An introduction to signal extraction in interferometric gravitational wave detectors

Eric D. Black and Ryan N. Gutenkunst
LIGO Project, California Institute of Technology, Mail Code 264-33, Pasadena, California 91125

In the very near future gravitational wave astronomy is expected to become a reality, giving us a completely new tool for exploring the universe around us. We provide an introduction to how interferometric gravitational wave detectors work, suitable for students entering the field and teachers who wish to cover the subject matter in an advanced undergraduate or beginning graduate level course. © *2003 American Association of Physics Teachers.*
[DOI: 10.1119/1.1531578]

I. INTRODUCTION

It would have been difficult to imagine the wonders that would later be revealed when Karl Jansky identified the first astronomical source of radio waves.[1] Since that first astonishing discovery, astronomers have shown us views of the universe through a wide range of the electromagnetic spectrum. This new vision has revealed to us wonders such as x-ray sources (black hole candidates; see for example, Ref. 2), the cosmic microwave background (a relic of the birth of the universe, see for example, Refs. 3 and 4), and gamma-ray bursts (unknown origin[5]). Not only have we explored vast regions of the electromagnetic spectrum, we have probed neutrinos from a supernova to learn about stellar collapse[6] and from the sun to learn fundamental physics.[7] As our ability to detect different kinds of signals expands, so does our understanding of the universe around us. It is then natural to ask if there are other windows on the universe waiting to be discovered. Gravitational waves may provide us with just such a window. As their name implies, gravitational waves are propagating gravitational fields, analogous to the propagating electromagnetic fields we have so effectively probed in recent years. We know that astronomical sources of gravitational waves exist, but, as of this writing, we have yet to achieve positive detection of gravitational waves on earth.

This is a field potentially rich with the promise of discovery. There are currently a number of experiments being performed and developed around the world to try and detect gravitational waves.[8–14] Much of this effort centers around interferometric detectors, and this paper introduces how these detectors work. Specifically, we will describe the optical configuration of an interferometric detector and how it converts a gravitational wave into a measurable signal. In the process, we will touch on several important topics in modern experimental physics, including nulled lock-in detection and optical Fabry–Perot cavities.

II. GRAVITATIONAL WAVES

A. Properties of gravitational waves

Before we jump into the physics of gravitational wave detectors, let us look at some of the properties of these waves. The existence of gravitational waves is predicted by general relativity, where they arise as wave-like solutions to Einstein's (linearized) field equations.[15] Gravitational waves are similar in many ways to electromagnetic waves, which are similarly predicted from wave-like solutions to Maxwell's equations. Both kinds of waves propagate at the speed of light; both are transverse, meaning that the forces they exert are perpendicular to the direction of propagation, and both exhibit two orthogonal polarizations. Because an electromagnetic wave carries propagating electric and magnetic fields, it should come as no surprise that a gravitational wave carries a gravitational field. Although the field in an electromagnetic wave exerts a force only on charged particles, the field in a gravitational wave exerts forces on all objects. For electromagnetic waves, the direction of the force exerted on a charged particle is relatively straightforward. A linearly polarized plane wave propagating along the z axis will exert an oscillating force on a charged particle along, say, the y axis. This force will push the (charged) particle up and down along the y axis. (We are neglecting the contribution of the magnetic field in this example, which is fine as long as the velocity of the charged particle remains small.)

The effect of a gravitational wave on matter is a little more subtle. Instead of simply exerting an oscillating force on an object along a fixed axis, a gravitational wave produces a force that stretches and squeezes the object along two axes, as shown in Fig. 1. If a gravitational wave propagating along the z axis (into the page) encounters such an object, the object will feel a compressive force along the y axis and simultaneously a stretching force along the x axis. Half a period later, the force along the y axis will have reversed sign and will stretch the object, while the x force will be compressive. This configuration describes one polarization, known as "+" relative to these axes. The other polarization, denoted "×," has the forces along a pair of axes rotated 45° from the + case.

The spatial pattern of this force resembles the tidal force exerted on a moon as it orbits a planet, and the force exerted by a gravitational wave is commonly referred to as tidal. If we replace our target object with a collection of smaller free masses not connected to each other, then the gravitational wave just moves those masses in the same pattern as described above: pulling them together along the y axis and pushing them apart along the x axis, then reversing sign in the next half period.

The spatial distribution of the force exerted by a gravitational wave is more complicated than that of an electromagnetic wave, so it is perhaps not surprising that its magnitude is more subtle as well. A gravitational wave induces a *strain* in an object. The amount of stretch or compression along the x or y axis is proportional to the length of the object along that axis. The larger the object, the more it stretches. Both
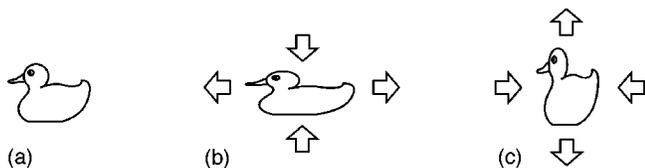
Fig. 1. An example of how a gravitational wave affects a compliant object, such as a rubber duck. The wave is propagating into the page and polarized in such a way that its tidal forces are oriented along the x and y axes. (a) The undisturbed object before the arrival of the wave; (b) a time when the tidal forces, represented by arrows, are at maximum amplitude; (c) the object one-half period later, when the tidal forces have reversed. Even cataclysmic astronomical events may only produce a few cycles detectable to ground-based instruments.

facts, that a gravitational wave induces a strain and that this strain is in a tidal pattern, influence the design of our detector.

To be precise, a gravitational wave actually produces a perturbation in the metric of space–time, but as long as our target object is small compared with the wavelength of the gravitational wave, our tidal forces view is equivalent.[16]

B. Gravitational radiation

Similar to the way electromagnetic waves are emitted by accelerating charges, gravitational waves are emitted by accelerating masses. However, the energy emitted in the form of gravitational waves for most objects in nonrelativistic interactions is quite small. This is not only due to the fundamentally weak nature of gravity, but also to the fact that mass only comes in one sign. The familiar and prominent dipole radiation of electrodynamics does not occur for gravitational radiation because it is impossible to produce an oscillating dipole moment about the center of mass in a system of interacting particles. The combination of the difficulty of generating gravitational waves and the weak nature of the gravitational field itself makes detecting gravitational waves quite difficult.

The gravitational waves have the best chance of being detected, and the ones that are perhaps the most interesting from an astronomer's point of view, are those produced by cataclysmic events, such as black hole collisions or type II supernovae. Such events involve gravity in the strong field, nonlinear regime, and the radiation emitted can carry information about what occurs under these conditions. This regime is new territory for physics as well as astronomy, as general relativity has so far only been studied experimentally in the weak-field limit.

C. Sources of gravitational radiation

It is generally accepted that gravitational waves exist. Not only are they predicted by general relativity, a well-established theory, we also have at least one known emitter of gravitational radiation: PSR1913+16, a binary system composed of two neutron stars, one of which is a pulsar.[17] The orbital period of this system is quite short, less than 8 hours, which means that two very massive, very compact objects (the two neutron stars) experience strong, regular accelerations. General relativity predicts that this system should be losing energy in the form of gravitational radiation. As this energy is lost, the two neutron stars should drift closer together, and as a result, the orbital period of the system should slowly decrease. We can calculate what the or-

bital decay of the system should be based on general relativity. Because one of the neutron stars is a pulsar, a fine natural clock, we can also accurately measure the orbital period and plot it as a function of time. (We can tell where the system is in its orbit by looking at the Doppler shift in the pulsar's signal.) This procedure is exactly what Russell Hulse and Joseph Taylor did as part of an extensive study of PSR1913+16. Over the course of years of observation, the system's slow orbital decay beautifully matched the predictions of energy loss from the emission of gravitational waves.[18,19] (Hulse and Taylor won a Nobel Prize for this work in 1993.)

Although the existence of gravitational waves was considered to be confirmed by this indirect observation, exploiting the information carried by the waves requires direct observation of the waves themselves. We do not want to observe them only to confirm their existence, we want to use them to get information about the systems that emit them. To be observable by an earth-based detector, a gravitational wave signal should be both strong and have a relatively high frequency. Strong signals are usually easier to detect than weak ones, and seismic disturbances are less likely to interfere with a measurement at high frequencies than at lower ones. The roughly 7.75 hour period of PSR1913+16 makes its gravitational radiation inaccessible to a ground-based detector, but a system with a period of 10 milliseconds or less would probably produce an observable signal in such a detector. Fortunately, the shorter the period of a binary system, the greater the acceleration involved, and hence the more energy emitted in the form of gravitational waves. The neutron stars in a binary system, such as PSR1913+16, spiral in toward each other as they lose energy, and the period of the system decreases. The two neutron stars will get closer together over time, and speed up, until they eventually collide, merging to form a single, massive object. Just before this collision, the orbital frequency can be quite high, and orbital periods of 10 ms or less are easily attainable. Moreover, the amplitude of the gravitational radiation emitted increases dramatically as this inspiral (or binary coalescence) progresses. The waveform of the signal produced is expected to provide a wealth of new information on the dynamics of the merger, and that is the scientific information we seek. Other possible sources of strong, high-frequency gravitational waves include black hole–black hole mergers and type II Supernovae. (For a review of possible sources, see Chap. 3 of Ref. 16.) These are just a few of the expected sources of gravitational radiation. A complete review of all of the possible sources belongs in a separate article devoted entirely to that subject. The most interesting sources, of course, will be the ones we did not anticipate.

III. A MICHELSON INTERFEROMETER AS A GRAVITATIONAL WAVE DETECTOR

Most modern interferometric detectors are based on the Michelson design. Michelson interferometry is particularly well-suited for detecting gravitational waves because of the geometry of the tidal force the waves produce. A classic Michelson interferometer, shown in Fig. 2, is sensitive to differential motion between the x- and y-arms, just the kind of strain produced by a gravitational wave. In this basic configuration, light enters the interferometer and is split in two by a beam splitter. The two resulting beams travel down the
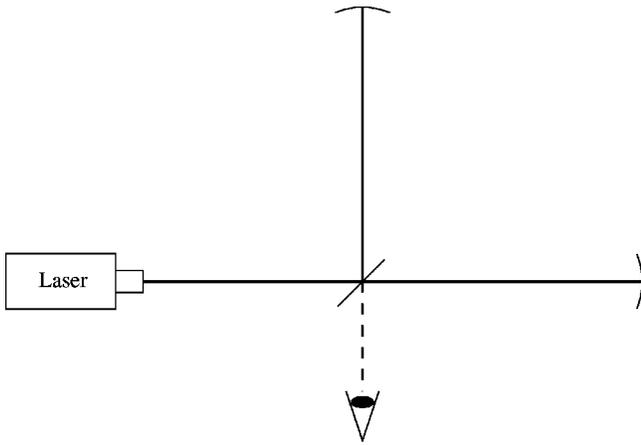
Fig. 2. A basic Michelson interferometer is sensitive to the kinds of strain a gravitational wave will produce. Incident laser light is split by a beam splitter, sent down orthogonal paths along the x and y axis, reflected from mirrors at the ends of these paths, and recombined back at the beam splitter. The interference between these two return beams produces a net intensity that is sensitive to differential changes in the lengths of the arms.
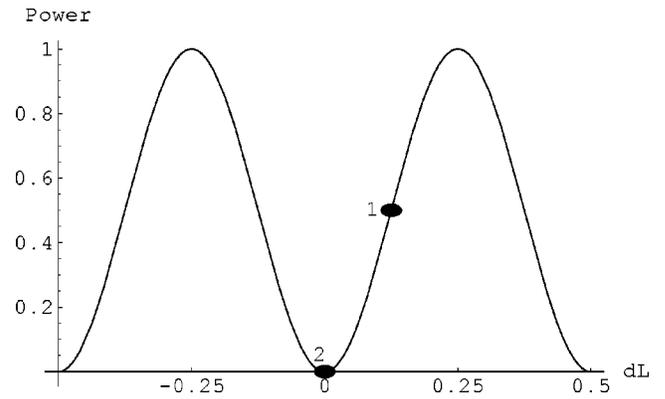

Fig. 3. The intensity of the light at the observation port versus the difference in arm lengths (units of λ, the wavelength of the light). Operating at point 1 maximizes the change in power for a given change in arm lengths, but also makes the instrument sensitive to intensity noise in the light source. Operating at point 2 eliminates this problem, but, in a simple Michelson interferometer, it reduces the signal to a second-order effect.

arms, reflect off the end mirrors, and recombine to interfere back at the beam splitter.[20] The light emitted at the observation or *antisymmetric* port provides a measure of the difference between the lengths of the interferometer's arms. (The *symmetric* port, from which light returns to the laser, also contains information about the relative arm lengths. Conservation of energy requires that the power coming out of the symmetric and antisymmetric ports, along with any power lost in the instrument, accounts for all of the input power.)

Let us consider quantitatively the response of a simple Michelson interferometer to a gravitational wave. It is not difficult to derive an expression for the electric field at the output of the interferometer, E_{out}, as a function of the electric field at its input, E_{in},

$$E_{\text{out}} = \tfrac{1}{2}(r_x e^{ik2\ell_x} - r_y e^{ik2\ell_y})E_{\text{in}}. \tag{1}$$

Here ℓ_x and ℓ_y are the lengths of the two arms, k is the wave number for the light we are using, and r_x and r_y are the amplitude reflectivities of the end mirrors. In our convention, a perfectly reflecting mirror has $r = -1$.

The power falling on the photodiode in Fig. 2 is the square of the magnitude of the electric field, $|E_{\text{out}}|^2$, or, for perfectly reflecting end mirrors,

$$P_{\text{out}} = P_{\text{in}} \cos^2[k(\ell_x - \ell_y)], \tag{2}$$

where $P_{\text{in}} = |E_{\text{in}}|^2$ is the power entering the interferometer, provided by the laser in Fig. 2. This output power, and hence the voltage produced by the photodiode, varies sinusoidally with the difference in arm lengths, as shown in Fig. 3. If we let the arm lengths in the absence of a gravitational wave be ℓ_x and ℓ_y, then we can write the total arm length as $\ell_x + \delta\ell_x$ and $\ell_y + \delta\ell_y$, where a gravitational wave induces the perturbations $\delta\ell_x$ and $\delta\ell_y$. If we write the strain induced by the gravitational wave as

$$h \equiv \frac{\delta\ell_x - \delta\ell_y}{\ell}, \tag{3}$$

then we can write the power at the output of the interferometer as

$$P_{\text{out}} = P_{\text{in}} \cos^2[k(\Delta\ell + \ell h)], \tag{4}$$

where $\Delta\ell = \ell_x - \ell_y$ is the asymmetry in the arm lengths in the absence of a signal, and the average arm length is $\ell = (\ell_x + \ell_y)/2$. In this paper we will assume that the gravitational wave strain is very small—small enough that $k\ell h \ll 1$. We can then choose an operating point at some $\Delta\ell$ and look at the small perturbations in the output power around that point that the gravitational wave produces. We can describe this small-signal response mathematically by a Taylor expansion about $\Delta\ell$.

$$P_{\text{out}} = P_{\text{in}} \cos^2(k\Delta\ell) + P_{\text{in}} \frac{\partial}{\partial u} \cos^2 u \bigg|_{u = k\Delta\ell} (k\ell h) + \cdots. \tag{5}$$

The response of our simple Michelson interferometer to a gravitational wave strain h is proportional to the derivative of the output power with respect to $\Delta\ell$, so the obvious thing to do is to operate at the point where that derivative is maximum, which is point 1 in Fig. 3. At this point $k\Delta\ell = \pi/4$, and

$$P_{\text{out}} \approx \frac{P_{\text{in}}}{2}[1 - 2k\ell h]. \tag{6}$$

Unfortunately, we are then left with a fairly large dc term, $P_{\text{in}} \cos^2(k\Delta\ell) = P_{\text{in}}/2$ in this case, which will fluctuate if $\Delta\ell$ varies due to any perturbations on the mirrors, whether it be a gravitational wave, or seismic disturbance, etc. More importantly, this dc term is proportional to P_{in}, which can fluctuate even if the mirrors remain still.

Measuring small changes in a large signal is seldom an effective way to do experimental physics. If the amplitude of the gravitational wave we want to study is very small, as is too often the case, fluctuations in the dc term described above can completely obscure our signal. What we need is a way to reduce or even eliminate the dc term while retaining and, if possible, boosting our signal. How we meet these two goals is the subject of this paper.

IV. TURNING A MICHELSON INTERFEROMETER INTO A PRACTICAL GRAVITATIONAL WAVE DETECTOR

There are four things we need to do to our simple Michelson interferometer to make it an effective gravitational wave detector.

(1) *Make it big*. Because a gravitational wave induces a strain $\delta\ell/\ell$ between two free masses, the total distance each end mirror moves, $\delta\ell$, will be proportional to the arm length, ℓ. The longest practical arm lengths for ground-based detectors are on the order of a few kilometers. We will show later that by *folding* the interferometer, we can squeeze effective arm lengths of several hundred kilometers or more into just a few kilometers.

(2) *Use a lot of laser power*. Not surprisingly, the brighter the light source we use, the stronger that our output signal will be. However, there is one additional benefit to using a lot of laser power that may not be obvious at first glance. The effects of shot noise are reduced as the laser power is increased. Using a very powerful laser for a light source is one obvious way to increase the power used, but there are clever tricks that we can use to increase the power going into the instrument even more. One such technique that is commonly used is power recycling, a scheme by which any unused light exiting the symmetric port (that is, going back to the laser instead of falling on the photodiode) is recycled back into the interferometer. We will discuss both power recycling and the effect of shot noise on the output of an interferometric gravitational wave detector.

(3) *Decouple external noise sources*. Some noise sources, such as seismic noise, act directly on the mirrors and thus compete directly with gravitational waves to produce an output signal. Other noise sources do not perturb the mirrors directly and can be thought of as external to the instrument. A cleverly designed interferometer can be made largely insensitive to a wide variety of these external noise sources, and we will look at one technique that is widely used in many areas of experimental physics: nulled lock-in detection. We will show how nulled lock-in detection is implemented in an interferometric gravitational wave detector to decouple its output from fluctuations in the input power P_{in}.

(4) *Reduce the noise in the mirrors*. The largest strain[16] that we expect to observe from a gravitational wave is on the order of 10^{-21}, which gives a mirror motion of 4×10^{-18} m in a detector with 4 km arms. Obviously we must reduce the level of ambient noise in the mirrors to a level comparable to or smaller than this, at the frequencies that we expect to observe. We will not go into physical noise reduction techniques here, except to note that the entire interferometer must be enclosed in vacuum and mounted on a high-performance seismic-isolation system. As of this writing, noise reduction is a field of active and ongoing research.

Making the interferometer large and using a lot of power enhances our signal, while decoupling external noise sources eliminates the dc term we found problematic in Sec. III. In this paper we will consider only these three items. A treat-
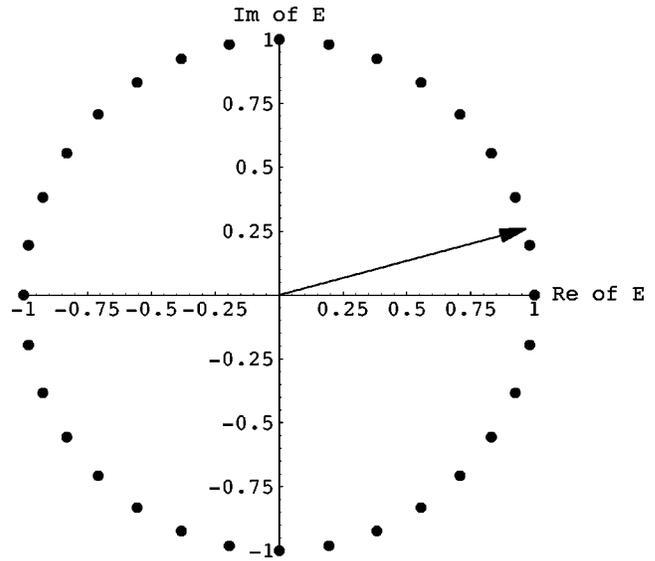

Fig. 4. An electromagnetic wave can be represented by a complex number. The length of the arrow corresponds to the amplitude of the wave, and the angle the arrow makes with the real axis corresponds to its phase.

ment of reducing the noise intrinsic to the mirrors is beyond the scope of this article, and an introduction to that subject is given in Ref. 16.

V. DECOUPLING INTENSITY NOISE: NULLED LOCK-IN DETECTION

A. Null-point operation

To eliminate the dc term in the output, interferometers for gravitational wave detection operate at the point labeled 2 in Fig. 3, known as the *null* point. Arm lengths are chosen so that, after passing back through the beam splitter, the two beams are 180° out of phase at the output (asymmetric) port. In the absence of a gravitational wave, they interfere perfectly destructively, and no light falls upon the photodetector; the port is "dark," even if the power delivered by the laser fluctuates.

The electric field at the output of a Michelson interferometer is given by Eq. (1). For a dark fringe Eq. (1) becomes

$$E_{out} = -iE_{in}e^{ik(\ell_x+\ell_y)}\sin(k\ell h), \qquad (7)$$

where we have assumed that $r_x = r_y = -1$. At this point, we introduce a graphical method of visualizing the calculation of the output of an interferometer. Because we are using complex numbers to represent electric fields, we may plot these numbers as vectors in the complex plane, as in Fig. 4. (This approach will be familiar to anyone who has studied phasors.) When we write the electric field in a light beam as $Ee^{i\omega t}$, we separate the time and spatial components of the phase into $e^{i\omega t}$ and E, respectively. In Fig. 4, we plot the real and imaginary parts of the spatial component E as a function of the position along the beam. As we advance along the beam, the spatial phase advances, and for each wavelength λ we travel, the vector representing E sweeps around a full circle in the complex plane. We can represent the calculation of the interferometer output graphically, as in Fig. 5. Here, we sketch the spatial components of the beam just before the light strikes the beam splitter the first time, where E is purely real, and after returning from each arm, where the differen-
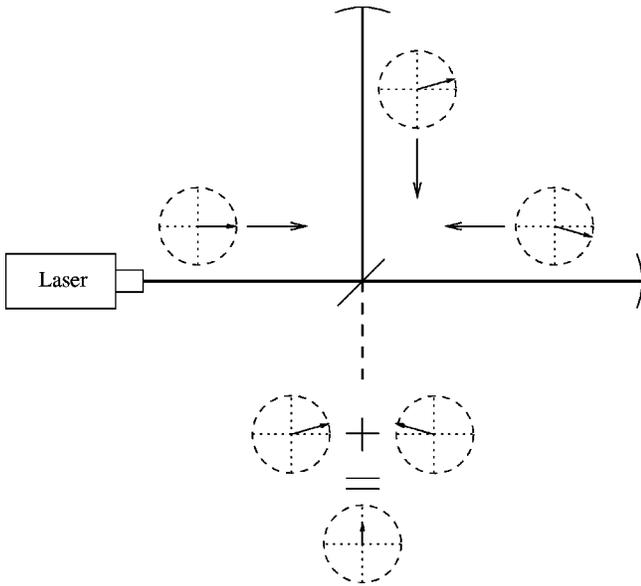
Fig. 5. When one arm is lengthened and the other shortened slightly by a gravitational wave, the two beams in the arms acquire equal but opposite small phase shifts. The beam splitter introduces an additional 180° shift between them. The two beams add so that the net result is a small amplitude beam phase shifted 90° from the input beam.

tial motion of the end mirrors has introduced a small phase in E. This phase is positive for the y arm (the arm along the y axis), but negative for the x arm because the mirrors move in opposite directions under the influence of the gravitational wave. When it reflects off of the beam splitter, the light from the x arm acquires a 180° phase shift, so that when it combines with the light from the y arm, destructive interference occurs and, in the absence of a gravitational wave, the port is dark. In the presence of a gravitational wave, the small imaginary components acquired in the x and y arms add constructively after the beam splitter, resulting in a small, purely imaginary E. That this field is 90° out of phase with the light incident on the beam splitter will be important later, when we talk about lock-in detection. The power falling on the photodetector is the square of the amplitude of this small, imaginary E, just after the photodetector.

$$P_{out} = P_{in} \sin^2(k\ell h) \approx P_{in} k^2 \ell^2 h^2, \qquad (8)$$

which is proportional to the square of the strain, h^2. We expect h to be very small, on the order of 10^{-21} or less, so an output proportional to h^2, rather than h, would be quite small and very difficult to detect. (For a kilometer-scale interferometer, P_{in} would have to be on the order of a kilowatt to produce more than one photon per second in P_{out}.) Operating at the null point has eliminated our intensity noise, but it has also nearly killed our signal. Fortunately, there is a way to recover the signal without coupling to the intensity noise, and that is the subject of Sec. V B.

B. Obtaining a linear signal: Lock-in detection

We can recover a signal that is linear in h without reintroducing intensity noise by using lock-in detection.[21] In lock-in detection, we modulate the signal and observe the resulting change in the output of the instrument. We then compare that change with our modulation signal to measure the derivative
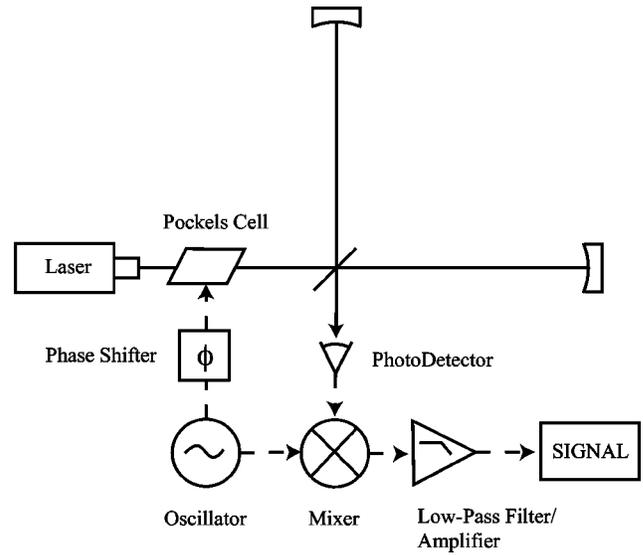

Fig. 6. We can use lock-in detection to recover a linear signal from a dark port.

of the instrument's output with respect to a signal. At the null point both the output power and its derivative are zero, but while a gravitational wave produces a second-order change in the output power, it produces a first-order change in its derivative. This first-order signal is something we have a chance of detecting.

1. Sidebands

Lock-in detection requires that we modulate the signal. The most obvious way to do that would be to purposefully move the mirrors in a way that mimics a gravitational wave signal. In practice, however, it is easier to both implement and describe this modulation by acting on the phase of the laser using a Pockels cell. A Pockels cell is essentially just a block of dielectric material with an index of refraction that depends on an applied electric field. If we pass the incident beam through a Pockels cell and apply an oscillating electric field to it, we will modulate the phase of the light going into the interferometer, as shown in Fig. 6. The electric field of the light going into the interferometer can then be written as

$$E_{in} = E_0 e^{i(\omega t + \beta \sin \Omega t)}$$

$$\approx E_0 [J_0(\beta) e^{i\omega t} + J_1(\beta) e^{i(\omega+\Omega)t} - J_1(\beta) e^{i(\omega-\Omega)t}], \qquad (9)$$

where J_0 and J_1 are zeroth and first order Bessel functions. The first term on the right-hand side of Eq. (9) is called the *carrier*; the next two are referred to as the *sidebands*.

We can calculate the electric field exiting the interferometer by considering the carrier and sidebands separately. We define the transfer function t of an interferometer for light of any given wavelength as the ratio of the output electric field to the incident field,

$$t \equiv \frac{E_{out}}{E_{in}}, \qquad (10)$$

where E_{out} is given by Eq. (1). Our modulated beam is composed of three different wavelengths, so we can find an expression for the light exiting the interferometer by applying the appropriate transfer function to each part, that is,

$$E_{\text{out}} = E_0[t_c J_0(\beta)e^{i\omega t} + t_+ J_1(\beta)e^{i(\omega+\Omega)t}$$
$$- t_- J_1(\beta)e^{i(\omega-\Omega)t}], \tag{11}$$

where t_c is the transfer function for the carrier, and $t_\pm$ are the transfer functions for the sidebands. We can calculate all of these transfer functions from Eq. (1), and in fact Eq. (1) gives the transfer function of the carrier immediately. The transfer functions for the sidebands can be obtained by the same procedure, if we use for their wave numbers

$$k_\pm = \frac{\omega \pm \Omega}{c} = 2\pi\left(\frac{1}{\lambda} \pm \frac{1}{\lambda_{\text{mod}}}\right). \tag{12}$$

Here λ_{mod} is the wavelength of an electromagnetic wave with frequency Ω. The transfer functions for the sidebands are then

$$t_\pm = i\sin\left[2\pi\left(\frac{\ell_x - \ell_y}{\lambda} \pm \frac{\ell_x - \ell_y}{\lambda_{\text{mod}}}\right)\right]e^{ik_\pm(\ell_x+\ell_y)}. \tag{13}$$

2. Schnupp asymmetry

Now comes a really clever part. We still want the output of the interferometer to be dark for the carrier, but if it is also dark for our sidebands, we still would have a second-order signal in h. If, however, we introduce a difference in the arm lengths $\Delta\ell \equiv \ell_x - \ell_y$, we can arrange for the output of the interferometer to be dark for the carrier but not dark for the sidebands. This trick is often attributed to Lise Schnupp, and the difference in arm lengths $\Delta\ell$ is known as the Schnupp asymmetry.[22,23]

With the Schnupp asymmetry and $r_x = r_y = -1$, the sidebands' transfer function reduces to

$$t_\pm = \mp i\sin\left[2\pi\left(\frac{\Delta\ell}{\lambda_{\text{mod}}}\right)\right]e^{i[(\omega\pm\Omega)/c](\ell_x+\ell_y)}. \tag{14}$$

We will neglect the change in $\Delta\ell$ produced by a passing gravitational wave.

3. Output of the instrument

The total electric field of the light exiting the interferometer is then

$$E_{\text{out}} = E_{\text{in}}e^{ik(\ell_x+\ell_y)}\left[iJ_0(\beta)2\pi\frac{\ell}{\lambda}h\right.$$
$$\left. + 2iJ_1(\beta)\sin\left(2\pi\frac{\Delta\ell}{\lambda_{\text{mod}}}\right)\cos\left(\Omega t + 4\pi\frac{\ell}{\lambda_{\text{mod}}}\right)\right], \tag{15}$$

where we have used Eq. (7) to calculate t_c. The power falling on our photodiode, as shown in Fig. 6, is then

$$P_{\text{out}} = P_{\text{in}}J_0^2(\beta)4\pi^2\frac{\ell^2}{\lambda}h^2 + 2P_{\text{in}}J_1^2(\beta)\sin^2\left(2\pi\frac{\Delta\ell}{\lambda_{\text{mod}}}\right)$$

$$+ 2P_{\text{in}}J_1^2(\beta)\sin^2\left(2\pi\frac{\Delta\ell}{\lambda_{\text{mod}}}\right)\cos\left(2\Omega t + 8\pi\frac{\ell}{\lambda_{\text{mod}}}\right)$$

$$+ P_{\text{in}}J_0(\beta)J_1(\beta)4\pi\frac{\ell}{\lambda}h\sin\left(2\pi\frac{\Delta\ell}{\lambda_{\text{mod}}}\right)$$

$$\times\cos\left(\Omega t + 4\pi\frac{\ell}{\lambda_{\text{mod}}}\right). \tag{16}$$

The voltage the photodiode produces is linearly proportional to the power falling on it, so for our purposes the signal we measure at the photodiode will be $V_{\text{pd}} = RP_{\text{out}}$, where R is the response of the photodiode.

There are four terms in this signal: two dc, one oscillating at 2Ω, and one oscillating at Ω. The one oscillating at Ω is proportional to the gravitational wave strain h and is the term we would like to isolate and measure. We isolate this term through a process that is standard in lock-in detection known as mixing, followed by low-pass filtering, or averaging over time. Mixing just means multiplication, and a mixer is a nonlinear device that takes two voltages on its inputs and produces a voltage at its output that is proportional to the product of the voltages at its inputs. If we feed V_{pd} into one input and $V_{\text{osc}}\cos(\Omega t + \phi)$ into the other, then we can write the time-averaged output of the mixer as

$$V_{\text{signal}} = \langle V_{\text{pd}}\cos(\Omega t + \phi)\rangle$$

$$= 2\pi(RP_{\text{in}})J_0(\beta)J_1(\beta)\frac{\ell}{\lambda}\sin\left(2\pi\frac{\Delta\ell}{\lambda_{\text{mod}}}\right)h, \tag{17}$$

where we have adjusted the user-specified phase ϕ to maximize the signal, and we have assumed that the amplitude V_{osc} is fixed and does not contribute to changes in the output. The signal we measure is now linear in h and zero in the absence of a gravitational wave. There is no dc offset to couple fluctuations in P_{in} to the output. A graphical representation of this calculation is illustrated in Fig. 7.

The sensitivity of our nulled lock-in scheme depends on the average length of the arms ℓ, the input laser power P_{in}, and the size of the Schnupp asymmetry $\Delta\ell$, all of which we would like to optimize to boost our signal. The best choice of the Schnupp asymmetry for this configuration is $\Delta\ell \approx \lambda_{\text{mod}}/4$, and this choice is neither difficult to achieve nor a very strict requirement. The size of the interferometer and the input power, however, should both be made as large as possible, and there are some interesting tricks for increasing both. We will consider techniques for increasing P_{in} and ℓ in the next two sections.
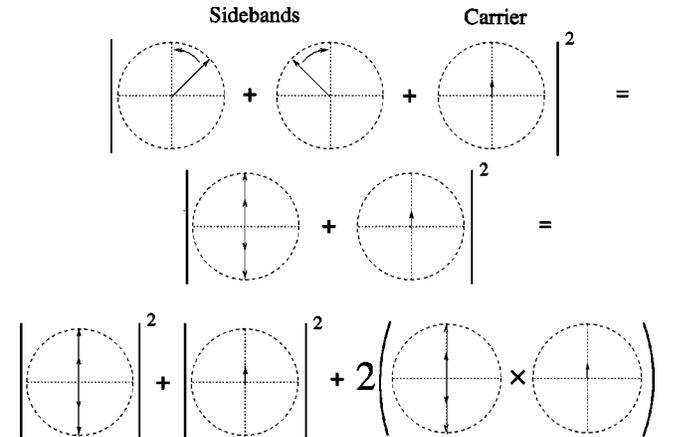

Fig. 7. The power out of the dark port is the modulus squared of the sum of all the interfering beams. The two sidebands add to produce an oscillation at the modulation frequency. There are three terms in the resulting power. The signal is extracted from the term that is the product of the sidebands and the carrier.
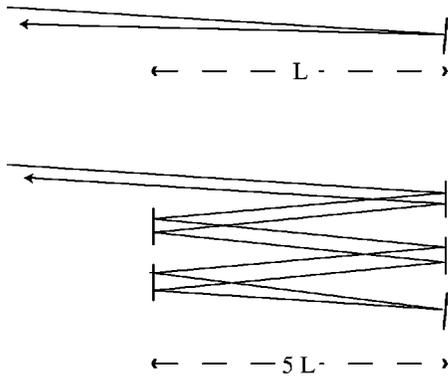
Fig. 8. A simple mirror versus an optical delay line. In a simple mirror, light travels a distance ℓ and reflects back, picking up a phase shift proportional to ℓ. If the mirror moves, the change in the phase shift is proportional to $\Delta\ell$. An optical path length of distance 5ℓ can be folded into a physical length ℓ with extra mirrors. If the end mirrors, three in this figure, all move by $\Delta\ell$, the change in the phase of the reflected light is proportional to 5ℓ. We would expect a gravitational wave to move all the end mirrors together.

VI. INCREASING THE ARM LENGTH: FABRY-PEROT CAVITIES IN THE ARMS

As we saw in Sec. V, longer arms yield larger signals. On Earth, we cannot afford to build arbitrarily long arms, but we can increase the effective length of the arms by bouncing the light back and forth within them, or *folding*. Figure 8 shows a simple folding scheme for one arm of an interferometer, where an optical length of 5ℓ is folded into a physical length of only ℓ. This is conceptually the easiest folding scheme to understand, and it is the one that Michelson and Morley used in their famous experiment. However, this scheme requires a number of different mirrors, one for each fold in the optical path. A simpler scheme to implement, one that involves only two mirrors regardless of the number of folds, involves the use of a Fabry–Perot cavity.[24]

A comprehensive treatment of Fabry–Perot cavities can be found in a good optics text, such as Hecht[20] or Siegman.[25] In this analysis we are only concerned with the transfer functions of these cavities, so let us briefly review what we need to know to proceed. A Fabry–Perot cavity is just two parallel, partially transmitting mirrors, as shown in Fig. 9. Most of the light falling on the input mirror reflects off of it, and this light is referred to as the *promptly reflected* beam. Some light, however, leaks through and circulates between the input and output mirrors. If this light returns from one round trip in-phase with new light leaking in, constructive interference will occur and a standing wave will build up, a condition known as resonance. The amplitude of this standing
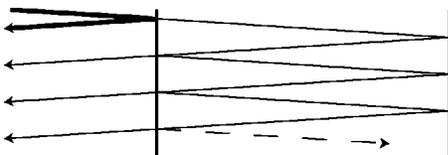

Fig. 9. A kind of delay line can be made with only two mirrors, if they are partially transmissive; in this configuration it is known as a Fabry–Perot cavity. The angle between the incident and reflected beams is greatly exaggerated. Normally this angle would be zero, that is, the reflected beams travel back along the incident, and the beams bouncing back and forth between the mirrors form a standing wave. In this example, only the input (left) mirror is partially transmissive.
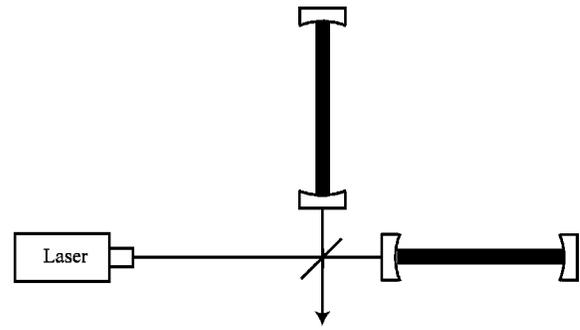

Fig. 10. We can substantially increase the effective arm length of a gravitational wave detector by folding the arms. In the figure the folding is done with Fabry–Perot cavities.

wave can become quite large, much larger than the amplitude of the incident light. When resonance occurs, the small fraction of that light that leaks out of the input and output mirrors can become comparable in intensity to the incident and promptly reflected beams. When this happens, the resulting leakage beam will interfere with the promptly reflected beam, and some interesting things will result.

Near resonance, the phase of this leakage beam is very sensitive to the distance between the mirrors, so a Fabry–Perot cavity can be used as a high-precision measuring device, capable of detecting very small deviations in the distance between its mirrors. If the resonance condition is not satisfied exactly, then any small phase shift that the light picks up in one round trip will get amplified by the total number of round trips it makes before leaking out. In this sense the standing wave is analogous to the multiple bounces in a delay line, where the number of bounces is determined by the average storage time of the cavity, that is, the average number of round trips a photon makes before leaking back out. If we replace the arms of a gravitational wave detector with long Fabry–Perot cavities, as shown in Fig. 10, we can achieve folding and increase the effective arm length by a factor proportional to the storage time, or the average number of bounces, of the cavity.

Thinking of a Fabry–Perot cavity as a delay line is useful conceptually, but for a quantitative model the analogy is of limited use. We will need a quantitative model of Fabry–Perot cavities, so we now establish a few important properties of Fabry–Perot cavities.

The ratio of the incident to reflected electric fields just before the input mirror is known as the reflection coefficient of the cavity and is easy to derive. The result is

$$r = \frac{-r_i + r_o(r_i^2 + t_i^2)e^{i4\pi L/\lambda}}{1 - r_i r_o e^{i4\pi L/\lambda}}, \tag{18}$$

where L is the length of the cavity, r_i and r_o are the amplitude reflection coefficients of the input and output mirrors, and t_i is the amplitude transmission coefficient of the input mirror. For lossless mirrors $|r|^2 + |t|^2 = 1$. Resonance occurs whenever $2L = N\lambda$, where N is an integer. How precisely this condition must be met for resonance to occur depends on the reflectivities of the input and output mirrors. If the input mirror has a relatively high transmission coefficient, that is, if it lets a fairly large amount of the incident light leak into the cavity, then the condition $2L = N\lambda$ does not have to be met very precisely for resonance to occur. If the input mirror is highly reflective, then the tolerances on $2L = N\lambda$ are much
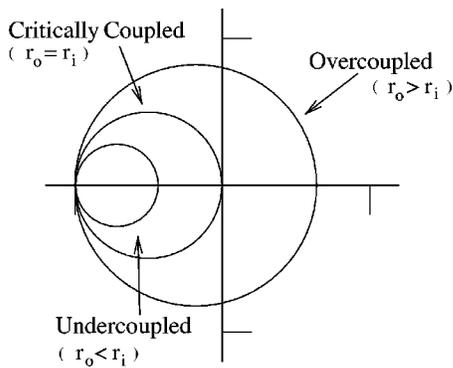
Fig. 11. The amplitude reflection coefficients for Fabry–Perot cavities are circles in the complex plane; r_i and r_o are the amplitude reflection coefficients of the input and output mirrors, respectively.

tighter. A measure of how sensitive the cavity is to changes in L or λ is the *finesse* of the cavity, $\mathcal{F}$, and is defined as the full width at half maximum of the amplitude of the standing wave inside the cavity, or linewidth, divided by the spacing between resonances, or free spectral range. If we consider λ fixed and allow L to vary, then the finesse is given by

$$\mathcal{F} \equiv \frac{\Delta L_{\mathrm{lw}}}{\lambda/2}, \tag{19}$$

where ΔL_{lw} is the linewidth. As we have said, this linewidth depends on the reflectivities of the mirrors, so we may as well define the finesse in terms of these reflectivities. For any Fabry–Perot cavity, the finesse can be written as[25]

$$\mathcal{F} = \frac{\pi \sqrt{r_i r_o}}{1 - r_i r_o}. \tag{20}$$

Now let's discuss how the reflectivities of the mirrors affect the behavior of the cavity. If we plot the reflection coefficient of a Fabry–Perot cavity in the complex plane, we find that it is always a circle.[26] The properties of this circle depend on the properties of the cavity, and there are three cases that we need to consider, $r_i = r_o$, $r_i < r_o$, and $r_i > r_o$. The reflection coefficients for each case are illustrated in Fig. 11. Far from resonance, the reflection coefficient for each case is close to -1, and its phase is relatively insensitive to changes in L or λ. As we approach resonance, say by sweeping L at a constant rate, the reflection coefficient advances around the circle in the counterclockwise direction, picking up speed as it approaches the rightmost edge of the circle. On resonance, the reflection coefficient lies on the rightmost edge of the circle, and its phase changes the most rapidly for a given change in L. After we pass through resonance, the reflection coefficient tracks along the top half of the circle and approaches -1 again, slowing down as it gets farther from resonance. It is this sensitivity in phase to changes in L near resonance that makes a Fabry–Perot cavity a useful device for measuring small changes in distance between the mirrors, and that is why we want to use it in our gravitational wave detector.

If both mirrors are lossless and have equal reflectivities, then on resonance the net reflected beam vanishes. Sufficient power builds up inside the cavity that the leakage beam exactly cancels the promptly reflected beam, and the reflection coefficient goes to zero. All of the incident power gets transmitted through the cavity. When this condition is satisfied
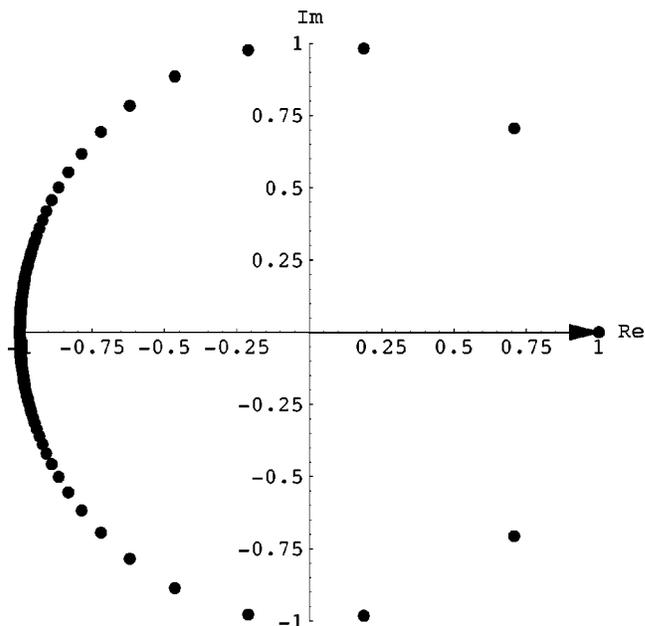

Fig. 12. On resonance, a small change in the length of a Fabry–Perot cavity dramatically changes its reflection coefficient. The amplitude reflection coefficient for an over-coupled Fabry–Perot cavity is plotted. Resonance is at $+1$, and the plot uses LIGO values for the mirrors (Ref. 27). 1000 points are plotted uniformly over a length of $\lambda/2$.

($r_i = r_o$ and both mirrors are lossless), the cavity is referred to as *critically coupled*. If the output mirror is much more reflective than the input mirror, and again both are lossless, then very little light gets transmitted through the cavity, even on resonance. The leakage beam (through the input mirror) has a larger amplitude than the promptly reflected beam, and there is some net reflected light even on resonance. In this case the cavity is referred to as *overcoupled*. If the output mirror is less reflective than the input mirror, then again little light gets transmitted through the cavity on resonance. This time, however, it is the promptly reflected beam that dominates, and the phase of the net reflected light is $-180°$, as shown in Fig. 11. (Note that the leakage beam is 180° out of phase with the promptly reflected beam on resonance, regardless of the coupling.)

If we want to use Fabry–Perot cavities as delay lines in the arms of a Michelson interferometer and recombine the light back at the beam splitter, and especially if we want to implement power recycling, it is best for us to use a strongly over-coupled cavity. In this case essentially all of the light is reflected when the cavity is on resonance, and the phase of this reflected light is very sensitive to deviations from resonance, as shown in Fig. 12.

Near resonance, the reflection coefficient for an over-coupled, lossy Fabry–Perot cavity of length L and finesse $\mathcal{F}_{\mathrm{ac}}$ is approximately

$$r_{x,y} = \left(1 - \frac{1}{\pi} \mathcal{F}_{\mathrm{ac}} \epsilon\right)\left[1 + i 8 \mathcal{F}_{\mathrm{ac}} \frac{\delta L_{x,y}}{\lambda}\right], \tag{21}$$

where $\delta L_{x,y}$ represents a small length deviation in either the x or y arm from perfect resonance, and we have approximated the finesse of an overcoupled cavity by

$$\mathcal{F}_{\mathrm{ac}} \approx \frac{2\pi}{t_i^2}. \tag{22}$$

The parameter ϵ is the fractional power lost in one round trip inside the cavity, which is typically small but not negligible for a 4 km Fabry–Perot cavity. For the arm cavity to be overcoupled, this loss must be less than the transmission of the arm cavity's input mirror, and in all our approximations for the arm cavities we will assume this to be the case.

We may use the reflection coefficients in Eq. (21) in place of the ordinary mirror reflection coefficients $r_{x,y}$ in Eq. (1) to find the transfer function of an interferometer with Fabry–Perot cavities in its arms. Only the carrier needs to resonate in the arm cavities. We will assume that the sidebands reflect off the arm cavities' input mirrors, and we will use a reflection coefficient of $r_{x,y} = -1$ for them. Then the Schnupp asymmetry only needs to be introduced between the beam splitter and the arm cavities' input masses, and these distances can be only a few meters, as opposed to several kilometers between the mirrors that form the Fabry–Perot cavities in the arms. From now on we will use a lower-case $\ell_{x,y}$ to refer to the distance between the beam splitter and the first, or input, mirror in the x and y arms. We will use an upper case $L_{x,y}$ to refer to the lengths of the Fabry–Perot cavities in the arms. Assuming that both arm cavities have the same length L in the absence of a gravitational wave and that $\ell_{x,y} \ll L$, we can get the transfer functions of an interferometer with Fabry–Perot cavity arms for both the carrier and the sidebands by combining Eqs. (1) and (21). They are, approximately,

$$t_c^{\text{ifo}} = i4 e^{ik(\ell_x + \ell_y)} \mathcal{F}_{\text{ac}} \frac{L}{\lambda} \left(1 - \frac{1}{\pi} \mathcal{F}_{\text{ac}} \epsilon \right) h, \tag{23}$$

and

$$t_\pm^{\text{ifo}} = \mp i e^{ik_\pm(\ell_x + \ell_y)} \sin\left(2\pi \frac{\Delta\ell}{\lambda} \right), \tag{24}$$

where t^{ifo} is the transfer function of the complete interferometer, and the subscripts c and $\pm$ refer to the carrier and sidebands, respectively. These transfer functions yield a demodulated signal of

$$V_{\text{signal}} = 4(RP_{\text{in}})J_0(\beta)J_1(\beta)\mathcal{F}_{\text{ac}} \frac{L}{\lambda} \sin\left(2\pi \frac{\Delta\ell}{\lambda_{\text{mod}}} \right)$$
$$\times \left(1 - \frac{1}{\pi} \mathcal{F}_{\text{ac}} \epsilon \right) h, \tag{25}$$

where we have again optimized the phase ϕ in the mixing process to maximize our signal. The introduction of folding using Fabry–Perot cavities in the arms has increased the effective length of the arms by about a factor of $\mathcal{F}_{\text{ac}}$, with a corresponding increase in the strength of the readout signal. A representative value of the finesse of an arm cavity in a real detector is $\mathcal{F}_{\text{ac}} \approx 130$.[27]

VII. BOOSTING THE EFFECTIVE POWER: POWER RECYCLING

We have seen how folding increases the effective length L of an interferometer without the need to make the instrument physically larger. Now we will turn to the input power P_{in} and see how that can be boosted, with a corresponding boost in the response of the instrument.

The most obvious solution is to use a stronger light source. As of this writing, most gravitational wave detectors use la-

sers that operate on the order of 10 W, and 100 W lasers are in development. However, in a process similar to folding, we can increase the effective power going into the instrument without changing the power of the laser. How is this possible? Well, if the output port is dark, energy conservation demands that the majority of the light gets reflected back toward the laser. If we could somehow recycle this wasted light and send it back into the interferometer, we could boost the sensitivity of the instrument without having to develop a more powerful laser. (And we could get even more out of such a laser when it becomes available.)

We saw in Sec. VI how a Fabry–Perot cavity could be used to implement folding of an optical path. A similar concept can also be used to implement power recycling.[28] In this case, we place a single, partially transmitting mirror between the laser and the beam splitter in our gravitational wave detector. The interferometer itself, the beam splitter and arms, acts like a partially transmitting mirror, as shown in Fig. 13. Some of the light incident on the beam splitter gets transmitted through the instrument, coming out the dark port, while most of it (most of the carrier, anyway) gets reflected back toward the laser. If we think of the interferometer as a compound mirror, we can use it as the output mirror of a Fabry–Perot cavity, placing a second, ordinary mirror between it and the laser. This second mirror then acts as the input mirror, and if we control the distance between it and the interferometer (the compound output mirror), then we can build up a standing wave between the two, effectively increasing the power incident on the interferometer by a factor proportional to the finesse of the resulting cavity. We will refer to this cavity as the *recycling cavity*, and we refer to the mirror we have introduced between the laser and the beam splitter as the *recycling mirror*.

The transmission coefficient of the complete interferometer, including power recycling and Fabry–Perot cavities in the arms, is the transmission coefficient of the recycling cavity, with the recycling mirror forming the input mirror and the rest of the interferometer acting as the output mirror. The recycling cavity reflection coefficient is just given by Eq. (18), with the transmission and reflection coefficients of the compound mirror, the interferometer with Fabry–Perot arm cavities, substituted for the output mirror coefficients t_o and r_o. The transmission coefficient of the recycling cavity is the same as that for any Fabry–Perot cavity, or

$$t_{\text{rc}} = \frac{t_{\text{rm}} t_{\text{ifo}} e^{i2\pi\ell_{\text{rm-bs}}/\lambda}}{1 - r_{\text{rm}} r_{\text{ifo}} e^{i4\pi\ell_{\text{rm-bs}}/\lambda}}, \tag{26}$$

where t_{rm} and r_{rm} are the transmission and reflection coefficients for the recycling mirror, t_{ifo} and r_{ifo} are the transmission and reflection coefficients for the rest of the interferometer, and $\ell_{\text{rm-bs}}$ is the distance from the recycling mirror to the beam splitter. Another way to think of $\ell_{\text{rm-bs}}$ is the distance between the input and (compound) output mirrors in the recycling cavity.

We calculated the transmission coefficients t_{ifo} for the carrier and sidebands in Sec. VI. The reflection coefficients are calculated in a similar manner:

$$r_{\text{ifo}} = \tfrac{1}{2} e^{ik(\ell_x + \ell_y)} [r_x e^{ik(\ell_x - \ell_y)} + r_y e^{-ik(\ell_x - \ell_y)}]. \tag{27}$$

Here, as before, ℓ_x is the distance from the beam splitter to the first mirror in the x arm (the input mirror to the x-arm's Fabry–Perot cavity), and ℓ_y is the distance from the beam splitter to the y-arm's input mirror.

A. Sidebands

For the sidebands that do not resonate in the arm cavities, the reflection coefficients are

$$r_x = r_y = -1. \tag{28}$$

This result yields

$$r_\pm^{\text{ifo}} = -e^{ik_\pm(\ell_x+\ell_y)}\cos\left(2\pi\frac{\Delta\ell}{\lambda_{\text{mod}}}\right) \tag{29}$$

and

$$t_\pm^{\text{ifo}} = \mp ie^{ik_\pm(\ell_x+\ell_y)}\sin\left(2\pi\frac{\Delta\ell}{\lambda_{\text{mod}}}\right). \tag{30}$$

If we substitute these into our expression for r_{rc} [Eq. (18) with the appropriate input and output mirror coefficients for the recycling cavity], we find that the resonance condition is now

$$e^{ik_\pm(2\ell_{\text{rm-bs}}+\ell_x+\ell_y)} = -1, \tag{31}$$

and critical coupling, that is, optimum recycling, requires that we adjust the Schnupp asymmetry so that

$$\cos\left(2\pi\frac{\Delta\ell}{\lambda_{\text{mod}}}\right) = r_{\text{rm}}. \tag{32}$$

With these conditions, the transmission coefficients for the sidebands through the recycling cavity (the complete interferometer) are

$$t_\pm^{\text{rc}} = \pm ie^{-ik_\pm\ell_{\text{rm-bs}}}. \tag{33}$$

The full power of the sidebands gets transmitted, and the field picks up a phase factor.

B. Carrier

We can calculate the transmission and reflection coefficients of the compound mirror for the carrier the same way we calculated them for the sidebands. Using the reflection coefficients $r_{x,y}$ from an overcoupled Fabry–Perot cavity near resonance, we find

$$t_c^{\text{ifo}} = ie^{ik(\ell_x+\ell_y)}4\mathcal{F}_{\text{ac}}\frac{L}{\lambda}h\left(1-\frac{1}{\pi}\mathcal{F}_{\text{ac}}\epsilon\right) \tag{34}$$

and

$$r_c^{\text{ifo}} = e^{ik(\ell_x+\ell_y)}\left(1-\frac{1}{\pi}\mathcal{F}_{\text{ac}}\epsilon\right). \tag{35}$$

If we use these results to find the carrier transmission and reflection coefficients for the recycling cavity, we find that the resonance condition is

$$e^{ik(2\ell_{\text{rm-bs}}+\ell_x+\ell_y)} = +1, \tag{36}$$

and the requirement for critical coupling is

$$r_{\text{rm}} = \left(1-\frac{1}{\pi}\mathcal{F}_{\text{ac}}\epsilon\right). \tag{37}$$

The requirements for resonance and critical coupling for both the carrier and the sidebands can be combined. They are

$$\cos\left(2\pi\frac{\Delta\ell}{\lambda_{\text{mod}}}\right) = r_{\text{rm}} = 1-\frac{1}{\pi}\mathcal{F}_{\text{ac}}\epsilon, \tag{38}$$

$$e^{ik(2\ell_{\text{rm-bs}}+\ell_x+\ell_y)} = +1, \tag{39}$$

and

$$e^{ik_{\text{mod}}(2\ell_{\text{rm-bs}}+\ell_x+\ell_y)} = -1. \tag{40}$$

The Schnupp asymmetry and the reflectance of the recycling mirror must both be adjusted to match the product of the arm cavities' loss and finesse, and the effective length of the recycling cavity becomes the sum of $\ell_{\text{rm-bs}}$ and the average length $(\ell_x+\ell_y)/2$. This effective length must be both a multiple of $\lambda/2$, where λ is the wavelength of the light used in the interferometer, and a multiple of $\lambda_{\text{mod}}/4$, where the modulation wavelength is given by λ_{mod}. The first condition tells us that the carrier must be resonant in the recycling cavity. Because the second condition is $\lambda_{\text{mod}}/4 = \ell_{\text{rm-bs}} + (\ell_x+\ell_y)/2$, we say that the sidebands must be antiresonant in the recycling cavity. Because $\lambda \approx 1\ \mu\text{m}$ and $\lambda_{\text{mod}} \approx 30$ m, it is not difficult to achieve both of these conditions at the same time.

With the conditions given by Eqs. (38), (39), and (40) met, the transfer function for the carrier through the recycling cavity (that is, the complete interferometer) becomes

$$t_c^{\text{rc}} = ie^{-ik\ell_{\text{rm-bs}}}\frac{4}{\sqrt{\pi}}\mathcal{F}_{\text{ac}}\sqrt{\mathcal{F}_{\text{rc}}}\frac{L}{\lambda}h, \tag{41}$$

where we have approximated the finesse of the optimally coupled recycling cavity as

$$\mathcal{F}_{\text{rc}} \approx \frac{\pi}{t_i^2}. \tag{42}$$

C. DC response of the interferometer

We are now in a position to calculate the response of the complete interferometer, including lock-in detection, Fabry–Perot arm cavities, and power recycling as shown in Fig. 14. All we need to do is calculate the electric field of the light falling on the photodiode,

$$E_{\text{out}} = E_{\text{in}}[t_c^{\text{rc}}J_0(\beta)e^{i\omega t} + t_+^{\text{rc}}J_1(\beta)e^{i(\omega+\Omega)t}$$
$$- t_-^{\text{rc}}J_1(\beta)e^{i(\omega-\Omega)t}], \tag{43}$$

then demodulate and average the resulting power.

$$V_{\text{signal}} = R\langle|E_{\text{out}}|^2\cos(\Omega t+\phi)\rangle, \tag{44}$$

where the angled brackets $\langle\cdots\rangle$ denote time averaging. The result is

$$V_{\text{signal}} = \frac{4}{\sqrt{\pi}}(RP_{\text{in}})J_0(\beta)J_1(\beta)\mathcal{F}_{\text{ac}}\sqrt{\mathcal{F}_{\text{rc}}}\frac{L}{\lambda}h. \tag{45}$$

Note that the response of the interferometer is enhanced by $\sqrt{\mathcal{F}_{\text{rc}}}$, a relatively minor improvement compared with the $\mathcal{F}_{\text{ac}}$ gain from using Fabry–Perot cavities as delay lines in the arms. A moment's thought shows that this weaker dependence should not surprise us. The response of the basic interferometer, given in Eq. (17), is proportional to the electric field in the carrier, $\sqrt{P_{\text{in}}}J_0(\beta)$, multiplied by the field in the sidebands, $\sqrt{P_{\text{in}}}J_1(\beta)$, multiplied by the length of the arms, ℓ. Folding the arms using Fabry–Perot cavities increases the effective length to $(2/\pi)\mathcal{F}_{\text{ac}}\ell$, because the total number of round trips in a Fabry–Perot cavity is proportional to its finesse. The field in a cavity, however, is proportional to the square root of its finesse. Because the sidebands do not reso-

nate in the arm cavities, only the carrier field is boosted by power recycling. The response of the instrument then is enhanced by the new field amplitude in the carrier, or

$$E_{\text{carrier}} \propto \sqrt{\mathcal{F}_{\text{ac}}}\sqrt{P_{\text{in}}}J_0(\beta). \tag{46}$$

D. AC response of the interferometer

So far we have performed all of our calculations under the assumption that the interferometer is in a quasistatic state. We have not taken into account dynamical effects. As long as any change in the instrument occurs on a time scale longer than the time it takes light to propagate completely through the instrument, this assumption is valid. However, if our gravitational wave signal h includes components at high enough frequencies, we cannot expect our quasistatic version of the response of the instrument to be valid. Before addressing the dynamic response of the interferometer, let us consider what these relevant time scales are, and what "high frequency" means.

How long does it take for the light in the interferometer to come to a steady state? Put another way, how long does it take for an appreciable field to build up inside a Fabry–Perot cavity so that the appropriate interference can occur? Light makes a single round trip inside a cavity in a time $2L/c$, and if we think of the finesse as a measure of the total number of round trips the light makes, then the relevant time scale for equilibrium in a Fabry–Perot cavity must be approximately

$$\tau \sim \frac{2L}{c}\mathcal{F}. \tag{47}$$

Any signals at frequencies of $f_0 \ll 1/\tau$ will satisfy the quasistatic requirement, but frequencies on the order of or greater than $1/\tau$ will be high enough to begin to probe the dynamic response of the interferometer.

If the arm cavities have finesses of 130 and lengths of 4 km, the characteristic frequency is

$$\frac{1}{\tau} \sim 300 \text{ Hz}, \tag{48}$$

which is low enough to be a problem if we are looking for signals at frequencies up to a kilohertz. The recycling cavity, on the other hand, also has $\mathcal{F}_{\text{rc}} \approx 100$, but its length is on the order of 10 m, giving it a cutoff frequency closer to 100 kHz. Our rough estimate shows that, for searching for gravitational waves at frequencies above a few hundred Hz, we must take into account the storage time τ of the arm cavities, but not that of the recycling cavity.

A careful derivation shows that the finite storage time of the arm cavities introduces a single pole in the response of the interferometer, and that the cutoff frequency f_0 is actually $1/2\tau$, or[29]

$$f_0 = \frac{c}{4L\mathcal{F}_{\text{ac}}}. \tag{49}$$

The dynamic response of the interferometer to a gravitational wave signal $h(f)$ at frequency f has the same frequency dependence as a simple, single-pole low-pass filter. For frequencies well below the recycling cavity's cutoff frequency, this response is[27]

$$V_{\text{signal}}(f)$$
$$= \frac{4}{\sqrt{\pi}}(RP_{\text{in}})J_0(\beta)J_1(\beta)\mathcal{F}_{\text{ac}}\sqrt{\mathcal{F}_{\text{rc}}}\frac{L}{\lambda}\frac{f_0}{f_0+if}h(f). \tag{50}$$

VIII. SHOT NOISE

No treatment of gravitational wave detectors would be complete without some consideration of noise. Even in a 4 km instrument, a strain of 10^{-21} produces motions of the end mirrors on the order of 10^{-18} m, or one-thousandth of the diameter of an atomic nucleus. Any instrument capable of detecting such motion must have a very low intrinsic noise level. The extraordinary sensitivity required for gravitational wave astronomy demands that our instrument have an extremely low noise floor, and to accomplish that we must have an understanding of the physics behind the noise. There are three principal sources of noise in an interferometric gravitational wave detector. They are

(1) *Seismic noise* which will perturb our mirrors if the ground shakes and can mimic a gravitational wave strain h.
(2) *Thermal noise* due to the motion of the individual molecules in the mirrors and suspensions due to heat.
(3) *Shot noise* due to the optical readout noise in our photodetector.

The first two sources can be reduced by clever design of the tables and suspensions used to hold the mirrors, and by an appropriate choice of the mirror material and geometry. The third, shot noise, is a fundamental consequence of the quantum nature of light.

In this section we will consider the limit that shot noise places on our ability to detect a gravitational wave strain h. The rms fluctuation in the power falling on the photodiode determines the shot noise seen in the readout of the instrument. We can use the standard formula for shot noise to calculate this fluctuation,[30]

$$\Delta P_{\text{shot,rms}} = \sqrt{2\hbar\omega P_{\text{dc}}\Delta f}, \tag{51}$$

where P_{dc} is the average dc power falling on the photodiode, and Δf is the measurement bandwidth. It will be useful for us to express this noise in terms of a voltage fluctuation per unit bandwidth, or

$$V_{\text{shot}}(f) = R\sqrt{2\hbar\omega P_{\text{dc}}}. \tag{52}$$

We can think of this expression as a Fourier transform of the shot noise voltage, expressing the noise density at a given frequency; V_{shot} in Eq. (52) has units of volts per square-root hertz. Note that the amplitude of the shot noise in the photodiode does not depend on frequency, only on the bandwidth of our measurement Δf.

This shot noise appears as a signal in our readout, and we would like to know how large h has to be to stand out above the noise. All we have to do to answer this question is set the output of the instrument, Eq. (50), equal to the signal due to shot noise.

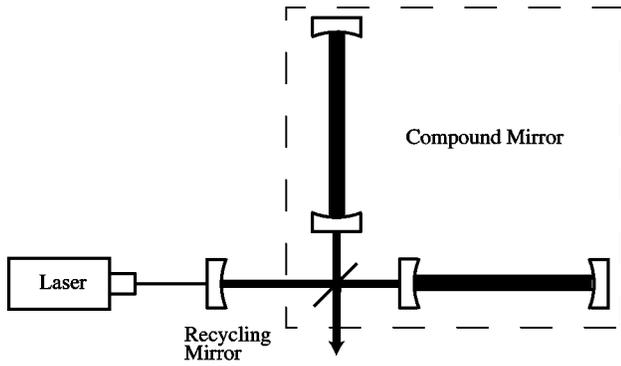$$V_{\text{signal}}(f) = V_{\text{shot}}(f) \tag{53}$$

or

Fig. 13. To understand how a single mirror can be used to implement power recycling, we can treat the interferometer as a compound mirror and consider it to be the output mirror of a Fabry–Perot cavity. The recycling mirror serves as the input mirror. If the resulting recycling cavity is critically coupled, then no light is reflected back to the laser.

$$\frac{4}{\sqrt{\pi}}(RP_{in})J_0(\beta)J_1(\beta)\mathcal{F}_{ac}\sqrt{\mathcal{F}_{rc}}\frac{L}{\lambda}\sqrt{\frac{f_0^2}{f_0^2+f^2}}|h_{shot}(f)|$$
$$= \sqrt{2\hbar\omega P_{dc}}R. \tag{54}$$

The average dc power is simply found from the output field, as we calculated in Sec. VII,

$$P_{out} = \langle|E_{out}|^2\rangle = 2J_1^2(\beta)P_{in}. \tag{55}$$

Equation (55) gives an expression for the shot noise limited sensitivity of an interferometric detector, or $|h_{shot}(f)|$:

$$|h_{shot}(f)| = \frac{\pi}{2}\frac{\sqrt{\hbar\lambda c}}{J_0(\beta)\sqrt{P_{in}}\mathcal{F}_{ac}\sqrt{\mathcal{F}_{rc}}L}\sqrt{\frac{f_0^2+f^2}{f_0^2}}. \tag{56}$$

Note that, even though the shot noise in the photodetector itself is frequency independent, the frequency-dependent response of the interferometer makes the shot noise limit frequency-dependent at high frequencies as illustrated in Fig. 15. This effect is merely a result of dividing the (flat) shot noise spectrum by the response of the interferometer.
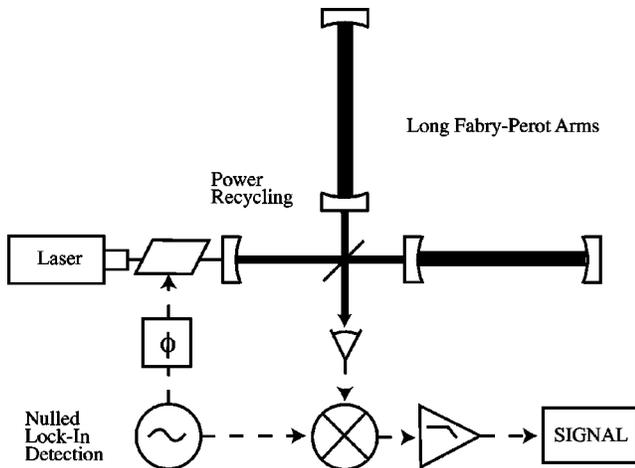

Fig. 14. The complete interferometric gravitational wave detector combines Fabry–Perot arms, power recycling, and nulled lock-in detection.
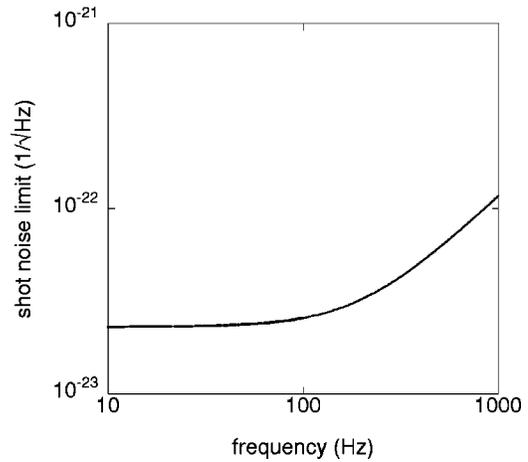

Fig. 15. The shot noise limited sensitivity for a interferometric gravitational wave detector with $L=4$ km, $P_{in}=10$ W, $\lambda=1.064$ μm, $\mathcal{F}_{ac}=130$, and $\mathcal{F}_{rc}=100$. The frequency dependence comes from dividing the shot noise in the photodiode by the response of the interferometer.

Representative parameters for a first-generation gravitational wave detector are $L=4$ km, $\lambda=1$ μm, $\mathcal{F}_{ac}=130$, $\mathcal{F}_{rc}=100$, and $P_{in}=10$ W. These parameters give a dc shot noise of approximately

$$|h_{shot}(0)| \approx 10^{-23} \text{ Hz}^{-1/2}. \tag{57}$$

IX. OPERATING AN INTERFEROMETRIC GRAVITATIONAL WAVE DETECTOR

Our analysis of an interferometric gravitational wave detector has relied on resonance conditions being satisfied, which involves maintaining precise distances between mirrors over long periods of time. The systems necessary to hold the Fabry–Perot cavities in the arms on resonance, to hold the recycling mirror at its optimum position, and to keep the output port dark for the carrier, are elaborate, sophisticated, and beyond the scope of this paper. For our purposes, the analysis of signal extraction and shot noise, we can just assume that all of these conditions are met. The automatic control systems that maintain them are considered transparent for this analysis, and their inner workings are a separate study of modern experimental techniques.

X. SUGGESTED PROBLEMS

(1) Derive the transmission coefficients for a simple Michelson interferometer at both the symmetric and antisymmetric ports, Eqs. (1) and (27), respectively. There is a subtlety here that is important to appreciate before going on to do any other calculations involving interferometry: The reflection coefficient of the beam splitter depends on the incoming direction of the incident light. In our case, we considered reflection from the left to have $r=-1/\sqrt{2}$ and from the right, $r=+1/\sqrt{2}$. This assumption is a slight oversimplification. The mirrors in modern interferometers are made up of multilayer dielectric coatings, with complicated transmission and reflection coatings. Nevertheless, this asymmetry in the sign of the reflection coefficient, and *not* the transmission coefficient, is required for the instrument to satisfy conservation of energy. This is a general property of mirrors and is not limited to beam splitters.

(2) Derive the reflection and transmission coefficients for a Fabry–Perot cavity, Eqs. (18) and (26), respectively. Again, conservation of energy requires that the reflection coefficient of the input mirror have a different sign for light incident from the left than for light incident from the right. The usual derivation of these coefficients involves summing an infinite series.[20] However, a more elegant method can be found in Siegman.[25]

(3) *Field lines 1*. Consider two points in space–time in the $x-y$ plane: an origin, described by $(0,0,0,0)$, and some other point $(ct,x,y,0)$. The interval s between these two points is given, in the ordinary case of flat space–time, by

$$s^2 = -c^2t^2 + x^2 + y^2. \tag{58}$$

The differential interval between two points in space–time can be expressed more generally as

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu, \tag{59}$$

where the metric $g_{\mu\nu}$ is known as the metric tensor. (Here we employ the usual Einstein summation convention of summation over repeated indices.) For the flat space–time of special relativity, the metric tensor is

$$g_{\mu\nu} = \eta_{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{60}$$

We can describe a gravitational wave propagating along the z axis, with (low) frequency $\omega/2\pi$ and $+$ polarization, by a perturbation of this metric,[15,16]

$$g_{\mu\nu} \rightarrow \eta_{\mu\nu} + h_{\mu\nu}, \tag{61}$$

where the additional, perturbative term is given by

$$h_{\mu\nu} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} h \cos[\omega(t-z/c)]. \tag{62}$$

This change in the metric is analogous to the movement of a free particle at coordinates x and y under the influence of a gravitational wave. Construct a vector field at a fixed time t in the $x-y$ plane that describes this change in the metric and plot it.

Solution: The metric between the origin and any point in the $x-y$ plane is given by

$$s^2 = -c^2t^2 + (x^2+y^2) + (x^2-y^2)h\cos(\omega t). \tag{63}$$

The change in the metric for a given point at coordinates x and y at time $t=0$ is then

$$\delta s^2 = (x^2-y^2)h. \tag{64}$$

We may visualize the pattern of motion induced by this change in the metric by looking at the spatial gradient of δs^2,

$$\vec{g} = \vec{\nabla} \delta s^2 = 2x\hat{i} - 2y\hat{j}. \tag{65}$$

See Chap. 5 of Ref. 15 for a rigorous treatment of the theory of gravitational waves.

(4) *Field lines 2*. Consider the case of $\times$ polarization, where the metric perturbation is given by

$$h_{\mu\nu} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} h \cos[\omega(t-z/c)]. \tag{66}$$

Again, construct and plot a vector field describing the lines of force. Show that a Michelson interferometer with arms oriented along the x and y axes is not sensitive to $\times$-polarized gravitational waves.

(5) The response of an interferometric gravitational wave detector with Fabry–Perot arm cavities depends on the product of the arm length and the arm cavity finesse, $\mathcal{F}_{ac}L$. Why, then, do we need to build detectors with kilometer-scale arms? Wouldn't it be easier to keep the arm lengths at a few tens of meters (laboratory scales) and compensate by increasing the arm-cavity finesse?

Solution: There are many answers to this question. The easiest to understand, at this level, is that the response of the instrument, as in Eq. (50), actually scales as the product $\mathcal{F}_{ac}Lh$. The amount of motion induced in the mirrors by a gravitational wave scales with the length of the instrument, Lh, whereas most external noise sources, such as seismic or thermal noise, move the mirrors by an amount that is independent of the arm length. Thus the signal-to-noise ratio of the instrument scales as the arm length L.

[1] Karl G. Jansky, "Electrical disturbances apparently of extraterrestrial origin," Proc. IRE **21**, 1387–1398 (1933).
[2] A. P. Cowley, "Evidence for black holes in stellar binary systems," Annu. Rev. Astron. Astrophys. **30**, 287–310 (1992).
[3] A. A. Penzias and R. W. Wilson, "A measurement of excess antenna temperature at 4080MC/S," Astrophys. J. **142**, 419–421 (1965).
[4] A. H. Jaffe *et al.*, "Cosmology from MAXIMA-1, BOOMERANG, and COBE DMR cosmic microwave background observations," Phys. Rev. Lett. **86**, 3475–3479 (2001).
[5] V. Schönfelder, *The Universe in Gamma-Rays* (Springer, Berlin, New York, 2001).
[6] K. Hirata *et al.*, "Observation of a neutrino burst from supernova SN1987A," Phys. Rev. Lett. **58**, 1490–1493 (1987).
[7] Q. R. Ahmad *et al.*, "Measurement of the rate of $nu(e)+d \rightarrow p+p +e(-)$ interactions produced by B-8 solar neutrinos at the Sudbury neutrino observatory," Phys. Rev. Lett. **87**, 071301-1–6 (2001).
[8] P. Astone *et al.*, "First cooling below 0.1 K of the new gravitational-wave antenna nautilus of the Rome group," Europhys. Lett. **16**, 231–235 (1991).
[9] Barry C. Barish and Rainer Weiss, "LIGO and the detection of gravitational waves," Phys. Today **52**, 44–50 (1999).
[10] B. Caron *et al.*, "The Virgo interferometer," Class. Quantum Grav. **14**, 1461–1469 (1997).
[11] M. Cerdonio *et al.*, "The ultracryogenic gravitational-wave detector AURIGA," Class. Quantum Grav. **14**, 1491–1494 (1997).
[12] Karsten Danzmann, "LISA: Laser interferometer space antenna for gravitational wave measurements," Class. Quantum Grav. **13**, A247–A250 (1996).
[13] Keita Kawabe and the TAMA collaboration, "Status of the TAMA project," Class. Quantum Grav. **14**, 1477–1480 (1997).

[14] H. Lück and the GEO600 team, "The GEO600 project," Class. Quantum Grav. **14**, 1471–1476 (1997).

[15] Hans Ohanian and Remo Ruffini, *Gravitation and Spacetime*, 2nd ed. (W. W. Norton, New York, 1994).

[16] Peter R. Saulson, *Fundamentals of Interferometric Gravitational Wave Detectors* (World Scientific, Singapore, 1994).

[17] R. A. Hulse and J. H. Taylor, "Discovery of a pulsar in a binary system," Astrophys. J., Lett. Ed. **195**, L51–L53 (1975).

[18] J. H. Taylor and J. M. Weisberg, "A new test of general relativity—gravitational-radiation and the binary pulsar PSR-1913+16," Astrophys. J. **253**, 908–920 (1982).

[19] J. H. Taylor and J. M. Weisberg, "Further experimental tests of relativistic gravity using the binary pulsar PSR-1913+16," Astrophys. J. **345**, 434–450 (1989).

[20] Eugene Hecht and Alfred Zajac, *Optics*, 3rd ed. (Addison-Wesley, Reading, MA, 1998).

[21] R. W. P. Drever *et al.*, "Laser phase and frequency stabilization using an optical resonator," Appl. Phys. B: Photophys. Laser Chem. **31**, 97–105 (1983).

[22] L. Schnupp, talk at a European Collaboration Meeting on Interferometric Detection of Gravitational Waves, Sorrento, 1988.

[23] R. Takahashi, J. Mizuno, S. Miyoki, and N. Kawashima, Phys. Lett. A **187**, 157 (1994).

[24] R. W. P. Drever *et al.*, "A gravity-wave detector using optical cavity sensing," in *Proceedings of the Ninth International Conference on General Relativity and Gravitation, Jena, 14–16 July 1980*, edited by E. Schmutzer (Cambridge U. P., Cambridge, 1983), pp. 265–267.

[25] A. E. Siegman, *Lasers* (University Science Books, Mill Valley, CA, 1986).

[26] Eric D. Black, "An introduction to Pound–Drever–Hall laser frequency stabilization," Am. J. Phys. **69**, 79–87 (2001).

[27] Jordan B. Camp, H. Yamamoto, S. E. Whitcomb, and D. E. McClelland, "Analysis of light noise sources in a recycled Michelson interferometer with Fabry–Perot arms," J. Opt. Soc. Am. A **17**, 120–128 (2000).

[28] R. W. P. Drever, "Interferometric detectors for gravitational radiation," in *Gravitational Radiation, Proceedings of the Les Houches Summer Institute, 2–21 June 1982*, edited by T. Piran and N. Deruelle (North-Holland, Amsterdam, 1983).

[29] Jun Mizuno, "Comparison of optical configurations for laser-interferometric gravitational-wave detectors," Ph.D. thesis, Max-Planck-Institut für Quantenoptik, July, 1995.

[30] Amnon Yariv, *Optical Electronics in Modern Communications*, 5th ed. (Oxford University Press, New York, 1997).