distribution of the signal-to-noise ratio, in the simulation of Fig. 7.14. (Courtesy of A. Viceré.)

Therefore, the maximization with respect to the remaining variables is equivalent to maximizing the sum in quadrature of the outputs of two matched filters. In the absence of signal, the signal-to-noise ratio $\rho$

is therefore a random variable which follows the Rayleigh distribution (7.87), while in the presence of signal it is a non-central $\chi^2$ distribution with two degrees of freedom. Indeed, we see in Fig. 7.15 that, in the simulation of Fig. 7.14, $\rho$ follows a Rayleigh distribution (except, of course, for the presence of the single spike with $S/N = 8$ at $t_* = 1$).[59]

[59] It should also be observed that, for coalescing binaries, non-Gaussian noise should be much less important than for short bursts, since it should be much easier to have an impulsive disturbance that simulates a short burst, rather than a noise that lasts for about 15 minutes, simulating for all this time the behavior of a chirping signal.

### 7.7.2 The sight distance to coalescing binaries

The Fourier transform of the chirp amplitude, to Newtonian order, has been computed in Problem 4.1, while the result in the restricted PN approximation, up to 2PN order, is given in eqs. (5.274) and (5.275). Then we find, for the Fourier transform of $h(t) = h_+F_+ + h_\times F_\times$,

$$\tilde{h}(f) = \left(\frac{5}{6}\right)^{1/2} \frac{1}{2\,\pi^{2/3}} \frac{c}{r} \left(\frac{GM_c}{c^3}\right)^{5/6} f^{-7/6} e^{i\Psi} Q(\theta,\phi;\iota)\,, \tag{7.177}$$

where

$$Q(\theta,\phi;\iota) = F_+(\theta,\phi)\,\frac{1+\cos^2\iota}{2} + iF_\times(\theta,\phi)\cos\iota\,. \tag{7.178}$$

The phase $\Psi$ is just the quantity denoted $\Psi_+$ in eq. (5.275), and the relative factor $i$ between the two terms in $Q$ is due to the fact that $\Psi_\times = \Psi_+ + (\pi/2)$.[60] Plugging this expression into eq. (7.51), we can write the signal-to-noise ratio for a coalescing binary as

[60] Actually, the expression that we used for $h_+$ and $h_\times$ assumes a given choice of the axes with respect to which the plus and cross polarizations are defined, which is related to the orientation of the orbit, see page 296. Since a priori we do not know the orientation of the orbit, this will in general differ by an unknown angle $\psi$ from the definition of axes that the experimenter uses to define the pattern functions. Correspondingly, the expressions for $h_+$ and $h_\times$ must be rotated as in eqs. (7.24) and (7.25), so the function $Q$ is actually $Q(\theta,\phi;\iota,\psi)$. However this $\psi$ dependence, being an orthogonal transformation, does not affect that computation of $\langle|Q|^2\rangle$ performed below.

$$\left(\frac{S}{N}\right)^2 = \frac{5}{6}\,\frac{1}{\pi^{4/3}}\frac{c^2}{r^2}\left(\frac{GM_c}{c^3}\right)^{5/3} |Q(\theta,\phi;\iota)|^2 \int_0^{f_{\rm max}} df\,\frac{f^{-7/3}}{S_n(f)}\,, \tag{7.179}$$

where $f_{\rm max}$ is the value of the GW frequency when the inspiral phase terminates and the two stars merge. An estimate of $f_{\rm max}$ is $f_{\rm max} = 2(f_s)_{\rm ISCO}$, where $(f_s)_{\rm ISCO}$ given in eq. (4.39), and the factor of 2 is valid as long as the emission is dominated by quadrupole radiation. For a wave coming from optimal direction (e.g. $F_+ = 1$ and $F_\times = 0$), and with optimal value of the inclination of the orbit ($\cos\iota = 1$), the function $Q(\theta,\phi;\iota) = 1$. However, a more appropriate reference value for $|Q(\theta,\phi;\iota)|^2$ is given by its average over all possible directions and inclinations. Using $\langle F_+^2\rangle = \langle F_\times^2\rangle = 1/5$ for interferometers (see Table 7.1), we get $\langle|Q(\theta,\phi;\iota)|^2\rangle = (1/5)g(\iota)$, where $g(\iota)$ was defined in eq. (3.338) and its average over the inclination $\iota$ is 4/5, see eq. (4.10). Therefore

$$\langle|Q(\theta,\phi;\iota)|^2\rangle^{1/2} = \frac{2}{5}\,, \tag{7.180}$$

where here $\langle\ldots\rangle$ denotes the average over the angles and over the inclination. Then we rewrite eq. (7.179) as

$$\frac{S}{N} = \frac{2}{5}\left(\frac{5}{6}\right)^{1/2} \frac{1}{\pi^{2/3}}\frac{c}{r}\left(\frac{GM_c}{c^3}\right)^{5/6} \frac{\langle|Q(\theta,\phi;\iota)|^2\rangle^{1/2}}{(2/5)} \times\left[\int_0^{f_{\rm max}} df\,\frac{f^{-7/3}}{S_n(f)}\right]^{1/2}\,. \tag{7.181}$$



This relation can be inverted to give the sight distance $d_{\rm sight}$, i.e. the maximum distance $r$ at which we can see a binary coalescence, once we have chosen a given threshold for $S/N$,[61] assuming an average direction and inclination,

[61] Recall however from page 359 that the signal can combine with the noise either in a constructive or in a destructive way, so the output $\rho$ of the interferometer is a random variable whose average is $S/N$ and which follows, in the presence of signal, a non-central $\chi^2$ distribution with two degrees of freedom. Therefore, at any distance $r$, there is a probability of missed detection, and the fact that a source is at $r < d_{\rm sight}$ does not mean that it will be certainly detected. Conversely, there is also a non-zero probability that the signal from a source at $r > d_{\rm sight}$ combines with the noise so that its $S/N$ raises above the threshold.

$$d_{\rm sight} = \frac{2}{5}\left(\frac{5}{6}\right)^{1/2} \frac{c}{\pi^{2/3}}\left(\frac{GM_c}{c^3}\right)^{5/6}\left[\int_0^{f_{\rm max}} df\,\frac{f^{-7/3}}{S_n(f)}\right]^{1/2} (S/N)^{-1}. \tag{7.182}$$

We will see in Chapter 9 the numerical values that can be obtained for $d_{\rm sight}$ at existing and advanced interferometers.

It is instructive to verify from these expressions that, in order of magnitude, for a coalescing binary the matched filtering procedure gives a gain $\sim \mathcal{N}_c^{1/2}$. To this end, we assume that $S_n$ has a constant value $S_0$ between a minimum frequency $f_0$ and $f_{\rm max}$, while it is essentially infinite for $f < f_0$. Then, neglecting all numerical factors (and using for simplicity units $c = 1$, and the notation $M = GM_c/c^3$), we can perform the integral in eq. (7.181), and we get

$$\frac{S}{N} \sim \frac{1}{r} M^{5/6} S_0^{-1/2} f_0^{-4/3}\,. \tag{7.183}$$

From eqs. (7.167) and (7.168) we see that the GW amplitude is of order

$$h_0 \sim \frac{1}{r} f_0^{2/3} M^{5/3}\,, \tag{7.184}$$

while, from eq. (4.23), the number of cycles spent in the interferometer bandwidth is

$$\mathcal{N}_c \sim M^{-5/3} f_0^{-5/3}\,. \tag{7.185}$$

Using eq. (7.184) to eliminate $r$ from eq. (7.183), and eq. (7.185) to eliminate $M$, we get

$$\frac{S}{N} \sim \frac{h_0}{(f_0 S_0)^{1/2}} \mathcal{N}_c^{1/2}\,, \tag{7.186}$$

which shows indeed that, in order of magnitude, the signal-to-noise ratio (in amplitude) is larger by a factor $\mathcal{N}_c^{1/2}$ than for a burst with a characteristic frequency $f_0$ (compare with eq. (7.107) with $\tau_g = 1/f_0$ and $f_{\rm max} = f_0$). Of course, a more precise estimate requires the real form of $S_n(f)$, as well as the exact computation of the integral in eq. (7.181). This shows explicitly how the matched filtering procedure allows us to dig deeply into the noise floor, as we discussed already on page 344. Consider in fact the situation in which, after tracking the signal by $\mathcal{N}_c \gg 1$ cycles, we finally get $S/N$ of order one, so we begin to see the signal. According to eq. (7.186), this means that $h_0/(f_0 S_0)^{1/2}$ was of order $1/\mathcal{N}_c^{1/2}$. However, $h_0/(f_0 S_0)^{1/2}$ is the "instantaneous" value of the signal-to-noise ratio, i.e. the value of $S/N$ over a single cycle.

Therefore, the integrated signal-to-noise ratio provided by the matched filtering procedure can be of order one or larger, even when the *instantaneous* signal is deeply buried into the noise.

Finally, an important issue is the precision that can be obtained in the reconstruction of the source parameters. In particular the chirp mass $M_c$, that appears in the phase of the waveform, can be estimated very precisely, since the phase can be followed accurately for $\mathcal{N}_c$ cycles. Thus, any mismatch $\Delta M_c$ between the true value of the source and the value used in our template will be amplified by a factor $\mathcal{N}_c$, and we could expect that

$$\frac{\Delta M_c}{M_c} \sim \frac{1}{\mathcal{N}_c}\,. \tag{7.187}$$

Given that at a ground-based interferometer $\mathcal{N}_c$ can be of order $10^3$–$10^4$, see eq. (4.23), this would give a rather remarkable accuracy $\Delta M_c/M_c \sim 10^{-4}$–$10^{-3}$. As for the reduced mass $\mu$, it appears in the 1PN corrections to the phase, which are smaller by a factor $O(v^2/c^2)$ than the leading term, so it can be measured less precisely.[62]

[62] The precise computation of the errors on the parameters can be done using the explicit expression of the waveform to evaluate the Fisher information matrix defined in eq. (7.74), and then we can compute the errors on the parameters as in eq. (7.75). Using the waveform with the post-Newtonian corrections to the phase and assuming a detection with $S/N = 10$ one finds that, if one knew that the spins of the star are negligible, then $M_c$ could indeed be measured with a precision of 0.01–0.1%, while the reduced mass $\mu$, which enters in the post-Newtonian corrections, could be measured to 1%. However, one in general has no a priori information on the spins, and the measurements of masses and spins happen to be strongly correlated. This degrades the accuracy on the mass reconstruction, so one finally obtains $\Delta M_c/M_c \sim 0.1$–$1\%$ (which, however, is still a quite remarkable accuracy) and $\Delta\mu/\mu \sim 10-15\%$ for NS-NS and NS-BH binaries, or $\Delta\mu/\mu \sim 50\%$ for BH-BH binaries with typical BH masses of order $10M_\odot$. Observe that, the larger the mass of the stars, the smaller is the number of cycles in the detector bandwidth, since the coalescence takes place earlier, see eq. (4.39), so the precision in the reconstruction of the parameters is less good. See Cutler and Flanagan (1994) for details.

## 7.8 Stochastic backgrounds

In 1965 Penzias and Wilson discovered that the Universe is permeated by the Cosmic Microwave Background (CMB) electromagnetic radiation. This radiation is a relic of the early Universe, and the microwave photons that compose it decoupled from the primordial plasma about $3 \times 10^5$ years after the Big Bang, and since then they have been propagating essentially freely. This discovery, providing direct evidence for the Big Bang, was one of the most significant in the history of cosmology.

Since then, the CMB has been subject to deep investigations. We now know that its spectrum is a perfect black-body (in fact, the most perfect black-body spectrum existing in nature). This background is, to a first approximation, isotropic. The observation by the COBE satellite of temperature fluctuations over the sky, at the level $\Delta T/T \sim 10^{-5}$, has been one of the most important discoveries in cosmology in the last decades, and the detailed investigation of the multipole moments of these anisotropies by COBE and various other experiments, and particularly by WMAP, has opened up the field of precision cosmology.

There are good reasons to expect that the Universe is permeated also by a stochastic background of GWs generated in the early Universe. Furthermore, a stochastic background of GWs can also emerge from the incoherent superposition of a large number of astrophysical sources, too weak to be detected separately, and such that the number of sources that contribute to each frequency bin is much larger than one.

The mechanisms that can lead to the production of stochastic GW backgrounds in cosmology and in astrophysics will be examined in detail in Vol. 2. Here we discuss how to characterize such a background in general, and what are the optimal strategies for its detection.



### 7.8.1 Characterization of stochastic backgrounds

Using the plane wave expansion (1.58), we can write

$$h_{ij}(t,\mathbf{x}) = \sum_{A=+,\times} \int_{-\infty}^{\infty} df \int d^2\hat{\mathbf{n}}\, \tilde{h}_A(f,\hat{\mathbf{n}})\, e^{A}_{ij}(\hat{\mathbf{n}})\, e^{-2\pi i f(t-\hat{\mathbf{n}}\cdot\mathbf{x}/c)}\,. \tag{7.188}$$

We work in the TT gauge, so $h^i_i = 0$ and $\partial^j h_{ij} = 0$. The tensors $e^A_{ij}(\hat{\mathbf{n}})$ are given in eq. (1.54). A stochastic background is a superposition of waves with all possible propagation directions $\hat{\mathbf{n}}$, therefore the indices $i,j$ above take the values $1,2,3$, contrary to the case of the GWs emitted from a single far source, where we could label the GW in the TT gauge as $h_{ab}$ with $a,b$, taking the values $1,2$ and labeling the two directions in the transverse plane. A stochastic background is defined by the fact that the amplitudes $\tilde{h}_A(f,\hat{\mathbf{n}})$ are random variables, characterized statistically by their ensemble averages.[63]

We will make the following assumptions on stochastic backgrounds of GWs.

- The background is stationary. This means that all correlators depend only on time differences, and not on the absolute values of time. So, for instance, the two-point correlator $\langle h_A(t)h_{A'}(t')\rangle$ can depend only on $t-t'$, and not separately on $t$ and $t'$. In Fourier space, this means that $\langle \tilde{h}^*_A(f)\tilde{h}_{A'}(f')\rangle$ must be proportional to $\delta(f-f')$. This assumption is certainly justified. For a background created in cosmological epochs, the typical time-scale on which it can change substantially is of the order of the age of the Universe (for instance, its spectrum changes because it is redshifted). During the duration of the experiment, which is at most a few years, it is very difficult to imagine that the properties of the background could change appreciably.[64]
- The background is Gaussian. This means that all $N$-point correlators are reduced to sum and products of the two-point correlator $\langle h_A(t)h_{A'}(t')\rangle$ (and of the vacuum expectation value $\langle h_A\rangle$ that however, as we have seen, can be set to zero). Gaussianity is rooted in the central limit theorem, that states that the sum of a large number of independent events produces a Gaussian stochastic process, whatever the probability distribution of the individual events. This assumption is therefore expected to hold to a very good accuracy for cosmological backgrounds. It would not hold for astrophysical backgrounds, if the number of sources that contribute is not that large, and we are on the verge of distinguishing the individual contributions. In this case, further information can be extracted from the higher-point correlators.
- The stochastic background is isotropic. Experience with CMB indicates that the early Universe was highly isotropic and, for the photons, temperature fluctuations across the sky are at the level $\Delta T/T \sim 10^{-5}$. It is reasonable to expect that a stochastic background of GWs of cosmological will also be in a first approximation

[63] An ensemble average is the average over many copies of the system. Our system is in this case the Universe and we do not have many copies of it! Of course, the ergodic assumptions must be used here, and the ensemble average is replaced by a temporal average, compare with Note 3 on page 337.

[64] Stationarity also implies that $\langle h_A(t)\rangle$ is a constant so, even if it were nonzero, it would just contribute to the vacuum energy density. As far as we are interested in GWs, that is in the time-dependent part, we can therefore set $\langle h_A\rangle = 0$.

isotropic. Of course, after a first detection of a GW background, it will be extremely interesting to investigate its anisotropies and therefore to give up this assumption. In particular, in a cosmological background we must expect a dipole term, dominated by the Earth motion in the rest frame of the CMB, while higher multipoles could give extremely interesting information on the early Universe.

We might have to give up completely the assumption of isotropy when we study stochastic backgrounds of astrophysical origin. In particular a background of galactic origin will not be isotropic, but rather it will be more intense when we look in the direction of the galactic plane, just as the electromagnetic background due to galactic sources gives its characteristic appearance to the Milky Way. We will in fact discuss in Vol. 2 an example of this type, the background created by galactic white dwarf binaries.

Waves coming from different directions should be uncorrelated, so $\langle \tilde{h}_A^*(f,\hat{\mathbf{n}})\tilde{h}_{A'}(f',\hat{\mathbf{n}}')\rangle$ should be proportional to a Dirac delta over the two-sphere, defined as

$$\delta^2(\hat{\mathbf{n}},\hat{\mathbf{n}}') = \delta(\phi-\phi')\delta(\cos\theta-\cos\theta')\,, \tag{7.189}$$

where $(\theta,\phi)$ are the polar angles that define $\hat{\mathbf{n}}$. Isotropy implies that the proportionality constant must be independent of $\hat{\mathbf{n}}$.

- Finally, we assume that the background is unpolarized, as it is natural both in a cosmological context and if it is the result of the superposition of many different astrophysical sources. This means that $\langle \tilde{h}_A^*(f,\hat{\mathbf{n}})\tilde{h}_{A'}(f',\hat{\mathbf{n}}')\rangle$ must be proportional to $\delta_{AA'}$ and the proportionality coefficient must be independent of the polarization index $A$.

Under these assumptions, a stochastic background of GWs is uniquely characterized by a single function $S_h(f)$, defined by

$$\boxed{\langle \tilde{h}_A^*(f,\hat{\mathbf{n}})\tilde{h}_{A'}(f',\hat{\mathbf{n}}')\rangle = \delta(f-f')\,\frac{\delta^2(\hat{\mathbf{n}},\hat{\mathbf{n}}')}{4\pi}\,\delta_{AA'}\,\frac{1}{2}S_h(f)\,.} \tag{7.190}$$

The function $S_h(f)$ is called the spectral density of the stochastic background, in analogy with the spectral density of the noise defined in Section 7.1. Just as for the noise spectral density, we use the convention that $S_h(f)$ is single-sided. It has dimensions $\mathrm{Hz}^{-1}$ and satisfies $S_h(f)=S_h(-f)$. The factor $1/(4\pi)$ in eq. (7.190) is a choice of normalization such that

$$\int d^2\hat{\mathbf{n}}\, d^2\hat{\mathbf{n}}'\, \langle \tilde{h}_A^*(f,\hat{\mathbf{n}})\tilde{h}_{A'}(f',\hat{\mathbf{n}}')\rangle = \delta(f-f')\delta_{AA'}\frac{1}{2}S_h(f)\,. \tag{7.191}$$

where, as usual, $d^2\hat{\mathbf{n}} = d\cos\theta d\phi$. We see that the factor 1/2 in the definition of $S_h(f)$ has been inserted so that $S_h(f)$ is normalized in the same way as the single-sided spectral density of the noise, see eq. (7.6).



Using eqs. (7.188) and (7.190), as well as $\sum_A e_{ij}^A e_{ij}^A = 4$, which follows from the normalization (1.55) of the polarization tensor $e_{ij}^A$, we get

$$\langle h_{ij}(t)h^{ij}(t)\rangle = 4\int_0^\infty df\, S_h(f)\,. \tag{7.192}$$

The sum over $i,j$ is understood, and $h_{ij}(t) = h_{ij}(t,\mathbf{x}=0)$. The spectral density of the signal, $S_h(f)$, is the quantity that allows us to perform a direct comparison with the noise in a detector, which is characterized by $S_n(f)$. However, to have a physical understanding it is much more convenient to think in terms of the energy density carried by the stochastic background. According to eq. (1.135), this is related to $h_{ij}$ by

$$\rho_{\rm gw} = \frac{c^2}{32\pi G}\langle \dot{h}_{ij}\dot{h}^{ij}\rangle\,. \tag{7.193}$$

In cosmology there is a very natural unit of energy density, that is, the energy density needed for closing the Universe. This critical energy density is

$$\rho_c = \frac{3c^2H_0^2}{8\pi G}\,, \tag{7.194}$$

where $H_0$ is the present value of the Hubble expansion rate. As we mentioned on page 194, the value of $H_0$ is usually written as $H_0 = h_0\times 100\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, where $h_0$ parametrizes the existing experimental uncertainty and is called the normalized Hubble expansion rate. The most recent determinations give $h_0 = 0.73(3)$. Numerically,

$$\rho_c \simeq 1.688\times 10^{-8}\, h_0^2 \quad \mathrm{erg\,cm^{-3}}\,. \tag{7.195}$$

Normalizing $\rho_{\rm gw}$ to $\rho_c$, the intensity of a stochastic background of gravitational waves can be characterized by the dimensionless quantity

$$\Omega_{\rm gw} \equiv \frac{\rho_{\rm gw}}{\rho_c}\,. \tag{7.196}$$

Using eqs. (7.192) and (7.193), the energy density can be written as an integral over $d\log f$ of some spectral density, that we denote by[65] $d\rho_{\rm gw}/d\log f$,

$$\rho_{\rm gw} \equiv \int_{f=0}^{f=\infty} d(\log f)\,\frac{d\rho_{\rm gw}}{d\log f}\,. \tag{7.197}$$

We also define

$$\boxed{\Omega_{\rm gw}(f) \equiv \frac{1}{\rho_c}\frac{d\rho_{\rm gw}}{d\log f}\,,} \tag{7.198}$$

so $\Omega_{\rm gw}$ in eq. (7.196) is related to $\Omega_{\rm gw}(f)$ by[66]

$$\Omega_{\rm gw} = \int_{f=0}^{f=\infty} d(\log f)\,\Omega_{\rm gw}(f)\,. \tag{7.199}$$

[65] There is a slight abuse of notation here. Of course $\rho_{\rm gw}$, on the left-hand side of eq. (7.197), is independent of the frequency, so its derivative with respect to $f$, or to $\log f$, vanishes. On the right-hand side, $d\rho_{\rm gw}/d\log f$ is not the derivative of $\rho_{\rm gw}$ with respect to $\log f$, but just a notation for the spectral density of $\rho_{\rm gw}$, which stresses that it is the energy density contained in a logarithmic interval of frequency.

[66] Here again there is a slight ambiguity in the notation, because we use the same symbol $\Omega_{\rm gw}$ for the normalized energy density, on the left-hand side of eq. (7.199), and for its spectral density, on the right-hand side. This notation is however standard in the GW literature, and we will conform to it.

The fact that we consider the energy per unit logarithmic interval of frequency, $d\rho_{gw}/d\log f$, rather than $d\rho_{gw}/df$, is useful because in this way $\Omega_{gw}(f)$ is dimensionless.

Even if the experimental error on the Hubble expansion rate is becoming smaller and smaller (just a few years ago values of $h_0$ between 0.4 and 1 where considered possible), still it is not very convenient to normalize $\rho_{gw}$ to a quantity, $\rho_c$, which is uncertain: this uncertainty would appear in all the subsequent formulas, although it has nothing to do with the uncertainties on the GW background itself. Therefore, one rather characterizes the stochastic GW background with the quantity $h_0^2\Omega_{gw}(f)$, which is independent of $h_0$.[67]

[67] Unfortunately, we sometime use $h_0$ also to denote a GW amplitude. Since the reduced Hubble constant $h_0$ will only appear in the combination $h_0^2\Omega_{gw}$, no confusion is possible.

We now compute the relation between $S_h(f)$ and $h_0^2\Omega_{gw}(f)$. As discussed in Section 1.4.3, the brackets in eq. (7.193) denote a time average. However (under the ergodic assumption, see Notes 3 and 63), this is just the ensemble average used above. We can then substitute the plane wave expansion (7.188) into eq. (7.193), and compute the ensemble average using eq. (7.190). The result is

$$\rho_{gw} = \frac{c^2}{8\pi G}\int_{f=0}^{f=\infty} d(\log f)\; f(2\pi f)^2 S_h(f)\,. \tag{7.200}$$

Comparing with the definition (7.197) we get

$$\frac{d\rho_{gw}}{d\log f} = \frac{\pi c^2}{2G} f^3 S_h(f)\,, \tag{7.201}$$

and

$$\boxed{\Omega_{gw}(f) = \frac{4\pi^2}{3H_0^2} f^3 S_h(f)\,.} \tag{7.202}$$

Finally, it is interesting to express $h_0^2\Omega_{gw}(f)$ in terms of the number of gravitons per cell of the phase space, $n(\mathbf{x},\mathbf{k})$. For an isotropic stochastic background $n(\mathbf{x},\mathbf{k}) = n_f$ depends only on the frequency (which is related to the momentum $\mathbf{k}$ by $|\mathbf{k}| = \hbar\omega/c = 2\pi\hbar f/c$), and not on the direction $\hat{\mathbf{k}}$. Then, writing $d^3k = |\mathbf{k}|^2 d|\mathbf{k}| d\Omega \to 4\pi(2\pi\hbar/c)^3 f^2 df$, and $df = f d\log f$, and considering that a graviton of frequency $f$ carries an energy $\hbar\omega = \hbar(2\pi f)$, we have

$$\begin{aligned}\rho_{gw} &= 2\int \frac{d^3k}{(2\pi\hbar)^3}\,\hbar(2\pi f)\, n_f \\ &= \frac{16\pi^2\hbar}{c^3}\int_0^\infty d(\log f)\, f^4 n_f\,,\end{aligned} \tag{7.203}$$

where the factor of 2 in front of the integral is due to the two helicity states of the graviton. Therefore

$$\frac{d\rho_{gw}}{d\log f} = \frac{16\pi^2\hbar}{c^3} n_f f^4\,, \tag{7.204}$$

and

$$h_0^2\Omega_{gw}(f) \simeq 3.6\left(\frac{n_f}{10^{37}}\right)\left(\frac{f}{1\text{kHz}}\right)^4. \tag{7.205}$$



As we will see in Vol. 2, this equation is useful in particular when one computes the production of a stochastic background of GWs due to amplification of vacuum fluctuations, since this computation gives directly $n_f$.

### 7.8.2 SNR for single detectors

The comparison of eqs. (7.6) and (7.191) makes it clear that an isotropic stochastic background of GWs is seen in a detector as an additional source of noise. This poses an important conceptual problem in the identification of a stochastic GW background. In practice what will happen is that, after a careful modeling of the detector and of its noise sources, one would expect to have a certain value of the spectral density of the noise, $S_n(f)$. When the detector is turned on, one measures $\langle s^2(t)\rangle$, where as usual $s(t) = n(t) + h(t)$, with $n(t)$ the noise and $h(t)$ the response of the detector to a GW signal. If one observes that $\langle s^2(t)\rangle$ is larger than expected, the crucial problem is how to tell whether this is really due to a GW background or, more trivially, to some source of noise that has not been adequately accounted for when estimating $S_n(f)$. Similar problems were faced in the discovery of the cosmic microwave background; Penzias and Wilson found an excess noise in their antenna (a horn reflector that was meant for satellite communications) and worked hard for one year in order to exclude all possible sources of terrestrial and astrophysical noise, before writing a short paper with the modest title "A Measurement of Excess Antenna Temperature at 4080 Mc/s", and concluding *"From a combination of the above, we compute the remaining unaccounted-for antenna temperature to be* 3.5 ± 1.0 *K at 4080 Mc/s"*.

To cope with this problem, it is clear that in the search for stochastic backgrounds of GWs with a single detector one must set at least a relatively high threshold on the signal-to-noise ratio; for instance, a signal-to-noise ratio $S/N = 5$ on the amplitude could be a typical choice (while lower values of $S/N$ could be used for the only purpose of putting upper bounds). Further handles could come from an anisotropy of the stochastic GW background, if it is due to unresolved galactic sources, since this would produce a sidereal time modulation due to the motion of the detector. Another handle is the possibility that the dependence of the excess noise on the frequency is found to be in agreement with some theoretical prediction from a given cosmological or astrophysical mechanism.

To compute the minimum value of $h_0^2\Omega_{gw}$ that can be measured at a given $S/N$, we observe that, if there is no signal, we have (see eq. (7.12))

$$\langle s^2(t)\rangle = \langle n^2(t)\rangle = \int_0^\infty df\, S_n(f)\,, \tag{7.206}$$

while, if a stochastic GW background is present, there is also a contribution from $h(t)$. For each propagation direction $\hat{\mathbf{n}}$ we can write $h(t) = h_+F_+ + h_\times F_\times$, and therefore, taking the ensemble average and averaging also over $\hat{\mathbf{n}}$ and over the polarization angle $\psi$,

$$\int \frac{d^2\hat{\mathbf{n}}}{4\pi}\frac{d\psi}{2\pi}\langle h^2\rangle = \left(\int \frac{d^2\hat{\mathbf{n}}}{4\pi}\frac{d\psi}{2\pi}F_+^2\right)\langle h_+^2 + h_\times^2\rangle\,, \tag{7.207}$$

where we used the fact that the angular averages of $F_+^2$ and of $F_\times^2$ are equal, see eq. (7.35). For an isotropic background, the ensemble average $\langle h^2\rangle$ that appears on the left-hand side of eq. (7.207) is independent of the angles $\hat{\mathbf{n}}$ and $\phi$, so the angular average gives one. The term on the right-hand side of eq. (7.207), instead, can be written in terms of $S_h(f)$ using eq. (7.192) and observing that, for any given propagation direction, we have $h_{ij}h^{ij} = 2(h_+^2 + h_\times^2)$. Then

$$\langle h^2\rangle = 2\langle F_+^2\rangle \int_0^\infty df\, S_h(f)\,, \tag{7.208}$$

where, with an abuse of notation, the brackets in $\langle h^2\rangle$ denote the ensemble average while the brackets in $\langle F_+^2\rangle$ denotes the average over $d^2\hat{\mathbf{n}}$ and $d\psi$. In eq. (7.37) we have defined the angular efficiency factor $F = \langle F_+^2\rangle + \langle F_\times^2\rangle = 2\langle F_+^2\rangle$, whose value for various detectors are given in Table 7.1. In particular, $F = 2/5$ for interferometers and $F = 8/15$ for resonant bars. Then

$$\langle h^2(t)\rangle = F\int_0^\infty df\, S_h(f)\,. \tag{7.209}$$

So, in the presence of signal,

$$\langle s^2(t)\rangle = \langle n^2(t)\rangle + \langle h^2(t)\rangle = \int_0^\infty df\, [S_n(f) + F S_h(f)]\,. \tag{7.210}$$

Therefore, if a stochastic background is present, one simply observes that $\langle s^2(t)\rangle$ is higher than the value expected from the noise, everywhere or just in some frequency range. More precisely, we can compare the output with the expected value of $S_n(f)$ in each frequency bin (with bins of size $\Delta f = 1/T$ after an observation time $T$). To take the binning into account, we replace

$$\int S_h(f)df \to \sum_i S_h(f_i)\Delta f\,, \tag{7.211}$$

and

$$\int S_n(f)df \to \sum_i S_n(f_i)\Delta f \tag{7.212}$$

The signal-to-noise ratio in each bin is therefore[68]

$$\left(\frac{S}{N}\right)^2 = F\frac{S_h(f_i)\Delta f}{S_n(f_i)\Delta f} = F\frac{S_h(f_i)}{S_n(f_i)}\,. \tag{7.213}$$

[68] Observe that here $S/N$ is the signal-to-noise ratio with respect to the GW amplitude, just as we have defined it for bursts, coalescence and periodic signals. For stochastic backgrounds, what is actually measured is an energy density, and it make sense to introduce the signal-to-noise ratio with respect to the energy density, which is quadratic in the amplitude. If one prefers to reserve the notation $S/N$ for the signal-to-noise ratio in energy, then on the left-hand side of eq. (7.213) one must write $S/N$ rather than $(S/N)^2$.



Of course the integration time $T$, which enters through $\Delta f$, canceled in eq. (7.213). Increasing the integration time, we decrease the size of the bins and therefore the noise in each bin, but we equally decrease the signal present in each bin. Therefore, in a single detector, as far as the signal-to-noise ratio is concerned, there is no gain in integrating the signal in time. Either the signal stands out immediately as soon as we switch on the detector, or it will always remain below the noise. If however the signal stands out, integrating it for a longer time we get a more detailed resolution of its frequency dependence.

In conclusion, the minimum value of $S_h(f)$ measurable with a single detector having a noise spectral density $S_n(f)$, at a given level $S/N$ of signal-to-noise ratio in amplitude, is

$$[S_h(f)]_{\rm min} = S_n(f)\frac{(S/N)^2}{F}\,, \tag{7.214}$$

and correspondingly the minimum detectable value of $\Omega_{\rm gw}$ is

$$\boxed{[\Omega_{\rm gw}(f)]_{\rm min} = \frac{4\pi^2}{3H_0^2}f^3 S_n(f)\,\frac{(S/N)^2}{F}\,.} \tag{7.215}$$

A very important feature of this expression is the factor $f^3$. It tells us that, if one is able to reach a given level in $S_n(f)$ at low frequencies, it will be possible to reach a much better sensitivity in $\Omega_{\rm gw}(f)$ compared to what can be obtained with a similar value of $S_n(f)$ at high frequencies. Of course, the experimental problems that one has to solve in order to reach a given value of $S_n(f)$ depend very strongly on the frequency $f$. However, at $f = 10^{-3}$ Hz, the space detector LISA aims at reaching a strain sensitivity $S_n^{1/2}(f) = 4\times 10^{-21}\,{\rm Hz}^{-1/2}$, while a ground-based interferometer at $f = 10^2$ Hz has $S_n^{1/2}(f) = 4\times 10^{-23}\,{\rm Hz}^{-1/2}$, as we will see in Chapter 9. Therefore, moving from $f = 10^2$ Hz to $f = 10^{-3}$ Hz, we lose only four orders in magnitude in $S_n(f)$, but we gain a factor $(10^2/10^{-3})^3 = 10^{15}$ thanks to $f^3$. Therefore, it is much easier to reach a small level for $[\Omega_{\rm gw}(f)]_{\rm min}$ at low $f$ rather than at high $f$. The other important question is in what frequency range should we expect that cosmological or astrophysical mechanisms produce an interesting value for $\Omega_{\rm gw}(f)$. As we will see in Vol. 2, there is a large variety of possible mechanisms, which can produce stochastic GW backgrounds everywhere from $f = 10^{-18}$ Hz up to $f = 10^9$ Hz. Their detection is therefore easier when they are large at low frequencies, since then comparatively high value of the noise spectral density $S_n(f)$ can be over-compensated by the factor $f^3$, and becomes more and more difficult as we go to high frequencies. Numerically, with normalizations useful for LISA, eq. (7.215) gives

$$[h_0^2\Omega_{\rm gw}(f)]_{\rm min} = 1.1\times 10^{-12}\left(\frac{f}{1\,{\rm mHz}}\right)^3\left(\frac{S_n^{1/2}}{4\times 10^{-21}\,{\rm Hz}^{-1/2}}\right)^2 \times\left(\frac{1/\sqrt{5}}{F}\right)\left(\frac{S/N}{5}\right)^2\,. \tag{7.216}$$

Using normalization factors appropriate for ground-based interferometers, we rather have

$$[h_0^2\Omega_{gw}(f)]_{\min} = 0.12\left(\frac{f}{100\,\mathrm{Hz}}\right)^3\left(\frac{S_n^{1/2}}{4\times10^{-23}\,\mathrm{Hz}^{-1/2}}\right)^2 \times\left(\frac{2/5}{F}\right)\left(\frac{S/N}{5}\right)^2. \qquad (7.217)$$

In both cases we used a rather high value of the signal-to-noise ratio as a reference value, $S/N = 5$, according to the discussion above. The huge difference between the value $h_0^2\Omega_{gw} \sim 10^{-12}$ in eq. (7.216) and the value $h_0^2\Omega_{gw} \sim 0.1$ in eq. (7.217) is due to the fact that LISA can reach a value of $S_n$ not far from that of ground-based interferometers, at a much lower frequency.

As we will see in Vol. 2, no cosmological or astrophysical background of GW is expected to exceed $h_0^2\Omega_{gw}(f) \sim 10^{-5}$, independently of the frequency. Therefore eqs. (7.216) and (7.217) tell us that LISA has an extremely good sensitivity for stochastic backgrounds of GWs, while ground-based interferometers, used as single detectors, do not reach an interesting level for stochastic backgrounds. However, having at our disposal more than one ground-based detector (interferometers or bars) we can correlate their outputs, and the sensitivity improves dramatically, as we discuss in the next section.

### 7.8.3 Two-detector correlation

#### Optimal signal-to-noise ratio

With a single detector, it is impossible to adapt to stochastic backgrounds the matched filtering technique that we studied in Section 7.3. The reason is that, to perform the matched filtering, we need to know the form of the signal, but for stochastic backgrounds the GW signal $h(t)$ is an unpredictable randomly fluctuating quantity, just like the noise $n(t)$. However, if we have two detectors, we can use a modified form of matched filtering in which, rather than trying to match the output of a single detector to a predetermined signal $h(t)$, we match the output of one detector to the output of the other.

To implement this idea we proceed as follows. We write the output $s_k(t)$ of the $k$-th detector as $s_k(t) = h_k(t) + n_k(t)$, where $k = 1, 2$ labels the detector. Observe that the scalar output $h_k(t)$ depends in general on the detector, because different detectors can have a different location and/or a different orientation and therefore a different pattern function. We are interested in the situation in which the GW signal $h_k(t)$ is much smaller than the noise $n_k(t)$, which is the realistic situation for all ground-based detectors, as we have seen in the previous section. Multiplying both sides of eq. (7.188) by the detector tensor $D^{ij}$ and using eq. (7.21), we can write the GW signal $h_k$ in the $k$-th detector as

$$h_k(t,\mathbf{x}_k) = \sum_{A=+,\times}\int_{-\infty}^{\infty} df \int d^2\hat{\mathbf{n}}\, \tilde{h}_A(f,\hat{\mathbf{n}})\, e^{-2\pi i f(t-\hat{\mathbf{n}}\cdot\mathbf{x}_k/c)} F_k^A(\hat{\mathbf{n}}), \qquad (7.218)$$

where $F_k^A$ are the pattern functions of the $k$-th detector and $\mathbf{x}_k$ is its location. As always, the size of the detector is taken to be much smaller than $\lambda\!\!\!^{-}$, so we can neglect the spatial variation of the GW over the extension of the detector. Passing to the Fourier transform, we have

$$\tilde{h}_k(f) = \sum_{A=+,\times}\int d^2\hat{\mathbf{n}}\, \tilde{h}_A(f,\hat{\mathbf{n}}) e^{2\pi i f\hat{\mathbf{n}}\cdot\mathbf{x}_k/c} F_k^A(\hat{\mathbf{n}}), \qquad (7.219)$$

where we denote $\tilde{h}_k(f,\mathbf{x}_k)$ simply as $\tilde{h}_k(f)$. To correlate the outputs $s_1(t)$ and $s_2(t)$ of the two detectors we define

$$Y = \int_{-T/2}^{T/2} dt \int_{-T/2}^{T/2} dt'\, s_1(t) s_2(t') Q(t-t'), \qquad (7.220)$$

where $T$ is the total observation time (e.g. one year) and $Q$ a real filter function, analogous to the function $K(t)$ in Section 7.3. $Y$ is our signal, and we want to maximize its signal-to-noise ratio.

We limit ourselves to functions $Q(t-t')$ that fall rapidly to zero for large $|t-t'|$. Passing to the Fourier transforms, we get

$$Y = \int_{-\infty}^{+\infty} df df' df''\, \delta_T(f-f'')\delta_T(f'-f'')\tilde{s}_1^*(f)\tilde{s}_2(f')\tilde{Q}(f''), \qquad (7.221)$$

where

$$\delta_T(f) \equiv \int_{-T/2}^{T/2} dt\, e^{i2\pi f t} = \frac{\sin(\pi f T)}{\pi f}, \qquad (7.222)$$

and becomes a delta function in the limit $fT \to \infty$. Even on a relatively short stretch of data with, say, $T = 10^3$ s, at $f = 10$ Hz we have $fT = 10^4$. Over the whole useful bandwidth of ground-based detectors we can therefore replace $\delta_T(f)$ by a Dirac delta, and eq. (7.220) becomes

$$Y \simeq \int_{-\infty}^{+\infty} df\, \tilde{s}_1^*(f)\tilde{s}_2(f)\tilde{Q}(f). \qquad (7.223)$$

Recall that, in the signal-to-noise ratio $S/N$, $S$ is defined as the ensemble average value of $Y$ when the signal is present, while $N$ is the rms value of $Y$ when the signal is absent. Then, assuming that the noise in the two detectors are not correlated (and averaging also over the polarization angle $\psi$),

$$S = \int_{-\infty}^{+\infty} df\, \langle \tilde{h}_1^*(f)\tilde{h}_2(f)\rangle\, \tilde{Q}(f)$$

$$= \int_{-\infty}^{+\infty} df \sum_{A,A'} \int d^2\hat{\mathbf{n}} d^2\hat{\mathbf{n}}' \int \frac{d\psi}{2\pi} e^{-2\pi i f(\hat{\mathbf{n}}\cdot\mathbf{x}_1 - \hat{\mathbf{n}}'\cdot\mathbf{x}_2)/c}$$
$$\times F_1^A(\hat{\mathbf{n}};\psi) F_2^{A'}(\hat{\mathbf{n}}';\psi) \langle \tilde{h}_A^*(f,\hat{\mathbf{n}}) \tilde{h}_{A'}(f,\hat{\mathbf{n}}')\rangle \tilde{Q}(f) . \qquad (7.224)$$

Using eq. (7.190), together with $\delta(0) = \int_{-T/2}^{T/2} dt = T$, this becomes

$$S = \frac{T}{2} \int_{-\infty}^{\infty} df\, S_h(f) \Gamma(f) \tilde{Q}(f) , \qquad (7.225)$$

where we have defined

$$\boxed{\Gamma(f) \equiv \int \frac{d^2\hat{\mathbf{n}}}{4\pi} \int \frac{d\psi}{2\pi} \left[ \sum_A F_1^A(\hat{\mathbf{n}}) F_2^A(\hat{\mathbf{n}}) \right] \exp\left\{ i 2\pi f \hat{\mathbf{n}} \cdot \frac{\Delta \mathbf{x}}{c} \right\} ,} \qquad (7.226)$$

and $\Delta\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1$ is the separation between the two detectors. The function $\Gamma$ in called the (unnormalized) *overlap reduction function*. It takes into account the fact that the two detectors can see a different gravitational signal, either because they are at different location or because they have a different angular sensitivity.

The difference in location is reflected in the exponential factor. In particular, if $2\pi f \Delta x/c \gg 1$, i.e. if the separation $\Delta x \gg \lambda\!\!\!\!-$, this exponential is rapidly oscillating and suppresses strongly the correlation. This reflect the fact that, when $\Delta x \gg \lambda\!\!\!\!-$, the two detectors are experiencing GW signals that are uncorrelated.

The different angular sensitivity of the two detectors is instead reflected in the term $\sum_A F_1^A(\hat{\mathbf{n}}) F_2^A(\hat{\mathbf{n}})$. It is also useful to introduce the quantity

$$F_{12} \equiv \int \frac{d^2\hat{\mathbf{n}}}{4\pi} \int \frac{d\psi}{2\pi} \sum_A F_1^A(\hat{\mathbf{n}}) F_2^A(\hat{\mathbf{n}}) \Big|_{\text{aligned}} , \qquad (7.227)$$

where the subscript means that we must compute $F_{12}$ taking the two detectors to be at the same location and oriented one relative to the other so that the quantity $F_{12}$ is maximized.[69] Observe that, if the two detectors are of the same type, e.g. two interferometers or two cylindrical bars, $F_{12}$ is the same as the constant $F$ defined in eq. (7.37). The (normalized) overlap reduction function $\gamma(f)$ is defined as

$$\gamma(f) = \frac{\Gamma(f)}{F_{12}} . \qquad (7.228)$$

For instance, for the correlation between two interferometers, $F_{12} = 2/5$. The factor $F_{12}$ takes into account the reduction in sensitivity due to the pattern functions, already present in the case of one interferometer, and therefore $\gamma(f)$ separately takes into account the effect of the separation $\Delta\mathbf{x}$ between the interferometers, and of their relative orientation. With this definition, $\gamma(f) = 1$ if the separation $\Delta x = 0$ and if the detectors are perfectly aligned. However, the use of $\Gamma(f)$ is more convenient when we want to write equations that hold independently of what detectors (interferometers, bars, or spheres) are used in the correlation.

We now find the optimal choice of the filter function $\tilde{Q}(f)$ that maximizes the signal-to-noise ratio. We need to compute

$$N^2 = \left[\langle Y^2\rangle - \langle Y\rangle^2\right]_{h=0} \qquad (7.229)$$
$$= \int_{-\infty}^{\infty} df df'\, \tilde{Q}(f) \tilde{Q}^*(f')$$
$$\times \left[ \langle \tilde{n}_1^*(f) \tilde{n}_2(f) \tilde{n}_1(f') \tilde{n}_2^*(f')\rangle - \langle \tilde{n}_1^*(f)\tilde{n}_2(f)\rangle \langle \tilde{n}_2^*(f')\tilde{n}_1(f')\rangle \right] .$$

If the noise in the two detectors are uncorrelated, the mixed correlator $\langle \tilde{n}_1^*(f)\tilde{n}_2(f)\rangle$ vanishes, so the second term in brackets is zero, while the first factorizes $\langle \tilde{n}_1^*(f) \tilde{n}_2(f) \tilde{n}_1(f') \tilde{n}_2^*(f')\rangle = \langle \tilde{n}_1^*(f)\tilde{n}_1(f')\rangle \langle \tilde{n}_2(f)\tilde{n}_2^*(f')\rangle$. Then we get

$$N^2 = \int_{-\infty}^{\infty} df df'\, \tilde{Q}(f)\tilde{Q}^*(f') \langle \tilde{n}_1^*(f)\tilde{n}_1(f')\rangle \langle \tilde{n}_2^*(f')\tilde{n}_2(f)\rangle . \qquad (7.230)$$

Using

$$\langle \tilde{n}_k^*(f) \tilde{n}_k(f')\rangle = \delta(f - f') \frac{1}{2} S_{n,k}(f) , \qquad (7.231)$$

where $S_{n,k}(f)$ is the noise spectral density of the $k$-th detector, and using $\delta(0) = T$, we finally get

$$N^2 = \frac{T}{4} \int_{-\infty}^{\infty} df\, |\tilde{Q}(f)|^2 S_n^2(f) , \qquad (7.232)$$

where we have defined the combined noise spectral density

$$S_n(f) = \left[S_{n,1}(f) S_{n,2}(f)\right]^{1/2} . \qquad (7.233)$$

Equations (7.225) and (7.232) show the same crucial feature that we already observed when we discussed the matched filtering for periodic signals: the signal $S$ increase linearly with the observation time $T$, while the noise $N$ increases only as $T^{1/2}$. Therefore, the signal-to-noise ratio increases with the observation time as $T^{1/2}$. Putting together eqs. (7.225) and (7.232) we have

$$\frac{S}{N} = T^{1/2} \frac{\int_{-\infty}^{\infty} df\, S_h(f) \Gamma(f) \tilde{Q}(f)}{\left[\int_{-\infty}^{\infty} df\, |\tilde{Q}(f)|^2 S_n^2(f)\right]^{1/2}} . \qquad (7.234)$$

We can now find the filter function $\tilde{Q}(f)$ that maximizes $S/N$. The procedure is analogous to what we have already done between eqs. (7.45) and (7.51). For any two complex functions $A(f), B(f)$ we define the positive definite scalar product

$$(A, B) = \int_{-\infty}^{\infty} df\, A^*(f) B(f) S_n^2(f) . \qquad (7.235)$$

[69] For two detectors of the same type this means to orient them in the same way, so in a two-interferometer correlation the arms are taken to be along the $x$ and $y$ axes for both interferometers, and for a two-bar correlation the longitudinal axes of the bars are taken parallel to each other. For the correlation between a bar and an interferometer, we see from the form of the pattern functions given in Table 7.1 that the optimal correlation is obtained aligning the longitudinal axis of the bar with one of the arms of the interferometer.

Then eq. (7.234) can be rewritten as

$$\frac{S}{N} = T^{1/2}\frac{\left(\tilde{Q}, \Gamma S_h/S_n^2\right)}{(\tilde{Q},\tilde{Q})^{1/2}}. \qquad (7.236)$$

As we already discussed below eq. (7.47), this expression is maximized choosing

$$\tilde{Q}(f) = \text{const.}\frac{\Gamma(f)S_h(f)}{S_n^2(f)}. \qquad (7.237)$$

It is important to observe that the optimal filter depends on the signal that we are looking for, since $S_h(f)$ enters eq. (7.237). Plugging eq. (7.237) into eq. (7.236) we find the optimal signal-to-noise ratio,

$$\frac{S}{N} = T^{1/2}\left(\frac{\Gamma S_h}{S_n^2}, \frac{\Gamma S_h}{S_n^2}\right)^{1/2}, \qquad (7.238)$$

or, writing explicitly the scalar product,[70]

$$\frac{S}{N} = \left[2T\int_0^\infty df\, \Gamma^2(f)\frac{S_h^2(f)}{S_n^2(f)}\right]^{1/2}. \qquad (7.239)$$

In particular, for a two-interferometer correlation, $\Gamma(f) = (2/5)\gamma(f)$ and

$$\left(\frac{S}{N}\right)_{\text{intf-intf}} = \left[\frac{8}{25}T\int_0^\infty df\, \gamma^2(f)\frac{S_h^2(f)}{S_n^2(f)}\right]^{1/2}. \qquad (7.240)$$

For two cylindrical bars, instead, $\Gamma(f) = (8/15)\gamma(f)$, while for the correlation between an interferometer and a cylindrical bar, from the explicit expressions of the pattern functions in Table 7.1, we get again $\Gamma(f) = (2/5)\gamma(f)$.

Using eqs. (7.233) and (7.202) we can also rewrite eq. (7.239) as

$$\frac{S}{N} = \frac{3H_0^2}{4\pi^2}\left[2T\int_0^\infty df\, \Gamma^2(f)\frac{\Omega_{\text{gw}}^2(f)}{f^6 S_{n,1}(f)S_{n,2}(f)}\right]^{1/2}, \qquad (7.241)$$

and in particular, for a two-interferometer correlation,

$$\left(\frac{S}{N}\right)_{\text{intf-intf}} = \frac{3H_0^2}{10\pi^2}\left[2T\int_0^\infty df\, \gamma^2(f)\frac{\Omega_{\text{gw}}^2(f)}{f^6 S_{n,1}(f)S_{n,2}(f)}\right]^{1/2}. \qquad (7.242)$$

We can now compare the measurements of stochastic backgrounds performed with the two-detector correlation, to the measurement which uses a single detector, both from the point of view of sensitivity, and of the ability to discriminate true GWs from noise.

[70] Observe that, for periodic signals and for bursts, as well as for a single-detector search of stochastic backgrounds, we defined the quantity $S/N$ as linear in the GW, i.e. if $\tilde{h}(f) \to \lambda\tilde{h}(f)$, then $(S/N) \to \lambda(S/N)$, see eq. (7.51) and eq. (7.213). For searches of stochastic backgrounds with two-detector correlations, we have rather defined $S/N$ as linear both in $h_1(t)$ and in $h_2(t)$ and therefore $S/N$ scales overall quadratically in the GW amplitude. If we prefer to use a quantity that is linear in the GW amplitude we can define $\text{SNR} = (S/N)^{1/2}$, so SNR is proportional to $T^{1/4}$. Of course, it is a matter of conventions whether to use SNR or $(S/N)$.



### Comparison of two-detector and single-detector sensitivities

To compare the sensitivity of a two-detector correlation with the sensitivity of a single detector we assume that we have two identical detectors at a very close distance and with the same orientation, so that $\Gamma(f)$ becomes equal to the angular efficiency factor $F_{12} = F$. (This is the most favorable situation; however in practice, if the detectors are too close, there will be correlated noise.) To perform an order-of-magnitude estimate, we approximate eq. (7.239) as

$$\left(\frac{S}{N}\right)^2 \sim (2T\Delta f)\, F^2\frac{S_h^2}{S_n^2}, \qquad (7.243)$$

where $\Delta f$ is the useful bandwidth of the detectors, centered around a frequency $f$, and $S_n$ and $S_h$ are typical values of $S_n(f)$ and $S_h(f)$, respectively, over this bandwidth. Then the minimum detectable value of $S_h$, at signal-to-noise level $S/N$, is

$$(S_h)_{\min} \sim \frac{S_n}{(2T\Delta f)^{1/2}}\frac{(S/N)}{F}, \qquad (7.244)$$

and therefore

$$[\Omega_{\text{gw}}]_{\min} \sim \frac{4\pi^2}{3H_0^2}\frac{f^3 S_n}{(2T\Delta f)^{1/2}}\frac{(S/N)}{F}. \qquad (7.245)$$

where $f^3$ is really a typical value of $f^3$ over the bandwidth. Comparing eq. (7.244) with eq. (7.214) we see that, correlating two detectors, we have gained a factor $(2T\Delta f)^{-1/2}$. Numerically,

$$\frac{1}{(2T\Delta f)^{1/2}} \simeq 1\times 10^{-5}\left(\frac{150\,\text{Hz}}{\Delta f}\right)^{1/2}\left(\frac{1\,\text{yr}}{T}\right)^{1/2}. \qquad (7.246)$$

Therefore, integrating for one year the output of two detectors with a bandwidth of 150 Hz, we can improve our sensitivity to $S_h$, and therefore to $h_0^2\Omega_{\text{gw}}$, by approximately five orders of magnitudes, with respect to the sensitivity of a single detector.[71] It is interesting to compare these results with the matched filtering procedure discussed in Section 7.3. In Section 7.3 we took advantage of the fact that we knew the form of the signal, in order to discriminate it from the noise. Here, instead, in a single detector both the signal and the noise have the same statistical properties, but we took advantage of the fact the signals in the two detectors are correlated, while the noise are decorrelated. In particular, the measure of the correlation between the signals in the two detectors is given by the overlap reduction function $\Gamma(f)$ of eq. (7.226), which shows that the signals are indeed well correlated if the separation between the detectors is much smaller than $\lambdabar$, and if the detectors are well oriented with respect to each other. Technically, the assumptions that the noise are uncorrelated entered in eq. (7.224), as well as when passing from eq. (7.229) to eq. (7.230), where we neglected the correlator $\langle\tilde{n}_1^*(f)\tilde{n}_2(f)\rangle$.

[71] The precise numbers, of course, can only be obtained once we have the form of $S_h(f)$ and of $S_n(f)$, carrying out the integral in eq. (7.239). Observe also that in eq. (7.214) appears $(S/N)^2$ while in eq. (7.244) appears $(S/N)$, but this is simply a consequence of the fact that, for the two-detector correlation, we have defined $S/N$ as a quantity quadratic in the GW amplitude, while for a single detector we defined it to be linear in the GW amplitude. Once we choose our criterion for fixing the confidence level, e.g. a signal-to-noise ratio 1.7 in amplitude, the quantity that we are denoting by $(S/N)^2$ here and the quantity denoted by $S/N$ in eq. (7.214) have the same numerical value.

Recall however that the optimal filter depends on the form of the signal. A stochastic background of cosmological origin, as we will see in Vol. 2, is not expected to show strong spectral features in the bandwidth $\Delta f \sim 1$ kHz of ground based interferometers, so it should be adequate a simple power-law parametrization,

$$h_0^2\Omega_{\rm gw}(f) = K f^{\alpha} \tag{7.247}$$

where $K$ and $\alpha$ are two parameters, and $\alpha$ could be positive or negative. For each value of $\alpha$ we can construct the optimal filter (the overall constant in the filter is irrelevant, as we have seen, so different values of $K$ give the same filter) and, given the noise spectral density $S_n(f)$, eq. (7.239) gives $S/N$ as a function of $K$ and $\alpha$, and therefore tells us what region of this parameter space can be explored, at a given confidence level. For astrophysical backgrounds, more elaborated parametrizations of $h_0^2\Omega_{\rm gw}(f)$ might be necessary at broadband detectors.

### Non-stationary noise

Until now, we have assumed that the noise in the detectors is stationary, and that it can be represented by a fixed function $S_n(f)$. However, such an assumption is not realistic, even more considering that we wish to use a very long observation time, of the order of months. Each detector has periods where it is more quiet and periods where, because of environmental or other disturbances, it is more noisy. Therefore the function $S_n(f)$ changes with time, and we must know how to combine periods in which the detectors had different noise. To study this issue we can subdivide the total observation time $T$ into $n$ intervals of length $T_I$, where $I = 1, \ldots, m$ labels the interval of data, and with $T = \sum_{I=1}^m T_I$. We choose the $T_I$ so that within each interval the noise of the two detectors can be considered stationary. To each of these intervals we can then apply eq. (7.239), so the value of the optimal signal-to-noise ratio from this interval is

$$\left(\frac{S}{N}\right)_I^2 = 2T_I \int_0^{\infty} df\, \Gamma^2(f) \frac{S_h^2(f)}{S_n^2(f;I)}\,. \tag{7.248}$$

Here $S_n(f;I)$ is the total noise spectral density during the $I$-th interval, $S_n^2(f;I) = S_{n,1}(f;I)S_{n,2}(f;I)$, where $S_{n,j}(f;I)$ is the noise spectral density of the $j$-th detector during the $I$-th interval. We now ask how we should combine the $(S/N)_I$ of the different intervals to form the total optimal signal-to-noise ratio. The correct answer can be guessed observing that the optimal $(S/N)_I^2$ is linear in $T_I$, see eq. (7.248) and, in the limit in which the noise is stationary over the whole observation time $T$, we must find that the total optimal signal-to-noise ratio $S/N$ satisfies $(S/N)^2 \sim T = \sum_I T_I$. This fixes uniquely the relation between the total optimal signal-to-noise ratio $S/N$ and the $(S/N)_I$,

$$\left(\frac{S}{N}\right)^2 = \sum_{I=1}^{m} \left(\frac{S}{N}\right)_I^2. \tag{7.249}$$



The same result can also be obtained more formally introducing the observable

$$Y_{\rm tot} = \frac{\sum_I \lambda_I Y_I}{\sum_I \lambda_I} \tag{7.250}$$

(where it is understood that the sums run over $I = 1, \ldots, m$) and choosing the variables $\lambda_I > 0$ so that the signal-to-noise ratio of $Y_{\rm tot}$ is maximized. From eq. (7.225), with $T$ replaced by $T_I$, we see that the $Y_I$ have a mean value

$$S_I \equiv \langle Y_I \rangle = \mu T_I\,, \tag{7.251}$$

where $\mu = \int_0^{\infty} df\, S_h(f)\Gamma(f)\tilde{Q}(f)$ is independent of $I$. For the noise, from eq. (7.232) we have

$$\begin{aligned} N_I^2 &= \frac{T_I}{4}\int_{-\infty}^{\infty} df\, |\tilde{Q}(f)|^2 S_n^2(f;I) \\ &\equiv T_I \sigma_I^2\,. \end{aligned} \tag{7.252}$$

The signal-to-noise ratio $S/N$ of $Y_{\rm tot}$ is obtained by writing

$$S = \langle Y_{\rm tot} \rangle = \mu \frac{\sum_I \lambda_I T_I}{\sum_I \lambda_I}\,, \tag{7.253}$$

and

$$\begin{aligned} N^2 &= \left[\langle Y_{\rm tot}^2 \rangle - \langle Y_{\rm tot} \rangle^2\right]_{\mu=0} \\ &= \frac{\sum_I \lambda_I^2 \sigma_I^2 T_I}{(\sum_I \lambda_I)^2}\,, \end{aligned} \tag{7.254}$$

where we assumed that noise in different intervals are uncorrelated, so $\langle Y_I Y_J \rangle = \delta_{IJ} N_I^2$. Therefore

$$\frac{S^2}{N^2} = \mu^2 \frac{(\sum_I \lambda_I T_I)^2}{\sum_I \lambda_I^2 \sigma_I^2 T_I}\,. \tag{7.255}$$

The maximization of this expression with respect to the $\lambda_I$ can be performed very simply, introducing the positive definite scalar product between two vectors with real components $a_I$ and $b_I$,

$$(a, b) \equiv \sum_I a_I b_I \sigma_I^2 T_I\,. \tag{7.256}$$

Then

$$\frac{S}{N} = \mu \frac{(\lambda_I, \sigma_I^{-2})}{(\lambda_I, \lambda_I)^{1/2}}\,. \tag{7.257}$$

This expression is maximized if the vectors with components $\lambda_I$ and $\sigma_I^{-2}$ are parallel, so $\lambda_I = 1/\sigma_I^2$ (apart from an irrelevant overall constant). Physically, this means that more noisy periods are weighted less. Then the variable $Y_{\rm opt}$, whose signal-to-noise ratio is optimal, is given by

$$Y_{\rm opt} = \frac{\sum_I \sigma_I^{-2} Y_I}{\sum_I \sigma_I^{-2}} \tag{7.258}$$

and the value of the optimal $S/N$ is given by

$$\left(\frac{S}{N}\right)^2 = \mu^2(\sigma_I^{-2}, \sigma_I^{-2}) = \mu^2 \sum_I \frac{T_I}{\sigma_I^2}, \tag{7.259}$$

which, using eqs. (7.251) and (7.252), is equivalent to eq. (7.249), as expected. Equation (7.239) then becomes

$$\frac{S}{N} = \left[2\int_0^\infty df\, \Gamma^2(f) S_h^2(f) \sum_{I=1}^m \frac{T_I}{S_n^2(f;I)}\right]^{1/2}. \tag{7.260}$$

This is equivalent to saying that, in eq. (7.239), we must make the replacement

$$\frac{T}{S_n^2(f)} \to \sum_{I=1}^m \frac{T_I}{S_n^2(f;I)}. \tag{7.261}$$

This way of composing the noise is very natural. It means that noisy periods contribute little to the total signal-to-noise ratio. If we perform the same order-of-magnitude estimate as in eq. (7.245), we conclude that

$$\frac{1}{[\Omega_{gw}(f)]^2_{\min}} = \sum_{I=1}^m \frac{1}{[\Omega_{gw}(f;I)]^2_{\min}}, \tag{7.262}$$

where $[\Omega_{gw}(f;I)]_{\min}$ is the minimum value of $\Omega_{gw}$ detectable using only the data in the $I$-th interval, and $[\Omega_{gw}(f)]_{\min}$ is the minimum value of $\Omega_{gw}$ detectable combining the $n$ intervals.

### How the background is actually measured

We can now give an example of an operative way of measuring the stochastic background. First of all, one divides the total observation time $T$ into intervals of length $T_I$, such that within each interval the detector noise is stationary. This scale is chosen based on observations of the detector noise variation, and could typically be of order of one to a few minutes. Within each interval, the spectral density $S_n(f;I)$ can be considered constant in time, and is determined experimentally. We can now compute the filter function, using the measured value of $S_n(f;I)$ and assuming a given form for $\Omega_{gw}$. For instance, $\Omega_{gw}$ = const. can be the simplest choice, or one can use the parametrization (7.247) and repeat the procedure for various values of $\alpha$.

To have an experimental determination of $S_I = \langle Y_I\rangle$ and of $N_I = [\langle Y_I^2\rangle - \langle Y_I\rangle^2]^{1/2}$ one further divides each interval into segments of length $\Delta t$, labeled by an index $J = 1,\ldots,n$, and with $T_I = n\Delta t$ (with $\Delta t$ much larger than the light travel time between the detectors, which for the two LIGO observatories is about 10 ms). The signal $Y_{IJ}$ relative to the $J$-th segment of the $I$-th interval is computed using eq. (7.220), with the time integration running only over the $J$-th segment of the $I$-th interval.[72] Observe that the filter function $Q(t-t')$ typically vanishes very fast for $|t-t'|$ larger than a few tens of ms, so in practice if $t$ belongs to the $J$-th interval, the support of $Q(t-t')$ is entirely contained in the $J$-th interval.

[72]In practice, it can be more convenient to perform a FFT over the segment and use the frequency space expression (7.223).

From the set of $Y_{IJ}$ at fixed $I$, one can construct the sample mean

$$S_I = \frac{1}{n}\sum_{J=1}^n Y_{IJ}, \tag{7.263}$$

and the sample variance

$$N_I^2 = \frac{1}{n-1}\sum_{J=1}^n (Y_{IJ} - S_I)^2, \tag{7.264}$$

of the $I$-th interval. We repeat this procedure for all intervals and, according to eq. (7.249), the total signal-to-noise ratio is

$$\left(\frac{S}{N}\right)^2 = \sum_{I=1}^m \left(\frac{S_I^2}{N_I^2}\right). \tag{7.265}$$

If this $S/N$ exceeds a predetermined threshold value one can state that a stochastic background is detected, with a confidence level which depends on the threshold used.[73]

[73]A subtle point is that it can be shown that, if we wait long enough, i.e. if the total observation time is sufficiently large, any predetermined fixed threshold will be exceeded. In other words, in the limit $T \to \infty$ the false alarm probability is 100%! To have a finite false alarm probability even in the limit $T \to \infty$, the value of the threshold must increase with the number of intervals $n$ faster than $\log\log n$.

### Multiple-detector correlation

Another interesting question is what happens if we correlate the outputs of $N$ detectors, with $N > 2$. For simplicity, we assume at first that we have $N$ identical detectors, with the same noise spectral density $S_n(f)$, and all running simultaneously for a time $T$.

With $N$ detectors we can form $N(N-1)/2$ independent two-point correlators

$$Y_{ij} = \int_{-T/2}^{T/2} dt \int_{-T/2}^{T/2} dt'\, s_i(t) s_j(t') Q(t-t'), \tag{7.266}$$

with $i < j$. (If the detectors have different noise spectral densities, then also the filter function depends on $i, j$, and we write it $Q_{ij}(t-t')$.) Conceptually, for a stationary stochastic background, there is no difference between the situation in which $N(N-1)/2$ identical pairs of detectors run for a time $T$, and the situation in which a single pair of detectors runs for a time $T_{total} = T \times N(N-1)/2$. In the former case, sampling the output of the detectors at times $t_k$, with $k = 1,\ldots,k_{\max}$, we get a set of values $Y_{ij}(t_k)$, for each of the $N(N-1)/2$ pairs $(i,j)$. In the latter case we directly get a set of values $Y(t_k)$ for the single pair considered, with $k$ taking values up to $k_{\max} \times N(N-1)/2$. In both cases we must then compute the average of $Y$ over all these values, so the result is the same and the difference is just a matter of notation. In conclusion, the signal-to-noise ratio with $N$ identical detectors can be obtained from eq. (7.239) making the replacement

$$T \to \frac{N(N-1)}{2} T. \tag{7.267}$$

If we denote by $[\Omega_{gw}]_{\min,N}$ the minimum value of $\Omega_{gw}$ measurable with $N$ identical detectors and by $[\Omega_{gw}]_{\min,2}$ the minimum value of $\Omega_{gw}$ detectable with two detectors, then

$$[\Omega_{gw}]_{\min,N} = \left[\frac{2}{N(N-1)}\right]^{1/2} [\Omega_{gw}]_{\min,2}\,. \tag{7.268}$$

In the more realistic case in which the detectors have different noise spectral densities, or have different common time of operation, the situation is formally identical to the case of non-stationary noise discussed above, where the observations taken during a time $T_{ij}$ by each pair of detectors $(i,j)$, with $i<j$, plays the role of the observations taken during the time intervals labeled by $I$ in eqs. (7.260) to (7.262). Therefore, the signal-to-noise ratio is obtained from eq. (7.239) with the replacement

$$\frac{T}{S_n^2(f)} \to \sum_{i<j} \frac{T_{ij}}{S_n^2(f;\langle ij\rangle)}\,, \tag{7.269}$$

where $T_{ij}$ is the common time of operation of the detectors $i$ and $j$, and $S_n^2(f;\langle ij\rangle) = S_n(f;i)S_n(f;j)$ is the product of the spectral densities of the $i$-th and $j$-th detector. The order-of-magnitude estimate of the minimum detectable value of $\Omega_{gw}$, eq. (7.262), becomes

$$\frac{1}{[\Omega_{gw}(f)]^2_{\min,N}} = \sum_{i<j} \frac{1}{[\Omega_{gw}(f;\langle ij\rangle)]^2_{\min}}\,. \tag{7.270}$$

When all detectors are equal and have the same common time of operation, $[\Omega_{gw}(f;\langle ij\rangle)]_{\min}$ becomes independent of the pair $i,j$ considered, and is the quantity that we denoted by $[\Omega_{gw}]_{\min,2}$, so we recover eq. (7.268).

In a sense, this result is disappointing. We have seen in eq. (7.245) that, passing from a single detector to a two-detector correlation, we gain a factor $1/(2T\Delta f)^{1/2}$ in the minimum detectable value of $\Omega_{gw}$. For $T=1$ yr and $f=100$ Hz, this means and improvement by a factor $10^5$ in sensitivity. Passing from $N=2$ to $N=3$ detectors, instead, we see from eq. (7.268) that we gain only a further factor $\sqrt{3}$.

This is very different from the situation for bursts discussed in Section 7.5.3. In the case of bursts, the noise that compete with the signal consists of large, relatively rare fluctuations. At any given moment the probability that, in a single detector and within a given time window, say of order few tens of ms, a fluctuation with a signal-to-noise ratio above a large threshold takes place, is a small number $\epsilon \ll 1$. The probability that a second detector has a simultaneous independent fluctuation above this threshold, within the same window, is $O(\epsilon^2)$, the probability of a three-detector coincidence is $O(\epsilon^3)$, etc. Then, for bursts, the gain in statistical significance passing from a single detector to a two-detector coincidence is that same as the gain passing from a two-detector to a three-detector coincidence. The crucial point is that for bursts, after matched filtering, we are left with short events with a large value of $S/N$, which are rare.



In contrast, for stochastic backgrounds we are never confronted with rare events. At any given moment the GW stochastic signal is always much below the noise, and is never responsible for large fluctuations of the output. There are no rare events to be searched in coincidence, and the only advantage of using more detector pairs is that the total amount of data available increases, which means that we have a longer effective observation time.

The situation does not change substantially if, rather than two-point correlators, we consider $M$-point correlators, with $M$ smaller than or equal the number of detectors $N$. For instance, with four detectors we can consider a four-point correlator $\langle s_1(f)s_2(f)s_3(f)s_4(f)\rangle$. Repeating the same steps as above, one finds again that the signal-to-noise ratio (always defined to be quadratic in the GW signal, in order to compare with the same quantity as in two-detector case) scales as $\sqrt{T}$.

On the other hand, an advantage of multiple-detector correlations is that it might be easier to suppress correlated environmental noise, especially if the various detectors are not close to each other.

### Correlated noise and signal chopping

Equation (7.239) shows that a true GW signal has a signature that in principle could allow us to distinguish it from the noise: increasing the observation time, the signal-to-noise ratio in the presence of a real GW signal must increase as $T^{1/2}$.

Actually, this is a signature that only allows us to distinguish a stochastic GW background from uncorrelated noise in the two detectors. Unfortunately, any residual correlated noise would still mimic the behavior of a real GW signal. The problem is therefore how to make sure that correlated noise are negligible, and this can be a hard task, particularly for very long integration times. If two detectors are at the same site, or very close, their overlap reduction function is maximized, but we will certainly have correlated environmental noise. We have seen that the overlap reduction function suppresses the GW correlation if the detector separation is $\Delta x \gg \lambda\!\!\!^{-}$. For instance, at $f=50$ Hz, $\lambda\!\!\!^{-} \simeq 1000$ km. Most environmental disturbances will decorrelate on a much shorter length-scale, so it is possible that two detectors at a suitable distance are still correlated as far as the GW background is concerned, but they have negligible correlated noise. However, beyond a given sensitivity level, seismic noise or propagating electromagnetic disturbances might still give important correlated noise, and this is a difficult issue that will have to be carefully studied experimentally.

An interesting option offered by the two-detector correlation is the possibility of *chopping* the signal. Chopping is a general term for measurements in which we switch our detector between the quantity that we want to measure and a reference quantity. It is a very powerful experimental technique, that exploits the fact that in many situations one can measure with a much better precision the variation of a quantity rather than the quantity itself because, taking the difference, many uncertainties, e.g. calibration uncertainties, cancel out.[74] In particular, one can compare the measurement in a situation where the signal is expected, to the situation where a null answer should come out.

At first sight, it appears that a measurement of this type is impossible for a stochastic backgrounds of GWs, since the background is always there, and gravitational forces cannot be screened. It seems therefore impossible to compare the output of a detector when no stochastic GW background acts on it, with the output when the background is acting on it. Remarkably, this is no longer true when we consider a two-detector correlation. In fact, changing the relative orientation of the two detectors, the factor $\sum_A F_1^A(\hat{\mathbf{n}})F_2^A(\hat{\mathbf{n}})$ in eq. (7.226) changes, and it is therefore possible to modulate the signal. To illustrate this point, we compute $F_{12}$ for a bar-interferometer correlation. Using Table 7.1 and eqs. (7.31) and (7.32) we see that, for $\psi$ generic, the pattern functions of an interferometer are

$$F_+^{(\text{intf})}(\theta,\phi;,\psi) = \frac{1}{2}(1+\cos^2\theta)\cos 2\phi\cos 2\psi - \cos\theta\sin 2\phi\sin 2\psi\,,$$
$$F_\times^{(\text{intf})}(\theta,\phi;\psi) = \frac{1}{2}(1+\cos^2\theta)\cos 2\phi\sin 2\psi + \cos\theta\sin 2\phi\cos 2\psi\,. \tag{7.271}$$

The pattern functions of the bar for $\psi$ generic can also be obtained from Table 7.1 and eqs. (7.31) and (7.32). We must however pay attention to the fact that in Table 7.1, the variable denoted by $\theta$ for resonant bars is the angle measured from its longitudinal axis, while for an interferometer with arms along the $x$ and $y$ axes, we denoted by $\theta$ the polar angle measured from the $z$ axis, so these two angles are not the same unless the bar is vertical. If instead the bar lies in the $x,y$ plane, at an angle $\alpha$ with the $y$ axis, and we denote by $\theta$ the polar angles measured from the $z$ axis, then the pattern functions of the bar become

$$\begin{aligned} F_+^{(\text{bar})}(\theta,\phi;,\psi) &= [-\cos^2\theta\cos^2(\phi-\alpha)+\sin^2(\phi-\alpha)]\cos 2\psi \\ &\quad +[\cos\theta\sin 2(\phi-\alpha)]\sin 2\psi \\ F_\times^{(\text{bar})}(\theta,\phi;,\psi) &= [-\cos^2\theta\cos^2(\phi-\alpha)+\sin^2(\phi-\alpha)]\sin 2\psi \\ &\quad -[\cos\theta\sin 2(\phi-\alpha)]\cos 2\psi\,. \end{aligned} \tag{7.272}$$

From this it follows that

$$\int\frac{d\hat{\mathbf{n}}}{4\pi}\frac{d\psi}{2\pi}\sum_A F_A^{(\text{bar})}F_A^{(\text{intf})} = -\frac{2}{5}\cos 2\alpha\,. \tag{7.273}$$

(The overall sign of $F_{12}$ is irrelevant since $\Gamma(f)$ enters quadratically in the signal-to-noise ratio.) We see that the correlation is maximum when the bar is aligned with one of the interferometer arms (i.e. when $\alpha=0$ or $\alpha=\pi/2$). In contrast, when $\alpha=\pi/4$ we have $F_{12}=0$. Therefore in this configuration the signal obtained from the interferometer-bar correlation vanishes. Even if GWs cannot be screened, the "composite detector" whose output is the correlation between a bar and an interferometer can be set in the "off source" position! We can then compare the result in this configuration with the result when the resonant bar is parallel to one of the interferometer arms, which is the position that maximizes the correlation. This chopping strategy has been used in the LIGO–ALLEGRO correlation. The ALLEGRO resonant bar (which has now terminated its activity) was located relatively close to the LIGO observatory in Livingston, and was mounted on a platform that allowed to rotate it easily. (After a rotation, data taking of good quality resumed in just half an hour.) The bar was therefore taken for a few months in the "off source" position, and then rotated to the "on source" position for a few more months.

[74] The classical example of this technique was the Dicke radiometer, which was developed by Dicke during World War II for application to microwave radars, and measured the radiation temperature of a radio source (i.e. the temperature of a black body having the same radio brightness). A direct measurement was difficult: the signal needed a large amplification, and fluctuations in the amplifier gain resulted in large errors. To overcome this difficulty, in the Dicke radiometer the receiver switches quickly between the source and a carefully calibrated black body, whose temperature was chosen to be of the order of the value expected for the source. To tell when these temperatures were equal was much easier than to obtain a direct determination of the source temperature. The same principle of comparing with a reference black body was used by the FIRAS spectrometer on board of the COBE satellite to measure the black-body spectrum of CMB. To measure the CMB anisotropies, i.e. the variation of the black-body temperature over the sky, the principle used by the DMR detector on COBE and by the subsequent high-precision experiments such as WMAP is to compare the temperatures between two points in the sky.



## Further reading

- For a textbook discussion of matched filtering and of detection of signals in noise see Wainstein and Zubakov (1962) and McDonough and Whalen (1995). For matched filtering and optimal signal-to-noise ratio for GW bursts see Thorne (1987), and Saulson (1994), Chapter 4. Statistical aspects of parameter estimation are discussed in Finn (1992) and in Cutler and Flanagan (1994), where the multiple detector case is also treated. For a review of data analysis for interferometric GW detector see Viceré (2000).
- Books on probability and statistics typically cover many shelves in any physics library, and recommendation are very much subjective. For an elementary but very practical introduction to statistics (tuned to the needs of particle physicists, but quite useful also in the GW context), see Lyons (1986). A concise and useful summary is given in the sections on probability and statistics of the Review of Particle Properties, in Yao *et al.* [Particle Data Group] (2006). A very nice discussion of Bayesian vs. frequentist method, in the context of particle physics, is given in Cousins (1995). A discussion of the frequentist vs. Bayesian approach in the GW context is given in appendix A of Cutler and Flanagan (1994).
- The analysis of bursts of unknown shape using band-pass filtering is discussed in Flanagan and Hughes (1998a, 1998b), in the context of the merging phase of black hole binaries. Time–frequency techniques are further discussed in Anderson and Balasubramanian (1999), Anderson, Brady, Creighton and Flanagan (2001) and Viceré (2002). An algorithm based on clusters of pixels in the time–frequency domain (termed TFCLUSTERS) is presented in Sylvestre (2002). A book on the use of wavelets in physics is van den Berg (1999). Application of wavelets to the analysis of GW bursts can be found in Klimenko, Yakushin, Rakhmanov and Mitselmakher (2004) and Klimenko and Mitselmakher (2004) (the WaveBurst algorithm).
- Some sources, such as accreting neutron or quark stars, as well as neutron stars stressed by large interior magnetic fields (magnetars), could emit repeatedly small bursts of GWs, with very characteristic correlations, both in energy and in time, among the different bursts, typical of systems displaying self-organized criticality. These correlations could give a further handle in their data analysis. These "GW bursters" are discussed in Coccia, Dubath and Maggiore (2004) and Dubath, Foffa, Gasparini, Maggiore and Sturani (2005).
- The search strategy for GW bursts using the three LIGO interferometers is discussed in Abbott *et al.* [LSC] (2004b). The sensitivity of a network of interferometers for reconstructing the source position is studied in Gürsel and Tinto (1989). Searches for GW bursts using coincidences between up to five resonant bars are described in Astone *et al.* [IGEC] (2003a). Results with correlations among three bars, with improved sensitivities, are reported in Astone *et al.* [IGEC2] (2007).
- Introductory discussions of the search strategies for periodic signals can be found in Saulson (1994), Section 14.6 and Schutz (1991). More detailed analysis are given in Brady, Creighton, Cutler and

Schutz (1998) and in Brady and Creighton (2000). The application of the Hough transform to periodic GWs is discussed in Krishnan *et al.* (2004). A search for periodic GWs from a single specific source, using the LIGO and GEO detectors, is described in Abbott *et al.* [LSC] (2004a). Limits on 28 isolated pulsar using the LIGO S2 run are given in Abbott *et al.* [LSC] (2005b).

- The importance of post-Newtonian corrections for the data analysis of coalescing binaries is emphasized in Cutler *et al.* (1993). Detailed discussions of data analysis procedure and parameter extraction for coalescences is given in Cutler and Flanagan (1994), Poisson and Will (1995), Królak, Kokkotas and Schäfer (1995) and Flanagan and Hughes (1998a). For computations of the waveform with the PN formalism, see the Further Reading section in Chapter 5.
- Optimal template placement for inspiraling compact binaries is discussed in Owen (1996) and Owen and Sathyaprakash (1999). A comparison of templates for binary inspiral is given in Damour, Iyer and Sathyaprakash (2001). A particularly useful family of templates for BH-BH inspiral have been proposed by Buonanno, Chen and Vallisneri (2003). A description of the LIGO search strategy for coalescences can be found in Abbott *et al.* [LSC] (2005a).
- The optimal SNR in a two-detector correlation and the overlap reduction function are discussed in Michelson (1987), Christensen (1992) and Flanagan (1993). A detailed discussion of signal processing strategies for stochastic backgrounds of GWs is given in Allen and Romano (1999). Signal chopping is discussed in Finn and Lazzarini (2001). Stochastic backgrounds of GWs are reviewed in Maggiore (2000). The search strategy of LIGO for stochastic backgrounds of GWs is discussed in Abbott *et al.* [LSC] (2004d) and (2005c).

# 8 Resonant-mass detectors



The history of experimental GW physics began with resonant-mass detectors. The pioneer was Joseph Weber who, in the 1960s, developed the concept and built the first resonant bars. In the course of the subsequent four decades, resonant-mass detectors operated by various groups have reached sensitivities better than Weber's original bars by about four orders of magnitudes in energy. Still, we will see in this chapter that these sensitivities could allow the detection of only relatively strong signals in our Galaxy or at most in our immediate galactic neighborhood, which are expected to be rare. To gain access to sources at large extragalactic distances it is necessary to build large interferometers, which will be the subject of the next chapter.

The passage from resonant detectors to interferometers implies a jump from "small-scale" experiments, performed by groups which can be as small as half a dozen people, to "Big Science", with collaborations of hundreds of people and financial costs which are higher by factors $O(10^2-10^3)$. As we will see in the next chapter, such a jump is justified by the formidable discovery potential of interferometers and especially advanced interferometers. We nevertheless begin our discussion of experiments with resonant-mass detectors, both because they still have the possibility of detecting rare or unexpected events, and also because their study is instructive in itself. Our emphasis will be on aspects that have an intrinsic conceptual interest, such as understanding how a GW interacts with a macroscopic piece of matter, and on how it is possible to detect vibrations of a macroscopic body which are incredibly small, with amplitude many orders of magnitude smaller than the size of a nucleus. We will see that, by themselves, resonant detectors are remarkable instruments; it is possible to measure vibrations in a two-ton object, such as a typical bar, which corresponds to just a few tens of phonons, and variations $\Delta L$ of their length $L$, with $\Delta L/L \sim 10^{-19}-10^{-18}$.

## 8.1 The interaction of GWs with an elastic body

### 8.1.1 The response to bursts

A typical bar is a cylinder of length $L \simeq 3$ m and radius $R \simeq 30$ cm, so in a first approximation we can treat its vibrations as one-dimensional. We orient the bar along the $x$ axis, with the end-faces at $\pm L/2$, and we study the dynamics of a volume element $dV$ of the bar originally located

# 9 Interferometers



The idea of interferometric detection of GWs is in principle simple and elegant, and goes back to 1962, when it was first considered by two Russian theorists, M. Gertsenshtein and V. I. Pustovoit. Weber also considered it, and it was then pushed in the late 1960s by R. Forward, R. Weiss, R. Drever, and others. In practice, however, a large GW interferometer is an extremely complex instrument, with many degrees of freedom that must be kept under control with extraordinary accuracy. Thus, their development up to the present scale has required the building up of large collaborations, comparable in size to modern particle physics experiments, as well as more than 30 years of preparation. Following the general approach of this book, as outlined in the Preface, we will not discuss the interesting history of the development of this idea, referring the reader to the Further Reading section for reviews, and we will rather focus on the present understanding of these detectors. We will begin in Section 9.1 with the most naive setting, a simple Michelson interferometer, and we will then add up successive layers of complexity in Sections 9.2 and 9.3. Having defined the experimental set-up, we will be able to discuss the principal noise sources in Section 9.4. The existing detectors (LIGO, VIRGO, GEO600 and TAMA) are discussed in Section 9.5.1 while advanced ground-based detectors, as well as the space-borne alternative, are discussed in Section 9.5.2.

## 9.1 A simple Michelson interferometer

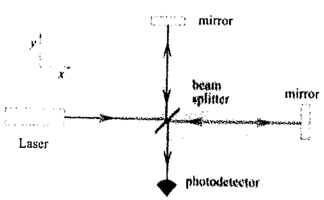


Fig. 9.1 The layout of a simple Michelson-type interferometer.

A Michelson interferometer, of the type used in the classical Michelson–Morley experiment in 1887 to show the non-existence of the ether, is an extraordinarily accurate instrument for measuring changes in the travel time of light in its arms. The simplest conceptual scheme (which is not exactly the one used historically by Michelson and Morley) is shown in Fig. 9.1. It consists of a monochromatic light source, which today is of course a laser, whose light is sent on a beam-splitter which separates the light, with equal probability amplitudes, into a beam traveling in one arms and a beam traveling in a second, orthogonal, arm. At the end of each arm we put totally reflecting mirrors. After traveling once back and forth, the two beams recombine at the beam-splitter, and part of the resulting beam goes to a photodetector, that measures its intensity (while a part goes back toward the laser). We denote by $\omega_L$ the frequency of the laser (the subscript L distinguishes it from the frequency $\omega_{gw}$ of the GWs that we want to detect), so $k_L = \omega_L/c$ and $\lambda_L = 2\pi/k_L$ are



the wavenumber and the wavelength of the laser light. It is convenient to use a complex notation for the electromagnetic field. Thus, a given spatial component of the electric field of the input laser light is written as

$$E_0 e^{-i\omega_L t + i\mathbf{k}_L\cdot\mathbf{x}}\,. \tag{9.1}$$

We denote by $L_x$ and $L_y$ the length of the two arms, where we have oriented the $x$ and $y$ axis as shown in Fig. 9.1. Consider a photon that arrives at the beam-splitter, coming from the laser, at some initial time $t_0$.[1] The part of the electric field that goes into the $x$ arm bounces on the mirror at a distance $L_x$ and arrives back at the beam-splitter at a time $t = t_0 + 2L_x/c$, while the part that went through the $y$ arm comes back at the beam-splitter at $t' = t_0 + 2L_y/c$. Thus, the beam that finally recombines at the beam-splitter at a given observation time $t$ is the superposition of a beam that entered the beam-splitter at a time $t_0^{(x)} = t - 2L_x/c$, and then went through the $x$ arm, and a beam that entered the beam-splitter at a *different* time $t_0^{(y)} = t - 2L_y/c$, and then went through the $y$ arm. Setting the beam-splitter at $\mathbf{x} = 0$, the former beam has an initial phase $\exp\{-i\omega_L t_0^{(x)}\} = \exp\{-i\omega_L t + 2ik_L L_x\}$, and the latter $\exp\{-i\omega_L t_0^{(y)}\} = \exp\{-i\omega_L t + 2ik_L L_y\}$. The phase of the field is conserved during the free propagation, while the fields acquire overall factors from reflections and transmission at the mirrors.[2] So, the two electric fields that recombine at time $t$ at the beam-splitter are given by

$$E_1 = -\frac{1}{2}E_0 e^{-i\omega_L t + 2ik_L L_x}\,. \tag{9.2}$$

and

$$E_2 = +\frac{1}{2}E_0 e^{-i\omega_L t + 2ik_L L_y}\,. \tag{9.3}$$

The total electric field is $E_{out} = E_1 + E_2$. Writing $2L_x = (L_x + L_y) + (L_x - L_y)$ and $2L_x = (L_x + L_y) - (L_x - L_y)$, we see that

$$E_{out} = -iE_0 e^{-i\omega_L t + ik_L(L_x+L_y)} \sin[k_L(L_y - L_x)]\,, \tag{9.4}$$

and the power measured by the photodetector is proportional to

$$|E_{out}|^2 = E_0^2 \sin^2[k_L(L_y - L_x)]\,. \tag{9.5}$$

Therefore any variation in the length of a arm results in a corresponding variation of the power at the photodetector. We now discuss how to apply this general idea to GW detection. We saw in Section 1.3.3 that the interaction of a GW with a detector can be described in two different languages, i.e. either using the TT frame, or using the proper detector frame. It is quite instructive to understand the functioning of an interferometer in both ways, as we do in the next two subsections.

### 9.1.1 The interaction with GWs in the TT gauge

Recall from Section 1.3.3 that, in the TT gauge, the coordinates are marked by the position of freely falling objects so, even when a GW

[1] Observe that, until we discuss shot noise, in Section 9.4.1, there is really no need to introduce photons, and the whole discussion could be done purely classically, replacing the word "photon" by "wave-packet".

[2] As we will discuss in a more general setting in Section 9.2.1, the reflection off a 50–50 beam splitter can be modeled multiplying the amplitude of the incoming electric field by a factor $r = +1/\sqrt{2}$ for reflection from one side and $r = -1/\sqrt{2}$ for reflection from the other side, while the transmission multiplies it by $t = 1/\sqrt{2}$, and reflection at the perfectly reflecting mirrors at the end of each arm multiplies the amplitude by $-1$. Thus, overall one beam acquires a factor $(1/\sqrt{2}) \times (-1) \times (1/\sqrt{2}) = -1/2$ and the other a factor $+1/2$.

is passing, the coordinates of freely falling masses by definition do not change. Of course, the mirrors of a ground-based interferometer are not freely falling; rather, the Earth's gravity is compensated by the suspensions. However, as we already discussed in Section 1.3.3, these forces are static, compared to the frequency of the GWs that we are searching and, as far as the motion in the horizontal plane is concerned, the mirrors can be taken to be in free fall, i.e. they follow the geodesics of the time-dependent part of the gravitational field.[3]

Thus, in the TT gauge description, the coordinates of the mirrors and of the beam-splitter are not affected by the passage of the wave. We define the origin of the coordinate system as the location of the beam-splitter, while the position of the mirror which terminates the $x$ arm defines the point with coordinates $(L_x, 0)$, and the position of the other mirror defines the point with coordinates $(0, L_y)$, and this remains true also when a GW is present.

In the TT gauge description, the physical effect of the GW is manifested in the fact that it affects the propagation of light between these fixed points. We assume for the moment that the GW has only the plus polarization, and comes from the $z$ direction.[4] In the $z = 0$ plane of the interferometer we therefore have

$$h_+(t) = h_0 \cos \omega_{\rm gw} t\,, \tag{9.6}$$

and the space-time interval in the TT frame is given by

$$ds^2 = -c^2 dt^2 + [1 + h_+(t)]dx^2 + [1 - h_+(t)]dy^2 + dz^2\,. \tag{9.7}$$

Photons travels along null geodesics, $ds^2 = 0$, so for the light in the $x$ arm we have, to first order in $h_0$,

$$dx = \pm\, c dt \left[1 - \frac{1}{2} h_+(t)\right]\,, \tag{9.8}$$

where the plus sign holds for the travel from the beam-splitter to the mirror and the minus sign for the return trip. Consider a photon that leaves the beam-splitter at a time $t_0$. It reaches the mirror, at the fixed coordinate $x = L_x$, at a time $t_1$ obtained integrating eq. (9.8) with the plus sign,

$$L_x = c(t_1 - t_0) - \frac{c}{2}\int_{t_0}^{t_1} dt'\, h_+(t')\,. \tag{9.9}$$

Then the photon is reflected and reaches again the beam-splitter at a time $t_2$ obtained integrating eq. (9.8) with the minus sign, between $x = L_x$ and $x = 0$,

$$\int_{L_x}^{0} dx = -c \int_{t_1}^{t_2} dt' \left[1 - \frac{1}{2} h_+(t')\right]\,, \tag{9.10}$$

i.e.

$$L_x = c(t_2 - t_1) - \frac{c}{2}\int_{t_1}^{t_2} dt'\, h_+(t')\,. \tag{9.11}$$

[3]Of course, there will also be some non-static forces, such as those due to suspension thermal noise or, more generally, to the coupling with the environment, which will provide the background noise, and that will be discussed in Section 9.4.

[4]The response to GWs with arbitrary direction and polarization will be studied in Section 9.2.3.



Summing eqs. (9.9) and (9.11) we get

$$t_2 - t_0 = \frac{2L_x}{c} + \frac{1}{2}\int_{t_0}^{t_2} dt'\, h_+(t')\,. \tag{9.12}$$

For a given value of $t_0$, the time of arrival $t_2$ after a round trip in the $x$ arm is therefore $t_0 + 2L_x/c$, plus a correction of order $h_0$. In the upper limit of the integral on the right-hand side we can replace $t_2$ by $t_0 + 2L_x/c$, since the integrand is already $O(h_0)$ and we are anyway neglecting terms $O(h_0^2)$, so we get

$$\begin{aligned} t_2 - t_0 &= \frac{2L_x}{c} + \frac{1}{2}\int_{t_0}^{t_0+2L_x/c} dt'\, h_0 \cos(\omega_{\rm gw} t') \\ &= \frac{2L_x}{c} + \frac{h_0}{2\omega_{\rm gw}} \left\{\sin[\omega_{\rm gw}(t_0 + 2L_x/c)] - \sin \omega_{\rm gw} t_0\right\}. \end{aligned} \tag{9.13}$$

Using the identity $\sin(\alpha+2\beta) - \sin\alpha = 2\sin\beta\cos(\alpha+\beta)$, we can rewrite this as

$$t_2 - t_0 = \frac{2L_x}{c} + \frac{h_0 L_x}{c}\frac{\sin(\omega_{\rm gw} L_x/c)}{(\omega_{\rm gw} L_x/c)}\cos[\omega_{\rm gw}(t_0 + L_x/c)]\,. \tag{9.14}$$

Observe that the difference $t_2 - t_0$ is a function of the time $t_0$ at which the photon left the beam-splitter, because of the term $\cos[\omega_{\rm gw}(t_0 + L_x/c)]$. Using eq. (9.6), we can also rewrite the above result as

$$\boxed{t_2 - t_0 = \frac{2L_x}{c} + \frac{L_x}{c}\, h(t_0 + L_x/c)\frac{\sin(\omega_{\rm gw} L_x/c)}{(\omega_{\rm gw} L_x/c)}\,.} \tag{9.15}$$

The quantity $t_0 + L_x/c$ which appears in the argument of $h(t)$ is, to zeroth order in $h_0$, the value of time $t_1$ at which the photon touches the far mirror on the $x$ arm. This result will be easily understood physically in the next subsection, thanks to the Newtonian intuition that we can use in the proper detector frame. The function

$$\text{sinc}\left(\frac{\omega_{\rm gw} L}{c}\right) \equiv \frac{\sin(\omega_{\rm gw} L/c)}{(\omega_{\rm gw} L/c)} \tag{9.16}$$

goes to one when $\omega_{\rm gw} L/c \to 0$. Therefore, when the period of the GW is large compared to $L_x/c$, the shift $\Delta t$ in the travel time $t_2 - t_0$, with respect to the unperturbed value $2L_x/c$, is simply $h(t_1)L_x/c$. If $\omega_{\rm gw} L_x/c \gg 1$, $\Delta t$ is suppressed. This is clearly understood physically: if $\omega_{\rm gw} L_x/c \gg 1$, during the travel time of the photon $h(t)$ changes sign many times, so it contributes sometimes positively and sometimes negatively to $\Delta t$, and these contributions partially cancel out. A plot of the function sinc$(x)$ is shown in Fig. 9.2.

In the $y$ arm the analysis is similar, but now the sign of $h(t)$ is reversed, as we see from eq. (9.7), so we now have

$$t_2 - t_0 = \frac{2L_y}{c} - \frac{L_y}{c}\, h(t_0 + L_y/c)\frac{\sin(\omega_{\rm gw} L_y/c)}{(\omega_{\rm gw} L_y/c)}\,. \tag{9.17}$$

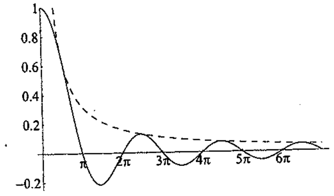


**Fig. 9.2** The function sinc$(x)$ = $(1/x)\sin x$ (solid line) and, for comparison, the function $1/x$ (dashed line).

In practice we will be interested in the light that comes out of the beam-splitter at a given value of the observation time $t$, so it is useful to rewrite these relations holding fixed the value of the time $t_2 \equiv t$ at which we observe the light that has recombined at the beam-splitter, and computing the corresponding value of $t_0$. In order to come back at the beam-splitter at time $t$, the light that went through the $x$ arm must have started its round-trip travel at a time $t_0^{(x)}$ obtained inverting eq. (9.15) to first order in $h_0$, which means that $h(t_0 + L_x/c)$ is replaced by $h(t - 2L_x/c + L_x/c) = h(t - L_x/c)$, so

$$t_0^{(x)} = t - \frac{2L_x}{c} - \frac{L_x}{c} h(t - L_x/c)\,\mathrm{sinc}\,(\omega_{\rm gw} L_x/c)\,, \tag{9.18}$$

and similarly the light that went through the $y$ arm, in order to arrive back at the beam-splitter at the same time $t$, must have started its round-trip travel at a different time $t_0^{(y)}$ given by

$$t_0^{(y)} = t - \frac{2L_y}{c} + \frac{L_y}{c} h(t - L_y/c)\,\mathrm{sinc}\,(\omega_{\rm gw} L_y/c)\,. \tag{9.19}$$

Again, we use the fact that the phase of the field is conserved during the free propagation. Setting the origin of the coordinate system at the beam-splitter, and writing the electric field of the light as in eq. (9.1), we see that the light that is at the beam splitter ($\mathbf{x} = 0$) at time $t_0^{(x)}$ has a phase $\exp\{-i\omega_{\rm L} t_0^{(x)}\}$. The free propagation along the arm does not change this phase, while reflections and transmission at the mirrors give an overall factor $\pm 1/2$, see Note 2 on page 471, so[5]

$$\begin{aligned} E^{(x)}(t) &= -\frac{1}{2} E_0 e^{-i\omega_{\rm L} t_0^{(x)}} \\ &= -\frac{1}{2} E_0 e^{-i\omega_{\rm L}(t - 2L_x/c) + i\Delta\phi_x(t)}\,, \end{aligned} \tag{9.20}$$

where

$$\Delta\phi_x(t) = h_0 \frac{\omega_{\rm L} L_x}{c} \mathrm{sinc}\,(\omega_{\rm gw} L_x/c)\, \cos[\omega_{\rm gw}(t - L_x/c)]\,, \tag{9.21}$$

and similarly the field that went through the $y$ arms, at time $t$ has the form

$$\begin{aligned} E^{(y)}(t) &= +\frac{1}{2} E_0 e^{-i\omega_{\rm L} t_0^{(y)}} \\ &= +\frac{1}{2} E_0 e^{-i\omega_{\rm L}(t - 2L_y/c) + i\Delta\phi_y(t)}\,, \end{aligned} \tag{9.22}$$

where

$$\Delta\phi_y(t) = -h_0 \frac{\omega_{\rm L} L_y}{c} \mathrm{sinc}\,(\omega_{\rm gw} L_y/c)\, \cos[\omega_{\rm gw}(t - L_y/c)]\,. \tag{9.23}$$

In general, $L_x$ and $L_y$ will be made as close as possible,[6] in order to cancel many common noise in the two arms. Thus, in $\Delta\phi_x$ and $\Delta\phi_y$, which are already of order $h_0$, we simply replace $L_x$ and $L_y$ by $L =$

[5] The superscript $(x)$ on $E^{(x)}(t)$ reminds us that this is the electric field of the light that went through the $x$-arm, and should not be confused with the $x$-component of the electric field vector. Here we are considering a given spatial component of the electric field.

[6] Except for a small asymmetry, the Schnupp asymmetry, that we will discuss in Section 9.3.2.



$(L_x + L_y)/2$, while in the terms $t - 2L_x/c$ and $t - 2L_y/c$ we still take into account any small difference between $L_x$ and $L_y$, writing $2L_x = 2L + (L_x - L_y)$ and $2L_y = 2L - (L_x - L_y)$. Then

$$E^{(x)}(t) = -\frac{1}{2} E_0 e^{-i\omega_{\rm L}(t - 2L/c) + i\phi_0 + i\Delta\phi_x(t)}\,, \tag{9.24}$$

$$E^{(y)}(t) = +\frac{1}{2} E_0 e^{-i\omega_{\rm L}(t - 2L/c) - i\phi_0 + i\Delta\phi_y(t)}\,, \tag{9.25}$$

where

$$\phi_0 = k_{\rm L}(L_x - L_y)\,, \tag{9.26}$$

$\Delta\phi_y = -\Delta\phi_x$, and

$$\begin{aligned} \Delta\phi_x(t) &= h_0\, k_{\rm L} L\, \mathrm{sinc}\,(\omega_{\rm gw} L/c)\, \cos[\omega_{\rm gw}(t - L/c)] \\ &\equiv |\Delta\phi_x| \cos(\omega_{\rm gw} t + \alpha)\,, \end{aligned} \tag{9.27}$$

with $\alpha = -\omega_{\rm gw} L/c$ a phase. The total phase difference induced by GWs in the Michelson interferometer is

$$\Delta\phi_{\rm Mich} \equiv \Delta\phi_x - \Delta\phi_y = 2\Delta\phi_x\,. \tag{9.28}$$

The total electric field at the output is

$$\begin{aligned} E_{\rm tot}(t) &= E^{(x)}(t) + E^{(y)}(t) \\ &= -iE_0 e^{-i\omega_{\rm L}(t - 2L/c)} \sin[\phi_0 + \Delta\phi_x(t)]\,. \end{aligned} \tag{9.29}$$

The phase $\phi_0$ is a parameter that the experimenter can adjust, choosing the best working point for the interferometer, as we will discuss in Section 9.3.2, while $\Delta\phi_x(t)$ contains the effect of the GW. In the limit $\omega_{\rm gw} L/c \ll 1$, eq. (9.27) reduces to

$$\Delta\phi_x(t) \simeq h(t - L/c)\, k_L L\,. \tag{9.30}$$

Comparing with eq. (9.26) we see that, in this limit, the effect of the GW on the phase shift is formally equivalent to a change of $L_x - L_y$ given by

$$\frac{\Delta(L_x - L_y)}{L} \simeq h(t - L/c)\,. \tag{9.31}$$

The total power $P \sim |E_{\rm tot}|^2$ observed at the photodetector is modulated by the GW signal as

$$\begin{aligned} P &= P_0 \sin^2[\phi_0 + \Delta\phi_x(t)] \\ &= \frac{P_0}{2} \{1 - \cos[2\phi_0 + 2\Delta\phi_x(t)]\} \\ &= \frac{P_0}{2} \{1 - \cos[2\phi_0 + \Delta\phi_{\rm Mich}(t)]\}\,. \end{aligned} \tag{9.32}$$

Clearly, we want to have $\Delta\phi_{\rm Mich}$ as large as possible. For a GW of a given frequency $\omega_{\rm gw}$, we see from eq. (9.27) that the dependence on $L$ is given by the factor $(\omega_{\rm L} L/c)\mathrm{sinc}\,(\omega_{\rm gw} L/c) = (\omega_{\rm L}/\omega_{\rm gw}) \sin(\omega_{\rm gw} L/c)$.

Thus the optimal length of the arms is given by $\omega_{gw}L/c = \pi/2$, i.e. $L = \lambda_{gw}/4$. In terms of $f_{gw} = \omega_{gw}/(2\pi)$, this gives

$$L \simeq 750\,\mathrm{km}\left(\frac{100\,\mathrm{Hz}}{f_{gw}}\right) . \tag{9.33}$$

For such a value of $L$, the time shift induced by the GW on the light has the same sign all along its round trip in a arm, so the effect adds up. For longer arms, the GW amplitude inverts its sign during the round trip, so past this moment it starts canceling the phase shift that the light already accumulated.[7] Arms of hundreds of kms are impossible to obtain in a ground-based interferometer, for practical and financial reasons. We will see in Section 9.2 how to "fold" this optimal pathlength of the light into an interferometers of manageable size.

It is useful to realize that the effect of the GW on the laser light is to generate sidebands in the light propagating in each of the two arms. Using eq. (9.27), and making use of the fact that $\Delta\phi_x$ is linear in $h_0$, we can expand $E^{(x)}(t)$ in eq. (9.24) to order $h_0$ as

$$\begin{aligned} E^{(x)}(t) &= -\frac{1}{2}E_0 e^{-i\omega_L(t-2L/c)+i\phi_0}\left[1 + i|\Delta\phi_x|\cos(\omega_{gw}t + \alpha)\right] \\ &= \frac{1}{2}E_0 e^{i\beta}\left[e^{-i\omega_L t} + \frac{i}{2}|\Delta\phi_x|\, e^{i\alpha}e^{-i(\omega_L-\omega_{gw})t} \right. \\ &\qquad \left. + \frac{i}{2}|\Delta\phi_x|\, e^{-i\alpha}e^{-i(\omega_L+\omega_{gw})t}\right] , \end{aligned} \tag{9.34}$$

with $\beta$ an irrelevant constant phase. Thus, beside the original electromagnetic wave at a frequency $\omega_L$ (the "carrier", in the language of radio engineering), we have two more electromagnetic waves, at the frequencies $\omega_L \pm \omega_{gw}$ (the "sidebands"). The modulus of the amplitude of the sidebands is $O(h_0)$ with respect to the carrier, and is given by $|\Delta\phi_x|/2$.

### 9.1.2 The interaction in the proper detector frame

It is instructive to compare the above results, obtained in the TT frame, with the description obtained using the language of the proper detector frame. Recall from Section 1.3.3 that the proper detector frame is the one implicitly used by the experimenter when he/she thinks about the apparatus. In particular, here coordinates are not marked by freely falling masses, as in the TT gauge, but rather are measured with a rigid ruler. We saw that in the proper detector frame the effect of the passage of a GW is a displacement of the test masses from their original position and, if these test masses are at a distance small compared to the reduced wavelength $\lambda\!\!\!^{-}_{gw}$ of the GW, this displacement is determined by the equation of the geodesic deviation (1.95). At the same time, the space-time metric can be taken as flat, at least in a region of space small compared to the scale of variation of the gravitational wave, which is its reduced wavelength $\lambda\!\!\!^{-}_{gw}$.[8]

[7]More precisely, for the value of $L$ given in eq. (9.33), the time shift always keeps the same sign for a photon whose time of time of entry inside the cavity is properly synchronized with the phase of the GW. For larger values, there is at least a partial cancellation, no matter what is the relation between the phase of the GW and the time of entry of the photon.

[8]This is correct as far as the fast-varying part of the gravitational field is concerned, while the static gravitational field of the Earth is compensated by the mirror suspensions, and other effects related to the laboratory frame (Coriolis forces, etc.) are negligible because slowly varying, see the detailed discussion in Section 1.3.3.

Thus, the proper detector frame description has the advantage of being very intuitive, since in a first approximation we can use the language of flat space-time, and the interaction of the mirrors with GWs is described by the equation of the geodesic deviation, i.e. in terms of Newtonian forces, so we can use our Newtonian intuition. However it must be kept in mind that, contrary to the TT gauge description, which is exact, the proper-frame description is approximate, and is valid only if the test masses are at a distance small compared to the reduced wavelength $\lambda\!\!\!^{-}_{gw}$ of the GW, see eq. (1.97). Since for a Michelson interferometer the distance between the beam-splitter and the end mirror of an arm is the arm-length $L$, the proper detector frame description assumes $L \ll \lambda\!\!\!^{-}_{gw}$, that is,

$$\frac{\omega_{gw}L}{c} \ll 1 . \tag{9.35}$$

Thus, we cannot expect to recover the full TT gauge result (9.27), which is exact, but only its limit for small values of $\omega_{gw}L/c$.

We first perform the computation in the proper detector frame to lowest-order in $\omega_{gw}L/c$. In this limit the space-time metric is exactly flat, see eq. (1.86), while the effect of the GW on the test masses is given by the equation of the geodesic deviation, eq. (1.95). Thus, the situation is reversed compared to the TT gauge description. In the TT gauge, the position of the mirrors is not affected by GWs, while the propagation of light between the mirrors is affected. In the proper detector frame, the mirrors are affected by the GWs, while light propagation is not. We fix the origin of the coordinate system on the beam-splitter so, by definition, the beam-splitter does not move, and we consider as before a GW with only the plus polarization coming from the $z$ direction, written as in eq. (9.6). The equation of the geodesic equation for the mirror on the $x$ arm, described by coordinates $(\xi_x, \xi_y)$, is then[9]

$$\ddot{\xi}_x = \frac{1}{2}\ddot{h}_+ \xi_x , \tag{9.36}$$

while $\xi_y(t)$ remains zero at all times if $\xi_y(0) = \dot{\xi}_y(0) = 0$. Equation (9.36) can be solved perturbatively in $h_0$; to zeroth order we have $\xi_x = L_x$, so to $O(h_0)$ we get $\ddot{\xi}_x = (1/2)\ddot{h}_+ L_x$, which has the solution

$$\xi_x(t) = L_x + \frac{h_0 L_x}{2}\cos\omega_{gw}t , \tag{9.37}$$

where we choose the integration constants so that the average value of $\xi_x$ over one period of the GW is equal to $L_x$, and the average value of the velocity $\dot{\xi}_x$ vanishes.

Since space-time is flat, a photon that starts at the beam-splitter at time $t_0$, moving along the positive $x$ axis, follows the trajectory $x(t) = c(t - t_0)$, so it reaches the mirror at a time $t_1$ given by $c(t_1 - t_0) = \xi_x(t_1)$. This equation is easily solved for $t_1$, perturbatively in $h_0$. To zeroth order in $h_0$ we get the trivial result $t_1 = t_0 + (L_x/c)$. Inserting this into $\cos\omega_{gw}t$ in eq. (9.37) (which is already multiplied by $h_0$), we get

$$c(t_1 - t_0) = L_x + \frac{h_0 L_x}{2}\cos[\omega_{gw}(t_0 + L_x/c)] . \tag{9.38}$$

[9]The fact that here appears the form of $h_{\mu\nu}$ in the TT gauge, even if we are working in the proper detector frame, is a consequence of the fact that the right-hand side of eq. (9.36) is really $-c^2 R_{i0j0}\xi^j$, see eq. (1.93). Recall from Section 1.1 that, in linearized theory, the Riemann tensor is invariant under coordinate transformations, so we are free to compute it in the frame that we wish, and in particular we can use the form of $h_{\mu\nu}$ in the TT gauge.

The round-trip time is twice as large, so the photon gets back at the beam-splitter at a time $t_2$ given by

$$t_2 - t_0 = \frac{2L_x}{c} + \frac{h_0 L_x}{c}\cos[\omega_{\rm gw}(t_0 + L_x/c)]\,. \tag{9.39}$$

This coincide with the result that we got in the TT gauge, eq. (9.14), *except* that the function sinc $(\omega_{\rm gw}L/c) = [\sin(\omega_{\rm gw}L/c)]/[\omega_{\rm gw}L/c]$ has been replaced by one, which is the lowest-order term of its Taylor expansion. This is as expected, since the proper-frame computation just performed is valid only to lowest order in $\omega_{\rm gw}L/c$.

It is instructive to compute also the next term in the expansion in $\omega_{\rm gw}L/c$ in the proper detector frame, and verify that we correctly recover the next term in the expansion of sinc $(\omega_{\rm gw}L/c)$. In principle, we have two kinds of corrections. (1) Corrections to the equation of motion of the mirrors, since the geodesic equation that we have used is the first term in an expansion in $L/\lambda\!\!\!^{-}_{\rm gw} = \omega_{\rm gw}L/c$, as it is clear from the derivation leading from eq. (1.66) to eq. (1.71). (2) Correction to the propagation of the photons, since the space-time metric is no longer flat.

Actually, the former type of correction in our problem vanishes at next-to-leading order.[10] The first correction to the photon propagation can be computed using the metric (1.87). For the propagation along a trajectory with $y = z = 0$ (and therefore with $dy = dz = 0$), recalling that the Riemann tensor is antisymmetric in the first and second pair of indices, eq. (1.87) reduces to

$$ds^2 = -c^2dt^2(1 + R_{0101}x^2) + dx^2\,. \tag{9.40}$$

We can compute the Riemann tensor using the form of $h_{\mu\nu}$ in the TT gauge (compare with Note 9) which gives, for a wave with only the plus polarization,

$$\begin{aligned} R_{0101} &= -\frac{1}{2c^2}\ddot{h}_+ \\ &= \frac{\omega_{\rm gw}^2}{2c^2}\,h_0\cos\omega_{\rm gw}t\,, \end{aligned} \tag{9.41}$$

see eq. (1.94). Light propagation is obtained imposing $ds^2 = 0$ in eq. (9.40); then, to next-to-leading order, the position $x(t)$ of a photon propagating along the $x$ arm is obtained integrating

$$dx = \pm c\,dt\left[1 + \frac{\omega_{\rm gw}^2}{4c^2}x^2(t)h_0\cos\omega_{\rm gw}t\right], \tag{9.42}$$

while the motion of the mirrors is still given by eq. (9.37).

Consider a photon that leaves the beam-splitter at time $t_0$ and propagates along the positive $x$ direction. To lowest order in $h_0$ we have the trivial result $x(t) = c(t - t_0)$. Inserting this into the right-hand side of eq. (9.42) we find the solution to order $h_0$,

$$x(t) = c(t - t_0) + h_0\frac{c\omega_{\rm gw}^2}{4}\int_{t_0}^{t} dt'\,(t' - t_0)^2\cos\omega_{\rm gw}t'\,. \tag{9.43}$$

[10] This can be shown by observing that the geodesic equation (1.66) for a mirror moving non-relativistically is simply $\ddot{\xi}^i = -\Gamma^i_{00}(\xi)$. Expanding it to second order in $\xi$, with $\Gamma^i_{00}(\xi = 0) = 0$, we get

$$\ddot{\xi}^i = -\xi^j\partial_j\Gamma^i_{00} - \frac{1}{2}\xi^j\xi^k\partial_k\partial_j\Gamma^i_{00} + O(\xi^3),$$

where the derivatives of $\Gamma^i_{00}$ are computed at $\xi = 0$. The first term gives the equation of motion that we already used. For a mirror along the $x$ arm $\xi^i = (\xi_x, 0, 0)$, so the second term is proportional to $\partial_1(\partial_1\Gamma^1_{00})$. A plane wave propagating along the $z$ direction is function only of $t$ and $z$, and so its Riemann tensor, as well as $(\partial_1\Gamma^1_{00})$, is independent of $x$, and $\partial_1(\partial_1\Gamma^1_{00})$ vanishes.



Writing

$$\begin{aligned} \cos\omega_{\rm gw}t' &= \cos[\omega_{\rm gw}(t' - t_0) + \omega_{\rm gw}t_0] \\ &= \cos[\omega_{\rm gw}(t' - t_0)]\cos\omega_{\rm gw}t_0 - \sin[\omega_{\rm gw}(t' - t_0)]\sin\omega_{\rm gw}t_0\,, \end{aligned} \tag{9.44}$$

the integral over $t'$ can be performed exactly. Consistently with the order to which we are working, we then expand the exact result to the first non-trivial order in $\omega_{\rm gw}(t - t_0)$ (which, in the final result, will correspond to the first non-trivial order in $\omega_{\rm gw}L/c$), and we get

$$x(t) \simeq c(t - t_0) + h_0\frac{c\omega_{\rm gw}^2}{12}(t - t_0)^3\cos\omega_{\rm gw}t_0\,. \tag{9.45}$$

The time $t_1$ at which the photon reaches the mirror is now obtained solving the equation $x(t_1) = \xi(t_1)$ iteratively in $h_0$. This gives

$$c(t_1 - t_0) = L_x + \frac{h_0L_x}{2}\cos[\omega_{\rm gw}(t_0 + L_x/c)] - h_0\frac{\omega_{\rm gw}^2}{12c^2}L_x^3\cos(\omega_{\rm gw}t_0)\,. \tag{9.46}$$

Observe that (writing $\epsilon \equiv \omega_{\rm gw}L_x/c$)

$$\begin{aligned} \cos[\omega_{\rm gw}(t_0 + L_x/c)] &= \cos(\omega_{\rm gw}t_0)\cos\epsilon - \sin(\omega_{\rm gw}t_0)\sin\epsilon \\ &= [1 + O(\epsilon^2)]\cos(\omega_{\rm gw}t_0) + O(\epsilon)\sin(\omega_{\rm gw}t_0)\,, \end{aligned} \tag{9.47}$$

so in the last term of eq. (9.46), which is already a factor $(\omega_{\rm gw}L_x/c)^2$ smaller than the second term, we can replace $\cos(\omega_{\rm gw}t_0)$ by $\cos[\omega_{\rm gw}(t_0 + L_x/c)]$, since the difference is of higher order in $\omega_{\rm gw}L_x/c$. Then we finally get

$$c(t_1 - t_0) = L_x + \frac{h_0L_x}{2}\cos[\omega_{\rm gw}(t_0 + L/c)]\left[1 - \frac{1}{6}\left(\frac{\omega_{\rm gw}L_x}{c}\right)^2\right]. \tag{9.48}$$

Writing similarly the equations for the round trip we find that, to this order in $\omega_{\rm gw}L_x/c$, the round-trip travel $t_2 - t_0$ is twice $t_1 - t_0$. In the last bracket we recognize the first two terms of the expansion

$$\frac{\sin x}{x} = 1 - \frac{x^2}{6} + O(x^4)\,. \tag{9.49}$$

We have therefore verified that the analysis in the proper detector frame correctly reproduces the leading and the next-to-leading terms of the TT gauge result given in eq. (9.15). It is also clear from this discussion that, while the description in the detector proper frame is more intuitive, since it allows us to think in terms of Newtonian forces acting on the mirrors, and of light propagating (in a first approximation) in the flat space-time of Newtonian physics, still the TT gauge description is much more powerful, since it allows us to get the exact closed form of the dependence on $\omega_{\rm gw}L_x/c$. In the detector proper frame the computation of still higher-order corrections becomes more and more involved and, without the hindsight from the TT gauge analysis, it would be difficult to imagine that the whole series resums to such a simple closed form.

## 9.2 Interferometers with Fabry–Perot cavities

We have seen in eq. (9.33) that, to measure GWs with frequencies of order of a few hundreds Hz, the optimal choice would be an arm-length $L$ of several hundreds kms. For Earth-based interferometers this is in practice impossible (consider, among other things, that the arms of the interferometers must be enclosed in a very high vacuum system, as we will see in the Section 9.4). Taking into account technological and financial constraints, LIGO has arms of length $L = 4$ km and VIRGO of 3 km, while GEO600 has $L = 600$ m and TAMA has $L = 300$ m. The idea is therefore to "fold" the optical path of light, making it bounce back and forth many times in each arm, before recombining the two beams. A solution that was first considered is the so-called "delay line". In this case, in each arm the light beam goes back and forth between two mirrors along trajectories that do not superimpose, and which make different spots on the mirrors. However, to reach an effective path length of order 750 km out of arms of order 3–4 km we need $O(100)$ bounces. In the delay line scheme, this leads to unpractically large mirrors.[11] Thus, the solution which has been adopted in LIGO and VIRGO is that of transforming each arm into a Fabry–Perot cavity. In the next subsection we will discuss the principles of operation of a Fabry–Perot (FP) cavity, and in Section 9.2.2 we will discuss its interaction with a GW, and we will see how it improves on the simple Michelson scheme.

[11] Furthermore, this scheme turns out to be quite sensitive to the problem that some part of the light is scattered inside the interferometer off the nominal path, and this light interferes with the main beam.

### 9.2.1 Electromagnetic fields in a FP cavity

#### Reflection and transmission coefficients

First of all we recall from elementary electromagnetism that, at the interface between two media with different index of refraction, the relation between the incoming field $E_{\rm in}$, the reflected field $E_{\rm refl}$ and the transmitted field $E_{\rm t}$ can be written as

$$E_{\rm refl} = rE_{\rm in}\,, \qquad E_{\rm t} = tE_{\rm in}\,, \tag{9.50}$$

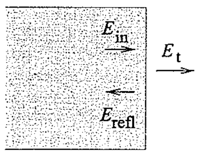


**Fig. 9.3** The situation in which the incoming field comes from the denser medium.

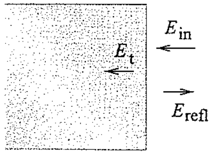


**Fig. 9.4** The situation in which the incoming field comes from the rarer medium.

where $r$ and $t$ are called the reflection and transmission coefficients, respectively, and are in general complex numbers. We consider for the moment the transmission and reflection across a sharp boundary. At a sharp boundary there is no physical mechanism that can produce a phase shift, so in this limit $r$ and $t$ are real. More precisely, $(r, t)$ are the reflection and transmission coefficients when $E_{\rm in}$ comes from the first medium, say the denser (from the left in Fig. 9.3). Similarly, we denote by $r'$ and $t'$ the reflection and transmission coefficients when $E_{\rm in}$ comes from the second medium, i.e. from the right in Fig. 9.4. Between these coefficients hold useful relations. In particular, since the energy associated to the electric field is proportional to $|E|^2$, and on a sharp boundary there are no losses and $r, t$ are real, energy conservation requires

$$r^2 + t^2 = 1\,, \tag{9.51}$$

and $r'^2 + t'^2 = 1$.[12] Between $(r, t)$ and $(r', t')$ we have so-called reciprocity relations, which can be obtained as follows. Consider the arrangement shown in Fig. 9.5, in which the incoming electric field arrives from the left, and there is a gap of width $d$ of a less dense medium between two layers of the more dense medium. We denote by $E_{\rm cav}$ and $E'_{\rm cav}$, respectively, the right-moving and left-moving electric fields in the gap, close to the first interface. Then, by definition of reflection and transmission coefficients, at the first interface we have the two relations

[12] More precisely, the energy density is actually proportional to $\mathbf{E}\cdot\mathbf{D}$, where $\mathbf{D}$ is the displacement vector. If we define the coefficients $r, t$ in terms of $\mathbf{E}$, we should then write $r^2 + t^2(n_1/n_2) = 1$. We can however simply reabsorb $n$ into the definition of $\mathbf{E}$ to keep the equations in simpler forms such as eq. (9.51). In any case, the issue is irrelevant for a situation such as that shown in Fig. 9.6, where we are interested in the fields in the vacuum, on both sides of the mirror.

$$E_{\rm cav} = tE_{\rm in} + r'E'_{\rm cav}\,, \tag{9.52}$$

$$E_{\rm refl} = rE_{\rm in} + t'E'_{\rm cav}\,. \tag{9.53}$$

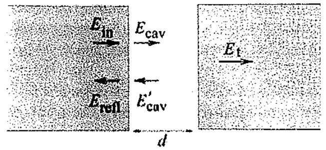


**Fig. 9.5** A gap of a less dense medium between two layers of denser media.

We now take the limit $d \to 0$. In this case $E_{\rm cav}$ and $E'_{\rm cav}$ are also the right and left-moving fields, respectively, at the second interface. Thus, we also have the relations

$$E_{\rm t} = t'E_{\rm cav}\,, \tag{9.54}$$

$$E'_{\rm cav} = r'E_{\rm cav}\,. \tag{9.55}$$

On the other hand, if $d \to 0$, there is no gap, and we must have

$$E_{\rm t} = E_{\rm in}\,, \tag{9.56}$$

$$E_{\rm refl} = 0\,. \tag{9.57}$$

Combining the six relations (9.52)–(9.57) we find the two conditions

$$r' = -r\,, \tag{9.58}$$

$$tt' - rr' = 1\,. \tag{9.59}$$

Inserting eqs. (9.58) and (9.51) into eq. (9.59) we get $t' = t$. In conclusion, we have

$$\boxed{r' = -r\,, \qquad t' = t\,.} \tag{9.60}$$

For a perfectly reflecting mirror, reflection from the less dense to the more dense medium is associated to a factor $r' = -1$, while from the denser to the less dense medium we have $r = 1$.

#### Reflected, transmitted and interior field in a FP cavity

We can now apply the above results to the study of a Fabry–Perot cavity. A Fabry–Perot cavity consists of two parallel mirrors, that for the time being we assume plane and of infinite transverse extent. We consider a component $E_{\rm in}$ of the incoming electric field. Part of the incoming field is reflected and partly transmitted, see Fig. 9.6. The

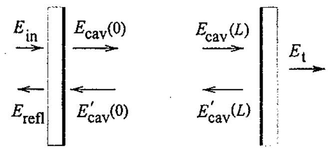


**Fig. 9.6** A schematic Fabry–Perot cavity.

transmitted field $E_{cav}(0)$ propagates to the far mirror, where it is partly reflected and partly transmitted. The reflected part goes back to the first mirror, where again it is partly reflected and partly transmitted, and so on. The total reflected, interior and transmitted fields are therefore determined by the superposition of many beams, corresponding to the multiple bounces.

The light from the laser comes from the left in Fig. 9.6. The mirrors are set with their high-reflectivity coating on the interior of the cavity. Before reaching the high-reflectivity coating, light enters from the left face of the mirror and passes through the substrate, so in general it acquires a further complex phase shift, both from the substrate and from the coating, and can also suffer losses. A beam which enters the cavity and, after a number of round-trips, is reflected back, traverses once more this substrate, acquiring a further phase. The important point, however, is that these phase shifts are the same for all beams, independently of the number of bounces made inside the FP cavity, so they just give an overall phase factor, independent of the length $L$ of the cavity, to the reflected (and transmitted) fields. This phase factor is compensated by the experimenter, moving the position of the mirrors until the interference pattern of the interferometer is on the desired working point (the dark fringe, as we will see), so we can simply forget about them.[13] We can therefore simply model the two mirrors of the FP cavity stating that, for the first mirror, we have real reflection and transmission coefficients $r_1$ and $t_1$ when the incoming field propagates from the interior of the mirror toward the cavity, and $r_1' = -r_1$ and $t_1' = t_1$ when it is going from the cavity toward the mirror. We then take into account the losses in the mirror writing

$$r_1^2 + t_1^2 = 1 - p_1\,, \tag{9.61}$$

where $p_1$ (typically of order of a few parts per million) represents the losses in the first mirror. We similarly introduce coefficients $(r_2, t_2)$ and $(r_2' = -r_2, t_2' = t_2)$ for the second mirror, with $r_2^2 + t_2^2 = 1 - p_2$, so again a field that propagates from the cavity toward the mirror and is reflected back gets a factor $-r_2$.

Other modelizations of the mirrors are possible. In particular, one could treat the reflection and transmission from the two sides of the mirror symmetrically so that, if a field $A$ is coming from the left, the reflected field is $A_R = z_R A$ and the transmitted field is $A_T = z_T A$, where $z_R$ and $z_T$ are the reflection and transmission coefficients, which now a priori can be complex because of the finite thickness of the mirror, and which satisfy $|z_R|^2 + |z_T|^2 = 1 - p$. Similarly, if a field $B$ is coming from the right, the reflected field is $B_R = z_R B$ and the transmitted field is $B_T = z_T B$, with the same $z_R, z_T$. In the presence of both a field $A$ coming from the left and a field $B$ coming from the right, as in Fig. 9.7, we have

$$A' = z_R A + z_T B \tag{9.62}$$

[13]Multiple reflections inside the substrate can be suppressed by using anti-reflection coatings, but are really eliminated only by giving to the external face of the mirror a wedge-shaped form, so after a few bounces the photon is simply lost, rather than reentering the main beam with a different phase (which would result in noise). It should however be observed that all these couplings of the phase of the light to the geometry of the optics (such as multiple reflections, phase shifts at the coatings, etc.) are channels from which noise can enter.

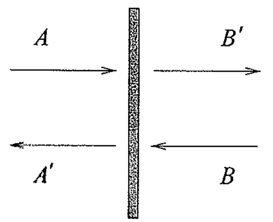


**Fig. 9.7** A symmetric mirror, with a field $A$ coming from the left and a field $B$ incoming from the right.



$$B' = z_T A + z_R B\,. \tag{9.63}$$

Requiring the energy balance $|A'|^2 + |B'|^2 = (1-p)(|A|^2 + |B|^2)$, we get the condition Re $(z_R z_T^*) = 0$. A possible solution is $z_R = ir$, $z_T = t$, where $r$ and $t$ are real and satisfy $r^2 + t^2 = 1 - p$.

These different modelizations of the mirrors of a cavity of length $L$ can however be compensated by a constant shift $\Delta L$ of some fraction of wavelength. For instance, with the modelization $r' = -r$, we will find below that a Fabry–Perot cavity resonates at $2k_L L = 2\pi n$, with $n$ integer. Repeating the computation for $z_R = ir$, equal for both sides, one would rather find resonances at $2k_L L = 2\pi(n + 1/2)$. In practice, the experimenter tunes the position of the mirrors until he/she finds a resonance, and all that matters is the behavior around resonance, which is the same in the two cases, so the modelization chosen for the mirrors becomes irrelevant. Similarly, using $z_R = ir$ and $z_T = t$, instead of eq. (9.5) one would find $|E_{out}|^2 = E_0^2 \cos^2[k_L(L_y - L_x)]$, but again the experimenter simply adjusts the lengths $L_x$ and $L_y$ until he/she finds the desired working point, such as the dark fringe. For definiteness, we will always use the modelization leading to eq. (9.60).

We can now compute the reflected and transmitted fields, and the field inside the cavity, as follows. We choose the coordinates so that the left mirror is at $x = 0$ and the right mirror at $x = L$. From the laser we send light with an electric field of the form $E_0 \exp\{-i\omega_L t + ik_L x\}$. Let $t = t_0$ be the value of time at which a given wave-packet reaches the mirror, at $x = 0$. Thus, the corresponding electric field is simply

$$E_0 e^{-i\omega_L t_0}\,. \tag{9.64}$$

Part of this beam will be immediately reflected back from the mirror, with amplitude $+r_1$, giving rise to a reflected beam with field

$$E_{refl}^{(0)} = r_1 E_0 e^{-i\omega_L t_0}\,. \tag{9.65}$$

This field will interfere with a beam that was send toward the mirror *earlier*, which entered the cavity, was reflected back at the second mirror, and then was transmitted from the first mirror, see Fig. 9.8. In order to arrive back at the first mirror at the same time $t_0$, it must have entered the cavity at time $t_0 - 2L/c$. So, its initial amplitude when it entered the mirror for the first time, arriving from the laser, was $E_0 e^{-i\omega_L(t_0 - 2L/c)}$, that is $E_0 e^{-i\omega_L t_0} e^{2ik_L L}$. After transmission from the first mirror it gets a factor $t_1$, reflection at the second mirror gives a factor $-r_2$ and finally transmission from the first mirror gives again $t_1$. Thus, at time $t_0$ the total reflected field gets also a contribution

$$E_{refl}^{(1)} = \left[-t_1^2 r_2\, e^{2ik_L L}\right] E_0 e^{-i\omega_L t_0}\,. \tag{9.66}$$

This beam has a relative amplitude $-r_2 t_1^2\, e^{2ik_L L}$ compared to the incoming laser field given in eq. (9.64).[14] Then, we have the field that entered the cavity at time $t_0 - 4L/c$, and went twice back and forth in the cavity. It comes out at time $t_0$ with an amplitude

$$E_{refl}^{(2)} = \left[-r_1 r_2^2 t_1^2\, e^{4ik_L L}\right] E_0 e^{-i\omega_L t_0}\,. \tag{9.67}$$

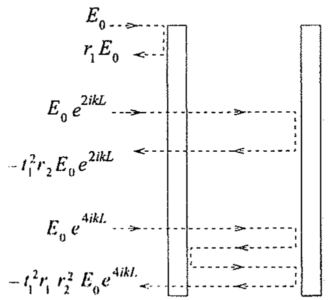


**Fig. 9.8** The building up of the reflected field from the interference of the directly reflected beams and of beams that entered the cavity at earlier times and made a number of bounces inside the cavity. For clarity, the various path have been drawn as if they were spatially separated.

[14]This result is often colloquially explained stating that the field has acquired a phase $e^{2ik_L L}$ from the free propagation, and the factors $-r_2 t_1^2$ from its various transmissions and reflection. This is misleading. Of course, a wave does not acquire any phase factor from its free propagation. Photons in free space propagate along a trajectory $x(t) = x_0 + c(t - t_0)$, so $x(t) - ct = x_0 - ct_0$ is a constant, and the phase factor $k_L x - \omega_L t = k_L(x - ct)$ is also constant along the free propagation. In flat space-time, multiplicative factors and phases are acquired by the amplitude only when there is interaction with matter, i.e. at reflection and transmission from the mirrors, and in our case this gives the factor $-r_2 t_1^2$. The factor $e^{2ik_L L}$ relative to eq. (9.64) is there because the beam that we are considering entered the cavity *earlier*, at a time $t_0 - 2L/c$, and spent a time $2L/c$ going back and forth in the cavity. Thus, this second beam already had from the start a phase different by a factor $e^{2ik_L L}$, compared to the field (9.64) that arrives at time $t_0$ directly from the laser.

More generally, the field that entered the cavity at time $t_0 - n(2L/c)$ and performed $n$ round trips comes out at time $t_0$ with an amplitude

$$E^{(n)}_{\rm refl} = \left[-r_1^{n-1} r_2^n t_1^2\, e^{2nik_{\rm L}L}\right] E_0 e^{-i\omega_{\rm L}t_0}\,. \tag{9.68}$$

The total reflected field is therefore given by

$$\begin{aligned} E_{\rm refl} &= E_0 e^{-i\omega_{\rm L}t_0}\left[r_1 - t_1^2 \sum_{n=1}^{\infty} r_1^{n-1} r_2^n e^{2nik_{\rm L}L}\right] \\ &= E_0 e^{-i\omega_{\rm L}t_0}\left[r_1 - t_1^2 r_2 e^{2ik_{\rm L}L} \sum_{m=0}^{\infty}\left(r_1 r_2 e^{2ik_{\rm L}L}\right)^m\right] \\ &= E_0 e^{-i\omega_{\rm L}t_0}\left[r_1 - t_1^2 r_2 \frac{e^{2ik_{\rm L}L}}{1 - r_1 r_2 e^{2ik_{\rm L}L}}\right], \end{aligned} \tag{9.69}$$

or, using $t_1^2 = 1 - p_1 - r_1^2$,

$$\boxed{E_{\rm refl} = E_0 e^{-i\omega_{\rm L}t_0}\,\frac{r_1 - r_2(1-p_1)e^{2ik_{\rm L}L}}{1 - r_1 r_2 e^{2ik_{\rm L}L}}\,.} \tag{9.70}$$

The transmitted field is computed similarly,

$$\begin{aligned} E_{\rm t} &= E_0 e^{-i\omega_{\rm L}t_0}\, t_1 t_2 \sum_{n=0}^{\infty} (r_1 r_2)^n e^{ik_{\rm L}L(2n+1)} \\ &= E_0 e^{-i\omega_{\rm L}t_0}\,\frac{t_1 t_2 e^{ik_{\rm L}L}}{1 - r_1 r_2 e^{2ik_{\rm L}L}}\,. \end{aligned} \tag{9.71}$$

The field inside the cavity, at the left mirror ($x = 0$), again at time $t_0$, is

$$\begin{aligned} E_{\rm cav}(0) &= E_0 e^{-i\omega_{\rm L}t_0}\, t_1 \sum_{n=0}^{\infty} (r_1 r_2)^n e^{2nik_{\rm L}L} \\ &= E_0 e^{-i\omega_{\rm L}t_0}\,\frac{t_1}{1 - r_1 r_2 e^{2ik_{\rm L}L}}\,, \end{aligned} \tag{9.72}$$

and for the field inside the cavity, at the other mirror, at time $t_0$, we have $E_{\rm cav}(L) = e^{ik_{\rm L}L}E_{\rm cav}(0)$. The same results can been obtained also in the following way, which is maybe less vivid physically, but will be easier to generalize to the situation in which GWs are present. We consider the total reflected, transmitted and cavity fields as shown in Fig. 9.6. Then, just as in eqs. (9.52) and (9.53), using $r_1' = -r_1$ and $t_1' = t_1$, at the first mirror we have

$$E_{\rm cav}(0) = t_1 E_{\rm in} - r_1 E'_{\rm cav}(0)\,, \tag{9.73}$$

$$E_{\rm refl} = r_1 E_{\rm in} + t_1 E'_{\rm cav}(0)\,. \tag{9.74}$$

Similarly, at the second mirror we have

$$E_{\rm t} = t_2 E_{\rm cav}(L)\,, \tag{9.75}$$

$$E'_{\rm cav}(L) = -r_2 E_{\rm cav}(L)\,. \tag{9.76}$$



Finally, since the solution inside the cavity is given by plane waves, the field $E_{\rm cav}(t,x)$, which represent a right-moving wave, is proportional to $\exp\{-i(\omega_{\rm L}t - k_{\rm L}x)\}$, while $E'_{\rm cav}(t,x)$, which represent a left-moving wave, is proportional to $\exp\{-i(\omega_{\rm L}t + k_{\rm L}x)\}$. Thus the cavity fields at $x = L$ and at $x = 0$, at equal value of time, are related by

$$E_{\rm cav}(L) = e^{ik_{\rm L}L}E_{\rm cav}(0)\,, \tag{9.77}$$

$$E'_{\rm cav}(L) = e^{-ik_{\rm L}L}E'_{\rm cav}(0)\,. \tag{9.78}$$

Then we have six equations, eqs. (9.73)–(9.78), that we can solve for the six quantities $E_{\rm refl}, E_{\rm t}, E_{\rm cav}(0), E_{\rm cav}(L), E'_{\rm cav}(0), E'_{\rm cav}(L)$, in terms of $E_{\rm in} = E_0 e^{-i\omega_{\rm L}t}$. With straightforward algebra we get back the solution found above. For instance, combining eqs. (9.76), (9.77) and (9.78) we get

$$E'_{\rm cav}(0) = -r_2 e^{2ik_{\rm L}L}E_{\rm cav}(0)\,. \tag{9.79}$$

Substituting this into eq. (9.73) we get

$$E_{\rm cav}(0) = t_1 E_{\rm in} + r_1 r_2 e^{2ik_{\rm L}L}E_{\rm cav}(0)\,, \tag{9.80}$$

from which the solution (9.72) for $E_{\rm cav}(0)$ follows, and similarly we get $E_{\rm refl}$ and $E_{\rm t}$.

### Resonant FP cavities

We see that the reflected, transmitted and interior fields are all proportional to the factor $1/[1 - r_1 r_2 e^{2ik_{\rm L}L}]$. When $2k_{\rm L}L = 2\pi n$, with $n = 0, \pm 1, \pm 2, \ldots$, this factor becomes $1/(1 - r_1 r_2)$ and, if the reflection coefficients $r_1$ and $r_2$ are close to one, this is large. We therefore have a set of resonances. Physically this means that, for $2k_{\rm L}L = 2\pi n$, the various beams that bounce back and forth interfere constructively, so the field inside the cavity raises to a very large value. Correspondingly, the transmitted field also gets large. As for the reflected field, for assessing its strength we must also take into account the dependence on $k_{\rm L}L$ of the numerator, which describes the interference between the field that is reflected after having entered the cavity and made one or more round trips, and the field that is immediately reflected. We first consider the power $P_t \sim |E_{\rm t}|^2$ of the transmitted field (or, equivalently, of the interior field, $E_{\rm cav}$, since $|E_{\rm t}|$ and $|E_{\rm cav}|$ differ just by a constant factor $t_2$). From eq. (9.71),

$$|E_{\rm t}|^2 = E_0^2\,\frac{t_1^2 t_2^2}{1 + (r_1 r_2)^2 - 2r_1 r_2 \cos 2k_{\rm L}L}\,. \tag{9.81}$$

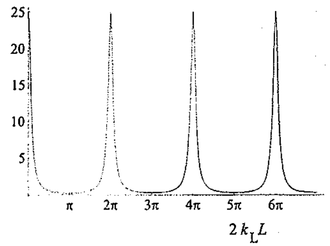


**Fig. 9.9** $|E_{\rm t}|^2$, in units of $(E_0 t_1 t_2)^2$, as a function of $2k_{\rm L}L$, for $r_1 r_2 = 0.8$.

This is plotted, as a function of $2k_{\rm L}L$, in Fig. 9.9. Writing $k_{\rm L} = \omega_{\rm L}/c$, the distance between the maxima is

$$\Delta\omega_{\rm L} = \frac{\pi c}{L}\,. \tag{9.82}$$

This is called the *free spectral range* of the cavity. Expanding the denominator in eq. (9.81) to quadratic order around a resonance, we find that the full width of the peaks at half maximum is

$$\delta\omega_L = \frac{c}{L}\frac{1-r_1r_2}{\sqrt{r_1r_2}}. \quad (9.83)$$

The *finesse* $\mathcal{F}$ of the cavity is defined as the ratio of the free spectral range to the full width at half maximum, $\mathcal{F} \equiv \Delta\omega_L/\delta\omega_L$, so

$$\mathcal{F} = \frac{\pi\sqrt{r_1r_2}}{1-r_1r_2}. \quad (9.84)$$

To understand the physical meaning of these results it is useful to compute the *storage time*, i.e. the average time spent by a photon inside the cavity. For simplicity we take $r_2 = 1$, so each photon has an amplitude probability $A(n)$ of making $n$ round trips, and finally comes out from the first mirror. Recall that the number density of photons is proportional to the modulus squared of the electric field, so the factors $-r_1$ and $-r_2$ acquired at the reflections from the mirrors are the quantum-mechanical probability amplitudes, while their squared modulus is a probability. Thus, the amplitude for performing $n$ round-trips and then coming out from the first mirror is given by $A(n) = t_1^2(-1)^n(-r_1)^{n-1} = \text{constant} \times r_1^n$, since each reflection at the far mirror has a probability amplitude $-1$ and at the first mirror $(-r_1)$. Thus, if a photon enters the cavity, the probability that it comes out after $n$ round-trips is

$$p(n) = \frac{r_1^{2n}}{\sum_{n=1}^{\infty} r_1^{2n}}, \quad (9.85)$$

where the denominator normalizes the total probability to one. The average number of round-trips is therefore

$$\sum_{n=1}^{\infty} n\,p(n) = \frac{1}{1-r_1^2}. \quad (9.86)$$

Since each round-trip lasts for a time $2L/c$, the storage time of the cavity, i.e. the average time spent inside by a photon, is

$$\tau_s = \frac{2L}{c}\frac{1}{1-r_1^2}. \quad (9.87)$$

If $r_1$ is close to one we can write $1 - r_1^2 = (1-r_1)(1+r_1) \simeq 2(1-r_1)$, and we can express the storage time in terms of the finesse, as

$$\tau_s \simeq \frac{L}{c}\frac{\mathcal{F}}{\pi}. \quad (9.88)$$

We see that, in the limit of high finesse, light is trapped in the FP cavity for a long time. If we illuminate the cavity and then we suddenly shut off the laser at $t = 0$, light will still continue to come out from the cavity for a long time. According to eq. (9.85), the intensity of the light coming out after $n$ round trips is proportional to $r_1^{2n} = \exp\{n \log r_1^2\}$. For $r_1$ close to one, $\log r_1^2 = \log[1-(1-r_1^2)] \simeq -(1-r_1^2)$. Therefore the intensity of light decreases with $n$ as $\exp\{-n(1-r_1^2)\}$. Since the light that performs $n$ round trips comes out at time $t = (2L/c)n$, for $r_1$ close to one and $r_2 = 1$ the intensity of the reflected light decreases with time as $\exp\{-t(c/2L)(1-r_1^2)\} = \exp\{-t/\tau_s\}$, with $\tau_s$ given in eq. (9.87), confirming the interpretation of $\tau_s$ as a storage time.



We consider now the reflected field. We write $E_{\rm refl} = |E_{\rm refl}|e^{-i\omega_L t}\,e^{i\phi}$, and we find from eq. (9.70) that the phase $\phi$ can be written as $\phi = \phi_1 - \phi_2$, where

$$\tan\phi_1 = -\frac{r_2(1-p_1)\sin(2k_L L)}{r_1 - r_2(1-p_1)\cos(2k_L L)}, \quad (9.89)$$

$$\tan\phi_2 = -\frac{r_1r_2\sin(2k_L L)}{1 - r_1r_2\cos(2k_L L)}. \quad (9.90)$$

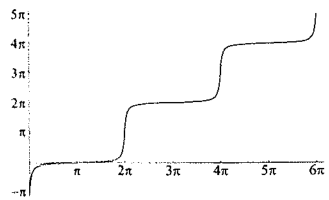


**Fig. 9.10** The phase $\phi$ of the reflected field, as a function of $2k_L L$, setting $r_1 = 0.9$, $r_2 = 1$, $p_1 = 0$. We have defined $\phi$ so that it is a continuous function of $2k_L L$, rather than reporting it always to the interval $[0, 2\pi]$.

A plot of $\phi$ as a function of $2k_L L$ is shown in Fig. 9.10. Two aspects of this graph are interesting. First, away from the resonances (which, as we have seen, are at $2k_L L = 2\pi n$), $\phi$ is almost flat as a function of $2k_L L$, and is basically equal to zero (mod $2\pi$). So, here the phase of the reflected light is insensitive to changes in the length $L$ of the cavity or of the frequency of the laser light. However, close to the resonances this dependence suddenly becomes very sharp. Writing $2k_L L = 2\pi n + \epsilon$ and expanding for small $\epsilon$, eqs. (9.89) and (9.90) give (setting for simplicity $r_2 = 1$ and $p_1 = 0$ and neglecting $O(\epsilon^2)$) $\partial\phi/\partial\epsilon = (1+r_1)/(1-r_1)$ or, taking $r_1$ close to one,

$$\boxed{\frac{\partial\phi}{\partial\epsilon} \simeq \frac{2\mathcal{F}}{\pi}.} \quad (9.91)$$

We can compare this with the result (9.2) for one arm of a simple Michelson interferometer which, in the present notation, reads $\phi = \epsilon$. When $r_1$ is close to one, the sensitivity of a FP cavity to changes in $2k_L L$ is enhanced by the large factor $(2/\pi)\mathcal{F}$, compared to the arm of a Michelson interferometer.

The result for generic values of $r_1, r_2$ (but still such that $\mathcal{F} \gg 1$) can be conveniently written observing that, for large $\mathcal{F}$, eq. (9.84) can be inverted to give

$$r_1r_2 = 1 - \frac{\pi}{\mathcal{F}} + O\left(\frac{\pi^2}{\mathcal{F}^2}\right). \quad (9.92)$$

We define $p$ from

$$(1-p_1)r_2^2 = (1-p), \quad (9.93)$$

and we introduce the *coupling rate* $\sigma$,

$$\boxed{\sigma = \frac{p\mathcal{F}}{\pi}.} \quad (9.94)$$

From the condition $r_1^2 = 1 - p_1^2 - t_1^2 < 1 - p_1$ it follows that $r_1^2r_2^2 < 1 - p$ and for small $p$ (typical values in VIRGO and LIGO are $p \sim 2 \times 10^{-5}$)

we have $r_1 r_2 < 1 - (p/2)$. Using eq. (9.92) we then obtain

$$1 - \frac{\pi}{\mathcal{F}} < 1 - \frac{p}{2}, \tag{9.95}$$

which, in terms of $\sigma$, gives $\sigma < 2$. Since of course $\sigma > 0$, we have

$$0 < \sigma < 2. \tag{9.96}$$

Writing $2k_L L = 2\pi n + \epsilon$ and expanding for small $\epsilon$, eqs. (9.89) and (9.90) become $\tan\phi_1 = (\mathcal{F}\epsilon/\pi)/(1-\sigma)$ and $\tan\phi_2 = -\mathcal{F}\epsilon/\pi$, so $\phi_2 \simeq -\pi + \arctan(\mathcal{F}\epsilon/\pi)$. For $\phi = \phi_1 - \phi_2$ we get

$$\boxed{\phi = \pi + \arctan\left[\frac{\mathcal{F}\epsilon}{\pi}\frac{1}{1-\sigma}\right] + \arctan\left[\frac{\mathcal{F}\epsilon}{\pi}\right].} \tag{9.97}$$

When $\sigma > 1$ this is rewritten more conveniently as

$$\phi = \arctan\left[\frac{\mathcal{F}\epsilon}{\pi}\frac{1}{\sigma-1}\right] + \arctan\left[\frac{\mathcal{F}\epsilon}{\pi}\right]. \tag{9.98}$$

Observe from eq. (9.70) that, at the resonances, the reflected electric field is

$$E_{\rm refl} = E_0 e^{-i\omega_L t_0}\,\frac{r_1 - r_2(1-p_1)}{1 - r_1 r_2}. \tag{9.99}$$

In particular, if $r_1 = r_2(1-p_1)$, at resonance there is no reflected light from the cavity. Physically, what happens is that the light that is immediately reflected back interferes destructively with the light that is reflected after one or more round trips in the cavity.[15] This situation is called the *optimal (or critical) cavity coupling.* Of course, it is optimal from the point of view of the transmitted field since, except for the losses, all incident light finally leaks out from the second mirror. For the arms of a GW interferometer, we will see that we are interested in the reflected signal and therefore we do not want this situation. If $r_1 > r_2(1-p_1)$ the cavity is undercoupled, while for $r_1 < r_2(1-p_1)$ the cavity is overcoupled. In terms of the coupling rate, using the definition (9.94) and neglecting $O(\pi^2/\mathcal{F}^2)$ in eq. (9.92), we have

$$\frac{r_1 - r_2(1-p_1)}{1 - r_1 r_2} = \frac{\sigma - 1}{r_2}, \tag{9.100}$$

so optimal coupling corresponds to $\sigma = 1$, while for $0 < \sigma < 1$ the cavity is overcoupled, and for $1 < \sigma < 2$ the cavity is undercoupled. Observe that Fig. 9.10 refers to an overcoupled cavity. For undercoupled cavities, instead, the region where the phase of the reflected field is very sensitive to changes in $2k_L L$ becomes smaller and smaller, and disappears completely when $\sigma \to 2$. A comparison of $\phi$, as a function of $\epsilon$, for $\sigma < 1$ and for $\sigma > 1$ is shown in Fig. 9.11. Clearly, the sensitivity to a change of $2k_L L$ is higher for an overcoupled cavity. For the arms of VIRGO and LIGO, the losses are such that $p \sim 2 \times 10^{-5}$ and the finesse is $\mathcal{F} \simeq 50$ for VIRGO and $\mathcal{F} \simeq 200$ for LIGO, so we have $\sigma \sim 3 \times 10^{-4}$ for VIRGO and $\sigma \sim 10^{-3}$ for LIGO. Therefore these cavities are well overcoupled.

[15]This at first sight can be surprising. If for instance $r_1 = 0.99$, almost all the incoming light is reflected back immediately and is not so intuitive that the total reflected field can be zero. What happens is that the small amount of light that enters the cavity eventually builds up a sufficiently strong interior cavity field, and the part of it that finally leaks back from the first mirror has a large enough amplitude, and the appropriate phase, to cancel the promptly reflected field.

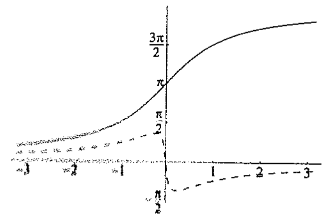


**Fig. 9.11** The phase $\phi$ of the reflected field, as a function of $\epsilon = 2k_L L - 2\pi n$, for an overcoupled cavity with $\sigma = 0.05$ (solid line) and for an undercoupled cavity with $\sigma = 1.05$ (dashed line).



### 9.2.2 Interaction of a FP cavity with GWs

We have seen that the effective storage time of light, which in the arm of a Michelson interferometer is $2L/c$, becomes $(L/c)\mathcal{F}/\pi$ in a Fabry–Perot cavity, i.e. is enhanced by a factor $\mathcal{F}/(2\pi)$, and the sensitivity to a phase shift is enhanced by a factor $(2/\pi)\mathcal{F}$. Since we finally measure a phase shift, we can expect that the same response to GWs of a Michelson interferometer with arm-length of hundreds of kms, as would be optimal for GWs with frequency $f_{\rm gw} = O(10^2)$ Hz, should be obtained replacing the arms by Fabry–Perot cavities with a length of a few kms, and a finesse $\mathcal{F} = O(10^2)$. Thus, our next approximation toward a realistic GW interferometer is as in Fig. 9.12. In this section we study the response of a FP interferometer to GWs, and we will see that the above expectation is indeed correct.

We want to compute how the reflected field of a FP cavity is affected by an incoming GW. We consider a FP cavity oriented along the $x$ axis and a GW with only the plus polarization propagating along $z$, as in eq. (9.6). We begin with a description in the proper detector frame. As we saw in Section 9.1.2, in this frame we can easily obtain the result to lowest order in $\omega_{\rm gw}L/c$ by making use of the fact that, even in the presence of GWs, light propagates along the geodesics of flat space-time, while the mirrors are shaken by a force exerted by GWs, so that their motion is given in eq. (9.37). Therefore the length $L$ of the cavity changes as

$$\Delta L_x(t) = \frac{Lh_0}{2}\cos\omega_{\rm gw}t. \tag{9.101}$$

This induces a change $\Delta\phi_x$ in the phase $\phi_x$ of the field reflected from the cavity along the $x$ arm, which is obtained from eq. (9.91), i.e. from $\Delta\phi_x = (2\mathcal{F}/\pi)\epsilon$, setting $\epsilon = 2k_L\Delta L$,

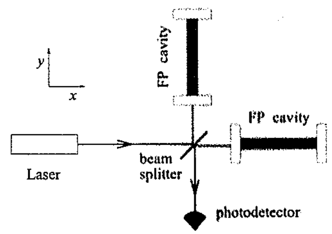


**Fig. 9.12** The layout of an interferometer with Fabry–Perot cavities.

$$\begin{aligned}\Delta\phi_x &\simeq \frac{4\mathcal{F}}{\pi}k_L\Delta L \\ &= \frac{2\mathcal{F}}{\pi}k_L L h_0\cos\omega_{\rm gw}t.\end{aligned} \tag{9.102}$$

The phase shift of a FP cavity along the $y$ arm is obtained reversing the sign of $h_0$ (see eq. (9.7)), so the total phase shift in the Fabry–Perot interferometer of Fig. 9.12 is $\Delta\phi_{\rm FP} = \Delta\phi_x - \Delta\phi_y = 2\Delta\phi_x$. We write $\Delta\phi_{\rm FP}(t) = |\Delta\phi_{\rm FP}|\cos\omega_{\rm gw}t$, so

$$|\Delta\phi_{\rm FP}| = \frac{4\mathcal{F}}{\pi}k_L L h_0. \tag{9.103}$$

This is the change of phase that would be induced in a Michelson interferometer with arm-length $(2/\pi)\mathcal{F}L$. Similar to what happens in a Michelson interferometer, we expect that, when the storage time $\tau_s$ given in eq. (9.88) becomes comparable to the period of the GW, the sensitivity degrades because we are summing over contributions with both positive and negative sign, so the above result is really the lowest order in an expansion in $\omega_{\rm gw}\tau_s$. To compute the result for $\omega_{\rm gw}\tau_s$ generic, we already know from our discussion in Section 9.1.2 that we cannot work in the proper detector frame, and we must rather switch to a TT gauge description.

First, it is useful to observe that, for a FP cavity, we can repeat without any modification the derivation done in eq. (9.34) for an arm of a Michelson interferometer, and we again conclude that, if a GW induces a phase shift $\Delta\phi_x(t) = |\Delta\phi_x| \cos\omega_{\rm gw}t$ in the field reflected from a cavity along the $x$ axis, this produces in the reflected field sidebands with frequencies $\omega_{\rm L} \pm \omega_{\rm gw}$ and an amplitude, relative to the carrier, whose modulus is $|\Delta\phi_x|/2$ in each sideband. Thus, to compute the phase shift $|\Delta\phi_x|$ of the reflected field, to all order in $\omega_{\rm gw}\tau_s$, we can compute in the TT gauge the amplitude of the sidebands of the reflected field. This can be done generalizing the computation of pages 483–485 as follows.

Consider the electric field coming on the first mirror of the cavity from the laser, as in Fig. 9.8. This incoming field is monochromatic, and oscillates as $e^{-i\omega_{\rm L}t}$. When it enters the cavity and bounces once back and forth, besides acquiring the usual transmission and reflection coefficients, when a GW is present it also acquires a phase modulation, so that when it comes back to the first mirror it consists of the carrier at frequency $\omega_{\rm L}$ plus the two sidebands at $\omega_{\rm L} \pm \omega_{\rm gw}$. These three monochromatic fields are partly reflected, with the usual coefficient $-r_1$, and can make one more round trip in the cavity, and so on. So, we need to know how a generic field with carrier plus sidebands is modified by a round trip. We therefore consider a right-moving electromagnetic field which, at the left mirror, has the time-dependence

$$A(t) = A_0 e^{-i\omega_{\rm L}t} + \frac{1}{2}h_0 A_1 e^{-i(\omega_{\rm L}-\omega_{\rm gw})t} + \frac{1}{2}h_0 A_2 e^{-i(\omega_{\rm L}+\omega_{\rm gw})t}, \quad (9.104)$$

while we denote by $B(t)$ the right-moving field at the end of the round trip,

$$B(t) = B_0 e^{-i\omega_{\rm L}t} + \frac{1}{2}h_0 B_1 e^{-i(\omega_{\rm L}-\omega_{\rm gw})t} + \frac{1}{2}h_0 B_2 e^{-i(\omega_{\rm L}+\omega_{\rm gw})t}. \quad (9.105)$$

If we denote by $t$ the time at which the field terminates its round-trip, the time $t_0$ at which it started is given by (compare with eq. (9.18))

$$t_0 = t - \frac{2L}{c} - \frac{L}{c} h_0 \cos[\omega_{\rm gw}(t - L/c)] \,\mathrm{sinc}\,(\omega_{\rm gw}L/c). \quad (9.106)$$

Since during free propagation the phase is unchanged, we must have (apart from the reflection coefficients at the mirrors that we will add separately) $B(t) = A(t_0)$,[16] that is

$$B(t) = A_0 e^{-i\omega_{\rm L}t_0} + \frac{1}{2}h_0 A_1 e^{-i(\omega_{\rm L}-\omega_{\rm gw})t_0} + \frac{1}{2}h_0 A_2 e^{-i(\omega_{\rm L}+\omega_{\rm gw})t_0}. \quad (9.107)$$

Using eq. (9.106) and developing to first order in $h_0$,

$$\begin{aligned} e^{-i\omega_{\rm L}t_0} &= e^{-i\omega_{\rm L}(t-2L/c)} \\ &+\frac{1}{2}h_0\, ik_{\rm L}L \,\mathrm{sinc}\,(\omega_{\rm gw}L/c) e^{i(2\omega_{\rm L}-\omega_{\rm gw})L/c} e^{-i(\omega_{\rm L}-\omega_{\rm gw})t} \\ &+\frac{1}{2}h_0\, ik_{\rm L}L \,\mathrm{sinc}\,(\omega_{\rm gw}L/c) e^{i(2\omega_{\rm L}+\omega_{\rm gw})L/c} e^{-i(\omega_{\rm L}+\omega_{\rm gw})t}. \end{aligned} \quad (9.108)$$

[16] Recall that the superposition of carrier and sidebands given in eqs. (9.104) and (9.105) derives from the expansion of a phase factor, see eqs. (9.24) and (9.34).



Again to order $h_0$, we can simply replace the terms $h_0 e^{-i(\omega_{\rm L}\pm\omega_{\rm gw})t_0}$ in eq. (9.107) by $h_0 e^{-i(\omega_{\rm L}\pm\omega_{\rm gw})(t-2L/c)}$. Collecting terms with the same time dependence in eq. (9.107) and comparing with eq. (9.105) we get a matrix relation $B_i = X_{ij}A_j$ (with $i = 0, 1, 2$), where

$$\mathbf{X} = \begin{pmatrix} X_{00} & 0 & 0 \\ X_{10} & X_{11} & 0 \\ X_{20} & 0 & X_{22} \end{pmatrix}. \quad (9.109)$$

The diagonal elements describe the free propagation of the carrier and of the sidebands, while the $X_{10}$ and $X_{20}$ term describe the fact that a round-trip of the carrier produces further contributions to the sidebands. Using eq. (9.108), the explicit expression of the matrix elements is

$$\begin{aligned} X_{00} &= e^{2i\omega_{\rm L}L/c}, \\ X_{11} &= e^{2i(\omega_{\rm L}-\omega_{\rm gw})L/c}, \\ X_{22} &= e^{2i(\omega_{\rm L}+\omega_{\rm gw})L/c}. \\ X_{10} &= ik_{\rm L}L \,\mathrm{sinc}\,(\omega_{\rm gw}L/c) e^{i(2\omega_{\rm L}-\omega_{\rm gw})L/c}, \\ X_{20} &= ik_{\rm L}L \,\mathrm{sinc}\,(\omega_{\rm gw}L/c) e^{i(2\omega_{\rm L}+\omega_{\rm gw})L/c}. \end{aligned} \quad (9.110)$$

For a Fabry–Perot cavity along the $y$ axis the same expressions hold, inverting the sign of $h_0$ (see eq. (9.7)) or, equivalently, inverting the sign of $X_{10}$ and of $X_{20}$.

This result allows us to generalize eq. (9.80) to the case when GWs are present, simply replacing the factors $e^{2ik_{\rm L}L}$ with the matrix $\mathbf{X}$. Thus, we can write the fields $\mathcal{B} = (\mathcal{B}_0, \mathcal{B}_1, \mathcal{B}_2)$ inside the cavity, at $x = 0$, in matrix form as

$$\mathcal{B} = t_1\mathcal{A}_{\rm in} + r_1 r_2 \mathbf{X}\mathcal{B}, \quad (9.111)$$

where $\mathcal{A}_{\rm in} = (E_0, 0, 0)$. The solution is

$$\mathcal{B} = (\mathbf{1} - r_1 r_2 \mathbf{X})^{-1} t_1 \mathcal{A}_{\rm in}. \quad (9.112)$$

This is the right-moving field at the first mirror (the equivalent of what we denoted by $E_{\rm cav}(0)$ in the absence of GWs, see Fig. 9.6). The left-moving field ($E'_{\rm cav}(0)$ in Fig. 9.6) in the absence of GWs is obtained from $E_{\rm cav}(0)$ using eq. (9.79). In the presence of GWs, we have seen that the factor $e^{2ik_{\rm L}L}$ is replaced by the matrix $\mathbf{X}$, acting on the vector space of the amplitude of the carrier and of the sidebands, so the left-moving field is now

$$\mathcal{B}' = -r_2\mathbf{X}\mathcal{B}, \quad (9.113)$$

and the total reflected field, which includes also the promptly reflected part, is given by

$$\begin{aligned} \mathcal{A}_{\rm refl} &= r_1\mathcal{A}_{\rm in} - t_1 r_2 \mathbf{X}\mathcal{B} \\ &= [r_1 - r_2(1-p_1)\mathbf{X}](\mathbf{1} - r_1 r_2\mathbf{X})^{-1}\mathcal{A}_{\rm in}, \end{aligned} \quad (9.114)$$

which replaces eq. (9.70). Setting $\mathcal{A}_{\rm in} = (1, 0, 0)$, we can now compute $\mathcal{A}_{\rm refl} \equiv (\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2)$. According to eq. (9.34), and taking into account the factor $h_0/2$ in the definition (9.104), the phase shift $|\Delta\phi_x|$ in a single Fabry–Perot cavity along the $x$ axis, is given by

$$\frac{1}{2}|\Delta\phi_x| = \frac{1}{2}h_0 \left|\frac{\mathcal{A}_1}{\mathcal{A}_0}\right| . \tag{9.115}$$

We are interested in particular in the situation when the FP cavity is locked on resonance, so $e^{2ik_L L} = 1$. In this case, with straightforward matrix algebra (easily performed with the help of any symbolic manipulation program) we get

$$\frac{\mathcal{A}_1}{\mathcal{A}_0} = X_{10}\, e^{2i\omega_{\rm gw}L/c} \frac{r_2(1-p) - r_1^2 r_2}{(e^{2i\omega_{\rm gw}L/c} - r_1 r_2)[r_2(1-p) - r_1]} , \tag{9.116}$$

so

$$\begin{aligned}|\Delta\phi_x| &= h_0 k_L L\, \text{sinc}\,(\omega_{\rm gw}L/c)\, \frac{r_2(1-r_1^2-p)}{[r_2(1-p)-r_1]}\, \frac{1}{|e^{2i\omega_{\rm gw}L/c} - r_1 r_2|} \\ &= h_0 k_L L\, \text{sinc}\,(\omega_{\rm gw}L/c)\, \frac{r_2(1-r_1^2-p)}{[r_2(1-p)-r_1]} \\ &\times \frac{1}{[1+(r_1 r_2)^2 - 2 r_1 r_2 \cos(2\omega_{\rm gw}L/c)]^{1/2}} .\end{aligned} \tag{9.117}$$

If we set $p = 0$ and $r_2 = 1$ (e.g. the present value for VIRGO is $r_2 \simeq 0.99995$) and we take $r_1$ close to one, the first fraction becomes simply $1 + r_1 \simeq 2$. So, we write

$$\frac{r_2(1-r_1^2-p)}{[r_2(1-p)-r_1]} = 2[1 + \epsilon(r_1, r_2, p)] , \tag{9.118}$$

where, in the typical experimental situation, $\epsilon(r_1, r_2, p) \ll 1$. Then

$$|\Delta\phi_x| = h_0\, 2k_L L\, [1 + \epsilon(r_1, r_2, p)] \frac{\text{sinc}\,(\omega_{\rm gw}L/c)}{[1+(r_1 r_2)^2 - 2 r_1 r_2 \cos(2\omega_{\rm gw}L/c)]^{1/2}} . \tag{9.119}$$

The dependence on $\omega_{\rm gw}L/c$ can be simplified observing that we want to have $\mathcal{F}L/c$ comparable to the wavelength of the GW, so $\mathcal{F}\omega_{\rm gw}L/c = O(1)$. However, we achieve this by using a large value of $\mathcal{F}$, so $\omega_{\rm gw}L/c$ is much smaller than one in the region where the interferometer operates. For instance, if $f_{\rm gw} = 100$ Hz and $L = 4$ km, $\omega_{\rm gw}L/c \sim 10^{-2}$. We can therefore replace sinc $(\omega_{\rm gw}L/c) \simeq 1$ in the numerator, and we expand $\cos(2\omega_{\rm gw}L/c)$ in the denominator. Then we get

$$\begin{aligned}|\Delta\phi_x| &\simeq h_0\, 2k_L L \frac{1+\epsilon(r_1,r_2,p)}{1 - r_1 r_2} \frac{1}{[1 + \frac{r_1 r_2}{(1-r_1 r_2)^2}(2\omega_{\rm gw}L/c)^2]^{1/2}} \\ &\simeq h_0\, 2k_L L \frac{\mathcal{F}}{\pi} \frac{1}{[1 + (4\pi f_{\rm gw}\tau_s)^2]^{1/2}} ,\end{aligned} \tag{9.120}$$

where, in the last line, we wrote the result in terms of the finesse $\mathcal{F}$, given in eq. (9.84), and of the storage time $\tau_s$ of the cavity, given in eq. (9.87), and we neglected in the numerator terms that are small when $r_2, r_1 \to 1$. The phase shift of a FP cavity along the $y$ arm is obtained changing the sign of $h_0$, so $\Delta\phi_y = -\Delta\phi_x$, and the difference between them is $\Delta\phi_{\rm FP} = \Delta\phi_x - \Delta\phi_y = 2\Delta\phi_x$.



We rewrite the result introducing the so-called *pole frequency*,

$$\boxed{f_p \equiv \frac{1}{4\pi\tau_s} ,} \tag{9.121}$$

or, from eq. (9.88),

$$f_p \simeq \frac{c}{4\mathcal{F}L} . \tag{9.122}$$

For initial LIGO, $L = 4$ km and $\mathcal{F} \simeq 200$ this gives $f_p \simeq 90$ Hz. For VIRGO, $L = 3$ km and $\mathcal{F} \simeq 50$, so $f_p \simeq 500$ Hz. The phase shift in a Fabry–Perot interferometer can then be written as

$$\boxed{|\Delta\phi_{\rm FP}| \simeq h_0 \frac{4\mathcal{F}}{\pi} k_L L \frac{1}{\sqrt{1 + (f_{\rm gw}/f_p)^2}} .} \tag{9.123}$$

For $f_{\rm gw} \ll f_p$ we recover the result found in the proper detector frame, eq. (9.103), as expected.[17] At $f_{\rm gw} \gg f_p$, eq. (9.123) shows that the sensitivity degrades linearly with $f_{\rm gw}$. This formula holds as long as $\omega_{\rm gw}L/c \ll 1$, i.e

$$\begin{aligned} f_{\rm gw} &\ll \frac{c}{2\pi L} \\ &\simeq 12 \text{ kHz} \left(\frac{4\,\text{km}}{L}\right) . \end{aligned} \tag{9.124}$$

Above this frequency the factor sinc $(\omega_{\rm gw}L/c)$ in eq. (9.119) can no longer be approximated by one, and cuts the response further, reflecting the fact that in each round-trip the GW changes sign.

In Fig. 9.13 we show the function $1/[1 + (f_{\rm gw}/f_p)^2]^{1/2}$, and we compare it with the function $|\text{sinc}\,(f/f_p)|$, which is the corresponding quantity for a Michelson interferometer whose length $L_{\rm Mich} = (2/\pi)\mathcal{F}L$ is chosen so that, in the limit $f_{\rm gw} \to 0$, its response function is the same as a FP cavity of length $L$ and finesse $\mathcal{F}$.

It is useful to write eq. (9.123) in the form

$$|\Delta\phi_{\rm FP}| = h_0 T_{\rm FP}(f) , \tag{9.125}$$

where (writing $k_L = 2\pi/\lambda_L$)

$$\boxed{T_{\rm FP}(f) \simeq \frac{8\mathcal{F}L}{\lambda_L} \frac{1}{\sqrt{1 + (f_{\rm gw}/f_p)^2}} ,} \tag{9.126}$$

is the transfer function of an interferometer with Fabry–Perot cavities.

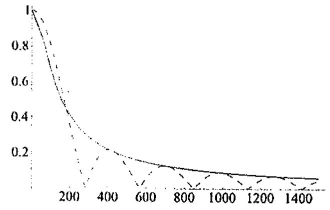


**Fig. 9.13** A plot of the function $[1 + (f/f_p)^2]^{-1/2}$, (solid line), compared to the function $|\text{sinc}\,(f/f_p)|$ (dashed line). We have taken $f_p = 90$ Hz.

[17] Recall that eq. (9.91), and therefore eq. (9.103), were obtained in the limit $r_2 = 1$, $p = 0$ and $r_1$ close to one. If we keep $r_1$ generic, still setting $r_2 = 1$ and $p = 0$, in eq. (9.91) the overall factor of 2 is replaced by $1 + r_1$, and the same result is obtained from eq. (9.117).

## 9.2.3 Angular sensitivity and pattern functions

Until now we have restricted ourselves to a GW with plus polarization, propagating along the $z$ axis. We now compute the response of an interferometer to GWs with arbitrary direction and polarization. As discussed in Section 7.2, this is encoded in the pattern functions $F_+(\theta,\phi)$ and $F_\times(\theta,\phi)$. We first consider the limit $\omega_{\rm gw}L/c \ll 1$. In this case we can use the proper detector frame, so the motion of the mirrors is governed by the geodesic equation,

$$\ddot{\xi}^i = \frac{1}{2}\ddot{h}_{ij}\xi^j\,. \tag{9.127}$$

For the mirror located at $\xi^j = (L,0,0)$, we are interested in its displacement along the $x$ direction, which is given by

$$\ddot{\xi}_x = \frac{1}{2}\ddot{h}_{xx}L\,. \tag{9.128}$$

This equation governs the change in the length of the $x$-arm of a Michelson interferometer, as well as the change in the length of a FP cavity lying along the $x$ axis. For the mirror located at $\xi^j = (0,L,0)$, we are rather interested in its displacement along the $y$ direction, which is given by

$$\ddot{\xi}_y = \frac{1}{2}\ddot{h}_{yy}L\,. \tag{9.129}$$

The relative phase shift between the $x$ and $y$ arms is therefore driven by $(1/2)(\ddot{h}_{xx} - \ddot{h}_{yy})$. When the wave comes from the $z$ direction we have $h_{xx} = h_+$ and $h_{xx} = -h_+$, so $(1/2)(h_{xx} - h_{yy}) = h_+$, but in the most general situation we must replace $h_+$ by $(1/2)(h_{xx} - h_{yy})$ in the computations of the phase shift in a Michelson or in a FP interferometer performed in the previous sections. In other words, the detector tensor (defined in eq. (7.1)) for an interferometer with arms along the $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ directions is

$$D_{ij} = \frac{1}{2}(\hat{\mathbf{x}}_i\hat{\mathbf{x}}_j - \hat{\mathbf{y}}_i\hat{\mathbf{y}}_j)\,. \tag{9.130}$$

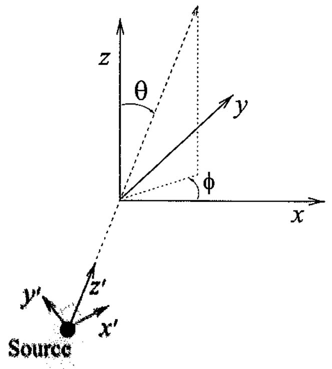


**Fig. 9.14** The geometry used in the computation of the pattern functions. The arms of the interferometer are along the $x$ and $y$ axes.

We compute $h_{xx}$ and $h_{yy}$ in terms of $h_+, h_\times$ for a wave coming from arbitrary direction. The computation is similar to that performed for resonant bars on page 425. The geometry is illustrated in Fig. 9.14: we have a frame $(x,y,z)$ such that the arms of the interferometer are along the $x$ and $y$ axes. We introduce a second reference frame $(x',y',z')$ such that the propagation direction of the GW coincides with the $z'$ axis. With respect to the $(x,y,z)$ frame, the $z'$ axis has polar angles $\theta$ and $\phi$, defined as in the figure.[18]

The polarizations $h_+$ and $h_\times$ are defined with respect to the $(x',y')$ axes, so in the $(x',y',z')$ frame the GW has the form

$$h'_{ij} = \begin{pmatrix} h_+ & h_\times & 0 \\ h_\times & -h_+ & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij}\,. \tag{9.131}$$

[18] When comparing with the calculation for resonant bars on page 425, observe that here we define $\theta$ as the angle from the $z$ axis, rather than from the $x$ axis.



The rotation that brings the $(x',y',z')$ frame onto the $(x,y,z)$ frame is given by a rotation by an angle $\theta$ around the $y$ axis followed by a rotation by an angle $\phi$ around the $z$ axis, i.e.

$$\mathcal{R} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}\,. \tag{9.132}$$

The GW in the $(x,y,z)$ frame is then given by the transformation law of a tensor with two indices, $h_{ij} = \mathcal{R}_{ik}\mathcal{R}_{jl}h'_{kl}$. From this we obtain

$$h_{xx} = h_+(\cos^2\theta\cos^2\phi - \sin^2\phi) + 2h_\times\cos\theta\sin\phi\cos\phi\,, \tag{9.133}$$

$$h_{yy} = h_+(\cos^2\theta\sin^2\phi - \cos^2\phi) - 2h_\times\cos\theta\sin\phi\cos\phi\,, \tag{9.134}$$

so

$$\frac{1}{2}(h_{xx} - h_{yy}) = \frac{1}{2}h_+(1+\cos^2\theta)\cos 2\phi + h_\times\cos\theta\sin 2\phi\,, \tag{9.135}$$

and therefore

$$F_+(\theta,\phi) = \frac{1}{2}(1+\cos^2\theta)\cos 2\phi\,,$$
$$F_\times(\theta,\phi) = \cos\theta\sin 2\phi\,. \tag{9.136}$$

We see that GW interferometers have blind directions. For instance, for a GW with plus polarization, the direction with $\phi = \pi/4$ is blind, since $F_+ = 0$. This is due to the fact that this wave produces the same displacement in the $x$ and in the $y$ arm, so the differential phase shift vanishes. If we change the definition of the axes with respect to which the polarizations $h_+$ and $h_\times$ are defined, rotating them by an angle $\phi$ in the $(x',y')$ plane, the pattern functions transform as in eq. (7.30).

Equation (9.136) has been obtained in the limit $\omega_{\rm gw}L/c \ll 1$. To compute the pattern functions for $\omega_{\rm gw}L/c$ generic we must perform the computation in the TT gauge, so we should repeat the computation leading to eq. (9.15) for a GW coming from arbitrary direction. Consider the arm of a simple Michelson interferometers, with the beam splitter at $x = 0$ and the far mirror at $x = L_x$ (or a FP cavity with mirrors at $x = 0$ and $x = L_x$). Then eq. (9.9) is replaced by

$$L_x = c(t_1 - t_0) - \frac{c}{2}\int_{t_0}^{t_1} dt'\, h_{xx}(t', \mathbf{x})\,. \tag{9.137}$$

If we denote by $\hat{\mathbf{n}}$ the propagation direction of the GW, we have $h_{xx}(t) = h_{xx}\cos[\omega_{\rm gw}(t - \hat{\mathbf{n}}\cdot\mathbf{x}/c)]$, and we must evaluate $\mathbf{x}$ on the trajectory $\mathbf{x}(t)$ of the photon, so along the $x$ arm we have

$$h_{xx}(t) = h_{xx}\cos\left[\omega_{\rm gw}\left(t - \frac{n_x x(t)}{c}\right)\right]\,, \tag{9.138}$$

which replaces eq. (9.6). To lowest order in $h_{xx}$ the trajectory of a photon is just the unperturbed one, so inside the cosine we can set $x(t) = c(t - t_0)$, while $\hat{n}$ can be written in terms of the angles $\theta, \phi$ as $\hat{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. Therefore, beside the dependence on $(\theta, \phi)$ in $h_{xx}$ and $h_{yy}$, that we already computed in eqs. (9.133) and (9.134), there is also an angular dependence through the term $n_x$ in eq. (9.138), and a similar term $n_y$ for the $y$ arm. In particular, eq. (9.137) becomes

$$L_x = c(t_1 - t_0) - \frac{c}{2} h_{xx} \int_{t_0}^{t_1} dt' \times \cos\left[(1 - \sin\theta\cos\phi)\omega_{\rm gw} t' + \omega_{\rm gw} t_0 \sin\theta\cos\phi\right] . \tag{9.139}$$

The return trip can be treated similarly, with the unperturbed photon trajectory given now by $x(t) = L - c(t - t_1)$, so eq. (9.11) is replaced by

$$L_x = c(t_2 - t_1) - \frac{c}{2} h_{xx} \int_{t_1}^{t_2} dt' \times \cos\left[(1 + \sin\theta\cos\phi)\omega_{\rm gw} t' - \omega_{\rm gw}(t_1 + L/c) \sin\theta\cos\phi\right] . \tag{9.140}$$

Summing the two equations we get

$$t_2 = t_0 + \frac{2L_x}{c} + \frac{1}{2} h_{xx} \int_{t_0}^{t_0 + L_x/c} dt' \cos[\omega_- t' + \phi_0] + \frac{1}{2} h_{xx} \int_{t_0 + L_x/c}^{t_0 + 2L_x/c} dt' \cos[\omega_+ t' - \phi_2] , \tag{9.141}$$

where we introduced the short-hand notation

$$\omega_{\pm} = \omega_{\rm gw}(1 \pm \sin\theta\cos\phi) , \tag{9.142}$$

$$\phi_0 = \omega_{\rm gw} t_0 \sin\theta\cos\phi , \tag{9.143}$$

$$\phi_2 = \omega_{\rm gw} t_2 \sin\theta\cos\phi , \tag{9.144}$$

and in the limits of the integral, as well as in $\phi_2$, we can use $t_1 = t_0 + L_x/c$ and $t_2 = t_0 + 2L_x/c$. For the $y$ arm we have similar expressions, with $L_y$ replacing $L_x$ and $n_y = \sin\theta\sin\phi$ replacing $n_x = \sin\theta\cos\phi$.

It is now in principle straightforward to perform the integrals and compute how $t_2 - t_0$ depends on the propagation direction of the GW. Carrying out the integrals, however, we see that all terms which depend on $\theta, \phi$ are multiplied by the factor $\omega_{\rm gw} L_x/c$. For instance, $\theta$ and $\phi$ enter in terms such as

$$\text{sinc}\left[\frac{\omega_{\rm gw} L_x}{2c}(1 \pm \sin\theta\cos\phi)\right] . \tag{9.145}$$

For a FP interferometer we saw that $\omega_{\rm gw} L_x/c$ is small, typically $O(10^{-2})$ in LIGO and VIRGO, and therefore the function sinc in eq. (9.145) is essentially unity, and its dependence on $\theta$ and $\phi$ is negligible, at least as long as the condition (9.124) is satisfied. Then, we can neglect the dependence on the GW direction in the travel time $t_2 - t_0$ and the only angular dependence comes from $h_{xx}$ and $h_{yy}$, as computed in eqs. (9.133) and (9.134), so for the pattern function we can use, to a very good approximation, the expressions given in eq. (9.136).



## 9.3 Toward a real GW interferometer

In this section we discuss a number of issues that are more technical, but are important for understanding how a real interferometer works.

### 9.3.1 Diffraction and Gaussian beams

Until now we have considered idealized FP cavities with mirrors of infinite transverse extent, so we could neglect any dependence of the electric field on the transverse coordinates. For a cavity along the $x$ axis, we have then treated the interior electric field as a plane wave, with a dependence on $x, t$ of the form $\exp\{-i\omega_L(t \pm x/c)\}$, and no dependence on the transverse coordinates $\mathbf{x}_\perp = (y, z)$. Of course, in practice the mirrors have a finite extent, and the beam has a profile in the transverse direction.

A beam of finite transverse extent is subject to diffraction. If, at some point in space, a photon of wavelength $\lambda_{\rm L}$ (and therefore longitudinal momentum $p = \hbar/\lambdabar_L$) is localized within a transverse width $\Delta \mathbf{x}_\perp = a$, by the Heisenberg principle it has an uncertainty on the transverse momentum $\Delta p_\perp \sim \hbar/a$, so the beam will widen, filling a cone of angle $\Delta\theta = \Delta p_\perp/p \sim \lambdabar_L/a$. After traveling a longitudinal distance $x$ the beam has become larger, in the transverse direction, by $x\Delta\theta \sim x\lambdabar_L/a$. As long as $x\lambdabar_L/a \ll a$ we are in the regime of *Fresnel diffraction*, and the broadening of the beam is negligible. When $x\lambdabar_L/a \gg a$, or, in terms of $k_L = 1/\lambdabar_L$,

$$x \gg k_{\rm L} a^2 , \tag{9.146}$$

we are in the regime of *Fraunhofer diffraction*, and the beam has become much broader than its original size. For interferometers such as LIGO and VIRGO, the wavelength of the laser is typically

$$\lambda_{\rm L} \simeq 1\,\mu m . \tag{9.147}$$

The border between these regimes is at $a = (x\lambdabar)^{1/2}$ which, for $x = 4$ km and $\lambda_{\rm L} = 1\,\mu$m, gives $a \simeq 2.5$ cm. This means that, for a laser beam whose initial width is smaller that 2.5 cm, the broadening of the beam becomes important already after a single one-way trip through the cavity. Furthermore for cavities with a finesse $O(100)$, as we need for GW detection, the beam is supposed to perform $O(100)$ round trips and, if the mirrors were flat, at each one-way trip the beam would widen further, as illustrated in Fig. 9.15, and would be finally dispersed on a region of transverse size larger than the mirrors.

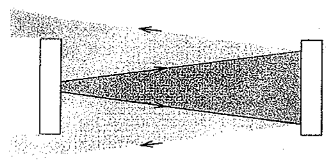

Fig. 9.15 The widening of a beam due to diffraction as it bounces between two flat mirrors.

Thus, it is clear that diffraction effects are important, and the naive scheme of a narrow beam (as typically obtained from a laser) bouncing between two flat mirrors cannot work. As a first step, we must understand in more detail the propagation of a beam of finite transverse extent over large distances. The tool that we need is the *paraxial propagator*, that we introduce in the next subsection.

## The paraxial propagator

A given spatial component of the electric field, propagating in the vacuum, obeys the equation

$$\left[-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}+\nabla^2\right]E(t,\mathbf{x})=0\,. \tag{9.148}$$

We take a single monochromatic component, which we write in complex notation as $E(t,\mathbf{x})=E(\mathbf{x})e^{-i\omega_{\rm L}t}$, so $E(\mathbf{x})$ satisfies

$$[\nabla^2+k_{\rm L}^2]E(\mathbf{x})=0\,. \tag{9.149}$$

We want to compute the propagation across a long distance in the $x$ direction, so $x$ is the longitudinal coordinate and $\mathbf{x}_\perp=(y,z)$ are the transverse coordinate, and we search for solutions of the form

$$E(\mathbf{x})=\mathcal{E}(x;y,z)e^{ik_{\rm L}x}\,, \tag{9.150}$$

where $\mathcal{E}(x;y,z)$ is a slowly varying function of $x$, in the sense that

$$|\partial_x\mathcal{E}|\ll k_{\rm L}|\mathcal{E}|\,. \tag{9.151}$$

Therefore $E(t,\mathbf{x})=\mathcal{E}(x;y,z)\exp\{-i\omega_{\rm L}t+ik_{\rm L}x\}$ is in a first approximation a plane wave, with a slower dependence on $x$, which manifests itself only on scales $x\gg\lambda_{\rm L}$. Plugging the ansatz (9.150) into eq. (9.149) we get

$$\nabla_\perp^2\mathcal{E}+2ik_{\rm L}\partial_x\mathcal{E}+\partial_x^2\mathcal{E}=0\,, \tag{9.152}$$

where $\nabla_\perp^2=\partial_y^2+\partial_z^2$. Because of the condition (9.151), we can neglect $\partial_x^2\mathcal{E}$ with respect to $k_{\rm L}\partial_x\mathcal{E}$, so in this approximation we write

$$\nabla_\perp^2\mathcal{E}+2ik_{\rm L}\partial_x\mathcal{E}=0\,. \tag{9.153}$$

We now perform the Fourier transform with respect to the transverse variables,

$$\mathcal{E}(x;y,z)=\int\frac{dp_y}{2\pi}\frac{dp_z}{2\pi}\,\tilde{\mathcal{E}}(x;p_y,p_z)\,e^{ip_yy+ip_zz}\,. \tag{9.154}$$

In terms of $\tilde{\mathcal{E}}(x;p_y,p_z)$, eq. (9.153) reads

$$-(p_y^2+p_z^2)\tilde{\mathcal{E}}(x;p_y,p_z)+2ik_{\rm L}\partial_x\tilde{\mathcal{E}}(x;p_y,p_z)=0\,. \tag{9.155}$$

The $x$ dependence can be integrated, and we get

$$\tilde{\mathcal{E}}(x;p_y,p_z)=\tilde{\mathcal{E}}(x=0;p_y,p_z)\exp\left\{-i\frac{p_y^2+p_z^2}{2k_{\rm L}}x\right\}\,. \tag{9.156}$$

Then eq. (9.154) becomes

$$\begin{aligned}\mathcal{E}(x;y,z)&=\int\frac{dp_y}{2\pi}\frac{dp_z}{2\pi}\,\tilde{\mathcal{E}}(x=0;p_y,p_z)\,e^{ip_yy+ip_zz-i\frac{p_y^2+p_z^2}{2k_{\rm L}}x}\\
&=\int\frac{dp_y}{2\pi}\frac{dp_z}{2\pi}\left[\int dy'dz'\,\tilde{\mathcal{E}}(x=0;y',z')e^{-ip_yy'+ip_zz'}\right]\\
&\quad\times e^{ip_yy+ip_zz-i\frac{p_y^2+p_z^2}{2k_{\rm L}}x}\\
&=\int dy'dz'\,\tilde{\mathcal{E}}(x=0;y',z')\\
&\quad\times\int\frac{dp_y}{2\pi}\frac{dp_z}{2\pi}\,e^{ip_y(y-y')+ip_z(z-z')-i\frac{p_y^2+p_z^2}{2k_{\rm L}}x}\,.\end{aligned} \tag{9.157}$$



The integrals over $dp_y$ and $dp_z$ are Fresnel integrals, that we already met in eq. (4.365), so we finally get

$$\boxed{\mathcal{E}(x;y,z)=\int dy'dz'\,G(x;y-y',z-z')\,\tilde{\mathcal{E}}(x=0;y',z')\,,} \tag{9.158}$$

where

$$\boxed{G(x;y-y',z-z')=\frac{-ik_{\rm L}}{2\pi x}\,\exp\left\{i\frac{k_{\rm L}}{2x}\left[(y-y')^2+(z-z')^2\right]\right\}} \tag{9.159}$$

is called the paraxial propagator. Equations (9.158) and (9.159) allow us to compute the field at $x$ generic, once we have its value on a transverse surface $x=0$.

### Fraunhofer diffraction

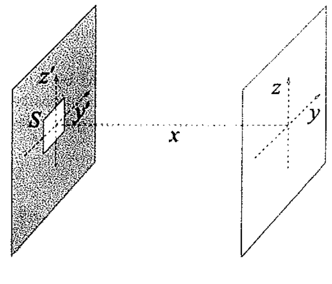


**Fig. 9.16** An aperture $S$ on a opaque screen. The plane of the opaque screen is parametrized by coordinates $(y',z')$. The image is observed on a screen at a distance $x$, parametrized by coordinates $(y,z)$.

As a first application, we consider a plane wave of infinite transverse extent that arrives on an aperture $S$ on a plane opaque screen and we compute the image on another screen at a large distance $x$, and at transverse coordinates $(y,z)$, see Fig. 9.16. Then, at $x=0$, we have $\tilde{\mathcal{E}}(x=0;y',z')=\tilde{\mathcal{E}}_0$ if $(y',z')$ are inside the aperture $S$, and zero otherwise, so

$$\begin{aligned}E(x,y,z)&=\frac{-ik_{\rm L}}{2\pi x}\,\tilde{\mathcal{E}}_0\,e^{ik_{\rm L}x}\int_S dy'dz'\,\exp\left\{i\frac{k_{\rm L}}{2x}[(y-y')^2+(z-z')^2]\right\}\\
&=\frac{-ik_{\rm L}}{2\pi x}\,\tilde{\mathcal{E}}_0\exp\left\{ik_{\rm L}\left[x+\frac{y^2+z^2}{2x}\right]\right\}\\
&\quad\times\int_S dy'dz'\,\exp\left\{-i\frac{k_{\rm L}}{x}(yy'+zz')+i\frac{k_{\rm L}}{x}(y'^2+z'^2)\right\}\,.\end{aligned} \tag{9.160}$$

Fraunhofer diffraction is defined by the condition (9.146), where $a$ is the size of the aperture. In this limit, we can neglect the term $k_{\rm L}(y'^2+z'^2)/x$ in the exponential. Furthermore we observe that, if $y^2+z^2\ll x^2$, the term $x+(y^2+z^2)/2x$ in the first exponential is just the first-order expansion of the distance $r = (x^2 + y^2 + z^2)^{1/2}$ from the observation point to the center of the aperture, since

$$(x^2 + y^2 + z^2)^{1/2} = x\left(1 + \frac{y^2+z^2}{x^2}\right)^{1/2} \simeq x\left(1 + \frac{y^2+z^2}{2x^2}\right). \qquad (9.161)$$

Similarly, to lowest order we can replace $1/x$ with $1/r$ in eq. (9.160). Then we get the well-known formula for the Fraunhofer diffraction by an aperture,

$$E(x,y,z) = \frac{-i\tilde{\mathcal{E}}_0 k_L}{2\pi}\,\frac{e^{ik_L r}}{r}\int_S dy'dz'\, e^{-ik_L(yy'+zz')/r}. \qquad (9.162)$$

Consider for example a circular aperture of radius $a$. In this case the integral can be performed exactly in terms of the Bessel function $J_1$. Writing $y = \rho\cos\varphi, z = \rho\sin\varphi$, and similarly $y' = \rho'\cos\varphi', z' = \rho'\sin\varphi'$, we get

$$\begin{aligned}\int_S dy'dz'\, e^{-ik_L(yy'+zz')/r} &= \int_0^a \rho' d\rho' \int_0^{2\pi} d\varphi'\, e^{-i(k_L/r)\rho\rho'\cos(\varphi-\varphi')}\\ &= 2\pi\int_0^a \rho' d\rho'\, J_0(k_L\rho\rho'/r)\\ &= \frac{2\pi a r}{k_L\rho} J_1(k_L\rho a/r). \end{aligned} \qquad (9.163)$$

Writing $\rho/r = \sin\theta$ and recalling that $\lim_{u\to 0} J_1(u)/u = 1/2$, we see that the intensity of light, which is proportional to the squared modulus of the electric field, is distributed in the scattering angle $\theta$ as[19]

$$I(\theta) = I(0)\left[\frac{2J_1(k_L a\sin\theta)}{k_L a\sin\theta}\right]^2, \qquad (9.164)$$

A plot of the function $2J_1(x)/x$ is shown in Fig. 9.17. $I(\theta)$ has its first zero at $k_L a\sin\theta \simeq 3.8$. Taking this as an estimate of the angular width $\Delta\theta$ of the beam we get (for $ƛ \ll a$) $\Delta\theta \simeq 3.8ƛ/a$, which is consistent with the uncertainty principle bound, but does not saturate it.

[19] This result was first derived by Airy in the 19th century, and this intensity distribution is known as the Airy pattern.

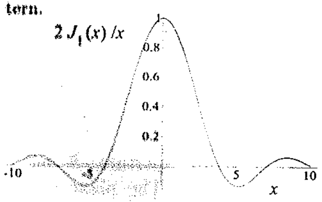


Fig. 9.17 The function $2J_1(x)/x$.

### Propagation of Gaussian beams

Consider now a beam that, at $x = 0$, has a Gaussian profile in the transverse direction,

$$\mathcal{E}(x=0;y,z) = \mathcal{E}_0 e^{-(y^2+z^2)/w_0^2}. \qquad (9.165)$$

Its profile at $x$ generic can be computed by inserting this initial value into eqs. (9.158) and (9.159). The resulting integrals can be computed exactly, without resorting to the Fraunhofer approximation, using

$$\int_{-\infty}^{\infty} e^{-(1+ia^2)y^2} = \frac{\sqrt{\pi}}{(1+a^2)^{1/4}}\exp\left\{-\frac{i}{2}\arctan a\right\}, \qquad (9.166)$$



where $a$ is a real constant. The result, written in terms of $E(x,y,z) = e^{ik_L x}\mathcal{E}(x;y,z)$, is

$$\begin{aligned}E(x,y,z) = &\frac{\mathcal{E}_0}{\sqrt{1+x^2/b^2}} e^{-(y^2+z^2)/w^2(x)}\\ &\times\exp\left\{ik_L\left[x + \frac{y^2+z^2}{2R(x)}\right] - i\arctan(x/b)\right\},\end{aligned} \qquad (9.167)$$

where the *Rayleigh range* $b$ is defined by

$$b = \frac{1}{2}k_L w_0^2, \qquad (9.168)$$

the width $w(x)$ is given by

$$w(x) = w_0\sqrt{1 + x^2/b^2}, \qquad (9.169)$$

and the curvature radius $R(x)$ is

$$R(x) = x + \frac{b^2}{x}. \qquad (9.170)$$

This shows that a beam which at $x = 0$ has a Gaussian profile, remains Gaussian at all $x$, with a $x$-dependent width given by eq. (9.169). Observe that, since $w_0$ is the initial transverse size, $b$ given in eq. (9.168) is the parameter that separates the Fresnel regime (at $x \ll b$) from the Fraunhofer regime (at $x \gg b$), compare with eq. (9.146). In agreement with the discussion above eq. (9.146), at $|x| \ll b$ we find that there is no appreciable widening of the beam, while at $|x| \gg b$ the width increases linearly, $w(x) \simeq w_0|x|/b$, as demanded by the uncertainty principle. Using the definition (9.168) of the Rayleigh range $b$, we get

$$w(x) \simeq \frac{|x|\lambda_L}{\pi w_0}, \qquad (|x| \gg b). \qquad (9.171)$$

Actually Gaussian beams saturate the uncertainty principle, i.e. they have the minimum possible spreading.

The term $\arctan(x/b)$ in eq. (9.167) is called the Gouy phase, and is an extra phase factor compared to the plane wave propagation. The surfaces of constant phase are obtained requiring that

$$k_L\left[x + \frac{y^2+z^2}{2R(x)}\right] - \arctan(x/b)$$

be constant. For a typical GW interferometer, with $\lambda_L = 1\,\mu$m and $w_0$ of order of a few cm, $b$ is of order of several hundred meters. Thus, if we want to compute the surfaces of constant phase in a region close to a point on the optical axis, i.e. in a region with coordinates ($x = x_0 + \delta x, y = \delta y, z = \delta z$), with $\delta x, \delta y, \delta z$ of the order of a few cm, we can

Fig. 9.18 A surface of constant intensity of the Gaussian beam (solid line) and surfaces of constant phase (dashed lines).

neglect the variation of $R(x)$ and of the Gouy phase, so at a given $x_0$ we simply have the condition

$$\delta x + \frac{\delta y^2 + \delta z^2}{2R(x_0)} = \text{constant}. \tag{9.172}$$

This equation describes a portion of spherical surface with radius $R(x_0)$, as we can check immediately by expanding the equation $x^2+y^2+z^2 = R^2$ around $x = R + \delta x, y = \delta y, z = \delta z$. Therefore the wavefronts of a Gaussian beam are spherical to an excellent approximation (as long as the transverse distances are much smaller than $b = O(10^2)$ m), and $R(x)$ is their curvature radius. The shape of the beam is therefore as shown in Fig. 9.18.[20] The characteristic length $w_0$, which determines the transverse size at $x = 0$, is called the *waist* of the beam.

[20] The solution (9.167) holds also at $x < 0$; in this case both the term $k_L x$ and $R(x)$ in the exponential change sign, and the radius of curvature is as shown in the figure.

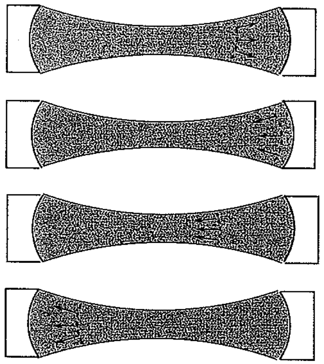

**Fig. 9.19** A wavefront (dashed lines) that propagates toward a spherical mirror is reflected back and focused toward the waist. After passing the waist it expands again toward the other mirror.

When the beam bounces many times between two mirrors, we want to avoid that at each trip it widens further, as in Fig. 9.15. This can be obtained shaping the mirrors so that their surfaces match exactly surfaces of constant phase of the beam. For Gaussian beams we have seen that the wavefronts are spherical, so we must use spherical mirrors. When the expanding wavefront of a Gaussian beam reaches a spherical mirror located at a position $x_0$ and with radius of curvature $R(x_0)$, its direction of propagation is reversed, and the beam is focused back and converges toward the waist at $x = 0$, before re-expanding again for $x < 0$. If we have another spherical mirror at $x = -x_0$ the beam bounces back and forth between them, and at each reflection its wavefronts are forced to converge back toward the waist, as shown in Fig. 9.19, so the beam does not increment its transverse size at each bounce.

Gaussian beams have two advantages over other shapes. First, they have the minimum spreading compatible with the uncertainty principle. Second, their wavefronts are spherical, and mirrors with a spherical shape are easy to manufacture. For these reasons, they are the choice used in present GW interferometers. Thus, we can replace the scheme given in Fig. 9.12 with the more realistic scheme of Fig. 9.20. Alternatively, rather than using two spherical mirrors with the waist in the middle of the cavity, we can put a flat mirror in the position of the waist and, at a distance $L$, a spherical mirror with curvature radius $R = L + b^2/L$. Presently, the former option is used in LIGO and the latter in VIRGO. With the waist of the beam chosen in the middle of the cavity, the value of $w(x)$ at the position of the mirrors $x = \pm L/2$ is given by

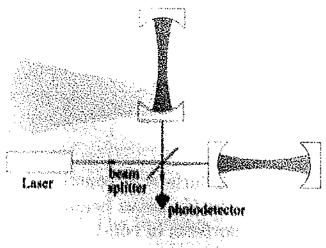


**Fig. 9.20** The scheme of an interferometer with Fabry–Perot cavities, with Gaussian beams and spherical mirrors.

$$w^2(\pm L/2) = w_0^2 + \frac{\lambda\!\!\!^{-}{}_L^2 L^2}{w_0^2}. \tag{9.173}$$

In order to be able to use mirrors of manageable size, we want to have $w(\pm L/2)$ small. Minimizing eq. (9.173) with respect to $w_0$ we find the optimal value of the waist,

$$w_0^{\text{optimal}} = (\lambda\!\!\!^{-}{}_L L)^{1/2}. \tag{9.174}$$



For arms of length $L = 4$ km and a wavelength of the laser light $\lambda_L \simeq 1.0\,\mu$m this gives $w_0 \simeq 2.5$ cm, to which corresponds a value $w(L/2) = (2\lambda\!\!\!^{-}{}_L L)^{1/2} \simeq 3.6$ cm. A suitable mirror radius for such a beam can therefore be $O(10)$ cm.[21] Observe that, since the waist $w_0$ is much larger than the wavelength $\lambda\!\!\!^{-}{}_L$, the paraxial approximation that we have used is well justified.

[21] A small spot allows us to keep down the mirror size. However, it also results in large intensity gradients, in region of high intensity of the electromagnetic field, inducing thermal deformations of the mirrors that must be compensated. An alternative possibility, that has been studied for advanced interferometers, is the use of so-called "mesa beams", i.e. beams with a flat profile, which average more effectively over these thermoelastic fluctuations, see Bondarescu and Thorne (2006).

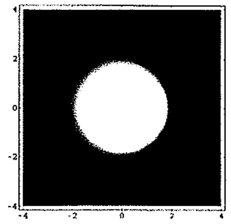

**Fig. 9.21** The intensity of the mode $TEM_{00}$ as a function of the transverse variables $(y, z)$, at a given $x$. (in units such that $w(x) = \sqrt{2}$).

The Gaussian beams that we have considered are by definition solutions of the paraxial evolution equation (9.153), since we obtained them evolving an initial condition on the surface $x = 0$ with the paraxial propagator. Of course, we can also verify this by direct substitution in the equation. Actually, the Gaussian beam is just one of many possible solutions. As can be checked by direct substitution into eq. (9.153), there is a complete orthonormal set of solutions called the *Hermite–Gauss modes*, given by

$$u_{mn}(x,y,z) = \frac{c_{mn}}{\sqrt{1+x^2/b^2}}\, e^{-(y^2+z^2)/w^2(x)}\, H_m\!\left(\frac{y\sqrt{2}}{w(x)}\right) H_n\!\left(\frac{z\sqrt{2}}{w(x)}\right)$$
$$\times \exp\left\{ik_L\left[x + \frac{y^2+z^2}{2R(x)}\right] - i(m+n+1)\arctan(x/b)\right\}, \tag{9.175}$$

where $c_{mn}$ are normalization constants and $H_n(\xi)$ are the Hermite polynomials, defined by

$$H_n(\xi) = e^{\xi^2}\left(-\frac{d}{d\xi}\right)^n e^{-\xi^2}. \tag{9.176}$$

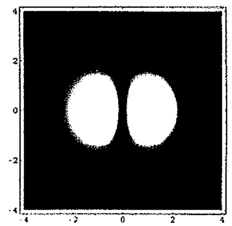

**Fig. 9.22** The same as Fig. 9.21, for the mode $TEM_{01}$.

In particular, $H_0(\xi) = 1$, $H_1(\xi) = 2\xi$, and $H_2(\xi) = 4\xi^2 - 2$. For these modes, both the electric and magnetic fields are transverse to the propagation direction, just as plane wave in free space, so they are also denoted as $TEM_{mn}$ modes. Comparing with eq. (9.167) we see that the Gaussian beam is just the mode $TEM_{00}$. In Figs. 9.21–9.23 we show the intensity $|u_{mn}|^2$ of the modes $TEM_{00}$, $TEM_{01}$ and $TEM_{11}$. Alternatively, one can use as a basis the so-called Laguerre–Gauss modes $LG_{mn}$, which are written in terms of Laguerre polynomials. The fundamental mode $LG_{00}$ is again the Gaussian beam.

Since the Gouy phase for the mode $TEM_{mn}$ is $(m+n+1)\arctan(x/b)$, see eq. (9.175), the resonance condition in a FP cavity depends on $(m,n)$. The laser emits predominantly in the $TEM_{00}$ mode, with a contamination typically less than 10% from higher modes (mostly $TEM_{01}$ and $TEM_{10}$). To eliminate these residual higher modes, which would not be resonant and would just produce noise, before sending it to the beam splitter the laser beam is sent into a Fabry–Perot cavity operated in transmission, called the *mode-cleaner*. Since the Gouy phase for the mode $TEM_{mn}$ depends on $(m,n)$, we can choose the length of the mode cleaner so that only the $(0,0)$ mode is in resonance and is efficiently transmitted.

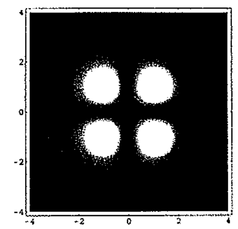

**Fig. 9.23** The same as Fig. 9.21, for the mode $TEM_{11}$.

## 9.3.2 Detection at the dark fringe

### Michelson interferometer

We have seen that the passage of a GW in an interferometer, whether of the simple Michelson type or with Fabry–Perot cavities in the arms, produces a phase shift $\Delta\phi_{gw}(t)$. We now ask how to extract this phase from the output of the detector. The issue, as we will discuss in this section, is quite non-trivial. The origin of the problem can be seen as follows. We consider first for simplicity a Michelson interferometer. We saw in eq. (9.32) that the power at its output is given by $P(\phi) = P_0 \sin^2 \phi$ where $\phi = \phi_0 + \Delta\phi_{gw}(t)$, and $\phi_0$ is a phase that can be adjusted at will by the experimenter. A plot of $P(\phi)/P_0$ is shown in Fig. 9.24. Naively, one might think from this figure that the best working point for the interferometer is at $\phi_0 = \pi/4$, since there the derivative $\partial P/\partial\phi_0$ is maximum, and the sensitivity to a small displacement $\phi_0 \to \phi_0 + \Delta\phi_{gw}(t)$ due to the passage of a GW is highest. Unfortunately, such a strategy would be doomed to failure. In fact, at this working point we are also very sensitive to fluctuations in the power $P_0$ of the laser. Since all that we measure is the power $P = P_0 \sin^2 \phi$ at the photodetector, it is impossible to tell whether a given variation in the measured power is due a variation $\phi_0 \to \phi_0 + \Delta\phi_{gw}(t)$ induced by the passage of a GW, or to a variation $P_0 \to P_0 + \Delta P_0(t)$ due to a fluctuation in the laser power. In particular, a GW with frequency $f_{gw} = O(10^2\text{–}10^3)$ Hz induces variations in the power $P$ with a frequency $f = 2f_{gw}$, which therefore must be compared with the power fluctuations of the laser in the same frequency range. With present lasers, the latter turns out to be much larger than the signal that we expect from GWs.

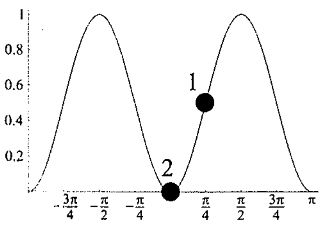


**Fig. 9.24** The power $P(\phi)/P_0$. The naive working point is marked as 1, and the dark fringe as 2.

From a more general point of view, whenever we are looking for very small effects a sound experimental strategy is to build a *null instrument*, that is an instrument that, when the signal is absent, records a zero output. This makes the instrument insensitive to calibration uncertainties that would otherwise overwhelm the tiny signal that we are searching. A prototype of a null instrument is the Dicke radiometer that we discussed in Note 74 on page 412. At the naive operation point marked as 1 in Fig. 9.24, the interferometer is not a null instrument. Even in the absence of a GW, the photodetector measures a large power. In this state, small variations in the power due to GWs should be read against this large DC contribution, and would be overwhelmed by its fluctuations.

This suggests that the best working point should be the dark fringe, marked as the point 2 in Fig. 9.24. There the output in the absence of GWs is zero, and we are insensitive to fluctuations in the laser power. Unfortunately, at the dark fringe not only $P = 0$, but even $\partial P/\partial\phi = 0$. Since $\Delta\phi_{gw} = O(h)$, this means that at the dark fringe the change in the output power induced by GWs is $\Delta P = O(h^2)$. Given that we expect GWs with amplitude $h$ at most $O(10^{-21})$, an effect quadratic in $h$ is of course invisible. So, apparently the choice is between operation points where the response of the interferometer is linear in $h$, but we have a large DC contribution whose fluctuations overwhelm the signal, and an operation point where we have no DC contribution, but no sensitivity to GWs either.



There is however a very elegant way out of this dilemma. The idea is to apply a *phase modulation* to the input laser light.[22] This can be obtained by passing the incident beam through a Pockels cell, which is a crystal or a block of dielectric material whose index of refraction depends on an applied electric field, $E_{appl} = |E_{appl}| \cos \Omega_{mod} t$. The speed of the response that can be obtained with appropriate materials is quite high, and the index of refraction oscillates with the frequency $f_{mod} = \Omega_{mod}/(2\pi)$, for values of $f_{mod}$ up to tens of MHz. Passing through a material with a time-varying index of refraction, the laser beam acquires a time-varying phase, so the beam which reaches the beam-splitter has the form

$$E_{in} = E_0 e^{-i(\omega_L t + \Gamma \sin \Omega_{mod} t)}, \tag{9.177}$$

where $\Gamma$ is called the modulation index, or the modulation depth. This expression can be expanded in Fourier modes as

$$E_{in} = E_0[J_0(\Gamma)e^{-i\omega_L t} + J_1(\Gamma)e^{-i(\omega_L+\Omega_{mod})t} - J_1(\Gamma)e^{-i(\omega_L-\Omega_{mod})t} + \ldots], \tag{9.178}$$

where $J_n$ are Bessel functions and the dots denote terms with frequencies $\omega_L \pm n\Omega_{mod}$, with $n = 2, 3, \ldots$. For $\Gamma \ll 1$ this expression can be simplified using $J_0(\Gamma) \simeq 1 - (\Gamma^2/4)$ and $J_1(\Gamma) \simeq \Gamma/2$. (In the limit $\Gamma \ll 1$ this expansion is obtained more simply expanding directly eq. (9.177) in powers of $\Gamma$). Therefore, the effect of the phase modulation is to generate sidebands.[23] For small $\Gamma$, higher sidebands are suppressed by higher powers of $\Gamma$, so we will limit ourselves to the carrier, which has frequency $\omega_L$ and wavenumber $k_L = \omega_L/c$, and to the first two sidebands, with frequencies

$$\omega_\pm = \omega_L \pm \Omega_{mod}, \tag{9.179}$$

and wavenumbers

$$k_\pm = \frac{\omega_\pm}{c} = 2\pi\left(\frac{1}{\lambda_L} \pm \frac{1}{\lambda_{mod}}\right). \tag{9.180}$$

Consider now what happens to the carrier and to the sidebands in a Michelson interferometer with arms of length $L_x$ and $L_y$. For the carrier the incoming electric field has amplitude $E_0 J_0(\Gamma)$ so, from the discussion in Section 9.1, the electric field at the output of the interferometer is

$$(E_{out})_c = \frac{1}{2}(r_1 e^{2ik_L L_x} - r_2 e^{2ik_L L_y})\, E_0 J_0(\Gamma) e^{-i\omega_L t}, \tag{9.181}$$

where $r_1, r_2$ are the reflectivities of the two end-mirrors. Taking perfectly reflecting mirrors, $r_1 = r_2 = -1$, we have

$$\begin{aligned}(E_{out})_c &= -iE_0 J_0(\Gamma)\, e^{-i\omega_L t + ik_L(L_x+L_y)} \sin[k_L(L_x - L_y)] \\ &= -iE_0 J_0(\Gamma)\, e^{-i\omega_L t + ik_L(L_x+L_y)} \sin\left[2\pi\frac{L_x - L_y}{\lambda_L}\right],\end{aligned} \tag{9.182}$$

[22] Another possible solution would be to control so well the laser fluctuations, that a detection scheme of the type discussed above becomes possible (typically at a working point which is slightly displaced from the dark fringe). This solution is under investigation for Advanced LIGO.

[23] For $\lambda_L = 1\,\mu$m we have $\omega_L/(2\pi) \simeq 300$ THz, while typically $\Omega_{mod}/(2\pi) \simeq 30$ MHz, so $\Omega_{mod} \ll \omega_L$.

compare with eq. (9.4). For the sidebands the calculation is the same, but of course now $k_L$ is replaced by $k_\pm$ and $\omega_L$ by $\omega_\pm$, and the amplitude of the incident field is $\pm J_1(\Gamma)E_0$. Thus, writing $L_x - L_y = \Delta L$, the electric field of the sidebands at the output is

$$(E_{\rm out})_\pm = \mp iE_0 J_1(\Gamma)\, e^{-i\omega_\pm t + ik_\pm(L_x+L_y)} \sin(k_\pm \Delta L) \tag{9.183}$$
$$= \mp iE_0 J_1(\Gamma)\, e^{-i\omega_\pm t + ik_\pm(L_x+L_y)} \sin\left[2\pi\left(\frac{\Delta L}{\lambda_L} \pm \frac{\Delta L}{\lambda_{\rm mod}}\right)\right].$$

Now comes the crucial point. If we take $L_x = L_y$, both the carrier and the sidebands are on the dark fringe, $(E_{\rm out})_c = (E_{\rm out})_\pm = 0$. However, instead of choosing $L_x = L_y$, we can set $L_x - L_y$ equal to an integer number of laser wavelengths, i.e. $\Delta L = n\lambda_L$. Then, as far as the carrier is concerned, we are still on the dark fringe, while the sidebands are no longer on the dark fringe. Rather,

$$(E_{\rm out})_\pm = -iE_0 J_1(\Gamma)\, e^{-i\omega_\pm t + ik_\pm(L_x+L_y)} \sin(2\pi\Delta L/\lambda_{\rm mod}). \tag{9.184}$$

This choice of asymmetric arms is called the *Schnupp asymmetry*. Consider now what happens when a GW arrives, taking for simplicity a plus polarization with optimal direction, and $\omega_{\rm gw}L/c \ll 1$. Then eq. (9.31) gives $L_x \to L_x + hL_x/2$ and $L_y \to L_y - hL_y/2$, so

$$(L_x - L_y) \to (L_x - L_y) + L\, h(t), \tag{9.185}$$

where $L = (L_x + L_y)/2$ and, to lowest order in $\omega_{\rm gw}L/c$, we could replace $h(t - L_x/c)$ and $h(t - L_y/c)$ by $h(t)$. Then we see from eq. (9.182) that the electric field of the carrier is shifted from the value $(E_{\rm out})_c = 0$ on the dark fringe to the value

$$(E_{\rm out})_c = -iE_0 J_0(\Gamma)\, e^{-i\omega_L t + 2ik_L L}\, k_L L h(t). \tag{9.186}$$

This is linear in $h$ and, if this were the total electric field, the power $|(E_{\rm out})_c|^2$ would be quadratic in $h$, as we saw above. However, now we also have the field of the sidebands, and the total electric field is

$$(E_{\rm out})_{\rm tot} = (E_{\rm out})_c + (E_{\rm out})_+ + (E_{\rm out})_- . \tag{9.187}$$

From eq. (9.184), in the absence of GW we have

$$(E_{\rm out})_+ + (E_{\rm out})_- = -2iE_0 J_1(\Gamma)\, e^{-i\omega_L t + 2ik_L L} \times \sin(2\pi\Delta L/\lambda_{\rm mod}) \cos(\Omega_{\rm mod}t - \alpha), \tag{9.188}$$

where $\alpha = 4\pi L/\lambda_{\rm mod}$ is a phase. In the presence of GWs this is modified by the fact that $\Delta L \to [1 + O(h)]\Delta L$. However, here we can neglect the term $O(h)$ because, as we will see below, it is the term $O(1)$ that combines with the carrier, giving a term proportional to $h$ in $|(E_{\rm out})_{\rm tot}|^2$ that will encode the GW signal.[24] Thus, the total electric field at the output, in the presence of GWs, is

$$(E_{\rm out})_{\rm tot} = -iE_0 e^{-i\omega_L t + 2ik_L L}[J_0(\Gamma)k_L L h(t) + 2J_1(\Gamma)\sin(2\pi\Delta L/\lambda_{\rm mod})\cos(\Omega_{\rm mod}t - \alpha)]. \tag{9.189}$$

[24]Furthermore, the term $O(h)$ is multiplied here by $\Delta L$, and $\Delta L \ll L$.



When we compute $|(E_{\rm out})_{\rm tot}|^2$ we therefore have three terms. (1) The squared modulus of the first term, which is $O(h^2)$, and therefore unobservable. (2) The squared modulus of the second, which is independent of $h$, and proportional to

$$\cos^2(\Omega_{\rm mod}t - \alpha) = \frac{1}{2}[1 + \cos(2\Omega_{\rm mod}t - 2\alpha)]. \tag{9.190}$$

Therefore it is the sum of a DC term and a term which oscillates with a frequency $2\Omega_{\rm mod}$. (3) Finally we have the mixed term, i.e. the beatings between the carrier and the sidebands, which is

$$4E_0^2 J_0(\Gamma)J_1(\Gamma)k_L L h(t)\sin(2\pi\Delta L/\lambda_{\rm mod})\cos(\Omega_{\rm mod}t - \alpha). \tag{9.191}$$

This term is linear in $h$ and oscillates with a frequency $\Omega_{\rm mod}$.[25] Therefore in the output we have a term linear in $h$, even if the carrier is on the dark fringe. This term can be extracted from the total output $|(E_{\rm out})_{\rm tot}|^2$ using a mixer, which is a non-linear device which takes at its input two voltages, and produces an output voltage proportional to the product of the two input voltages. Then, we can multiply the voltage produced by $|(E_{\rm out})_{\rm tot}|^2$ in the photodetector by a voltage $V_{\rm osc}\cos(\Omega_{\rm mod}t - \alpha)$. The time-averaged output of the mixer selects the part of $|(E_{\rm out})_{\rm tot}|^2$ which oscillates as $\cos(\Omega_{\rm mod}t - \alpha)$, while the DC part and the part oscillating as $\cos(2\Omega_{\rm mod}t - 2\alpha)$ average to zero. The result (9.191) can be optimized choosing $\Delta L/\lambda_{\rm mod} = m + 1/4$, with $m$ any integer.

In this way we have an output which is linear in $h$, and is insensitive to the power fluctuations of the carrier, which is on the dark fringe. In principle, we are still sensitive to power fluctuations of the laser because the sidebands are not on the dark fringe. However, apart from the fact that the electric field of the sidebands is smaller since it is $O(\Gamma)$, the crucial point is that now the signal has been encoded in a term which oscillates as $\cos(\Omega_{\rm mod}t - \alpha)$, so it must no longer compete with the fluctuations of the laser at a frequency $f_{\rm gw}$ of the GWs that we are searching, but rather with the fluctuations of the laser at a frequency $f_{\rm mod}$ which is much higher, typically 30 MHz. The power fluctuations of the laser is an example of a $1/f$ noise (see page 339), and at high frequencies it is small. In conclusion, we have achieved two results with this technique: (1) We are using the interferometer as a null instrument, since when $h = 0$ the output of the mixer, i.e. the term in the output power oscillating as $\cos(\Omega_{\rm mod}t - \alpha)$, vanishes. (2) The signal is linear in the GW amplitude and is encoded into a high-frequency term, so that it must now compete with much smaller $1/f$ noise.

[25]More precisely, since $h(t)$ is proportional to $\cos\omega_{\rm gw}t$, it oscillates at frequencies $\Omega_{\rm mod} \pm \omega_{\rm gw}$. Since $f_{\rm mod} = O(10)$ MHz and $f_{\rm gw} < O(1)$ kHz, $\omega_{\rm gw} \ll \Omega_{\rm mod}$.

### Interferometers with Fabry–Perot cavities

We now discuss how to apply this technique to an interferometer with Fabry–Perot cavities in the arms. In this case we consider two FP cavities both with the same length $L$, and the Schnupp asymmetry consists in the fact that the distances of their respective input mirrors (i.e. the mirror first encountered by the beam) from the beam-splitter are $l_x$ and $l_y$ respectively, with $l_x \neq l_y$.

The field at the output of the photodetector can be computed using eq. (9.70), which states that, as far as the reflected field is concerned, for light with wavenumber $k$ a FP cavity is equivalent to a mirror with a reflectivity

$$\mathcal{R}(k) = \frac{r_1 - r_2(1-p_1)e^{2ikL}}{1 - r_1 r_2 e^{2ikL}}. \tag{9.192}$$

Again we modulate the laser light with a Pockels cell, so the light incident on the beam-splitter is composed of a carrier at the laser frequency $\omega_L$ and two sidebands at $\omega_L \pm \Omega_{\rm mod}$. We choose the cavity length $L$ so that the carrier is resonant. The modulation frequency $f_{\rm mod} = \Omega_{\rm mod}/2\pi$ is much larger than the width of the resonances of the FP cavities. For instance, for $L = 4$ km and $\mathcal{F} = 200$, eqs. (9.82) and (9.83) give a width at half maximum $\delta f = O(200)$ Hz, while the modulation frequency is in the MHz region. Therefore, the modulation frequency $\Omega_{\rm mod}$ can be chosen so that the sidebands are not resonant, and fall roughly in between resonant peaks. From eqs. (9.99) and (9.100) we see that, for an overcoupled cavity, setting $r_2 = 1$, at the resonance $\mathcal{R} = -(1-\sigma)$, so the phase $\phi \equiv \arg(\mathcal{R})$ is equal to $\pi$. In contrast, for a generic value of $2kL$ far from the resonance (using $r_2 = 1$ and $\sigma \ll 1$), eq. (9.192) gives $\mathcal{R} = 1 + O(\sigma^2)$, so in particular $\arg(\mathcal{R}) = 0 \pmod{2\pi}$, as we see also from Fig. 9.11. Thus, as far as the reflected field is concerned, a FP cavity is equivalent to a mirror with a reflectivity $\mathcal{R}$ which is different for the carrier and for the sidebands,

$$\mathcal{R}(k) \simeq \begin{cases} -(1-\sigma) & (\text{if } k = k_L) \\ +1 & (\text{if } k \text{ is not close to } k_L). \end{cases} \tag{9.193}$$

The total electric field at the photodetector is the superposition of the field that propagates in the $x$ arm for a length $l_x$, and is then reflected by the Fabry–Perot cavity with a reflection coefficient $\mathcal{R}_x$, and of the field that propagates in the $y$ arm for a length $l_y$, and is then reflected by the Fabry–Perot cavity, with a reflection coefficient $\mathcal{R}_y$. In the absence of GWs we have $\mathcal{R}_x = \mathcal{R}_y = \mathcal{R}$, with the appropriate value of $\mathcal{R}$ depending on whether we consider the carrier or the sidebands. Then, as in eq. (9.181), the field at the output is

$$\begin{aligned} E_{\rm out} &= \frac{1}{2}(\mathcal{R}_x e^{2ikl_x} - \mathcal{R}_y e^{2ikl_y})\, E_{\rm in} e^{-i\omega t} \\ &= i\mathcal{R} E_{\rm in} e^{-i\omega t + 2ikl} \sin(k\Delta l), \end{aligned} \tag{9.194}$$

where $2l = l_x + l_y$, $\Delta l = l_x - l_y$, and $(k,\omega)$ are equal to $(k_L, \omega_L)$ for the carrier and to $(k_\pm, \omega_\pm)$ for the sidebands. Thus, for the carrier we have

$$(E_{\rm out})_c = -i(1-\sigma)E_0 J_0(\Gamma) e^{-i\omega_L t + 2ik_L l} \sin(2\pi\Delta l/\lambda_L), \tag{9.195}$$

and for the sidebands

$$(E_{\rm out})_\pm = \pm i E_0 J_1(\Gamma)\, e^{-i\omega_\pm t + 2ik_\pm l} \sin\left[2\pi\left(\frac{\Delta l}{\lambda_L} \pm \frac{\Delta l}{\lambda_{\rm mod}}\right)\right]. \tag{9.196}$$

Just as we did for the Michelson interferometer, we choose as working point the dark fringe of the carrier, that is we choose $l_x$ and $l_y$ so that the Schnupp asymmetry $\Delta l$ is equal to an integer times $\lambda_L$, and also $\Delta l/\lambda_{\rm mod} = m + 1/4$ for some integer $m$, so that $\sin(2\pi\Delta l/\lambda_{\rm mod}) = 1$. Thus, in the absence of GWs, $(E_{\rm out})_c = 0$ while



$$(E_{\rm out})_+ + (E_{\rm out})_- = 2iE_0 J_1(\Gamma)\, e^{-i\omega_L t + 2ik_L l} \cos(\Omega_{\rm mod} t - \alpha), \tag{9.197}$$

with $\alpha = 4\pi l/\lambda_{\rm mod}$. Consider now what happens in the presence of a GW, as usual with optimal orientation and $\omega_{\rm gw} L/c \ll 1$. The effect of the GW is to change the reflectivities $\mathcal{R}_x$ and $\mathcal{R}_y$ of the Fabry–Perot cavities. Consider first the carrier. The passage of the GW induces a phase shift $\Delta\phi_x$ in the reflected field given in eq. (9.102), so we get (for $\sigma \ll 1$)

$$\mathcal{R}_x(k_L) = -(1-\sigma)e^{i\Delta\phi_x}, \tag{9.198}$$

with

$$\Delta\phi_x = \frac{2\mathcal{F}}{\pi} k_L L h(t). \tag{9.199}$$

For the cavity along the $y$ arm we have an opposite phase shift, $\Delta\phi_y = -\Delta\phi_x$, so $\mathcal{R}_y(k_L) = -(1-\sigma)e^{-i\Delta\phi_x}$. Then eq. (9.194) gives

$$\begin{aligned} (E_{\rm out})_c &= -\frac{1}{2}(1-\sigma)(e^{i\Delta\phi_x} e^{2ik_L l_x} - e^{-i\Delta\phi_x} e^{2ik_L l_y}) E_{\rm in} e^{-i\omega_L t} \\ &= -i(1-\sigma)E_0 J_0(\Gamma) e^{-i\omega_L t + 2ik_L l} \sin\left[2\pi\frac{\Delta l}{\lambda_L} + \Delta\phi_x\right]. \end{aligned} \tag{9.200}$$

On the dark fringe $\Delta l/\lambda_L = n$ and $\sin\left[2\pi\frac{\Delta l}{\lambda_L} + \Delta\phi_x\right] = \sin(\Delta\phi_x) \simeq \Delta\phi_x$. Thus, in the presence of a GW the electric field of the carrier at the photodetector shifts from the value $(E_{\rm out})_c = 0$ to the value

$$(E_{\rm out})_c = -i(1-\sigma)E_0 J_0(\Gamma) e^{-i\omega_L t + 2ik_L l} \frac{2\mathcal{F}}{\pi} k_L L h(t). \tag{9.201}$$

As we already saw for the Michelson interferometer, the modification of the electric field of the sidebands due to the GWs is negligible, since it gives a corrections $1 + O(h)$, but it is the term $O(1)$ which, beating against the term $O(h)$ in the carrier, gives a term linear in $h$ in the output. In conclusion, we can write the total (carrier plus sidebands) electric field at the output as

$$\begin{aligned} (E_{\rm out})_{\rm tot} &= -2iE_0 e^{-i\omega_L t + 2ik_L l} \\ &\times \left[(1-\sigma)J_0(\Gamma)\frac{\mathcal{F}}{\pi} k_L L h(t) - J_1(\Gamma)\cos(\Omega_{\rm mod} t - \alpha)\right]. \end{aligned} \tag{9.202}$$

The situation is now the same that we already discussed for a Michelson interferometer: taking the modulus squared, the term which oscillates as $\cos(\Omega_{\rm mod} t - \alpha)$ is linear in $h$ and is demodulated with a mixer. We therefore have an output that is linear in $h(t)$ even if the carrier is on the dark fringe, and is encoded in a term which oscillates at the frequency $\Omega_{\rm mod}$, so $1/f$ noise such as laser power fluctuations are small. This procedure is a special case of the Pound–Drever–Hall locking, see Section 9.3.4.

### 9.3.3 Basic optical layout

We can now complete the description of a realistic GW interferometer. One further improvement with respect to the scheme that we have discussed is the *power recycling*. The basic observation is that, since we have chosen as working point the dark fringe for the carrier, in the absence of GWs no light at all emerges from the beam-splitter in the direction of the photodetector, at the carrier frequency. This means that all the light at frequency $\omega_L$ that circulates in the arms is eventually reflected by the beam-splitter back toward the laser and, in this sense, is wasted. When we discuss the noise sources in the next section, and in particular the shot noise, we will see that we want to have the highest possible laser intensity circulating in the arms. However, the power of a continuous (and very stable) laser is currently limited to $O(10)$ watts, which could become $O(100)$ W in the near future. To increase the power circulating in the interferometer, the idea is to "recycle" the light that comes back toward the laser, placing a mirror (the power-recycling mirror) that reflects the light back toward the beam-splitter. As far as the light reflected toward the laser is concerned, we can model the whole interferometer as an equivalent mirror, with a reflectivity that accounts for the total reflected field. The addition of the power-recycling mirror between the laser and the beam-splitter creates a new Fabry–Perot cavity, made of the power-recycling mirror and the "equivalent interferometer mirror". If this cavity is arranged so that it is resonant for the input laser light, the total intensity of the light that circulates in the interferometer is enhanced. Indeed, in this way a gain of $O(100)$ can be obtained (the maximum gain that can be reached is inversely proportional to the losses inside the interferometer), so the power circulating between the power-recycling mirror and the beam splitter raises to about 1 kW. Inside the Fabry–Perot cavities in the arms, this power increases further because it is resonant, and in initial LIGO and VIRGO it reaches a value of order 15 kW.

A second feature of a real interferometer is an output mode cleaner. Even if the initial beam has been accurately prepared in the $TEM_{00}$ mode thanks to the input mode cleaner, various imperfections in the mirrors, as well as misalignments, regenerate higher modes inside the interferometer. These higher modes are not on the dark fringe, and therefore simply produce a noise that lower the contrast at the output. The output mode cleaner, placed between the beam-splitter and the photodetectors, filters out these higher modes, enhancing the contrast and therefore the sensitivity.

Putting together all these elements, we arrive at the optical layout shown in Fig. 9.25 where, for definiteness, we have used the parameters of VIRGO. The laser, a continuous Nd:Yag with wavelength $\lambda_L = 1.064\,\mu$m, provides 20 W of power. The laser beam passes through an electro-optic modulator, i.e. a Pockels cell, which generates sidebands at $\Omega_{mod}/(2\pi) = 6.2$ MHz. The beam is then passed through the input mode cleaner. This is a long cavity with very high finesse, and a triangular shape that forbids reflection back toward the laser. The beam that comes out of the input mode cleaner is very nearly a $TEM_{00}$ mode, both in the carrier and in the sidebands. It is transmitted through the power recycling mirror and enters the interferometer. The Schnupp asymmetry is realized choosing $l_x \simeq 6.4$ m and $l_y \simeq 5.5$ m for the distances between the beam splitter and the input mirrors of the two Fabry–Perot cavities. After going back and forth in the Fabry–Perot cavities, with a length of 3 km and a finesse $\mathcal{F} = 50$, the beams are recombined on the beam-splitter. Since we work on the dark fringe, at the beam-splitter the carrier is entirely reflected back toward the laser, and then finds the power-recycling mirror, that sends it back to the interferometer. When the carrier is displaced from the dark fringe, for instance because of the passage of a GW, the beating between the carrier and the sidebands goes toward the photodetector. It first passes through the output mode cleaner, a single crystal of 3.6 cm, where again makes a triangular path, and then goes to an array of photodetectors, and it is finally demodulated and detected.[26]



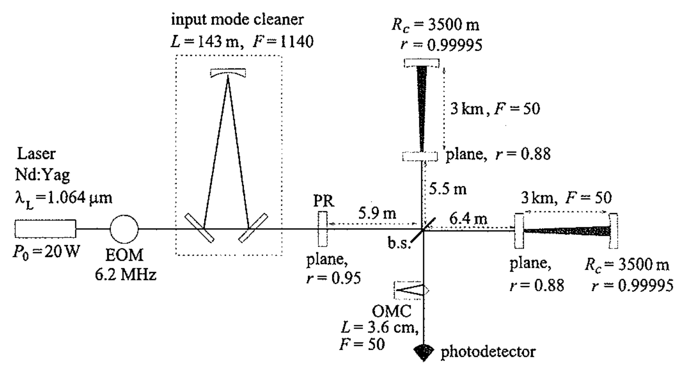


**Fig. 9.25** The basic layout of a GW interferometer. EOM = Electro-optic modulator (Pockels cell); PR = power recycling mirror; OMC = output mode cleaner; b.s. = beam splitter. The curvature radius $R_c$ and reflectivity $r$ of the various mirrors are indicated. For definiteness, we used the values for the initial VIRGO interferometer.

[26] Figure 9.25 is still somewhat simplified. For instance, there are also lenses that are used to match the laser beam into the mode cleaner. A Faraday isolator is used to protect the laser from back-tracking light from the interferometer. A mode-matching telescope is used to blow the input laser beam, which has an initial transverse size of just a few millimeters, to the waist appropriate for the Fabry–Perot cavities which, as we have seen, is rather of order 2–3 cm. Further mirrors are used to pick up signals that are needed for control purposes.

### 9.3.4 Controls and locking

This section is slightly technical, but is meant to give at least a flavor of the problems that must be overcome to turn the beautiful theoretical idea of a GW interferometer into a working instrument.

The scheme that we have discussed above reaches its high sensitivity because the laser light is resonant in the Fabry–Perot cavities. On resonance, a FP cavity is extremely sensitive to changes in its length $L$, but as soon as we move away from the resonance, it becomes "dead", and the phase of the reflected field loses essentially any dependence on

$L$, see Fig. 9.10. This means that the mirrors of the FP cavity must be hold still, and in the right position, so that $k_L L = \pi n$ for some integer $n$. To estimate the precision needed we observe from eq. (9.81) that the half-width of the resonance peak, as a function of $L$ at fixed $k_L$, is reached if $L$ shifts from the value that fulfills the resonance condition to a value $L + \delta L$, with

$$\delta L = \frac{\lambda_L}{4\mathcal{F}} \,. \tag{9.203}$$

(Compare with eqs. (9.83) and (9.84), where we computed the full width at half maximum in $\omega_L$, at fixed $L$.) With a finesse $\mathcal{F} = O(200)$ this means that we need to keep the length $L$ of the Fabry–Perot cavities fixed, within a precision better than $\delta L \sim 10^{-3}\lambda_L$. Similarly, the power-recycling technique that we have discussed allows us to gain a factor $O(100)$ in laser power, but again the power-recycling mirror must be located in a precise position, in order to satisfy the resonance condition in the power-recycling cavity, i.e. in the cavity made by the power-recycling mirror and the equivalent interferometer mirror. Again, this must hold to a precision much smaller than $\lambda_L$. Finally, our detection scheme requires that the interferometer be on the dark fringe, again within a small fraction of wavelength. Typically, a value $\delta L \sim (10^{-6}\text{–}10^{-4})\lambda_L$ is required for good performances. With $\lambda_L = 1\,\mu$m, this means that the relative position of the mirrors must be kept fixed, at a distance $L$ of order a few kms, within a precision

$$\delta L \sim (10^{-12}\text{–}10^{-10})\,\text{m}\,, \tag{9.204}$$

which is less than the size of an atom! Last but not least, all these lengths are measured with respect to $\lambda_L$, so we also need a laser whose frequency is stabilized to great precision. At first sight, the idea of controlling the length of a 4-km cavity down to an accuracy of $10^{-10}$ m might seem preposterous. Indeed, it is here that a large part of the complexity of GW interferometers resides. However, by now this is routinely achieved in the large GW interferometers, and it is quite interesting to understand how this is possible.

First of all, one could make the possibly naive remark that a mirror does not have a smooth surface down to $10^{-10}$ m, since at this level we resolve the individual atoms, and even more so at $10^{-12}$ m; thus, one might object that the notion of the length $L$ of the cavity is not even well defined down to these scales. However, we must keep in mind that the laser beam at the mirror locations has a transverse size of a few centimeters. This means that what the laser beams actually senses is the position of the surface of the mirror, *averaged over a macroscopic scale*, of order a few cms. Thus, the individual atomic fluctuations cancel out, at least to a first approximation[27] and, in this averaged sense, the notion of the length $L$ of the cavity *is* well defined, even down to such small scales. This is a simple but fundamental point to keep in mind to understand the statement that interferometers (or, as we saw in Chapter 8, resonant bars) are finally able to detect displacements which are much smaller than the size of a nucleus.

[27] Of course, any imperfection in the mirrors will create noise, such as scattered light in the interferometer, so their micro-roughness cannot exceed too much the figure $O(10^{-10})$ m found above. The mirrors of LIGO and VIRGO are polished so that their rms micro-roughness is about 0.5Å, that is 50 times larger than $10^{-10}$ m.



Another simple but important conceptual point is that, actually, we do not need to know the value of the length $L$ of a FP cavity down to a precision $10^{-10}$ m. All that we need, in order to have a FP cavity which works properly, is that it is on one of its resonances, i.e. $2k_L L = 2\pi n$, or $L = \lambda_L n/2$, for some integer $n$. Since $2L/\lambda_L = O(10^{10})$, the corresponding value of $n$ is very large, but we do not need to know it. All that we want is that the FP cavity be on some resonance, corresponding to some unspecified value of $n$, and does not move away from it by more than $\delta L \sim 10^{-4}\lambda_L$. Once the two FP cavities in the arms are resonant, we must arrange their relative position so that the interferometer is on some dark fringe. Again, we do not need to know on which one.

So, what we need is to "trap", or lock, each FP cavity in some of its resonances, and to lock the interferometer on some dark fringe. The general strategy is the one common to all *feedback control systems*. In general terms, this consists of a sensor and an actuator. The sensor detects the value of the quantity of interest and produces an error signal, which measures the difference between the actual and the desired value. The actuator then provides a feedback, which corrects the error, driving the observed value closer to the desired one.

For a FP cavity, the error signal is obtained using the Pound–Drever–Hall locking scheme. This is a widely used technique, originally invented for stabilizing the wavelength of a laser, using as a reference the length $L$ of a Fabry–Perot cavity. Suppose that we have a FP cavity whose length is fixed, to a sufficiently good precision. The wavelength of any laser has in general fluctuations, and if we want to stabilize it a simple idea is to shine it on a FP cavity of the appropriate length $L$, chosen so that the desired value of $\lambda_L$ is resonant, and look at the transmitted light. As shown in Fig. 9.9, we have a series of narrow peaks as a function of $\lambda_L$. If the wavelength of the laser has a slight mismatch with respect to the resonant value, we are just slightly displaced from the peak, so the transmitted intensity is lower. We could then use this as an error signal, and correct for this error with a feedback mechanism. This scheme however has two drawbacks. First, from the fact that the power decreases we cannot tell in which direction the wavelength fluctuated, and therefore we do not know the sign of the correction to be applied. Second, we cannot disentangle wavelength fluctuations from intrinsic power fluctuations of the laser.

The solution is to use modulated light, so the electric field entering the FP cavity has the form (9.178), with the carrier at the resonant frequency. Upon reflection, taking $\sigma \ll 1$, we see from eq. (9.193) that the carrier takes a minus sign while the sidebands get a plus sign. Using the fact that $J_0(-\Gamma) = J_0(\Gamma)$ and $J_1(-\Gamma) = -J_1(\Gamma)$, we see that this still has the form of modulated light, with modulation index $-\Gamma$. Thus the reflected power is simply $|E_0|^2$, i.e. it contains a DC term and no term at the modulation frequency $\Omega_{\rm mod}$. In this sense, we have a null instrument: the signal that we use, which is the part oscillating as $\Omega_{\rm mod}$, vanishes in the absence of perturbations. Suppose now that a fluctuation changes the wavelength $\lambda_L$ of the laser, with respect to the length $L$

of the cavity. The carrier, which is on resonance, is very sensitive to this perturbation, while the sidebands, which are far from resonance, are completely insensitive, see Fig. 9.10. Thus the reflected field of the carrier is multiplied by a factor $\exp\{i\Delta\phi\}$, while the sidebands are unchanged. In the power $|E_{\rm refl}|^2$, the beating between the carrier and the sidebands now produces a term oscillating at a frequency $\Omega_{\rm mod}$ and linear in $\Delta\phi$, which can be demodulated with a mixer. This is a way to obtain an error signal which, at least close to the resonance, is linear in the deviation $\Delta\phi$. We can therefore use it to lock the laser wavelength to the length $L$ of the cavity.

If we assume for a moment that the laser frequency is already sufficiently stable (we will come back below to this point) and the cavity is not rigid, as in a GW interferometer, we can turn the argument around, using $\lambda_{\rm L}$ as our standard of length, and lock the cavity length $L$ to the wavelength $\lambda_{\rm L}$ of the laser. We now realize that the detection scheme on the dark fringe that we discussed in Section 9.3.2 is nothing but a variant of this Pound–Drever–Hall locking scheme. In the original Pound–Drever–Hall method the need for a signal linear in $\Delta\phi$ arises because we want to know the sign of $\Delta\phi$ in order to correct for it in the proper direction, while in the detection scheme for GWs it arises first of all from the fact that the shift $\Delta\phi$ is $O(h)$, and a quantity $O(h^2)$ would be undetectably small. Observe that, even if the interferometer is always on the dark fringe, the information on the value of $h$ can then be read from the fact that we know the feedback that we had to apply to keep it there.

For the control of the interferometer, one generally collects all the beams that come out from it (including the little light that comes out in transmission from the end-mirrors of the arms) and uses all the information contained in this modulated light to perform a Pound–Drever–Hall locking of the three FP cavities of the interferometer (the two arms and the power-recycling cavity). Actually, if we had a laser sufficiently stable in frequency, we could use as GW detector a single arm with a FP cavity, and the sensitivity of the detector would degrade only by a factor of two (given that, for optimal orientation, the contribution of the GWs in the two arms is summed up). However, lasers of the required stability do not exist. So, the above procedure can be seen as locking first the laser to the length of the FP cavity in one arm, and then using the laser wavelength so stabilized, to lock the second arm cavity to it, so we are really measuring the displacement in one arm in units of the length of the other arm. More precisely, we lock the laser to the common mode, where the two arms move symmetrically, and we detect GWs in the differential mode, where the two arms move anti-symmetrically.

The Pound–Drever–Hall locking ensures that all the FP cavities are operating on resonance. Then, we must ensure that the two beams combine at the output of the interferometer so that they are on the dark fringe. This is done using as error signal the one generated by the Schnupp asymmetry. We discussed it on page 506, considering the phase shift $\Delta\phi$ induced by a GW, but the same argument can be repeated for the phase shift induced by the noise. Again, using modulated light and asymmetric arm lengths, we get a signal linear in $\Delta\phi$ at the frequency $\Omega_{\rm mod}$, that we can use as error signal. This technique is called Schnupp locking.

Experimentally, once the interferometer is at its working point, it is not so difficult to keep it there for very long time. The most difficult part is the so-called lock acquisition, i.e. bringing the instrument from the free state down to a controlled state. In the absence of controls, the mirrors are typically swinging with an amplitude of a few microns, therefore a factor $O(10^4)$ larger than what we can tolerate, and have typical speeds of a few microns per second, so they are sweeping across many resonances. Thus the control system must be fast enough to "grab" a mirror when it passes close to a resonance, and keep it there, using magnetic actuators. Moreover, we have stringent conditions on the alignment of the mirrors; for instance the input mirrors of the Fabry–Perot cavities must be aligned within $\delta\theta < O(10^{-8})$ rad. Such efficient control systems have by now been developed by the collaborations running the large GW interferometers, and locking and correct alignment are by now obtained quite routinely.



## 9.4 Noise sources

Having defined the experimental setup, we can now investigate the sensitivity that can be obtained. The sensitivity at which a GW interferometer must aim, to have good chances of detection, is extremely ambitious. We saw in Chapter 7 that the GW amplitude that can be detected depends crucially on the kind of signal (burst, periodic, coalescence or stochastic) that we are searching. As a first benchmark, we can consider a burst that releases in GWs an energy of $10^{-2}$ solar masses, taking place in the Virgo cluster of galaxies. As we saw on page 365, this gives a GW amplitude on Earth of just $h_0 \sim 10^{-21}$. As we have seen in this chapter, the corresponding displacement of the mirror of the interferometer is $\Delta L = (1/2)h_0 L$ (for $\omega_{\rm gw}L/c \ll 1$), so for $L = 4$ km, we have

$$\Delta L \sim 2 \times 10^{-18}\,{\rm m}\,, \qquad (9.205)$$

which is smaller that the size of a nucleus by a factor $10^3$! Impressive as it might be, this figure is however somewhat misleading because, as we have repeatedly emphasized, we must not forget that this is a coherent displacement of all the atoms of a macroscopic body such as a mirror. A better figure is given by the corresponding phase shift, which for a simple Michelson interferometer is $\Delta\phi_{\rm Mich} = (4\pi/\lambda_{\rm L})\,h_0 L$, see eqs. (9.27) and (9.28). Setting $\lambda_{\rm L} = 1\,\mu$m gives $\Delta\phi_{\rm Mich} \sim 5 \times 10^{-11}$ rad. We have seen however that for an interferometer with Fabry–Perot cavities we gain a factor $2\mathcal{F}/\pi$ in $\Delta\phi$, see eq. (9.102). For $\mathcal{F} = 200$ this is a factor $\simeq 130$, which means that we aim at measuring a phase shift

$$\Delta\phi_{\rm FP} \sim 10^{-8}\,{\rm rad}\,. \qquad (9.206)$$

In the following subsections we examine the dominant noise sources, to see what sensitivity can be reached. The sensitivity is conveniently expressed in terms of the strain sensitivity $S_n^{1/2}(f)$, with dimensions $\mathrm{Hz}^{-1/2}$. From its value we can then obtain the sensitivity to all type of signals, such as bursts, periodic signals, etc., as discussed in Chapter 7.

### 9.4.1 Shot noise

The first source of noise that we consider is the shot noise of the laser. This originates from the fact that the laser light comes in discrete quanta, the photons. Let $N_\gamma$ be the number of photons that arrives on the photodetector in an observation time $T$. Then the average power measured at the photodetector during this observation time is

$$P = \frac{1}{T} N_\gamma \hbar\omega_L \,. \tag{9.207}$$

When we measure the average output power, we are actually counting the number of photons that arrived in a time $T$. Whenever we count a number of *discrete* independent events the set of outcomes follows the Poisson distribution,

$$p(N;\bar{N}) = \frac{1}{N!}\bar{N}^N e^{-\bar{N}} \,, \tag{9.208}$$

where $\bar{N}$ is the average value of $N$. Since this is the probability distribution when we count a number of independent events, it is also known as the *counting statistics*. For large $N$ the Poisson distribution becomes a Gaussian, with standard deviation equal to $\sqrt{N}$. Therefore, the fluctuation in the number of photons is given by

$$\Delta N_\gamma = \sqrt{N_\gamma} \,. \tag{9.209}$$

It is worth stressing that this is a fundamental limitation due to the corpuscular nature of light. This produces a fluctuation in the observed power given by

$$\begin{aligned}(\Delta P)_{\rm shot} &= \frac{1}{T} N_\gamma^{1/2} \hbar\omega_L \\ &= \left(\frac{\hbar\omega_L}{T} P\right)^{1/2} ,\end{aligned} \tag{9.210}$$

where in the second line we eliminated $N_\gamma^{1/2}$ using eq. (9.207). We want to compare this result with the power fluctuations induced by a GW.

To make the setting simpler, we first consider a Michelson interferometer, with no Fabry–Perot cavities in the arms. We neglect the modulation of the laser light and for the moment we work at a generic point $\phi_0$. Then, according to eq. (9.32), in the absence of GWs the output power $P$ is related to the input power $P_0$ by $P = P_0 \sin^2\phi_0$, so eq. (9.210) becomes

$$(\Delta P)_{\rm shot} = \left(\frac{\hbar\omega_L}{T} P_0\right)^{1/2} |\sin\phi_0| \,. \tag{9.211}$$



On the other hand, again from eq. (9.32), the variation in power due to a GW is

$$(\Delta P)_{\rm GW} = \frac{P_0}{2} |\sin 2\phi_0|\, (\Delta\phi)_{\rm Mich} \,. \tag{9.212}$$

We consider a periodic GW with frequency $f$, with only the plus polarization and coming from optimal orientation, and at first we take for simplicity $2\pi f L/c \ll 1$. According to eqs. (9.21) and (9.28), the amplitude of the phase shift $\Delta\phi_{\rm Mich}$ is then

$$|\Delta\phi_{\rm Mich}| = \frac{4\pi L}{\lambda_L} h_0 \,, \tag{9.213}$$

so the power fluctuations induced by this GW have an amplitude

$$(\Delta P)_{\rm GW} = \frac{P_0}{2} |\sin 2\phi_0| \frac{4\pi L}{\lambda_L} h_0 \,. \tag{9.214}$$

The signal-to-noise ratio (defined, as in Chapter 7 to be linear in the amplitude $h_0$ of the GW) for this periodic GW, when the only source of noise is the shot noise, is then

$$\begin{aligned}\frac{S}{N} &= \frac{(\Delta P)_{\rm GW}}{(\Delta P)_{\rm shot}} \\ &= \left(\frac{P_0 T}{\hbar\omega_L}\right)^{1/2} \frac{4\pi L}{\lambda_L} h_0 |\cos\phi_0| \,.\end{aligned} \tag{9.215}$$

For definiteness, we compute the shot noise at the naive working point where $\cos\phi_0 = 1/\sqrt{2}$ (the point 1 in Fig. 9.24),[28] so

$$\frac{S}{N} = \left(\frac{P_0 T}{2\hbar\omega_L}\right)^{1/2} \frac{4\pi L}{\lambda_L} h_0 \,. \tag{9.216}$$

On the other hand, we see from eq. (7.129) that, for a periodic GW of frequency $f$, coming from optimal direction and observed for a time $T$, the signal-to-noise ratio is written in terms of the strain sensitivity $S_n^{1/2}(f)$ as

$$\frac{S}{N} = \left[\frac{T}{S_n(f)}\right]^{1/2} h_0 \,. \tag{9.217}$$

Comparing eqs. (9.216) and (9.217) we see that $T^{1/2}$ and $h_0$ cancel, and we get the strain sensitivity due to the shot noise,

$$S_n^{1/2}(f)\Big|_{\rm shot} = \frac{\lambda_L}{4\pi L} \left(\frac{2\hbar\omega_L}{P_0}\right)^{1/2} . \tag{9.218}$$

For an interferometer with Fabry–Perot cavities, the result can be obtained replacing $|\Delta\phi_{\rm Mich}|$ in eq. (9.213) by $|\Delta\phi_{\rm FP}|$. For an interferometer with power recycling, the input laser power $P_0$ in eq. (9.220) must be replaced with the power circulating in the recycling cavity, so $P_0 \to CP_0$, where $C$ is the factor gained with power recycling (typically $C = O(50\text{–}100)$ with present detectors, so that $CP_0 = O(1)$ kW). We

[28] Actually, eq. (9.215) is maximized when $\phi_0 = 0$, i.e. on the dark fringe. Thus, even in the absence of phase modulation, the dark fringe would be the optimal working point, if the only noise were the shot noise. However, this comes out because both the GW signal and the shot noise vanish at the dark fringe, with a finite ratio which optimizes $S/N$. Since there are other noise, such as test mass movements, that do not vanish at the dark fringe, in the absence of phase modulation the dark fringe is not an acceptable working point.

also take into account the efficiency of the photodetector, which reduces the effective power used to extract electrons at the photodiode by a factor $\eta$ (a typical value can be $\eta = 0.93$), so $P_0 \to \eta P_0$.

Furthermore, we do not want to limit ourselves to the static limit, but we take into account the dependence on the GW frequency. Thus we use eq. (9.123) for $|\Delta\phi_{FP}|$, and eq. (9.217) is replaced by

$$\frac{S}{N} = \left(\frac{\eta C P_0 T}{2\hbar\omega_L}\right)^{1/2} \frac{8\mathcal{F}L}{\lambda_L} h_0 \frac{1}{\sqrt{1+(f/f_p)^2}}, \tag{9.219}$$

and (writing $\omega_L = 2\pi c/\lambda_L$) eq. (9.218) becomes

$$S_n^{1/2}(f)\Big|_{\rm shot} = \frac{1}{8\mathcal{F}L}\left(\frac{4\pi\hbar\lambda_L c}{\eta P_{\rm bs}}\right)^{1/2} \sqrt{1+(f/f_p)^2}, \tag{9.220}$$

where $P_{\rm bs} \equiv CP_0$ is the power on the beam-splitter after recycling.

An instructive way to rephrase the above computation is as follows. According to eq. (9.32), the variation in the output power of a Michelson interferometer induced by a GW, choosing as working point $\phi_0 = \pi/4$, is $\Delta P = P_0\,\Delta\phi_{\rm Mich}/2$. Since all we measure is the power at the photodetector, the power fluctuation eq. (9.210) has the same effect as a phase shift $\Delta\phi_{\rm Mich}$ induced by a GW, with

$$\frac{1}{2}P_0\,\Delta\phi_{\rm Mich} = \frac{1}{T}N_\gamma^{1/2}\hbar\omega_L. \tag{9.221}$$

On the other hand, at $\phi_0 = \pi/4$, we have $P = P_0/2$, so $P_0 = 2P = 2N_\gamma\hbar\omega_L/T$, which, inserted in eq. (9.221), gives

$$\Delta\phi_{\rm Mich} = \frac{1}{\sqrt{N_\gamma}}. \tag{9.222}$$

This is the rms value of the equivalent phase shift. To compute its spectral density $S_{\Delta\phi}(f)$ we proceed as follows. Let $A(t)$ be any random variable, such that

$$\langle A(t)A(t')\rangle = A_0\delta(t-t'). \tag{9.223}$$

This is the case of shot noise, since there is no correlation between the fluctuations of photon number at different times. As we saw in eq. (7.16), the (single-sided) spectral density $S_A(f)$ of any quantity $A$ is in general defined from

$$\langle A(t)A(t')\rangle = \frac{1}{2}\int_{-\infty}^{\infty} df\, S_A(f) e^{-i2\pi f(t-t')}. \tag{9.224}$$

When eq. (9.223) holds, we see that $S_A(f)$ is independent of $f$, and has the value $S_A = 2A_0$ (as we already saw below eq. (8.122)). On the other hand, setting $t = t'$ in eq. (9.223), we get $\langle A^2(t)\rangle = A_0\delta(0) = (1/2)S_A\delta(0)$. If we do not have an instantaneous resolution in time, but rather we perform a coarse graining over an observation time $T$, the Dirac delta must be replaced by its regularized version (with unit area), defined by $\delta(t) = 1/T$ if $-T/2 < t < T/2$ and $\delta(t) = 0$ if $|t| > T/2$, so $\delta(0) = 1/T$. Therefore



$$\langle A^2(t)\rangle = \frac{1}{2T}S_A. \tag{9.225}$$

Thus the strain sensitivity $S_A^{1/2}$ can be obtained from the rms value of $A$, $\langle A^2(t)\rangle^{1/2}$, multiplying it by $(2T)^{1/2}$. In particular the spectral density of the phase shift, $S_{\Delta\phi}$, is given by

$$S_{\Delta\phi}^{1/2} = \sqrt{\frac{2T}{N_\gamma}} = \sqrt{\frac{2\hbar\omega_L}{P}}. \tag{9.226}$$

To pass from the spectral density of $\Delta\phi$ to the spectral density of the noise $n(t)$, which is the quantity to be compared with the GW signal $h(t)$, we use eq. (9.125), i.e. we divide by the transfer function (9.126). (In the language explained at the beginning of Chapter 7, dividing by the transfer function we are referring the noise to the input of the detector). This gives back eq. (9.220).

Observe that $S_n^{1/2}(f)|_{\rm shot}$ is flat up to the pole frequency, and then raises linearly in $f$. This is due to the fact that the shot noise in itself is independent of the frequency, while the transfer function, i.e. the sensitivity of a FP interferometer to GWs, degrades linearly with $f$ beyond $f_p$. Inserting the numerical values we get

$$S_n^{1/2}(f)\Big|_{\rm shot} \simeq 1.5\times 10^{-23}\,{\rm Hz}^{-1/2}\left(\frac{50}{\mathcal{F}}\right)\left(\frac{3\,{\rm km}}{L}\right)\left(\frac{1\,{\rm kW}}{P_{\rm bs}}\right)^{1/2} \times\sqrt{1+(f/f_p)^2}, \tag{9.227}$$

where we set $\lambda_L = 1\,\mu$m, and we used reference values appropriate for the initial VIRGO. Recall that (at the initial detector stage) for VIRGO $f_p \simeq 500$ Hz and for LIGO $f_p \simeq 90$ Hz.[29]

[29] Our discussion is simplified, since we have not taken into account the effect of the phase modulation of the laser light. Numerically, this gives a strain sensitivity higher by approximately a factor of $(3/2)^{1/2}$ compared to the one obtained in eq. (9.220), see the Further Reading section.

### 9.4.2 Radiation pressure

Equation (9.220) indicates that, to beat the shot noise, we should increase the power $P_{\rm bs}$, either increasing the input laser power or increasing the recycling factor $C$. However, a beam of photons that impinges on a mirror and is reflected back exerts a pressure on the mirror itself. If this radiation pressure were constant, it could simply be compensated by the mechanism that holds the mirrors in place. However, since the number of photons arriving on the mirror fluctuates as in eq. (9.209), the radiation pressure fluctuates, too, and generates a stochastic force that shakes the mirrors. We see from eq. (9.210) that this stochastic force grows as $\sqrt{P_{\rm bs}}$ while, from eq. (9.220), shot noise decreases as $1/\sqrt{P_{\rm bs}}$. If, in order to beat the shot noise, we increase the power $P_{\rm bs}$, beyond a certain limiting value the fluctuations in the radiation pressure will become important, and will dominate over the shot noise.

To compute the strain sensitivity due to radiation pressure we proceed as follows. Consider a laser beam with power $P$ that impinges perpendicularly on a mirror. At reflection each photon changes its momentum from $+\mathbf{p}$ to $-\mathbf{p}$, so it transfers a momentum $2|\mathbf{p}|$ to the mirror. Since the photon energy is $E_\gamma = |\mathbf{p}|/c$, the force that a beam of power $P$ exerts on the mirror is $F = 2P/c$. The rms fluctuations of the force in a time $T$ are therefore related to the power fluctuations by $\Delta F = 2\Delta P/c$. Using eq. (9.210),

$$\Delta F = 2\sqrt{\frac{\hbar\omega_{\rm L} P}{c^2 T}}\,. \qquad (9.228)$$

The fluctuation in the number of photons is independent of the frequency, so the spectral density of the force, $S_F(f)$, must be flat in frequency and, using eq. (9.225), is given by

$$S_F^{1/2} = 2\sqrt{\frac{2\hbar\omega_{\rm L} P}{c^2}}\,. \qquad (9.229)$$

This stochastic force $F$ acts on a mirror that, in the horizontal plane, is otherwise free, so we have $F = M\ddot{x}$, where $M$ is the mass of the mirror and $x$ its coordinate. In Fourier space, this means $\tilde{F}(f) = -M(2\pi f)^2\tilde{x}$, so the spectral density of the displacement of the mirror is[30]

$$S_x^{1/2}(f) = \frac{2}{M(2\pi f)^2}\sqrt{\frac{2\hbar\omega_{\rm L} P}{c^2}}\,. \qquad (9.230)$$

Consider now what happens in an interferometer. We consider first a simple Michelson interferometer, taking the beam-splitter much heavier than the end-mirrors (so we can neglect its recoil). When a photon arrives on the beam-splitter, it is scattered randomly into one or the other arm. As a result, in each arm the distribution of photons is a Poissonian. The important point is that the distributions in the two arms are *anti-correlated.* One more photon into one arm means one less in the other. In the differential mode of the interferometer the contributions due to radiation pressure in the two arms therefore add up, so the radiation pressure in an interferometer is obtained multiplying eq. (9.230) by a factor of two (while correlated fluctuations in the two arms, such as intrinsic laser power fluctuations, cancel out).[31]

To express the result in terms of the equivalent noise spectral density, we must divide by the transfer function that relates $\Delta L$ to the GW amplitude $h$. For a simple Michelson interferometer, at $f \ll f_p$ we have $\Delta L = hL$, so the transfer function is simply $L$, and the strain sensitivity $S_n^{1/2}(f)$ due to radiation pressure is

$$S_n^{1/2}(f)\Big|_{\rm rad\ pres} = \frac{4}{ML(2\pi f)^2}\sqrt{\frac{2\hbar\omega_{\rm L} P}{c^2}}\,. \qquad (9.231)$$

[30] More accurately, when a suspended mirror oscillates there is a restoring force due to gravity, so it should be really treated as a harmonic oscillator, with resonance frequency $\omega_0$ and dissipation coefficient $\gamma$ (defined as in eq. (8.20)). Then the factor $(2\pi f)^2 = \omega^2$ in the denominator of eq. (9.230) must be replaced by $|\omega^2 - \omega_0^2 + i\gamma\omega_0|$, compare with eq. (8.23). However, the resonance frequency $\omega_0$ and the dissipation coefficient $\gamma$ are smaller than the frequency $\omega$ at which we are interested, and in a first approximation can be neglected.

[31] Another way to describe the same phenomenon is in terms of vacuum fluctuations of the electromagnetic field entering the interferometer from the output port, see Caves (1980, 1981).



Consider next an interferometer with Fabry–Perot cavities. In this case the result depends on the finesse $\mathcal{F}$ of the arm cavities. Physically, this dependence can be understood observing that, in a FP cavity with finesse $\mathcal{F}$, light is performing on average $\mathcal{N} = 2\mathcal{F}/\pi$ bounces. Then each photon hits the mirrors $O(\mathcal{N})$ times, so the rms value $\Delta L$ of the length of the cavity is $O(\mathcal{N})$ times larger than the value when the light make only one bounce. Furthermore, when the cavity is at resonance, a given value of $\Delta L$ produces a phase shift in the reflected light larger by a factor $T_{\rm FP}(f) = O(\mathcal{N})$, compared to the one-bounce case. Overall, the total effect on the phase shift induced by the radiation pressure is therefore $O(\mathcal{N}^2)$. However, to compare with the effect of a GW (i.e. to refer the noise to the detector input, in the language of Chapter 7) we must divide by the transfer function of an interferometer with Fabry–Perot cavities, which is again $T_{\rm FP}(f)$, and we are left with a single factor $O(\mathcal{N})$.

In other words, a given displacement $\Delta\tilde{x}(f)$ of a mirror due to radiation pressure results in a phase shift $\Delta\phi_{\rm FP}(f)$ which is much larger than for the single-bounces case, since now the transfer function is $T_{\rm FP}(f)$, given in eq. (9.126), which is proportional to $\mathcal{F}$ or, equivalently, to the number of bounces. However, in order to refer the noise to the detector input, we must divide by the same transfer function, so the two effects cancel. The fact that each photon performs $O(\mathcal{N})$ bounces results in the fact that the power inside the cavity is larger by a factor $O(\mathcal{N})$ than the power $P_{\rm bs}$ at the beam-splitter. Indeed, from eq. (9.72), at resonance the power inside the cavity is

$$P_{\rm cav} = P_{\rm bs}\frac{t_1^2}{(1-r_1 r_2)^2}\,. \qquad (9.232)$$

Setting for simplicity $r_2 = 1$ and $t_1^2 = 1 - r_1^2 - p_1 \simeq 1 - r_1^2$, this gives $P_{\rm cav} = P_{\rm bs}(1+r_1)/(1-r_1)$ which, for $r_1$ close to one, can be written as

$$P_{\rm cav} \simeq P_{\rm bs}\frac{2\mathcal{F}}{\pi}\,. \qquad (9.233)$$

Therefore a fluctuation $\Delta P_{\rm bs}$ of the light arriving on the input mirror induces a fluctuation of the field inside the cavity $\Delta P_{\rm cav} \simeq \Delta P_{\rm bs}(2\mathcal{F}/\pi)$. Actually, if the mirror vibrates at a frequency $f$, the cavity is displaced off resonance, and the power inside the cavity is reduced by a factor $[1 + (f/f_p)^2]$, as we see from eq. (9.81), together with the definitions (9.84) and (9.122). As a result (writing $\omega_{\rm L} = 2\pi c/\lambda_{\rm L}$ in eq. (9.231))

$$S_n^{1/2}(f)\Big|_{\rm rad} = \frac{16\sqrt{2}\,\mathcal{F}}{ML(2\pi f)^2}\sqrt{\frac{\hbar}{2\pi}\frac{P_{\rm bs}}{\lambda_L c}}\,\frac{1}{\sqrt{1+(f/f_p)^2}}\,. \qquad (9.234)$$

This result answers a question that might have been asked when we realized that the response to GWs of an interferometer with Fabry–Perot cavities, with arm-length $L$ and finesse $\mathcal{F}$, is equivalent to that of a simple Michelson interferometer with arm-length $(2/\pi)\mathcal{F}L$. Given that very high finesse cavities are not difficult to build (e.g. the mode cleaner has a finesse $O(10^3)$) why bother to construct a km-sized arm, with all the financial and technical problems that this implies (e.g. the very high vacuum in such a long arm, see below)? We could think that it is sufficient to build a table-top interferometer with a sufficiently large finesse.

The answer is that the response of the detector to GWs, encoded in the transfer function (9.126), is only one side of the issue. What really matters is the signal-to-noise ratio, and we must also ask how the various noise scale with $\mathcal{N}$ and $L$. For instance, shot noise is independent of $\mathcal{N}$. When divided by the transfer function, which is $O(\mathcal{N})$, the signal-to-noise ratio therefore scales as $1/\mathcal{N} = O(1/\mathcal{F})$. To beat down such a noise we could in principle keep $L$ small, as long as we use a cavity with a sufficiently high finesse. However, we have seen that radiation pressure noise rather scales as $\mathcal{N}^2$, so after dividing by the transfer function we have a dependence proportional to $\mathcal{N}$, i.e. to $\mathcal{F}$. In this case a very large finesse would be harmful. Below, we will see that displacement noise, such as mirror thermal noise, scale as $\mathcal{N}$, so after dividing by the transfer function we get a the signal-to-noise ratio which is independent of $\mathcal{F}$ (but still proportional to $1/L$), so in this case a high finesse does not help to beat it down, and we need a large arm-length $L$. So, in general we still want to keep $L$ as large as possible, compatibly with technological and financial constraints.

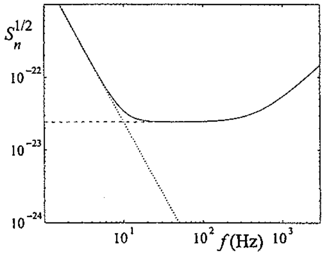


**Fig. 9.26** The strain sensitivity $S_n^{1/2}(f)$ (in units $\mathrm{Hz}^{-1/2}$) due to shot noise (dashed), to radiation pressure (dotted) and the total optical read-out noise (solid line). For definiteness, we have used numerical values of the various parameters typical of the initial VIRGO interferometer.

### 9.4.3 The standard quantum limit

Consider now the combined effect of shot noise and radiation pressure, that we denote as *optical read-out noise*. Its spectral density is

$$S_n(f)|_{\rm opt} = S_n(f)|_{\rm shot} + S_n(f)|_{\rm rad} \,. \qquad (9.235)$$

A plot of this expression, and of the separate shot noise and radiation pressure contributions, is shown in Fig. 9.26. The shot noise contribution is proportional to $P_{\rm bs}^{-1/2}$ while the radiation pressure to $P_{\rm bs}^{1/2}$. We see here the uncertainty principle in action. The situation is conceptually similar to the Heisenberg microscope. We are using photons to measure the position of an object. The photons impart non-deterministically a recoil to the object, here in the form of fluctuations of the radiation pressure, and this recoil disturbs the measure that we are performing. It is amazing that a quantum effect due to the uncertainty principle can be important in the measurement of the position of a macroscopic body, like the mirror of an interferometer, which typically weights $O(20)$ kg. However, for GW detection we need such an extreme accuracy in the determination of the mirror position that, as we will see in this section, the uncertainty principle can indeed become important.[32]

Using the explicit expressions and defining

$$f_0 = \frac{8\mathcal{F}}{2\pi}\sqrt{\frac{P_{\rm bs}}{\pi\lambda_{\rm L} cM}} \,, \qquad (9.236)$$

[32] We already met a similar situation for resonant bars in Sections 8.3.3 and 8.3.4, where we found that, without quantum non-demolition techniques, the best one can do is to detect vibrations in the bar corresponding to $O(1)$ phonon. We have seen that present bar sensitivities are not that far from this limit.



eq. (9.235) can be written as

$$S_n^{1/2}(f)\Big|_{\rm opt} = \frac{1}{L\pi f_0}\sqrt{\frac{\hbar}{M}}\left[\left(1+\frac{f^2}{f_p^2}\right) + \frac{f_0^4}{f^4}\frac{1}{1+f^2/f_p^2}\right]^{1/2} . \qquad (9.237)$$

For a given value of $f$ we can minimize $S_n^{1/2}(f)|_{\rm opt}$ with respect to $f_0$. (In particular, $f_0$ is varied changing the power $P_{\rm bs}$, so this amounts to finding the optimal value of $P_{\rm bs}$.) The optimal value of $f_0$ is the one for which the shot noise and radiation pressure contributions are equal, and is given by

$$1 + \frac{f^2}{f_p} = \frac{f_0^2}{f^2} \,. \qquad (9.238)$$

The corresponding optimal value of $S_n^{1/2}(f)$ defines the *standard quantum limit* (SQL),

$$S_{\rm SQL}^{1/2}(f) = \frac{1}{2\pi f L}\sqrt{\frac{8\hbar}{M}} \,. \qquad (9.239)$$

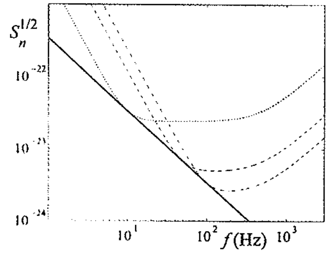


**Fig. 9.27** The optical read-out strain sensitivity $S_n^{1/2}(f)|_{\rm opt}$ (in units $\mathrm{Hz}^{-1/2}$) for $f_0 = 10$ Hz (dotted line), $f_0 = 50$ Hz (dashed line) and $f_0 = 100$ Hz (dot-dashed line), compared to the SQL pseudo-spectral density (solid line). The other parameters are $L = 3$ km, $M = 20.3$ kg, $f_p = 500$ Hz.

It should be stressed that $S_{\rm SQL}(f)$, even if written as a function of $f$, cannot be interpreted as the minimum noise spectral density that can be reached with this type of optical read-out. In fact, the value of $f_0$, i.e. of the laser power, has been optimized keeping fixed the value of $f$. It therefore represents the minimum value of the spectral density which can be obtained (as long as only optical read-out noise is concerned) at that value of $f$. Once we have chosen the power so to optimize the sensitivity at a given frequency $f$, at all other values of the frequency we are not in the optimal situation, and the strain sensitivity is higher than the standard quantum limit. So, eq. (9.239) rather gives the envelope of the minima of the family of functions $S_n^{1/2}(f;f_0)|_{\rm opt}$, parametrized by $f_0$, as shown in Fig. 9.27. (For this reason, it is called a "pseudo-spectral density".)

It is useful to define the dimensionless quantity

$$\mathcal{K}(f) \equiv \frac{8\omega_{\rm L} P_{\rm bs}}{ML^2}\frac{1}{\omega^2(\omega_p^2+\omega^2)} \,, \qquad (9.240)$$

where $\omega_p = 2\pi f_p$ and $\omega = 2\pi f$. Then eq. (9.237) can be rewritten as[33]

$$S_n(f)|_{\rm opt} = \frac{1}{2} S_{\rm SQL}(f)\left[\frac{1}{\mathcal{K}(f)} + \mathcal{K}(f)\right] . \qquad (9.241)$$

We have seen that the existence of the limiting value $S_{\rm SQL}(f)$ is a manifestation of the Heisenberg uncertainty principle. However the uncertainty principle does not put a limit on the accuracy of measurements of position, but only on the accuracy of simultaneous measurements of conjugate variables, and it is possible to go beyond the standard quantum

[33] This result can also be obtained with an elegant formalism, in which radiation pressure and shot noise are related to the quantum vacuum fluctuations entering the interferometer from the output port, see Kimble, Levin, Matsko, Thorne and Vyatchanin (2000). The quantity that we denote by $\omega_p$ corresponds to $\gamma$ in this reference. More precisely, $\gamma = ct_1^2/(4L)$, where $t_1$ is the transmissivity of the first mirror. In our computation we have assumed negligible losses, so $t_1^2 = 1 - r_1^2$, and $r_1, r_2$ close to one, so $\gamma \simeq \pi c/(2L\mathcal{F})$. Observe also that we have set to one the efficiency $\eta$ of the photodiode and we have neglected the effect of light modulation. For instance, in the present configuration of VIRGO, this results in the replacement $1/\mathcal{K}(f) \to 3/[2\eta\mathcal{K}(f)]$ in the first term of eq. (9.241). Different modulation schemes can give rise to different numerical factor for this shot noise term.

limit using quantum non-demolition (QND) techniques. The general principles of QND measurements have already been presented in Section 8.3.4. We refer the reader to the Further Reading section for the application of these techniques to GW interferometers. These techniques can become important for advanced interferometers.

### 9.4.4 Displacement noise

The optical read-out noise discussed above is intrinsic to the way that we use to detect the displacement of the test masses induced by GWs, using a laser beam that bounces between them. Of course, the test masses also move because of many other effects that have nothing to do with GWs. We generically denote all these other effects as "displacement noise", and we characterize them with a strain spectral density of the displacement $S_x^{1/2}(f)$, that we denote simply as $x(f)$.

Recall that the effect of a GW on the length $L$ of a FP cavity is to change it by the amount $\Delta L = hL$ (as long as $\omega_{gw}L/c \ll 1$, i.e. as long as eq. (9.124) holds). Thus, if the length of the cavity changes by an amount $\Delta x$ because of one of these displacement noise, the corresponding equivalent GW amplitude is $\Delta x/L$. So, to refer the noise at the detector input (i.e. to compute the equivalent GW that would induce the same phase shift), we must divide the strain sensitivity of the displacement by the arm-length $L$. The finesse of the cavity, or equivalently the number of bounces $\mathcal{N}$ performed by the laser beam inside the FP cavity, does not enter here. This can also be understood in terms of the phase shift $\Delta\phi_{FP}$, observing that the phase shift induced by a GW and that induced by the displacement noise of a mirror are both multiplied by the number of bounces of the light inside the cavity, so when we refer the noise to the detector input the factor $\mathcal{N}$ cancels.

The computation of these displacement noise depends on many technical issues such as properties of materials, details of the suspension mechanisms, etc. and a complete discussion is beyond the scope of this book. We limit ourselves to mentioning the most important displacement noise below. Graphs showing their separate effect on the strain sensitivity are shown in Section 9.5 below, in Fig. 9.31 for VIRGO and in Fig. 9.32 for LIGO.

#### Seismic and Newtonian noise

The Earth's ground is in continual motion, with amplitudes of order a few microns. In the region 1–10 Hz this is mostly due to human activity such as local traffic, trains, etc. as well as to local phenomena such as winds. Furthermore, there is a micro-seismic background, which affects a GW interferometer mostly in the form of surface waves that shake the suspension mechanisms and, finally, the mirrors.[34] Its strain sensitivity has in general the form

$$x(f) \simeq A \left(\frac{1\,\text{Hz}}{f^\nu}\right) \text{m}\,\text{Hz}^{-1/2}\,, \qquad (9.242)$$

[34] This backgrounds originates mostly from atmospheric cyclonic systems over the oceans. Energy is transferred from the atmosphere to the ocean, and then to the ocean floor. From there it is transmitted through the crust for long distances, $O(10^3)$ km, mostly in the form of surface waves. Its amplitude presents a peak corresponding to the period of the ocean waves (12 s) and another one at twice this frequency, and decreases as a power law at higher frequencies. Near the costs, there is also a contribution from sea waves breaking on the shore.



where (above about 1 Hz) the index $\nu \simeq 2$ while, at a typical quiet location, $A$ can be of order $10^{-7}$. Dividing $x(f)$ by the length $L =$ 3–4 km, we are left with a noise strain sensitivity at least 10 orders of magnitude larger than the values at which we are aiming. The seismic noise must therefore be attenuated by a huge factor. This is in general obtained with a set of pendulums in cascade.[35] A single pendulum with resonance frequency $f_0$, at frequencies $f \gg f_0$ attenuates the strain sensitivity $x(f)$ by a factor $f_0^2/f^2$, and a multistage filter made by $N$ stages provides an attenuation factor $(f_0^2/f^2)^N$. Therefore one must choose $f_0$ much smaller than the GW frequency of interest. In practice, this means that the seismic noise can be reduced below a level interesting for GW detection only at frequencies above, say, 10 Hz. This is the main reason while a ground-based interferometer cannot search for GWs below the $\sim 10$ Hz region.

Newtonian noise, also known as "gravity gradient noise", is due to the Newtonian gravitational forces of objects that are moving, which results in a time-varying gravitational force.[36] The most important Newtonian effect is induced by micro-seismic noise, which produce mass density fluctuations and therefore a fluctuation of the gravitational field of the Earth, which couples directly to the test masses of the interferometer. One can get a feeling for the extreme sensitivity of a GW interferometer, when one realizes that even the changing gravitational attraction due to atmospheric turbulence gives a non-negligible contribution to the Newtonian noise below $O(1)$ Hz.

While the seismic noise can be attenuated arbitrarily (at least in principle, if one were able to build an arbitrarily good attenuator), the Newtonian noise cannot be eliminated, since the gravitational force cannot be screened. In present GW interferometers the Newtonian noise is not the dominant effect (below a few Hz it is overwhelmed by the seismic noise and above a few Hz by the pendulum thermal noise, see Fig. 9.31). However, even if one were able to push further down the seismic and thermal noise, which in principle can be done with technological improvements, still one would remain with the Newtonian noise which, for a ground-based detector, would anyway provide the ultimate limitation at low frequencies (although some noise reduction might be possible monitoring it with a complex network of accelerometers, and then subtracting it).

#### Thermal noise

Thermal noise induce vibrations both in the mirrors and in the suspensions. As discussed in Section 8.3.1, its effect can be computed using the fluctuation–dissipation theorem. We saw that, for a linear system subject to a force $F$, we can always write the equation of motion in the form

$$\tilde{F}(\omega) = -i\omega Z(\omega)\tilde{x}(\omega)\,, \qquad (9.243)$$

where $Z(\omega)$ is called the impedance. The fluctuation–dissipation theorem gives the spectral density of the force responsible for thermal fluc-

[35] Such an attenuation system is in itself a remarkable technological achievement. In particular, the VIRGO superattenuator, made of 8 m tall towers, is the most performing device of this kind ever built. The construction of these towers also present non-trivial problems in material science. To get a feeling for the kind of issues involved, consider that the slippage of two grains of steel under stress releases an energy sufficient to shake the mirror at the level of $10^{-12}$ m, about a million times larger than the expected GW signal.

[36] Obviously, these are quasi-static gravitational fields in the near region of their sources, and are distinct from GWs, that are time-varying gravitational fields in the far region of their sources.

tuations, $S_F(\omega)$, in terms of the real part of $Z$, see eq. (8.125). The displacement spectral density is then given by eq. (8.128), that we write as

$$x(\omega) = \frac{1}{\omega|Z(\omega)|}\left[4kT\,\mathrm{Re}\,Z(\omega)\right]^{1/2}. \tag{9.244}$$

Therefore, $x(\omega)$ is known once we have $Z(\omega)$. This has the great advantage that we do not need to have a detailed microscopic model of the dissipation mechanism. For a simple damped harmonic oscillator, $Z(\omega)$ is given by eq. (8.126). For a more complex extended object, the impedance associated to a normal mode with frequency $\omega_0$ can be modeled more generally as

$$Z = -\frac{im}{\omega}\left[\omega^2 - \omega_0^2 + i\omega^2\phi(\omega)\right], \tag{9.245}$$

where the dimensionless function $\phi(\omega)$ is called the loss angle. The most important thermal noise are the following.

*Suspension thermal noise.* Any vibration induced in the suspension of the test masses results in a displacement noise. In particular, we have

- Pendulum thermal fluctuations. These are thermal fluctuations that induce a swinging motion in the suspensions, and therefore a horizontal displacement of the mirrors. In the present detectors this noise is the dominant one between a few Hz and $O(50)$ Hz, see Fig. 9.31.
- Vertical thermal fluctuations. Thermal noise induce also a vertical motion of the suspensions. In a GW interferometer, we are only interested in the horizontal distance between the mirrors. However, because of the curvature of the Earth, the direction of the vertical at the two mirror locations, which are separated by a distance $L = 3$–$4$ km, is not the same. This results in a vertical–horizontal coupling of the order of the angle $\theta = L/(2R_\oplus) \simeq 2\times 10^{-4}$.
- Violin modes. These are vibrations that can be described in terms of fluctuations of the normal modes of the wire. They are responsible for the set of spikes between 300 Hz and a few kHz in Fig. 9.31. The width of these resonances is however very narrow, so they affect the sensitivity only in very small intervals of frequencies.

*Test-mass thermal noise.* These are thermal fluctuations within the test masses themselves. We can distinguish the following effects.

- Brownian motion of the mirrors. The atoms of a mirror at temperature $T$ have a Brownian motion due to their kinetic energy, which gives rise to mirror thermal noise. Just as with the violin modes, its effect can be computed performing a normal-mode decomposition. This is presently the dominant noise between a few tens and a few hundred Hz, see Fig. 9.31.



- Thermo-elastic fluctuations. These are due to the fact that, in a finite volume $V$, the temperature fluctuates, with a variance $(\delta T)^2 = k_B T^2/(\rho C_V V)$, where $C_V$ is the specific heat and $\rho$ the density of the material. These temperature fluctuations generate displacement noise through the expansion of the material. Thermo-elastic fluctuations take place both in the mirror bulk and in the mirror coatings.
- Thermo-refractive fluctuations. The refraction index of the coatings is a function of the temperature. Thus, the same temperature fluctuations responsible for the thermo-elastic noise also induce fluctuations in the refraction index of the mirrors.

Of course, thermal noise is proportional to the dissipations present in the system, which depend strongly on the material used, and therefore there is an ongoing search for materials with optimal properties.

### Other noise

Beside read-out and displacement noise, other noise are relevant, and keeping them under control require advanced technologies. We mention some of them, to give a feeling for the complexity of a GW interferometer.

- The laser beam must travel in a ultra-high vacuum pipe, in order to keep the noise induced by fluctuations in the index of refraction below the design sensitivity. For the initial interferometers the pressure must be lower than $10^{-7}$ mbar while, for advanced interferometers, it must be lower than $10^{-9}$ mbar.[37] Furthermore, the residual gas must be free of condensable organic molecules (hydrocarbons), in order to keep the optical surfaces clean. It is estimated that a hydrocarbon partial pressure of $10^{-13}$ mbar is required if one wants to avoid the cumulative deposition of one monolayer of molecules on the optical elements in 4 years.
- To limit diffuse light scattering in the interferometer, the mirrors are polished to a rms micro-roughness of about 0.5Å, over a diameter of order 20 cm, and have losses of order a few parts per million.[38]
- Fluctuations of the laser in power and in frequency must be kept under control to great accuracy.
- Other important concerns are so-called technical noise, often related to the servo loops that keep the many degrees of freedom of an interferometer under control.
- Seismic noise can be reinjected in the detector because the enclosure walls couple to the mirror magnets both directly, because of diamagnetism, and through eddy currents.
- The suspension wires undergo creep, i.e. sudden grain-boundaries slipping, which at this sensitivity level are so frequent that they finally constitute a Gaussian noise.

[37] These vacuum tubes have a diameter of about 1.2 m in order to contain the diffraction-limited laser beam, and a 3- or 4-km length, resulting in a total high-vacuum volume of about 9000 m$^3$. For comparison, this is much larger than the vacuum volume of the LEP particle accelerator, where the ring is almost 27 kms in length, but the transverse section of the vacuum pipe was an ellipse with semiaxes of about 6 cm and 3 cm, respectively.

[38] Nevertheless, the remaining diffused light still generates important noise because it can interact with the pipe walls, thereby getting modulated by its seismic noise, and then it can be rediffused back in the beam by reflection on a mirror. Even if only a few parts per million of the circulating light is diffused, an unacceptably high noise results. For this reason, in each of the 3 km arms of VIRGO have been mounted about 100 circular rings, obtained from a conical surface, that trap and absorb most of the residual diffused light.

- Non-Gaussian noise is also present. For instance, the release of residual gas pockets from the tube walls can generate sudden bursts.

So, many subtle effects can become important at the sensitivity level at which a GW interferometer aims. In spite of the apparent simplicity of the original idea, a GW interferometer is clearly a very complex instrument.

## 9.5 Existing and planned detectors

### 9.5.1 Initial interferometers

At time of writing (2007) there are various collaborations running GW interferometers. In the US, the LIGO collaboration runs two interferometers with arms of 4 km, one in Hanford (Washington State) and one in Livingstone (Louisiana). The two detectors have been placed at a large distance (the light travel time between them is about 10 ms), so that their noise should be uncorrelated, and are used to search for coincidences. In the Hanford site there is also a second smaller interferometer, with 2 km arms, in the same vacuum system. While of course there will be correlated noise between the shorter and the longer interferometer, still the presence of the smaller interferometer gives a further handle that helps discriminating real signals from spurious noise, making use of the fact that many common noise in the two detectors are independent of $L$, while the effect of the GW scales with $L$. A view of the Hanford detector is shown in Fig. 9.28. The scientists collaborating to the project are members of the LIGO Scientific Community (LSC).

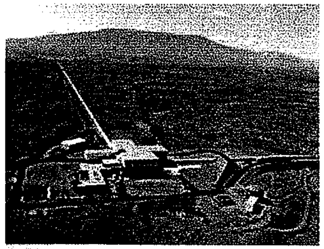

Fig. 9.28 A view of the LIGO detector in Hanford, Washington State. (Courtesy of the LIGO collaboration.)

The VIRGO interferometer, located near Pisa, Italy, is a collaboration between Italy and France, and has arms of 3 km. A view of the detector is shown in Fig. 9.29.

Fig. 9.29 A view of the VIRGO interferometer in Cascina, near Pisa. (Courtesy of the VIRGO collaboration.)

Beside these large GW interferometers, there are two smaller ones: GEO600, with arms of 600 m, is located near Hannover and is a German-British collaboration. GEO600 works in close collaboration with LIGO, and its members are also members of the LSC. TAMA is located in Tokyo, and has arms of 300 m. These smaller detectors are useful also for developing techniques that will be used by LIGO and VIRGO in their advanced stage.

In Fig. 9.30 we show a simplified model of the strain sensitivity of these detectors, in their initial stage. The best sensitivities are reached by the two detectors with longer arms, LIGO and VIRGO. In the low-frequency regime, the dominant noise is the seismic. For LIGO and GEO this results in a "seismic wall" below about 30–40 Hz. VIRGO has developed an advanced super-attenuator, so its target sensitivity is better at low frequencies. In the intermediate region the dominant noise is the mirror thermal noise. In this region GEO600 compensates the smaller arm length using fused silica suspensions, an advanced technique that reduces suspension thermal noise, and that will be adopted in advanced



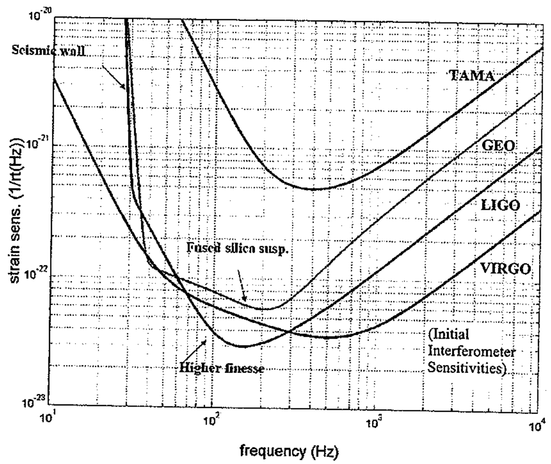


Fig. 9.30 A simplified model of the strain sensitivities of the initial interferometers.

interferometers. Then, shot noise takes over and becomes the limiting noise at high frequencies. In this regime, the difference between the sensitivities of LIGO and VIRGO is due to the fact that LIGO has longer arms (4 km instead of 3 km for VIRGO) and a higher finesse ($\mathcal{F} = 200$ for LIGO and $\mathcal{F} = 50$ for VIRGO). On the one hand this means that, at $f \ll f_{\rm pole}$, the shot noise of LIGO is smaller than that of VIRGO, see eq. (9.220), which helps to give a better sensitivity in the 100 Hz region. On the other hand, we see from eq. (9.88) that the pole frequency of LIGO is smaller, $f_p \simeq 90$ Hz for LIGO and $f_p \simeq 500$ Hz for VIRGO. This means that in LIGO the shot noise curve begins to raise linearly earlier, so it finally get higher than in VIRGO.

A more accurate plot of the sensitivity, including the separate contributions from the various noise sources, is shown in Fig. 9.31 for VIRGO, while actual data from LIGO are shown in Fig. 9.32, and in Fig. 9.33 for GEO.[39]

Once we have the strain sensitivity, the signal-to-noise ratio of GW interferometers for different type of signals (coalescences, bursts, periodic signals and stochastic backgrounds) can be computed using the results of Chapter 7, similarly to what we did in Section 8.3.5 for resonant bars.

For a broadband GW detector such as an interferometer, a useful measure of the sensitivity is given in terms of the sight distance to coalescing binaries, that we introduced in Section 7.7.2. Inserting in eq. (7.182) the strain sensitivity of the initial LIGO and VIRGO one finds that, for a NS-NS binary, with NS masses $m_1 = m_2 = 1.4 M_\odot$, initial interferometers have a range

$$d_{\rm NS-NS} = O(20)\,{\rm Mpc}, \tag{9.246}$$

[39] At time of writing (2007) initial LIGO has reached its target sensitivity and is completing a long science run, termed S5, with one year of coincident data between its detectors. As shown in Fig. 9.32, the noise budget is very well understood, and reproduces precisely the theoretical curves. VIRGO is presently close to reaching its target sensitivity, and is starting its first science runs. VIRGO and the LSC have signed an agreement for joint data taking and data exchange. The sensitivity of GW interferometers will probably be in rapid evolution in the near future, with various improvements and upgrades leading from the initial detectors to advanced interferometers.

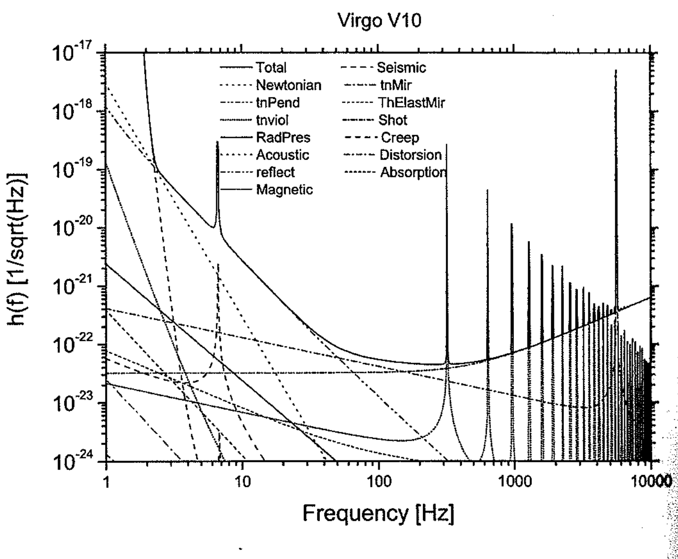


**Fig. 9.31** The predicted strain sensitivity $S_n^{1/2}(f)$ (here denoted $h(f)$) of the initial VIRGO detector, and the various noise contributions. (Courtesy of the VIRGO collaboration.)

which barely includes the Virgo cluster of galaxies. For BH-BH binaries, assuming a black-hole mass of $10M_{\odot}$, as suggested by the observation of typical stellar-mass galactic black holes, gives a sight distance at initial interferometers

$$d_{\rm BH-BH} = O(100)\,{\rm Mpc}\,. \tag{9.247}$$

Estimates of the rate are uncertain and will be discussed in detail in Vol. 2, where we will see that, if our theoretical understanding of the formation and evolution of compact binaries is correct, at these distances the chances of a detection are small, with $O(10^{-4})$–$O(10^{-1})$ events per year for BH-BH coalescences, and $O(10^{-3})$–$O(10^{-2})$ for NS-NS binaries. We will see however in Section 9.5.2 that these rates improves drastically for advanced interferometer.

For burst searches, the sensitivity of a broadband detector depends strongly on where, in frequency, the burst is peaked, and on its temporal duration. Assuming for definiteness a flat spectrum over the frequency bandwidth, a burst that radiates an energy $10^{-6}M_{\odot}c^2$ in GWs would be visible, at SNR = 8 and assuming optimal orientation, up to $O(10)$ kpc.

For spinning pulsars, the sight distance is obtained from eq. (7.166) as a function of the ellipticity $\epsilon$ and of its frequency $f_0$.



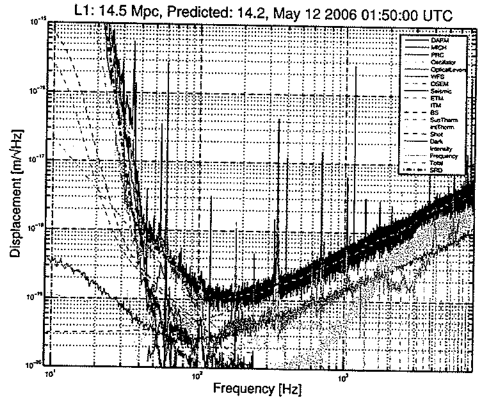


**Fig. 9.32** The strain sensitivity, in $\rm m/\sqrt{Hz}$, of the LIGO detector in Livingston. The strain sensitivity in $\rm Hz^{-1/2}$ is obtained dividing by the arm length $L = 4000$ m. The noise budget is very well understood. (Courtesy of the LIGO collaboration.)

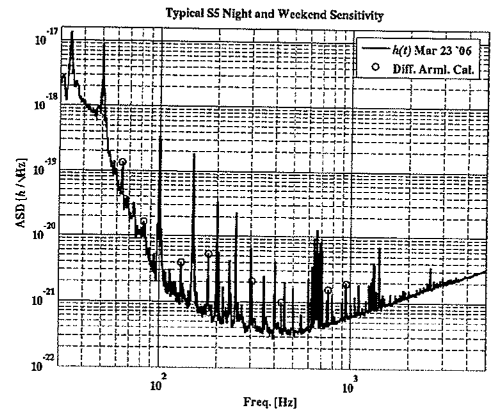


**Fig. 9.33** The strain sensitivity, in $1/\sqrt{\rm Hz}$, of the GEO detector during the S5 run. (Courtesy of the GEO collaboration.)

## 9.5.2 Advanced interferometers

### Ground-based detectors

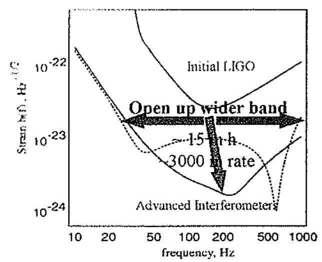


Fig. 9.34 The planned sensitivity of Advanced LIGO compared to initial LIGO. A wideband and a narrow-band configuration are shown. (Courtesy of the LIGO collaboration.)

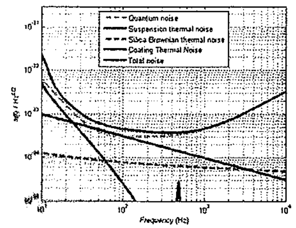

Fig. 9.35 The main noise contributions in a possible Advanced LIGO configuration. (Courtesy of the LIGO collaboration.)

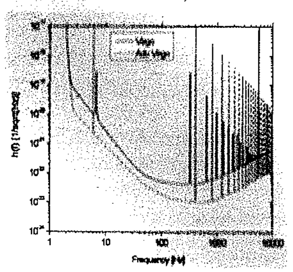

Fig. 9.36 A possible sensitivity curve of Advanced Virgo, compared to the initial Virgo. (Courtesy of the VIRGO collaboration.)

For the LIGO and VIRGO interferometers there is a well-defined plan for upgrades which should lead, in a few years, to second-generation detectors with much better sensitivities, Advanced LIGO (to which will contribute also GEO600 and the Australian consortium ACIGA) and Advanced VIRGO. GEO600 should evolve into a tunable narrow-band detector in the high-frequency region (GEO-HF), and in a facility for testing technologies for future interferometers. A cryogenic detector, LCGT, is under study in Japan, and there are plans for an interferometer in Australia, AIGO. Examples of planned sensitivities are shown in Figs. 9.34–9.36.

While a number of details might still change, the baseline for the main improvements that are planned is the following.

- An increase in the input laser power from the present value of order 10–20 W up to 100–200 W. After power recycling, this would lead to a laser power in the Fabry–Perot arm cavities of order 1 MW. This has the effect of reducing the shot noise, improving the sensitivity in the high-frequency region. Such a huge power will however induce thermal lensing in the test mass optics, due to absorption in the substrate and coatings, and compensation effects will be added.
- As discussed in Section 9.4.2, the increased laser power will produce a larger radiation pressure noise, up to the point that this becomes a dominant noise at low frequency. This is compensated increasing the mirror masses, up to about 40 kg.
- LIGO will introduce much better seismic isolation, improving the sensitivity at low frequencies and bringing the "seismic wall" from 40 Hz down to about 10 Hz. VIRGO already has a seismic isolation appropriate for an advanced interferometer.
- The test-mass suspensions, presently made of steel wires, will be replaced by silica (which has lower losses), fused to the mirror with silicate bonding to create a single monolithic object, thereby reducing suspension thermal noise. This technique has already been developed in GEO600. Further improvement can be obtained shaping the suspension fibers in the form of a ribbon. The resulting suspension thermal noise will be lower than the radiation pressure noise (in broad-band observation mode, see below), and comparable to the Newtonian background at 10 Hz.
- New mirror coatings, with lower thermal noise and lower losses (e.g. thanks to the insertion of dopants) are investigated.

The basic optical configuration is still a power-recycled interferometer with Fabry–Perot cavities in the arms. To this configuration is added *signal recycling*. This consists of adding a new mirror, the signal-recycling mirror, at the output port of the interferometer, i.e. between the beam-splitter and the photodetectors in Fig. 9.25. As in our discussion of the power recycling cavity, the addition of this mirror creates a new cavity, the signal-recycling cavity, composed by the "effective interferometer mirror" and the signal-recycling mirror.



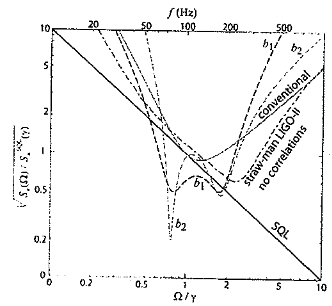


Fig. 9.37 Plot of $\sqrt{S_n(\Omega)/S_n^{\rm SQL}(\gamma)}$ (where $\Omega = 2\pi f$ and $\gamma = 2\pi f_{\rm pole}$) as a function of $\Omega/\gamma$, for different optical configurations. Reprinted with permission from Buonanno and Chen, *Class. Quantum Grav.* **18** L95 (2001b).

Recall from Section 9.2.2 that a GW of frequency $\omega_{\rm gw}$ generates in the interferometer sidebands at $\omega_L \pm \omega_{\rm gw}$. If the signal-recycling cavity is tuned in resonance with a sideband corresponding to some given value of $\omega_{\rm gw}$, the sensitivity of the interferometer for this GW frequency is enhanced, at the cost of the bandwidth. This enhancement depends on the finesse of the signal-recycling cavity. If the signal-recycling cavity is tuned to anti-resonance, the sidebands are extracted and the bandwidth of the detector is increased, with respect to the case where no signal-recycling cavity is present. This technique is know as *resonant sideband extraction*. As a result, with tiny adjustments of the position of the signal-recycling mirror, of order of fractions of $\lambda_L$, we can either tune the interferometer to some specific source or increase the bandwidth. These two options are illustrated in Fig. 9.34. The signal-recycling technique is already implemented in GEO600, although without Fabry–Perot cavities in the arms.

Thanks to these and to other improvements, an advanced interferometer in wideband mode will be limited, over a bandwidth from about 10 Hz to a few kHz, mostly by the optical read-out noise, as we see from Fig. 9.35, and therefore by the quantum limit. In the signal-recycling configuration it is however possible to perform quantum non-demolition measurements, hence going beyond the standard quantum limit, with a simple modification of the input–output optics, see Fig. 9.37 and the papers by Buonanno and Chen in the Further Reading section.

With the strain sensitivity of an advanced interferometer, the perspective for detection and for opening the field of GW astrophysics are quite good. A detailed discussion of the sources and their strength will be the subject of Vol. 2. Here we observe that the sight distance to NS-NS binaries becomes

$$d_{\rm NS-NS} = O(300)\,{\rm Mpc}\,. \qquad (9.248)$$

At this distance, the expected rate is between $O(10)$ per year and $O(100)$ per year. For BH-BH with masses $10M_\odot$, the sight distance becomes

$$d_{\rm BH-BH} = O(1.5)\,{\rm Gpc}\,, \qquad (9.249)$$

and expected rates are between one signal per year and $O(500)$ per year.[40] For BH-NS binaries,

$$d_{\rm BH-NS} = O(750)\,{\rm Mpc}\,, \qquad (9.250)$$

with an expected rate between one signal per year and $O(30)$ per year.

Looking further ahead, there are ideas for "third generation" interferometers. Among the features that are being considered, is the possibility of building an underground detector. As discussed in Section 9.4.4 (compare with Note 34 on page 524) the micro-seismic noise is mostly propagated through surface waves, so underground it is sensibly reduced, which also results in a reduction of the Newtonian noise induced by the

[40] A major source of theoretical uncertainty is related to the fact that the potential progenitors of a BH-BH system can go through a phase of common envelope evolution, that can lead to the merging of the progenitor stars rather than to the formation of a BH-BH binary. See Belczynski, Taam, Kalogera, Rasio and Bulik (2006) for a discussion of these rates and of their theoretical uncertainties.

micro-seismic motion. Another third-generation feature could be the use of cryogeny. A Japanese collaboration, building on the experience gained with TAMA, has proposed the realization of the LCGT detector, made of two independent underground interferometers in the same vacuum system, with 3 kms arms and cryogenic mirrors, cooled at 20 K, to be located at Kamioka, an old mine transformed in an underground physics laboratory. This site is about 1000 m below the top of a mountain, and provides a very stable seismic and temperature environment. A prototype, CLIO, has already been installed and is under development.

### Interferometers in space

The region below about 10 Hz is unaccessible to ground-based interferometers, because of the wall due to seismic and Newtonian noise. Still, we will see in Vol. 2 that the mHz region is potentially very rich in GW sources, including particularly fascinating objects such as supermassive black holes. The only way to detect them is to go in space, where the seismic noise is absent. One such project is LISA. The LISA mission is a collaboration between the European Space Agency (ESA) and NASA. The concept of the LISA mission is quite impressive. It consists of three spacecrafts, separated by 5 million kms, in a equilateral triangle configuration, orbiting the Sun. The center of the triangle should be at a distance of about 50 million kms (i.e. about 20° degrees along the orbit) behind the Earth. The size of the arms is chosen to optimize the sensitivity for GWs in the 10 mHz region, and in general LISA would be sensitive to GW frequencies in the range 0.1 mHz–0.1 Hz.

For a detailed description of the mission concept we refer the reader to the Further Reading section. Here we briefly mention some aspects of this remarkable experiment.

- Inside each spacecraft there will be two test masses (one for each arm), freely floating. The spacecraft is kept centered on the test masses using a *drag free* technique, in which the position of the masses is sensed, and the spacecraft adjusts its position with respect to them, using micro-thrusters. The thrusts necessary to maintain drag-free operation are extremely small, less than 100 $\mu$N, and the required recoil is obtained emitting in space just a handful of fast ions. This compensates for external influences such as solar winds, micro-meteorites, etc. that in the long term would sensibly alter the nominal position of the spacecraft. The LISA Pathfinder is a ESA mission to demonstrate the drag-free control technique at the required accuracy.
- The free masses exchange among them laser signals. Over a distance of 5 million kms, reflection is impossible because of power losses due to diffraction; after a travel of 5 million kms, the laser beam is spread over a surface of radius 20 kms. So LISA uses a laser transponding scheme in which the incoming laser light is sensed, and another laser is phase-locked to it and sends back another beam.



- LISA has unequal arms, with arm-lengths known to ±20 m from the measurement of the round-trip time. Then, contrary to what happens in a Michelson interferometer with equal arms, laser frequency fluctuations do not cancel out when taking the difference between two arms. For this reason, the LISA concept uses time-delay interferometry, in which the outputs of the two arms are recombined with a time delay that takes care of the arm difference. In the process, laser frequency noise is canceled (while the signal from GWs with frequencies in the mHz region is unaffected).
- After minimizing spurious forces on test masses, the other most important issue is the need to keep the temperature distribution in the spacecraft as constant as possible, since the mass displacements due local temperature fluctuations would induce changes in the Newtonian gravitational forces on the test mass, as well as accelerations of the test masses due to thermal radiation pressure.

Clearly, LISA would be an extremely impressive instrument and its scientific achievements could be truly spectacular.

A number of other space missions are currently discussed, such as DECIGO, a Japanese space project with arm-lengths shorter than LISA, to bridge the gap between LISA and the ground-based detectors; and ASTROD, a Chinese space project with arm-lengths longer than LISA which, among other relativity experiments, would extend the search for GWs to lower frequencies. Follows-up to the LISA mission, such as the Big-Bang Observer (BBO), are also being investigated.

## Further reading

- For a lively discussion of the history of gravitational-wave research, as well as the development of GW interferometers, see the popular book Thorne (1994). See also the review Thorne (1987).
- A textbook devoted to the interferometric detection of GWs is Saulson (1994). For reviews, see also Giazotto (1989), Drever (1991), and Ju, Blair and Zhao (2000). A large bibliography on GW interferometers can be found in the review by Rowan and Hough (2000).
- A detailed discussion of the optics of GW interferometers is the "VIRGO Physics Book, Optics and related Topic", available at http://wwwcascina.virgo.infn.it/vpb/. The effect of scattered light in interferometers is discussed in Vinet, Visson and Braccini (1996) and Vinet *et al.* (1997). Mesa beams are proposed in Bondarescu and Thorne (2006).
- Computations of the sensitivities to GWs of Fabry–Perot cavities in various configurations can be found in Vinet, Meers, Man and Brillet (1988) and Meers (1988, 1989).
- A pedagogical discussion of lock-in detection is given in Black and Gutenkunst (2003). A nice discussion of null instrument is given in Saulson (1994), Chapter 10. For a discussion of Pound–Drever–Hall locking see Saulson (1994), Section 12.5, and Black (2001). A discussion of the global control of the VIRGO detector can be found in Arnoud *et al.* (2005). For resonant sideband extraction see Mizuno *et al.* (1993).
- Shot noise in modulated interferometers is discussed in Niebauer, Schilling, Danzmann, Rüdiger and Winkler (1991) and Bondu (2003). Radiation pressure is discussed in Edelstein, Hough, Pugh and Martin (1978) and Caves (1980, 1981).

- For quantum non-demolition measurements see Caves, Thorne, Drever, Sandberg and Zimmermann (1980) and Braginsky and Kalili (1992). For application to advanced interferometers see Kimble, Levin, Matsko, Thorne and Vyatchanin (2000) and Buonanno and Chen (2001a, 2001b, 2002). For a review of quantum noise in GW interferometers see Corbitt and Mavalvala (2004).
- The effect of seismic noise in GW interferometers is discussed in Saulson (1994), Chapter 8. Newtonian noise are studied in Saulson (1984) and Beccaria *et al.* (1998). Our discussion of thermal noise followed the internal VIRGO note Flaminio *et al.* (2005), where calculations of the various thermal noise are performed in detail. Thermo-elastic noise is discussed in Braginsky and Vyatchanin (2003).
- Updated information on the existing GW interferometers, as well as technical documents, PhD theses, etc. can be found at
http://www.ligo.caltech.edu/ (LIGO)
http://wwwcascina.virgo.infn.it/ (VIRGO)
http://www.geo600.uni-hannover.de/ (GEO600)
http://tamago.mtk.nao.ac.jp/ (TAMA)
- A detailed description of the LISA mission can be found in the LISA Pre-Phase A Report (1998). See also the reviews Bender (2001), Danzmann and Rüdiger (2003), and the web site http://lisa.jpl.nasa.gov.
- For lack of space, we have not discussed experiments searching for GWs using the Doppler tracking of spacecraft. For a recent review, see Armstrong (2006).

# Bibliography


Abbott, B. *et al.* [LIGO Science Community, LSC] (2004a). *Phys. Rev. D* **69**, 082004.

Abbott, B. *et al.* [LSC] (2004b). *Phys. Rev. D* **69**, 102001.

Abbott, B. *et al.* [LSC] (2004c). *Phys. Rev. D* **69**, 122001.

Abbott, B. *et al.* [LSC] (2004d). *Phys. Rev. D* **69**, 122004.

Abbott B. *et al.* [LSC] (2005a). *Phys. Rev. D* **72**, 082001.

Abbott, B. *et al.* [LSC] (2005b). *Phys. Rev. Lett.* **94** 181103.

Abbott, B. *et al.* [LSC] (2005c). *Phys. Rev. Lett.* **95**, 221101.

Abramowitz, M. and I. A. Stegun (1972). *Handbook of Mathematical Functions*, Dover, New York.

Aguiar, O. D. *et al.* (2004). *Class. Quant. Grav.* **21**, S457.

Allen, Z. A. *et al.* [IGEC] (2000). *Phys. Rev. Lett.* **85**, 5046 .

Allen, B. and J. D. Romano (1999). *Phys. Rev. D* **59**, 102001.

Anderson, W. G. and R. Balasubramanian (1999). *Phys. Rev. D* **60**, 102001.

Anderson, W. G., P. R. Brady, J. D. E. Creighton and E. E. Flanagan (2001). *Phys. Rev. D* **63**, 042003.

de Andrade, V. C., L. Blanchet and G. Faye (2001). *Class. Quant. Grav.* **18**, 753.

Apostolatos, T. A., C. Cutler, G. J. Sussman and K. S. Thorne (1994). *Phys. Rev. D* **49**, 6274.

Arnaud, N. *et al.* (2005). *Nucl. Instr. Meth. A* **550**, 467.

Arnowitt, R., S. Deser and C. W. Misner (1961). *Phys. Rev.* **121**, 1556.

Arkani-Hamed, N., H. Georgi and M. D. Schwartz (2003). *Ann. Phys.* **305**, 96.

Arkani-Hamed, N., H. C. Cheng, M. A. Luty and S. Mukohyama (2004). *JHEP* **0405**, 074.

Armstrong, J. W. (2006). *Living Rev. Relativity* **9**, 1. Online article http://www.livingreviews.org/lrr-2006-1.

Arun, K. G., L. Blanchet, B. R. Iyer and M. S. Qusailah (2004). *Class. Quant. Grav.* **21**, 3771. [Erratum-ibid. **22**, (2005) 3115].

Ashby, N. and J. Dreitlein (1975). *Phys. Rev. D* **12**, 336.

Astone, P., *et al.* [ROG] (1997b). *Astroparticle Phys.* **7**, 231.

Astone, P., *et al.* [ROG] (2002). *Phys. Rev. D* **65**, 022001.

Astone, P., *et al.* [IGEC] (2003a). *Phys. Rev. D* **68**, 022001

Astone, P., *et al.* [IGEC-2] (2007). arXiv:0705.0688[gr-qc].

Astone, P., *et al.* [ROG] (2003b). *Phys. Rev. Lett.* **91**, 111101.

Backer, D. C. and R. W. Hellings (1986). *Ann. Rev. Astron. Astrophys.*

**24**, 537.
Bardeen, J. M. and W. H. Press (1973). *J. Math. Phys.* **14**, 7.
Bassan, M. (2002). In *Proceeding of the 15th International Conference on General Relativity and Gravitation*, M. Francaviglia ed., World Scientific, Singapore.
Beccaria, M. *et al.* (1998). *Class. Quant. Grav.* **15**, 3339.
Belczynski, K., R. E. Taam, V. Kalogera, F. A. Rasio and T. Bulik (2006). arXiv:astro-ph/0612032.
Bender, P. L. (2001). In *Gravitational Waves*, I. Ciufolini *et al.* eds, IoP Publishing, Bristol.
van den Berg, J. C. ed. (1999). *Wavelets in Physics.* Cambridge University Press, Cambridge.
Bertotti, B. and A. M. Anile (1973). *Astron. Astrophys.* **28**, 429.
Bianchi, M., E. Coccia, C. N. Colacino, V. Fafone and F. Fucito (1996). *Class. Quant. Grav.* **13**, 2865.
Bianchi, M., M. Brunetti, E. Coccia, F. Fucito, and J. A. Lobo (1998). *Phys. Rev. D* **57**, 4525.
Binney, J. and M. Merrifield (1998). *Galactic Astronomy.* Princeton University Press, Princeton.
Binney, J. and S. Tremaine (1994). *Galactic Dynamics.* Princeton University Press, Princeton.
Black E. D. (2001). *Am. J. Phys.* **69**, 79.
Black E. D. and R. N. Gutenkunst (2003). *Am. J. Phys.* **71**, 365.
Blair, D., *et al.* (1995). *Phys. Rev. Lett.* **74**, 1908.
Blanford, R. D. and S. A. Teukolsky (1976). *Astrophys. J.* **205**, 580.
Blanchet, L. (1987). *Proc. Roy. Soc. Lond. A* **409**, 383.
Blanchet, L. (1993). *Phys. Rev. D* **47**, 4392.
Blanchet, L. (1995). *Phys. Rev. D* **51**, 2559.
Blanchet, L. (1996). *Phys. Rev. D* **54**, 1417.
Blanchet, L. (1997). *Phys. Rev. D* **55**, 714.
Blanchet, L. (1998a). *Class. Quant. Grav.* **15**, 89.
Blanchet, L. (1998b). *Class. Quant. Grav.* **15**, 113.
Blanchet, L. (1998c). *Class. Quant. Grav.* **15**, 1971.
Blanchet, L. (2006). *Living Rev. Relativity* **9**, 4. Online article http://www.livingreviews.org/lrr-2006-4.
Blanchet, L. and T. Damour (1986). *Phil. Trans. R. Soc. London* **A320**, 379.
Blanchet, L. and T. Damour (1988). *Phys. Rev. D* **37**, 1410.
Blanchet, L. and T. Damour (1989). *Ann. Inst. H. Poincaré* **50**, 377.
Blanchet, L. and G. Schäfer (1989). *Mon. Not. R. Astron. Soc.* **239**, 845.
Blanchet, L. and T. Damour (1992). *Phys. Rev. D* **46**, 4304.
Blanchet, L. and G. Schäfer (1993). *Class. Quant. Grav.* **10**, 2699.
Blanchet, L., T. Damour and B. R. Iyer (1995). *Phys. Rev. D* **51**, 5360. [Erratum-ibid. *D* **54**, (1996) 1860].
Blanchet, L. and B. S. Sathyaprakash (1995). *Phys. Rev. Lett.* **74**, 1067.
Blanchet, L., B. R. Iyer, C. M. Will and A. G. Wiseman (1996). *Class. Quant. Grav.* **13**, 575.



Blanchet, L., G. Faye and B. Ponsot (1998). *Phys. Rev. D* **58**, 124002.
Blanchet, L. and G. Faye (2001). *Phys. Rev. D* **63**, 062005.
Blanchet, L., B. R. Iyer and B. Joguet (2002). *Phys. Rev. D* **65**, 064005. [Erratum-ibid. *D* **71**, 129903 (2005)]
Blanchet, L., G. Faye, B. R. Iyer and B. Joguet (2002). *Phys. Rev. D* **65**, 061501. [Erratum-ibid. *D* **71**, 129902 (2005)].
Blanchet L. and B. R. Iyer (2003). *Class. Quant. Grav.* **20**, 755.
Blanchet, L., T. Damour and G. Esposito-Farèse (2004). *Phys. Rev. D* **69**, 124007.
Blanchet, L., T. Damour, G. Esposito-Farèse and B. R. Iyer (2004). *Phys. Rev. Lett.* **93**, 091101.
Blanchet, L., T. Damour, G. Esposito-Farèse and B. R. Iyer (2005). *Phys. Rev. D* **71**, 124004.
Blanchet, L., G. Faye, and S. Nissanke (2005). *Phys. Rev. D* **72**, 044024.
Blanchet, L., A. Buonanno and G. Faye (2006). *Phys. Rev. D* **74**, 104034.
Bonazzola, S. and E. Gourgoulhon (1996). *Astron. Astrophys.* **312**, 675.
Bondarescu, M. and K. S. Thorne (2006). *Phys. Rev. D* **74**, 082003.
Bondi, H., M. G. J. van der Burg and A. K. W. Metzner (1962). *Proc. Roy. Soc. London A* **269**, 21.
Bondu, F. (2003). VIRGO internal note VIR-NOT-OCA-1390-243, unpublished.
Bontz R.J. and M.P. Haugan (1981). *Astrophysics and Space Science* **78**, 199.
Boulware, D. G. and S. Deser (1972). *Phys. Rev. D* **6**, 3368.
Brady, P. R., T. Creighton, C. Cutler and B. F. Schutz (1998). *Phys. Rev. D* **57**, 2101.
Brady, P. R. and T. Creighton (2000). *Phys. Rev. D* **61**, 082001.
Braginsky, V. B. and L. P. Grishchuk (1986). *Sov. Phys. JETP* **62**, 427.
Braginsky, V. B. and K. S. Thorne (1987). *Nature* **327**, 123.
Braginsky, V. B. and F. Ya. Kalili (1992). *Quantum Measurement.* Cambridge University Press, Cambridge.
Braginsky, V. B. and S. P. Vyatchanin (2003). *Phys. Lett. A* **312**, 244.
Briant, T., *et al.* (2003). *Phys. Rev. D* **67**, 102005.
Bruhat, Y. (1962). In *Gravitation: An Introduction to Current Research*, L. Witten ed., Wiley, New York.
Buonanno, A. and Y. B. Chen (2001a). *Phys. Rev. D* **64**, 042006.
Buonanno, A. and Y. B. Chen (2001b). *Class. Quant. Grav.* **18**, L95.
Buonanno, A. and Y. B. Chen (2002). *Phys. Rev. D* **65**, 042001.
Buonanno, A., Y. B. Chen and M. Vallisneri, (2003). *Phys. Rev. D* **67**, 024016.
Burgay, M. *et al.* (2003). *Nature*, **426**, 531.
Burke, W. L. (1971). *J. Math. Phys.* **12**, 401.
Caves, C. M. (1980). *Phys. Rev. Lett.* **45**, 75.
Caves, C. M. (1981). *Phys. Rev. D* **23**, 1693.
Caves, C. M., K. S. Thorne, R. W. P. Drever, V. D. Sandberg, and M. Zimmermann (1980). *Rev. Mod. Phys.* **52**, 341.
Cerdonio, M. *et al.* [AURIGA] (1997). *Class. Quant. Grav.* **14**, 1491.

Cerdonio, M., L. Conti, J. A. Lobo, A. Ortolan and J. P. Zendri (2001). *Phys. Rev. Lett.* **87**, 031101.
Cerdonio, M. (2003). *Class. Quant. Grav.* **20**, S23.
Chandrasekhar, S. (1969). *Ellipsoidal Figures of Equilibrium.* Yale University Press, New Haven.
Chandrasekhar, S. and F. P. Esposito (1970). *Astrophys. J.* **160**, 153.
Christensen, N. (1992). *Phys. Rev. D* **46**, 5250.
Christodoulou, D. (1991). *Phys. Rev. Lett.* **67**, 1486.
Christodoulou, D. and S. Klainerman (1993). *The global nonlinear stability of the Minkowski space.* Princeton University Press, Princeton.
Clark, J.P.A. and D. M. Eardley (1977). *Astrophys. J.* **215**, 315.
Coccia, E. (1997). In *Proceeding of the 14th International Conference on General Relativity and Gravitation*, M. Francaviglia ed., World Scientific, Singapore.
Coccia, E., J. A. Lobo, and J. A. Ortega (1995). *Phys. Rev. D* **52**, 3735.
Coccia, E. and V. Fafone (1996). *Phys. Lett. A* **213**, 16.
Coccia, E., V. Fafone, G. Frossati, J. A. Lobo and J. A. Ortega (1998). *Phys. Rev. D* **57**, 2051.
Coccia, E., F. Dubath and M. Maggiore (2004). *Phys. Rev. D* **70**, 084010.
Corbitt, T. and N. Mavalvala (2004). *J. Opt. B* **6**, S675.
Cousins, R. D. (1995). *Am. J. Phys.* **63**, 398.
Creminelli, P., A. Nicolis, M. Papucci and E. Trincherini (2005). *JHEP* **0509**, 003.
Cutler, C. *et al.* (1993). *Phys. Rev. Lett.* **70**, 2984.
Cutler, C. and E. E. Flanagan (1994). *Phys. Rev. D* **49**, 2658.
Cutler, C. and D. I. Jones (2001). *Phys. Rev. D* **63**, 024002.
van Dam, H. and M. Veltman (1970). *Nucl. Phys. B* **22**, 397.
Damour, T. (1983a). *Phys. Rev. Lett.* **51**, 1019.
Damour, T. (1983b). In *Gravitational Radiation*, N. Deruelle and T. Piran eds., Les Houches 1982, North-Holland, Amsterdam.
Damour, T. (1986). In *Proceedings of the 4th Marcel Grossmann Meeting on General Relativity* ed R. Ruffini, North-Holland, Amsterdam.
Damour, T. (1987). In *300 Years of Gravitation*, S. Hawking and W. Israel eds., Cambridge University Press, Cambridge.
Damour, T. and N. Deruelle (1981). *Phys. Lett. A* **87**, 81.
Damour, T. and N. Deruelle (1985). *Ann. Inst. H. Poincaré* **43**, 107.
Damour, T. and N. Deruelle (1986). *Ann. Inst. H. Poincaré* **44**, 263.
Damour, T. and G. Schäfer (1988). *Nuovo Cimento B*, **101**, 127.
Damour, T. and B. R. Iyer (1991a). *Phys. Rev. D* **43**, 3259.
Damour, T. and B. R. Iyer (1991b). *Ann. Inst. H. Poincaré* **54**, 115.
Damour, T. and J. H. Taylor (1991). *Astrophys. J.* **366**, 501.
Damour, T. and J. H. Taylor (1992). *Phys. Rev. D* **45**, 1840.
Damour, T., P. Jaranowski and G. Schäfer (2000). *Phys. Rev. D* **62**, 021501. [Erratum-ibid. *D* **63**, (2001) 029903].
Damour, T., P. Jaranowski and G. Schäfer (2001a). *Phys. Rev. D* **63**, 044021. [Erratum-ibid. *D* **66**, (2002) 029901].
Damour, T., P. Jaranowski and G. Schäfer (2001b). *Phys. Lett.* **513**, 147.
Damour, T., B. R. Iyer and B. S. Sathyaprakash (2001). *Phys. Rev. D* **63**, 044023. [Erratum-ibid. *D* **72** (2005) 029902].
Damour, T., I. I. Kogan and A. Papazoglou (2003). *Phys. Rev. D* **67**, 064009.
Damour, T., A. Gopakumar and B. R. Iyer (2004). *Phys. Rev. D* **70**, 064028.
Danzmann, K. and A. Rüdiger (2003). *Class. Quant. Grav.* **20**, S1.
Davies, M. R., R. Ruffini, W. H. Press and R.H. Price (1971). *Phys. Rev. Lett.* **27**, 1466.
Davies, M. R., R. Ruffini and J. Tiomno (1972). *Phys. Rev. D* **5**, 2932.
Deffayet, C., G. R. Dvali, G. Gabadadze and A. I. Vainshtein (2002). *Phys. Rev. D* **65**, 044026.
Detweiler, S. L. and E. Szedenits (1979). *Astrophys. J.* **231**, 211.
Deser, S. (1970). *Gen. Relativ. Gravit.* **1**, 9.
DeWitt, B. S. (1967). *Phys. Rev.* **162**, 1239.
Drever, R. W. P. (1991). In *The Detection of Gravitational Waves*, D. G. Blair ed., Cambridge University Press, Cambridge.
Dubath, F., S. Foffa, M. A. Gasparini, M. Maggiore and R. Sturani (2005). *Phys. Rev. D* **71**, 124003.
Dubovski, S.L., P. G. Tinyakov and I. I. Tkachev (2005). *Phys. Rev. Lett.* **94**, 181102.
Edelstein, W. A., J. Hough, J. R. Pugh and W. Martin (1978). *J. Phys. E* **11**, 710.
Ehlers, J., A. Rosenblum, J. N. Goldberg and P. Havas (1976). *Astrophys. J.* **208**, L77.
Einstein, A., L. Infeld and B. Hoffmann (1938). *Ann. Math.* **39**, 65.
Epstein, R. (1977). *Astrophys. J.* **216**, 92.
Epstein, R. (1979). *Astrophys. J.* **231**, 644.
Epstein, R. and R. V. Wagoner (1975). *Astrophys. J.* **197**, 717.
Faye, G., L. Blanchet and A. Buonanno (2006). *Phys. Rev. D* **74**, 104033.
Feldman G. J. and R. D. Cousins (1998). *Phys. Rev. D* **57**, 3873.
Feynman, R. P., F. B. Morinigo, and W. G. Wagner (1995). *Feynman Lectures on Gravitation*, Addison-Wesley, Reading, MA.
Finn, S. L. (1992). *Phys. Rev. D* **46**, 5236.
Finn, S. L. and D. F. Chernoff (1993). *Phys. Rev. D* **47**, 2198.
Finn, S. L. and A. Lazzarini (2001). *Phys. Rev. D* **64**, 082002.
Finn, S. L. and P. J. Sutton (2002). *Phys. Rev. D* **65**, 044022.
Flaminio, R. *et al.* (2005). Internal VIRGO note VIR-NOT-DIR-1390-304, unpublished.
Flanagan, E. E. (1993). *Phys. Rev. D* **48**, 2389.
Flanagan, E. E. and S. A. Hughes (1998a). *Phys. Rev. D* **57**, 4535.
Flanagan, E. E. and S. A. Hughes (1998b). *Phys. Rev. D* **57**, 4566.
Forward, R. (1971). *Gen. Relativ. Gravit.* **2**, 149.
Giazotto, A. (1989). *Phys. Rept.* **182**, 365.
Gottardi, L. (2004). *Transducers and low-noise two-stage SQUID amplifiers for the spherical gravitational wave antenna MiniGrail.* PhD

thesis, Leiden University, Leiden, The Netherlands.
Gibbons, G. W. and S. W. Hawking (1971). *Phys. Rev. D* **4**, 219.
Giffard, R. P. (1976). *Phys. Rev. D* **14**, 2478.
Goldhaber, A. S. and Nieto, M. M. (1974). *Phys. Rev. D* **9**, 1119.
Goldberger, W. D. and I. Z. Rothstein (2006). *Phys. Rev. D* **73**, 104029.
Gopakumar, A., B. R. Iyer and S. Iyer (1997). *Phys. Rev. D* **55**, 6030.
Gopakumar, A. and B. R. Iyer (2002). *Phys. Rev. D* **65**, 084011.
Gradshteyn I. S. and Ryzhik, I. M. (1980). *Table of Integrals, Series and Products.* Academic Press, London.
Grisaru, M. T., P. van Nieuwenhuizen, and C. C. Wu (1975). *Phys. Rev. D* **12**, 397.
Gupta S. N. (1954). *Phys. Rev.* **96**, 1683.
Gürsel, Y. and M. Tinto (1989). *Phys. Rev. D* **40**, 3884.
Hamermesh, M. (1962). *Group Theory.* Addison-Wesley, Redwood.
Hartle, J. B. (2003). *Gravity. An Introduction to Einstein's General Relativity*, Addison Wesley, San Francisco.
Haugan, M. P. (1985). *Astrophys. J.* **296**, 1.
Haugan, M. P., S. L. Shapiro, and I. Wasserman (1982). *Astrophys. J.* **257**, 283.
Hulse, R. A. (1994) *Rev. Mod. Phys.* **66**, 699.
Hulse, R. A. and J. H. Taylor (1975). *Astrophys. J.* **195**, L51.
Isaacson, R. A. (1968a). *Phys. Rev.* **166**, 1263.
Isaacson, R. A. (1968b). *Phys. Rev.* **166**, 1272.
Iwasaki, Y. (1970). *Phys. Rev. D* **2**, 2255.
Iyer, B. R. and C. M. Will (1993). *Phys. Rev. Lett.* **70**, 113.
Iyer, B. R. and C. M. Will (1995). *Phys. Rev. D* **52**, 6882.
Itoh, Y. (2004). *Phys. Rev. D* **69**, 064018.
Itoh, Y., Futamase, T. and Asada, H. (2000). *Phys. Rev. D* **62**, 064002.
Itoh, Y., Futamase, T. and Asada, H. (2001). *Phys. Rev. D* **63**, 064038.
Itoh, Y. and Futamase, T. (2003). *Phys. Rev. D* **68**, 121501(R).
Jackson, J. D. (1975). *Classical Electrodynamics.* Wiley, Chichester.
Jackson, J. D. and L. B. Okun (2001). *Rev. Mod. Phys.* **73**, 663.
Jaranowski, P. and G. Schäfer (1998). *Phys. Rev. D* **57**, 7274. [Erratum-ibid. *D* **63** (2001) 029902].
Jaranowski, P. and G. Schäfer (1999). *Phys. Rev. D* **60**, 124003.
Jaranowski, P. and G. Schäfer (2000). *Ann. Phys.* **9**, 378.
Johnson, W. W. and S. M. Merkowitz (1993). *Phys. Rev. Lett.* **70**, 2367.
Johnson, W. W. and S. M. Merkowitz (1997). *Phys. Rev. D* **56**, 7513.
Ju, L., D. G. Blair and C. Zhao (2000). *Rep. Prog. Phys.* **63**, 1317.
Kennefick, D. (1994). *Phys. Rev. D* **50**, 3587.
Kennefick, D. (1997). arXiv:gr-qc/9704002.
Kennefick, D. (2007). *Traveling at the Speed of Thought.* Princeton University Press, Princeton.
Kidder, L. E. (1995). *Phys. Rev. D* **52**, 821.
Kidder, L. E., C. M. Willand and A. G. Wiseman (1993). *Phys. Rev. D* **47**, R4183.
Klimenko, S., I. Yakushin, M. Rakhmanov and G. Mitselmakher (2004).



*Class. Quant. Grav.* **21**, S1685.
Klimenko, S. and G. Mitselmakher (2004). *Class. Quant. Grav.* **21**, S1819.
Kimble, H. J., Y. Levin, A. B. Matsko, K. S. Thorne and S. P. Vyatchanin (2002). *Phys. Rev. D* **65**, 022002.
Konigsdorffer, C., G. Faye and G. Schäfer (2003). *Phys. Rev. D* **68**, 044004.
Kolb, E. W. and M. S. Turner (1990). *The Early Universe*, Addison Wesley, Reading, MA.
Kraichnan, R. H. (1955). *Phys. Rev.* **98**, 1118; *ibid.* **101**, 482.
Kramer, M. *et al.*, (2005). "Testing GR with the Double pulsar: Recent Results", astro-ph/0503386.
Kramer, M. *et al.*, (2006). "Tests of general relativity from timing the double pulsar" astro-ph/0609417.
Krishnan, B. *et al.* (2004). *Phys. Rev. D* **70**, 082001.
Królak, A., K. D. Kokkotas and G. Schäfer (1995). *Phys. Rev. D* **52**, 2089.
Landau, L. D. and E. M. Lifshitz (1976). *Course of Theoretical Physics*, Vol. I: *Mechanics.* Pergamon Press, Oxford.
Landau, L. D. and E. M. Lifshitz (1979). *Course of Theoretical Physics*, Vol. II: *The Classical Theory of Fields.* Pergamon Press, Oxford.
Landau, L.D. and E. M. Lifshitz (1982). *Course of Theoretical Physics*, Vol. IV (by V.B. Berestetskij, E. M. Lifshitz and L. P. Pitaevskij): *Quantum Electrodynamics.* Pergamon Press, Oxford.
Landau, L. D. and E. M. Lifshitz (1970). *Course of Theoretical Physics*, Vol. VII: *Theory of Elasticity.* Pergamon Press, Oxford.
Larson, S. L. and W. A. Hiscock (2000). *Phys. Rev. D* **61**, 104008.
*LISA: Pre-Phase A Report*, MPQ 208, Max-Planck-Institute für Quantenoptik, Garching, Germany.
Lobo, J. A. (1995). *Phys. Rev. D* **52**, 591.
Lobo, J. A. (2002). *Class. Quant. Grav.* **19**, 2029.
Lorimer, D. R. (2005). Living Rev. Relativity **8**, 7. Online article http://www.livingreviews.org/lrr-2005-7.
Love, A. E. H. (1944). *A Treatise on the Mathematical Theory of Elasticity.* Dover, New York.
Lyne, A. G. *et al.* (2004). *Science*, **303**, 1153.
Lyne, A. G. and F. Graham-Smith (2005). *Pulsar Astronomy*, 3rd edition, Cambridge University Press, Cambridge.
Lyons, L. (1986). *Statistics for Nuclear and Particle Physicists*, Cambridge University Press, Cambridge.
McDonough, R. N. and A. D. Whalen (1995). *Detection of Signals in Noise*, 2nd edition. Academic Press, San Diego.
Maggiore, M. (2000). *Phys. Rept.* **331**, 283.
Maggiore, M. (2005). *A Modern Introduction to Quantum Field Theory.* Oxford University Press, Oxford.
Maggiore, M. and A. Nicolis, (2000). *Phys. Rev. D* **62**, 024004.
Mauceli, E., *et al.* [ALLEGRO] (1996). *Phys. Rev. D* **54**, 1264.
Meers, B. J. (1988). *Phys. Rev. D* **38**, 2317.

Meers, B. J. (1989). *Phys. Lett. A* **142**, 465.
Merkowitz, S. (1995). *Truncated icosahedral gravitational wave antenna*, PhD Thesis, Lousiana State University, available online at http://sam.phys.lsu.edu.
Merkowitz, S. M. and W. W. Johnson (1995). *Phys. Rev. D* **51**, 2546.
Michelson, P. F. (1987). *Mon. Not. R. Astron. Soc.* **227**, 933.
Michelson, P. F. and R. C. Taber (1981). *J. Appl. Phys.* **52**, 4313.
Michelson, P. F. and R. C. Taber (1984). *Phys. Rev. D* **29**, 2149.
Misner, C. W., K. S. Thorne and J. A. Wheeler (1973). *Gravitation*, Freeman, New York.
Mizuno, J. *et al.* (1993). *Phys. Lett. A* **175**, 273.
Newman, E. T. and R. Penrose (1968). *Proc. Roy. Soc. London A*, **305**, 175.
Ni, W.-T. and M. Zimmermann (1978). *Phys. Rev. D* **17**, 1473.
Niebauer, T.M., R. Schilling, K. Danzmann, A. Rüdiger and W. Winkler (1991). *Phys. Rev. A* **43**, 5022.
van Niewenhuizen, P. (1973). *Phys. Rev. D* **7**, 2300.
Nissanke, S. and L. Blanchet (2005). *Class. Quant. Grav.* **22**, 1007.
Ogievetsky, V. I. and I. V. Polubarinov (1965). *Ann. Phys.* **35**, 167.
Ohanian, H. C. and R. Ruffini (1994). *Gravitation and Spacetime*. Norton, New York.
Owen, B. J. (1996). *Phys. Rev. D* **53**, 6749.
Owen, B. J. and B. S. Sathyaprakash (1999). *Phys. Rev. D* **60**, 022002.
Padmanabhan, T. (2004). arXiv:gr-qc/0409089.
Paik, H. J. (1976). *J. Appl. Phys.* **47**, 1168.
Pallottino, G. V. and G. Pizzella (1981). *Nuovo Cim. C* **4**, 237.
Pallottino, G. V. and G. Pizzella (1984). *Nuovo Cim. C* **7**, 155.
Pallottino, G. V. and G. Pizzella (1991). In *The Detection of Gravitational Waves*, D. G. Blair ed., Cambridge University Press, Cambridge.
Pati, M. E. and C. M. Will (2000). *Phys. Rev. D* **62**, 124015.
Pati, M. E. and C. M. Will (2002). *Phys. Rev. D* **65**, 104008.
Penrose, R. (1963). *Phys. Rev. Lett.* **10**, 66.
Penrose, R. (1965). *Proc. Roy. Soc. London A* **284**, 159.
Peters, P. C. (1964). *Phys. Rev.* **136**, B1224.
Peters, P. C. and J. Mathews (1963). *Phys. Rev.* **131**, 435.
Pirani, F. A. E. (1964). In *Lectures on General Relativity*, A. Trautman *et al.* eds., Prentice-Hall, Englewood Cliffs.
Poisson, E. (1993). *Phys. Rev. D* **47**, 1497.
Poisson, E. (1999). arXiv:gr-qc/9912045.
Poisson, E. and C. M. Will (1995). *Phys. Rev. D* **49**, 848.
Poujade, O. and L. Blanchet (2002). *Phys. Rev. D* **65**, 124020.
Price, J. C. (1987). *Phys. Rev. D* **36**, 3555.
Rendall, A. D. (1992). *Proc. Roy. Soc. London A* **438**, 341.
Richard J.-P. and W. M. Folkner (1991). In *The Detection of Gravitational Waves*, D. G. Blair ed., Cambridge University Press, Cambridge.
Rowan, S. and J. Hough, (2000). *Living Rev. Rel.* **3**, 3. Online article http://www.livingreviews.org/lrr-2000-3.
Rubakov, V., arXiv:hep-th/0407104.
Sachs, R. K. (1961). *Proc. Roy. Soc. London A* **264**, 309.
Sachs, R. K. (1962). *Proc. Roy. Soc. London A* **270**, 103.
Sasaki, M. (1994). *Prog. Theor. Phys.* **92**, 17.
Saulson, P. R. (1984). *Phys. Rev. D* **30**, 732.
Saulson, P. R. (1994). *Fundamentals of Interferometric Gravitational Wave Detectors*, World Scientific, Singapore.
Schutz, B. F. (1986). *Nature*, **323**, 310.
Schutz, B. F. (1991). In *The Detection of Gravitational Waves*, D. G. Blair ed., Cambridge University Press, Cambridge.
Schutz, B. F. and F. Ricci (2001). In *Gravitational Waves*, I. Ciufolini *et al.* eds, IoP Publishing, Bristol.
Shapiro, S.L. and S.A. Teukolsky (1983). *Black Holes, White Dwarfs and Neutron Stars*, Wiley, New York.
Smarr, L. (1979). In *Sources of Gravitational Radiation*, ed. L. Smarr. Cambridge University Press, Cambridge.
Spallicci, A., G. Frossati and A. Królak (1997) *Class. Quant. Grav.* **14**, 577.
Stairs, I. H. (2003). *Living Rev. Relativity* **6**, 5. Online article http://www.livingreviews.org/lrr-2003-5.
Stairs, I. H., S. E. Thorsett, J. H. Taylor and A. Wolszczan (2002). *Astrophys. J.* **581**, 501.
Stevenson, T. R. (1997). *Phys. Rev. D* **56**, 564.
Straumann, N. (2004). *General Relativity. With Applications to Astrophysics*. Springer, Berlin.
Tagoshi, H. and M. Sasaki (1994). *Prog. Theor. Phys.* **92**, 745.
Tagoshi, H., A. Ohashi and B. J. Owen (2001). *Phys. Rev. D* **63**, 044006.
Tanaga, T., H. Tagoshi and M. Sasaki (1996). *Prog. Theor. Phys.* **96**, 1087.
Taylor, J. H. (1994). *Rev. Mod. Phys.* **66**, 711.
Taylor, J. H., L. A. Fowler and P. M. McCulloch (1979). *Nature* **277**, 437.
Taylor, J. H. and J. M. Weisberg (1982). *Astrophys. J.* **253**, 908.
Taylor, J. H. and J. M. Weisberg (1989). *Astrophys. J.* **345**, 434.
Thorne, K. S. (1980). *Rev. Mod. Phys.* **52**, 299.
Thorne, K. S. (1983). In *Gravitational Radiation*, N. Deruelle and T. Piran eds., Les Houches 1982, North-Holland, Amsterdam.
Thorne, K. S. (1987). In *300 Years of Gravitation*, S. Hawking and W. Israel eds., Cambridge University Press, Cambridge.
Thorne, K. S. (1992). *Phys. Rev. D* **45**, 520.
Thorne, K. S. (1994). *Black Holes and Time Warps: Einstein's Outrageous Legacy*, Norton, New York.
Thorne, K. S and Y. Gürsel (1983). *Mon. Not. R. Astron. Soc.* **205**, 809.
Vainshtein, A. I. (1972). *Phys Lett. B* **39**, 393.
Van Den Broeck, C. (2005). *Class. Quant. Grav.* **22**, 1825.

Veitch P. J. (1991). In *The Detection of Gravitational Waves*, D. G. Blair ed., Cambridge University Press, Cambridge.
Veltman, M. J. G. (1976). *Quantum Theory of Gravitation*, in *Methods in Field Theory*, Les Houches 1975, R. Balian and J. Zinn-Justin eds., North-Holland, Amsterdam.
Viceré, A. (2000). In *Experimental Physics of Gravitational Waves*, Urbino 1999 Summer School, M. Barone *et al.* eds., World Scientific, Singapore.
Viceré, A. (2002). *Phys. Rev. D* **66**, 062002.
Vinet, J.-Y., B. Meers, C. N. Man and A. Brillet (1988). *Phys. Rev. D* **38**, 433.
Vinet, J.-Y., V. Brisson and S. Braccini (1996). *Phys. Rev. D* **54**, 1276.
Vinet, J.-Y. *et al.* (1997). *Phys. Rev. D* **56**, 6085.
de Waard, A. (2003). *MiniGrail. The first spherical gravitational wave antenna.* PhD thesis, Leiden University, Leiden, The Netherlands.
de Waard, A., L. Gottardi, J. van Houwelingen, A. Schumack and G. Frossati (2003). *Class. Quant. Grav.* **20**, S143.
Wagoner, R. V. (1979). *Phys. Rev. D* **19**, 2897.
Wagoner, R. V. and C. M. Will (1976). *Astrophys. J.* **210**, 764.
Wagoner, R. V. and H. J. Paik (1977). In *Experimental Gravitation, Proceedings of the Pavia International Symposium*, Accademia Naz. dei Lincei, Roma.
Wainstein, L. A. and V. D. Zubakov (1965). *Extraction of Signals from Noise*, Dover Publications, New York.
Walker, M. and C. Will (1980a). *Phys. Rev. Lett.* **45**, 1741.
Walker, M. and C. Will (1980b). *Astrophys. J.* **242**, L129.
Watson, (1966). *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge.
Weinberg, S. (1972). *Gravitation and Cosmology.* Wiley, New York.
Weisberg, J. M. and J. H. Taylor, (2002). In Proceedings of "Binary Pulsars", Chania, Crete 2002, APS Conference Series, M. Bailes, D. J. Nice and S.E. Thorsett eds., astro-ph/0211217.
Weisberg, J. M. and J. H. Taylor, (2004). In Proceedings of "Aspen Winter Conference on Astrophysics: Binary Radio Pulsars", ASP Conference Series, F. A. Raisio and I. H. Stairs eds., astro-ph/0407149.
Will, C. M. (1984). *Ann. Phys.* **155**, 133.
Will, C. M. (1993). *Theory and Experiment in Gravitational Physics*, Revised edition, Cambridge University Press, Cambridge.
Will, C. M. (1998). *Phys. Rev. D* **57**, 2061.
Will, C. M. (2005). *Phys. Rev. D* **71**, 084027.
Will, C. M. (2006). *Living Rev. Relativity* **9**, 3. Online article http://www.livingreviews.org/lrr-2006-3.
Wiseman, A. G. (1992). *Phys. Rev. D* **46**, 1517.
Wiseman, A. G. (1993). *Phys. Rev. D* **48**, 4757.
Wiseman, A. G. and C. M. Will (1991). *Phys. Rev. D* **44**, R2945.
Wiseman, A. G. and C. M. Will (1996). *Phys. Rev. D* **54**, 4813.
Wolszczan, A. (1991). *Nature*, **350**, 688.



Yao, W.-M. *et al.*[Particle Data Group] (2006). *J. Phys. G* **33**, 1. Online version at http://pdg.lbl.gov/
Zakharov, V. I. (1970). *JETP Lett.* **12**, 312.
Zeldovich, Ya. B. and A. G. Polnarev (1974). *Sov. Astron.* **18**, 17.
Zendri, J.-P., *et al.* [AURIGA] (2002). *Class. Quant. Grav.* **19**, 1925.
Zhou, C. Z. and P. F. Michelson (1995). *Phys. Rev. D* **51**, 2517.
Zimmermann, M. (1980). *Phys. Rev. D* **20**, 351.
Zimmermann, M. and E. Szedenits (1979). *Phys. Rev. D* **21**, 891.