Interferometric gravitational wave detection: accomplishing the impossible

Class. Quantum Grav. **17** (2000) 2441–2448. Printed in the UK PII: S0264-9381(00)50637-6

Interferometric gravitational wave detection: accomplishing the impossible

Peter R Saulson
Department of Physics, Syracuse University, Syracuse, NY 13244-1130, USA

Received 28 April 2000

Abstract. Detection of gravitational waves involves technological challenges that appear almost insurmountable. I list the chief obstacles, and briefly explain how they are being overcome.

PACS number: 0480N

1. Introduction

In this article I would like to take a very particular point of view on gravitational wave detection, that of an experimental physicist who marvels at the audacity of the attempt. The detection of gravitational waves is *nearly* impossible. That we should be so close to detecting them is remarkable. It will become truly wonderful when we succeed, sometime in the next few years.

The plan of this article is as follows. I will set the stage by describing the sensitivity that astrophysicists tell us will be required if we are to detect gravitational waves, and by commenting on Einstein's apparent lack of interest in gravitational wave detection. Then I will enumerate the particular aspects of the measurement challenge that ought to strike reasonable observers as being nearly impossible to meet. I will devote the remainder of the article to a brief demonstration of how each of those challenges are now being met by the interferometric gravitational wave detectors now nearing completion around the globe.

2. The challenge

2.1. How weak are gravitational waves?

Very weak, by any ordinary standard. Gravitational waves move free masses, just as electromagnetic waves move free charges. There is a distinctive pattern to the motions induced by a gravitational wave, as illustrated in figure 1. Three free masses are illustrated, arranged in the shape of the letter L. A gravitational wave travelling in a direction normal to the plane containing the masses will cause equal and opposite changes in the proper distance between the mass at the centre and the two masses at the ends of the L. (This is what is meant physically by describing the waves as transverse and traceless.) The effect has the character of a tidal force, that is to say that all proper distances are changed by the same ratio. Larger distances

0264-9381/00/122441+08$30.00

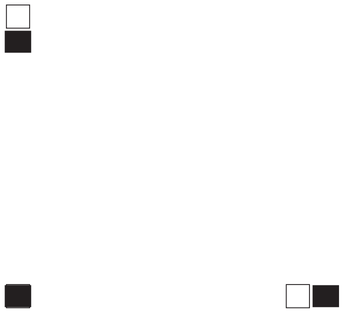

Figure 1. Three test masses respond to a (very strong) gravitational wave. The open squares represent the original positions of the masses, while the full squares show the pattern of motion induced by the wave.

will thus change by larger absolute amounts. For this reason one typically describes the effect as a *strain*, usually denoted by

$$h \equiv 2\Delta L/L,$$

where ΔL is the change in proper length of a separation of nominal length L.

Detecting gravitational waves would be trivial if we ever had an opportunity to observe one that moved free masses by the amount illustrated in the figure. Unfortunately, the figure was drawn wildly out of scale, to make the effect visible to the unaided eye. The change in length of one arm of the L, indicated by the displacement of the mass from its original position (marked by the open square) to its new position (marked by the full square), is shown in the figure as about $h \approx 2 \times 10^{-1}$. The bitter astrophysical truth is that what we must be prepared to look for are instead waves with a maximum amplitude of

$$h_{strongest} \sim 10^{-21},$$

or perhaps an order of magnitude weaker. This will be *very* difficult to see.

2.2. What did Einstein think about gravitational wave detection?

Apparently, very little. I am not an Einstein scholar, but I have tried to find every word that our founder wrote on the subject of gravitational waves. I know of three articles, including the famous papers of 1916 and 1918 that lay out the basic theory of the waves [1]. The only other Einstein paper I know that discusses gravitational waves is a rather dull affair, written in 1937 with Nathan Rosen and published in *J. Franklin Inst.*, presenting an exact solution [2].

Even more remarkable is the record of comment on gravitational waves in Einstein's voluminous popular writings on relativity: as far as I know, it is non-existent. It is not as if Einstein was a theoretical physics chauvinist uninterested in the experimental consequences of his ideas. After all, he proposed the three classic tests of general relativity. They were an interesting mix of measurements with a wide range of difficulty. The precession of Mercury's orbit (by 43 arcsec/century) had already been seen with high precision. The deflection of starlight (of the order of 1 arcsec near the limb of the Sun) was hard to measure, but was seen (more or less) by Eddington in 1919, in the process making Einstein a world celebrity. The gravitational redshift was strong enough in stars to yield a detectable signal of about 10^{-6}, but was devilishly hard to distinguish from systematic errors; it took someone with the talent of Robert Pound to make a clean laboratory measurement, decades later [3].

Einstein occasionally felt free to speculate about possible future measurements of other relativistic effects. In one of his popular works, Einstein expressed the opinion that it might someday become possible to detect the dragging of inertial frames by the rotation of the Earth (around 40 marcsec/year); within the past few years we have heard about one possible marginal detection of the effect [4], and are about to witness the launch of Gravity Probe B, which should make a definitive measurement [5].

However, nowhere does Einstein even remember to mention that detection of gravitational waves would constitute a wonderfully dramatic validation of relativistic causality in the realm of gravity. In fact, it appears that Einstein forgot (or wished to forget) that he had predicted the existence of gravitational waves. Why? One reason may be that Einstein was at times afflicted with doubt about the physical reality (i.e. gauge independence) of gravitational waves. He may have been one of the objects of Eddington's remark that some people thought that gravitational waves 'travel at the speed of thought'.

It is also possible, though, that Einstein was actually making a reasonable judgement about the chances that gravitational waves might ever be detected. He was, after all, a former patent examiner, someone who had spent years being paid to determine whether new machines would work. If Einstein had set out to guess whether gravitational waves could be detected, what would he have thought? He might have thought about interferometric measurements of changes in perpendicular distances, but if he read Michelson and Morley's 1887 paper [6] he would have seen that their sensitivity was about $h \sim 10^{-9}$. He might have tried to estimate what might be the strongest waves to reach the Earth from the stars, but recall that the most violent astronomical quadrupoles he could have known about were binary stars with periods of years; the discovery of neutron star binaries was decades away [7]. Only a clairvoyant could have had (in the first half of the 20th century) any reason for optimism about the eventual detection of gravitational waves.

3. What makes gravitational wave detection difficult?

The interferometric detectors of gravitational waves, such as LIGO in the USA [8], the French–Italian VIRGO [9], British–German GEO600 [10] and Japanese TAMA300 [11], appear to be on the verge of detecting gravitational waves. Most are expected to reach a strain sensitivity of the order of 10^{-21} for brief (millisecond-scale) bursts, sometime within the next few years. Yet on the face of it, this might appear to be fundamentally implausible, if not downright impossible. Among the things that an experienced experimental physicist might reasonably doubt could be achieved are:

- interferometry with free masses,
- with a strain sensitivity of 10^{-21} or better,
- equivalent to position sensitivity far below nuclear dimensions,
- in the presence of much larger noise.

To understand why experimenters in the field are optimistic, one has to understand how each of these seemingly impossible tasks is being accomplished.

3.1. Interferometry with free masses

3.1.1. Why it seems impossible. To someone with experience in optics, even the most basic notion of an interferometric gravitational wave detector might seem farfetched. To respond to the gravitational effect, the mirrors must be some reasonable approximation to free masses.

Even we are not crazy enough to let the mirrors actually fall freely†; instead, they are hung from fine wires as pendulums. Well above the pendulum resonant frequency, the motion of a mirror suspended in this way is the same (in the horizontal direction, at any rate) as if it were free.

However, even this concession to practicality would hardly calm an optics expert. Take a look at almost any other interferometer, and you will see mirrors, beamsplitters and other parts firmly bolted to a rigid table. And for good reason! An interferometer only works in even the crudest sense if the mirrors are accurately enough aligned so that the beams from the two arms overlap at the output of the beamsplitter. Hanging mirrors from fine wires hardly seems like the best way to guarantee good alignment. Beyond that, an interferometer can only be used as a precise measuring device if the mirrors move very little indeed, even in the presence of mechanical and acoustic inputs from the noisy environment. This is because unless excursions of the mirrors are limited to small distances from their operating point, the output light from the beamsplitter will not yield a linear function of the arm length difference. Instead, an interferometer with large excursions exhibits a sinusoidal relationship between arm length difference and output power, not an easy function to invert.

The perspective drawing of the apparatus in Michelson and Morley's 1887 paper lovingly records the standard solution to these issues: mirrors firmly bolted to rigid brackets, which are in turn screwed down to a massive and stiff table. Equivalent parts are available for purchase to this day from a number of commercial vendors (with a few changes in material, such as composite honeycomb tables in place of stone slabs). Optical scientists and engineers patronize these vendors for the most sensible of reasons.

3.1.2. How it is possible nevertheless. The requirement of free masses and the requirement of rigidly aligned and spaced optics seem as contradictory as it is possible to imagine. Yet a solution exists: suspend the mirrors freely, but control their orientation and position by means of servomechanisms.

The archetype of a gravitational physics experiment using feedback to tame the required dynamics is the equivalence principle test of Roll *et al* [13]. Their torsion balance needed to be free to track the Sun if it wanted (at a frequency of 1 d^{-1}), but also needed to stay aligned so that the optical lever readout could function at high precision. When this experiment was first written up, it was necessary to include a primer of servo theory to explain how it all worked. The technique has now become a standard element of the experimental physicist's toolkit. Weiss, one of the originators of the interferometric gravitational wave detector and a Dicke postdoctoral student in the early 1960s, showed how this technique should be applied to an interferometer in his earliest design study [14]. Weiss set to work implementing his ideas, but it should be noted that the first published account of the construction of a prototype interferometric gravitational wave detector came from Robert Forward (himself a former student of Joseph Weber [15]).

The magic of the servo solution lies in the fact that a system built around feedback control can easily have its dynamic properties tailored as a function of frequency. The proper mix for our function is to allow the mirrors to swing freely at high frequencies, while at low frequencies keeping them aimed properly and holding them at the right interferometric operating point. This is among the more natural tasks to ask of a servo, since it is easy to design feedback systems that have high gain at low frequencies but negligible gain at high frequencies.

There are several catches. (There always are.) One is that a complete interferometer requires dozens of servo loops, since there are many mirrors, each of which has a number

† The advantages of doing so were considered seriously by James Faller and Robert Spero [12].

of degrees of freedom to control. Typical interferometer designs have six critical mirrors and several others that do not need quite such good control. Depending on the suspension design, there may be four or five degrees of freedom per mirror that need to be controlled.

Feedback always comes with costs. Each degree of freedom requires a sensor and an actuator, which must be well enough engineered so that they do not inject significant amounts of noise. Each loop must stay locked robustly at the chosen operating point, and must not lapse into instability. For most of the loops in an interferometer, these requirements are relatively straightforward to achieve. The exceptions are explained in more detail below.

3.2. Strain sensitivity of 10^{-21}

3.2.1. Why it seems impossible. The precision with which an interferometer can measure a strain is equal to the precision with which it can make length comparisons between the two arms, divided by the total optical length of an arm, or

$$h_{rms} \sim \Delta L_{rms}/L_{opt}.$$

One natural advantage of an interferometer is that the alternation between bright and dark outputs occurs for motions of the order of the wavelength of light. This is rather fine compared with the tick marks on an ordinary ruler. But, even with such precision, achieving a sensitivity h_{rms} of 10^{-21} is a daunting challenge. Michelson and Morley were able to estimate the phase difference between their two interfering beams to $\frac{1}{20}$ of a wave, and yet they only achieved a strain sensitivity of a few times 10^{-9}.

3.2.2. How it is possible nevertheless

Make L_{opt} large. A given length comparison precision corresponds to a better strain measurement, the longer is the optical path length taken by light in an arm. This is why the new generation of interferometers are all so long, up to LIGO's 4 km.

The optical path length in an arm can also be made much longer than the 'out and back' distance between the beamsplitter and the end mirror. The Michelson–Morley interferometer of 1887 was an improvement over Michelson's 1881 device by virtue of the multiple bounces across the table that were added, creating arms of a total length 22 m on a single table. The price was many more mirrors than the minimum, all needing to be rigidly held in the proper alignment. Interferometric gravitational wave detectors use a similar trick. GEO600 uses the Michelson–Morley version almost exactly, adding a second mirror adjacent to the beamsplitter to cause the light to make two round trips through an arm.

All of the other current projects use a clever idea that gives even greater benefits. Each arm is made up of a highly reflective end mirror lined up with a partially transmitting mirror adjacent to the beamsplitter. This forms a Fabry–Perot cavity which, when its length is adjusted to resonance with the wavelength of the light, holds the light for a time corresponding to many round trips. It truly functions as if the light were travelling a distance of the order of the cavity finesse times its length. This idea is due to Drever [16]. One of its key virtues is that it only requires one mirror at each end of each arm to give the benefit of many round trips, much more elegant than the Michelson–Morley solution. Given the tight specifications (and attendant high cost) associated with these mirrors, elegance translates rather directly into the difference between a practical idea and a dream.

One extra servo challenge arises when this solution is employed. The connection between mirror separation and the phase of the exiting light is like the analogous long optical path only when the mirror separation is close to resonance. This means that the cavity must be held in resonance to work at all. Drever, Hall and their collaborators worked out the servo-locking scheme [17], based on an analogous servo for microwave cavities invented by Pound [18]. A challenge arises from the fact that far from resonance, the cavity gives a uselessly insensitive signal to the servo. Acquiring lock is thus a highly nonlinear dynamical process, more interesting to design than one might wish.

Make precise arm length comparisons. Michelson and Morley arranged that the beams from the two interferometer arms overlapped at a slight angle to one another. This meant that the relative phase between the two beams changed across the face of the beam, causing the recombined light to exhibit the alternating bright and dark stripes that came to be known as 'fringes'. They did this because it made it convenient for them to use their eyes to observe any shifts in relative arm lengths—they would show up as side-to-side motion of the fringe pattern. Michelson and Morley estimated they could see relative shifts of $\frac{1}{20}$ of a wavelength.

To make better arm length comparisons, one has to make better use of the light The modern choice is to align the two beams, giving a fixed phase relationship across the whole beam. Then, depending on the phase difference between the two beams, the entire spot is either bright, dark or something in between. Electronic measurement of the light power in the output spot can yield much more precise measurements than $\lambda/20$. The fundamental limit is the shot noise in the arrival of photons, which can be quite small if the photon flux (i.e. the power) is large.

One way to obtain the desired high light power is to use a powerful laser. Solid-state Nd:YAG lasers can now be relied on to supply about 10 W of highly stable light at a wavelength of 1.06 μm. It is also possible to make this work as if it were a laser of substantially higher power than that. Weiss, following Dicke, had taught us to operate the interferometer so that its operating point was to send no light to the output port [14]. (One adds a high-frequency dither, demodulated with a Dicke-style lock-in amplifier, to generate a well behaved output that is proportional to the arm length difference.) Drever then pointed out that one could make use of the light that leaves the beamsplitter headed back toward the laser, returning it in phase coherence with the 'fresh' light [16]. Given the low-loss optics available today, the light can be 'recycled' dozens of times. When all of these tricks are applied, it is possible to make arm length comparisons at a level of 10^{-10} rad [19].

3.3. Position sensitivity far below nuclear dimensions

3.3.1. Why it seems impossible. The mirrors that serve as our test masses are made of atoms, after all. In particular, their surfaces are made of atoms, and are rough on an atomic scale (if not much rougher than that.) Not only that, but the individual atoms jitter about their equilibrium positions by large amounts. However, to achieve strain sensitivity of 10^{-21}, we need position noise in each mirror no greater than $x_{rms} = h_{rms}L/2 \sim 10^{-18}$ m, where L is the separation between the mirrors. How can this possibly be achieved?

3.3.2. How it is possible nevertheless. Fortunately, we do not really care about the structure of the mirrors or their surfaces on the atomic scale. The output beams in an interferometer have their phases determined by the average of the positions of all of the atoms in the part of the surface illuminated by the light beam. In a large interferometer, with corresponding centimetre-scale beam diameters, one is dealing with $\sim 10^{17}$ atoms. For the same reason, individual atomic motions are not relevant, but instead all that we care about are the coherent motions of the mirror surface, due to the lowest few dozen normal modes of vibration of the mirror. This is bad enough, but still lets us achieve precisions of better than 10^{-18} m.

3.4. Measurement in the presence of larger noise

3.4.1. Why it seems impossible. Random motion of the test masses due to seismic noise will have a magnitude of the order of 1 μm or greater. Even if that could be strongly filtered, Brownian motion of the mirror's centre of mass will have a magnitude of the order of 3×10^{-12} m, and excitations of the mirror's surface with respect to its centre of mass will be of the order of 3×10^{-16} m.

To detect a signal, one must have a signal-to-noise ratio which is much larger than unity. If the position noise needs to be of the order of 10^{-18} m in order to see signals with strain amplitudes of 10^{-21}, how can interferometric gravitational wave detection succeed?

3.4.2. How it is possible nevertheless

Filter external noise. Seismic noise can be blocked from affecting the test masses, by supporting the mirrors through a long chain of alternating masses and springs. This can be very effective at frequencies substantially above the resonance frequencies of the mass–spring 'stack'.

Take full advantage of the frequency dependence of the noise. Many of the most important noise sources have spectra with strong frequency dependences. Seismic noise falls steeply with frequency, especially after being filtered by the isolation stack, so it is only important up to a cut-off frequency that will be somewhere in the decade 3–30 Hz, depending on the isolator design. Thermal noise is strongly peaked at the resonance frequencies of the test mass and its suspension, most dramatically so when the quality factors of the resonances are high. The resonances of greatest importance are the pendulum mode at about 1 Hz, and internal resonances of the interferometer mirrors, above 10 kHz.

This means that there is a band in frequency space that is comparatively low in noise. For gravitational wave signals whose frequency content lies primarily in this band, the noise against which they compete for visibility can be very much smaller than the numbers quoted above. Of course, this reasoning assumes that the interferometer responds in an accurately linear fashion, otherwise filtering would not be possible. The servos that hold the interferometer locked at its operating point are what are responsible for this invaluable feature.

Thinking about the frequency content of the signal and noise points out that the signal-to-noise ratio should not be evaluated simply by comparing the root-mean-square magnitudes of each. For every combination of a signal $s(t)$ (with Fourier transform $\widetilde{s}(\omega)$) and an instrumental noise spectrum $\Phi(\omega)$, there is an optimal filter given by

$$Y(\omega) = \mathrm{e}^{-\mathrm{i}\omega T}\widetilde{s}^*(\omega)/\Phi(\omega)$$

that maximizes the signal-to-noise ratio†. Attaining this optimum requires precise knowledge of the time dependence of the signal. Much of the benefit can be achieved even when there is less perfect knowledge, by the simple procedure of suppressing frequency bands where the noise is strong and the signal is relatively weak.

4. Prospects

Within a few years, a worldwide network of interferometers will be scanning the skies for gravitational wave signals with amplitudes of the order of 10^{-21}, in a band two decades wide

† See, for example, [20].

centred on a frequency of a few hundred Hertz. By the second half of the decade, sensitivities should approach 10^{-22} [21]. It is likely that signals will be detected somewhere in that range. Then our seemingly impossible task will have been achieved.

Acknowledgement

Research on gravitational wave detection in the United States has been generously supported by the National Science Foundation.

References

[1] Einstein A 1916 *Sitzungsberichte der Koniglich Preussischen Akademie der Wissenschaften* pp 688–96
Einstein A 1918 *Sitzungsberichte der Koniglich Preussischen Akademie der Wissenschaften* pp 154–67
[2] Einstein A and Rosen N 1937 *J. Franklin Inst.* **223** 43
[3] Pound R V and Rebka G A Jr 1960 *Phys. Rev. Lett.* **4** 337
Pound R V and Snider J L 1965 *Phys. Rev.* B **140** 788
[4] Ciufolini I *et al* 1998 *Science* **279** 2100
[5] Buchman S *et al* 2000 *Adv. Space Research* **25** 1177
[6] Michelson A A and Morley E W 1887 *Am. J. Sci.* **34** 333
[7] Hulse R A and Taylor J H 1975 *Astrophys. J. Lett.* **195** L51
[8] Abramovici A *et al* 1992 *Science* **256** 325
[9] Bradaschia C *et al* 1990 *Nucl. Instrum. Methods* A **289** 518
[10] Danzmann K *et al* 1992 *Lecture Notes in Physics* vol 410, p 184
[11] Tsubono K 1997 *Proc. TAMA Workshop on Gravitational Wave Detection* ed K Tsubono (Tokyo: Universal Academy Press)
[12] Robert Spero 1991 Personal communication
[13] Roll P G, Krotkov R and Dicke R H 1964 *Ann. Phys., NY* **26** 442
[14] Weiss R 1972 *MIT Quart. Prog. Rep.* **105** 54
[15] Moss G E, Miller L R and Forward R L 1971 *Appl. Opt.* **10** 2495
[16] Drever R W P 1983 *Gravitational Radiation* ed N Deruelle and T Piran (Amsterdam: North-Holland) p 321
[17] Drever R W *et al* 1983 *Appl. Phys.* B **31** 97
[18] Pound R V 1946 *Rev. Sci. Instrum.* **17** 490
[19] Fritschel P *et al* 1998 *Phys. Rev. Lett.* **80** 3181
[20] Helstrom C W 1968 *Statistical Theory of Signal Detection* (Oxford: Pergamon) p 123
[21] Gustafson E, Shoemaker D, Strain K and Weiss R 1999 *LSC White Paper on Detector Research and Development* 11 September, LIGO Technical Document T990080-00-D