Modulation Locking Subsystems for Gravitational Wave Detectors

Benedict John Cusack

September 2004

Thesis submitted for the degree of Master of Philosophy
at the Australian National University

Declaration

I certify that the work contained in this thesis is, to the best of my knowledge, my own original research. All material taken from other references is explicitly acknowledged as such. I certify that the work contained in this thesis has not been submitted for any other degree.

Benedict Cusack
September, 2004

The Road Not Taken

Two roads diverged in a yellow wood,
And sorry I could not travel both
And be one traveller, long I stood
And looked down one as far as I could
To where it bent in the undergrowth;

Then took the other, as just as fair,
And having perhaps the better claim,
Because it was grassy and wanted wear;
Though as for that the passing there
Had worn them really about the same,

And both that morning equally lay
In leaves no step had trodden black.
Oh, I kept the first for another day!
Yet knowing how way leads on to way,
I doubted if I should ever come back.

I shall be telling this with a sigh
Somewhere ages and ages hence;
Two roads diverged in a wood, and I —
I took the one less travelled by,
And that has made all the difference.

Robert Frost, 1920.

Acknowledgements

I don't think I ever doubted that this document would one day come into existence, it is more that I couldn't see for a long time what would fill it. Finding a research project from scratch is really hard! There are many people whom I wish to thank because, in the end, I made it.

In terms of the science, I owe a great debt of gratitude to the gifted Ping Koy Lam for encouraging me that the research I have produced has value. He has gone beyond the call of duty in helping me reach my goal, and he has always set a wise and ethical example as a researcher. I would also like to thank Stan Whitcomb, Daniel Shaddock, Malcolm Gray and Benjamin Sheard, who all contributed in some significant way to the quality of my experimental results. I am grateful to David McClelland without whose tireless effort writing grant applications the ANU Centre of Gravitational Physics, and more to the point the laboratory and all its equipment, would not exist. A particularly big thanks go to Katie Pilypas, Ben Sheard, Ping Koy Lam and David McClelland for their timely proofreading of this document.

I would like to thank my fellow students at the ANU Physics Department, including Ben Sheard, Simon Haine, Katie Pilypas, Magnus Hsu, Andrew Lance, Vincent Delaubert, Kirk McKenzie, Cameron Fletcher, Nick Robins, Geoff Ericsson, Mike Ashley, Ben Whale, Ingrid Irmer, Joe Hope, Alan Griffith and others I'm sure I have forgotten, who are always ready with a smile or a joke in the hallway. A nod of thanks goes to the ACIGA Data Analysis group, particularly Susan Scott, Antony Searle and Andrew Moylan, whom I always felt at home working with, and likewise to John Smillie and Ben Evans who run the supercomputers and mass data stores. Thanks also to the administration and technical staff, Sharon Lopez, Damien Hughes, Zeta Hall, Andrew Papworth and David Bramwell, whom I have always found amenable and helpful. Perhaps most of all, I'd like to thank Aidan Byrne for being a part of the solution.

A number of people I met while overseas were very helpful with comments and tidbits of information, so thank you to Ken Strain, David Reitze, Peter Bender, Graham Woan, Benoit Mours, Albert Lazzarini, Julien Sylvestre, Mike Zucker, John McNabb, Patrick Sutton and Szabolcs Marka.

On a personal front, thank you to my close friends for timely diversions and companionship, especially Stefan Kaufman, Josh Morton, Ben Robbins, Andrew Ley, Scott Edwards, Scott Pacey, Aung Si, John Greenwood, Fergus McGlynn, Ben Sheard (again!), Peter Pullicino, Christine Carmody and Brett Cuthbertson. Thanks to my family, Frances, Kara, Leonie, Gregory, Julie and Justin for their support and for reminders that there is life outside a research office. Thanks to Laurenza, Jan and Mike for keeping me sane.

Most of all, thanks and lots of love to my wonderful Rachel, with whom I am happy without even trying.

Publications

Two papers have been published from the results of this thesis. Parts of Chapters 3, 5 and 7 are taken from the first, and parts of Chapter 4 are taken from the second.

- B J Cusack, B S Sheard, D A Shaddock, M B Gray, P K Lam, and S E Whitcomb. Electro-optic modulator capable of generating simultaneous amplitude and phase modulations. *Applied Optics*, 43(26):5079–5091, 2004.

- B J Cusack, D A Shaddock, B J J Slagmolen, G de Vine, M B Gray, and D E McClelland. Double pass locking and spatial mode locking for gravitational wave detectors. *Classical and Quantum Gravity*, 19(7):1819–1824, 2002.

In addition, B J Cusack contributed to the work of the LIGO Scientific Collaboration, and to the work of ACIGA, earning co-authorship on the following papers:

- B Abbott et al. Analysis of LIGO data for gravitational waves from binary neutron stars. *Physical Review D*, 69(12):2001, 2004.

- B Abbott et al. Analysis of first LIGO science data for stochastic gravitational waves. *Physical Review D*, 69(12):2004, 2004.

- B Abbott et al. First upper limits from LIGO on gravitational wave bursts. *Physical Review D*, 69(10):2001, 2004.

- B Abbott et al. Setting upper limits on the strength of periodic gravitational waves from PSR J1939+2134 using the first science data from the (geo) 600 and LIGO detectors. *Physical Review D*, 69(12):2004, 2004.

- B Allen and G Woan et al. Upper limits on the strength of periodic gravitational waves from PSR J1939+2134. *Classical & Quantum Gravity*, 21(5):S671–S676, 2004.

- B Abbott et al. Detector desription and performance for the first coincidence observations between LIGO and GEO. *Nuclear Instruments and Methods in Physics Research A*, 517(1-3):154–179, 2004.

- S M Scott, A C Searle, B J Cusack, and D E McClelland. The ACIGA data analysis programme. *Classical & Quantum Gravity*, 21(5):S853–S856, 2004.

- L Ju et al. ACIGA's high optical power test facility. *Classical & Quantum Gravity*, 21(5):S887–S893, 2004.

- J S Jacob et al. Australia's role in gravitational wave detection. *Publications Astronomical Society of Australia*, 20(3):223–241, 2003.

Abstract

This thesis deals with modulation-based optical subsystems in gravitational wave detectors. Modulation locking techniques used for longitudinal feedback control of gravitational wave detectors are reviewed and analysed. The prospect of tuning the frequency response of a signal-recycled gravitational wave detector, potentially in real-time, is discussed. The technique proposed, called *tune-locking* is theoretically investigated and shown to have potential for independently tuning the resonance width and peak frequency of a detector's response.

An electro-optic device called a *quadrature variable modulator* (QVM) is theoretically and experimentally proven to be capable of producing the range of modulations required for the tune-locking configurations. The QVM is also tested with a range of other potential applications in mind.

A new theory of modulation called the *modulation sphere* representation is presented, which clarifies the physical significance of amplitude modulation (AM) and phase modulation (PM) and also of the relative phase between AM and PM in a situation where both are present. The modulation sphere or *M-space* formulation is used throughout the thesis to further the understanding of the modulation-based subsystems presented. The modulation sphere is particularly helpful as a visual tool for navigating the use of the QVM device.

Fabry-Perot "mode-cleaner" cavities are discussed, and alternatives to the commonly used Pound-Drever-Hall locking are proposed. The alternative locking methods, called *tilt locking* and *flip locking* use spatial mode interference rather than modulation, to avoid thermal lensing problems with modulators used with high optical powers. Tilt locking was developed by Shaddock and Gray [1], and flip locking constitutes an original component of this thesis. *Double pass* locking is also presented independently as a means of reducing the maximum power that a photodetector will be exposed to during lock acquisition.

Contents

List of Figures

List of Tables

Introduction to Gravitational Waves and their Detection

This chapter outlines the history of gravitational waves, from theoretical discovery to current day detector technology. Except where otherwise referenced, the information in this chapter is distilled from [2–4]. These should be consulted for further references of individuals' work.

1.1 History of Gravity: Newton to Einstein

Isaac Newton's description of gravity as a force between two massive bodies, proportional to each of their respective masses, persisted as the best available theory of gravity until late in the 19th century. Newton's theory explained the elliptical motions of the planets around the sun (attributed to Johannes Kepler) and, with the exception of Mercury's precession, theory and experiment were in excellent agreement. Newton's theory also showed that the phenomena of objects falling to the Earth's surface and the motions of celestial objects were due to one and the same force: gravity.

Newton's theory wasn't entirely satisfactory; theorists still wondered about the dual role of mass in gravity. An object's mass determines the magnitude of gravitational force exerted on it, but the mass also determines the rate of acceleration due to that force. It seemed that a more enlightening theory existed if only one could find it.

Arguably, the seeds of change were planted when James Clerk Maxwell used his mathematical laws of electromagnetism (circa 1864) to show that light is an electromagnetic wave with a finite propagation speed. All waves known to the scientific community at that point propagated in some sort of medium (sea waves in water, sound waves in air, earthquake waves through the ground), and so it was thought that some sort of *ether* or interstellar medium existed in which light propagated in. The consequence of this was that the universe had a preferred frame of reference in which the ether was stationary and in which the light speed was exactly equal to the value calculated by Maxwell. When the speed of light was measured by an observer moving relative to the ether, the value would be different from Maxwell's value and could be determined by subtracting the velocity vector of the observer performing the measurement.

Many experiments were conducted in an attempt to detect the ether. The most famous is the interferometer experiment by Michelson and Morley (1887) where a series of mirrors, fixed to a platform and suspended on mercury, guided a light beam along two

orthogonal paths of several metres before interfering them together to form a fringe pattern. The Michelson's fringe pattern served as a readout of the length of time that the light took to traverse the two paths, and hence would show any relative difference in the light's speed. The idea was to conduct the experiment over six months in which time the Earth would complete a half-circle around the sun, and would hence be moving "with" the ether at some time, and "against" the ether at another time. The experiment was, to Michelson and Morley, a failure in that they did not measure the presence of any ether. However, the experiment was successful in another way as their apparatus was reportedly sensitive enough that it should have measured an ether drift if it existed. Due to this and other failed attempts to find the ether, the scientific community was forced to look for an alternative.

The solution came in 1905 from Albert Einstein, who postulated that the laws of physics be the same in all non-accelerating reference frames. Since Maxwell's electromagnetism equations constitute physical laws, it follows that the speed of light must be equal to Maxwell's value c for all observers no matter how fast they were travelling relative to one another, and regardless even of whether they were measuring the velocity of the same light beam. Moreover, Einstein's theory, now called his Special Theory of Relativity, showed that no particle can travel faster than the speed of light. To do so would see someone arrive at a time earlier than their departure, and would hence place the common-sense tenet of event causality in jeopardy.

Other scientists of the era were heading in the same direction as Einstein at the time of his discovery. Poincaré had made a similar relativistic suggestion to Einstein a year earlier but had not realised its significance. Lorentz, between 1892 and 1909, developed a theory of two ad-hoc hypotheses, *length contraction* and *time dilation*, in order to explain the failure of the ether experiments. These effects are now known to follow directly from Einstein's principle of relativity, but Lorentz is still credited with finding the mechanical transformation equations, the *Lorentz transformations*, that relate one reference frame to another.

Einstein next produced his General Theory of Relativity (1915) which overhauled the notion of a gravitational "force" in favour of *curvature* of space-time caused by heavy objects, analogous with the distortion of the surface of an inflated balloon in response to applied pressure. The concept of space-time curvature explained why all objects in a gravitational field are accelerated by the same amount, and also predicted that light would similarly bend around a gravitational field. The latter effect, now known as gravitational lensing, was experimentally verified in 1919, when Einstein was elevated to celebrity status. The theory also predicted that objects in near-field orbits would precess, and accurately explained the precession of Mercury, a mystery that had remained unsolved for centuries.

Since then, experiments to measure the time-dilation of clocks have been performed to great precision, with unanimous agreement with predictions based on Einstein's theories (taking into account *both* time-dilation effects: fast-travelling clocks run slower, and clocks run slower when closer to gravitational sources). In today's world, the use of high precision clocks and satellite communications involves making relativistic corrections in order to maintain accuracy. A prime example is the satellites of the US' Global Positioning System, whose onboard clocks constantly run at different speeds to identical
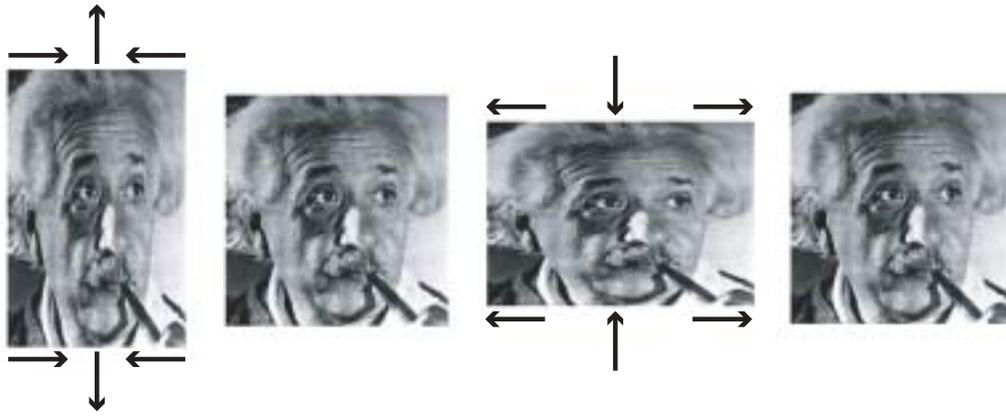
Fig. 1.1: A gravitational wave distorts spacetime by a stretching of the spacetime metric in one dimension at the same time as a compression in an orthogonal direction. A complete wave cycle of a gravitational wave includes both compressions and elongations of each axis alternately. Shown is a figurative representation of Albert Einstein as a gravitational wave (of grossly exaggerated strength) passes through.

Earth-based clocks.

1.2   Gravitational Waves

The purpose here, however, is to examine a consequence of Einstein's theory that has yet to be experimentally observed directly.

Gravity, as Newton described it, is in conflict with Einstein's theories. Newton's gravity acts instantaneously at a distance, whereas Einstein's principle of relativity outlaws this as being non-causal. General Relativity gets around this by predicting that gravitational influences will propagate at a finite speed (in fact at the speed of light c) in the form of waves. Thus, if a heavy object moves in such a way that its gravitational far-field changes, these changes do not "update" until the new gravitational influence arrives in the form of space-time ripples, known as *gravitational waves*.

A gravitational wave propagates in space itself, and is manifested as a time-oscillating length-contraction and time-dilation effect across the region it propagates. In other words, a gravitational wave stretches and squashes space as it travels through it, and the extent of this distortion corresponds to the amplitude of the gravitational wave. Gravitational waves are weak signals by nature, reflecting that spacetime is a very "stiff" medium, and the strongest wave scientists can hope to detect from a deep space source has a *strain* (relative length change) of order 1 part in 10^{21}.

While serious attempts at measuring gravitational waves directly are in progress at the time of writing, detection has not occurred yet. The closest verification of the existence of these waves comes from radio astronomical observations by Hulse and Taylor of the binary pulsar PSR1913+16. The pulsar's orbital period was observed to continually decrease at a rate consistent with relativistic predictions that the system will shed energy in the form of gravitational radiation, much as an accelerating charge will shed energy through electromagnetic radiation. Hence the experiment constituted an indirect detec-

tion of gravitational radiation, and Hulse and Taylor were awarded a Nobel prize for their work in 1993.

Gravitational waves, as with electromagnetic waves, are only qualified in the far-field, that is one must be a sufficient distance (many wavelengths) away from the source for the disturbance to "count" as a wave, or else it is just called a near-field disturbance. In the far-field of a gravitational wave emitting source, the General Relativistic metric $g_{\mu\nu}$ reduces to the Minkowski metric (flat space metric) plus a linear oscillating perturbation:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{1.1}$$

where the Minkowski metric is given by

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.2}$$

and the perturbation obeys the wave equation

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) h_{\mu\nu} = 0. \tag{1.3}$$

There is an extra requirement that the perturbation matrix be transverse and traceless[1]. The consequence of this is that, of the two axes orthogonal to the direction of propagation of the gravitational wave, the metric component of one will increase as the metric component of the other decreases. This is easily visualised as a simultaneous stretch of one axis and a compression of the other axis, as shown in Fig. 1.1.

The oscillation described here is a *quadrupole* oscillation, whose strength drops off quickly (equivalent force going inversely as the fourth power of distance) with distance from the source. This is a major reason why gravitational waves are so weak, in conjunction with the remoteness of deep space objects sufficiently large to gravitationally radiate. Monopole or dipole gravitational waves, which would produce stronger signals, are forbidden from occurring for fundamental reasons. If the "hand of god" plucked a star out of existence, ie if its mass could simply vanish, a monopole gravitational wave would result as the spacetime depression left by the star recovered to a level of lower curvature. Likewise, if the hand of god were to give a star a "push" of momentum, without some other matter gaining an equal and opposite impulse, dipole gravitational radiation would result. Both scenarios are, of course, impossible. The next strongest possibility then is quadrupole radiation, which typically requires some sort of "twist" in the motions of the emitting objects.

Quadrupole oscillations of this nature come in two polarisations; those with axes aligned cardinally, and with axes aligned diagonally. These are denoted with basis metric components $\hat{h}_+$ and $\hat{h}_\times$ respectively, such that a general gravitational wave perturbation,

---

[1] This requirement comes from the choice of expressing the perturbation in the *transverse traceless gauge* which is the simplest gauge, where coordinates are marked out by the world lines of freely-falling test masses.

propagating in some specified direction, has the form $h = a\hat{h}_+ + b\hat{h}_\times$.

1.3 The Prospect of Gravitational Astronomy

The analogy between gravitational radiation and electromagnetic radiation goes only so far. In the latter case, Heinrich Hertz was able to devise an experiment to transmit radio waves several metres across his laboratory between two giant electromagnetic coils, thus both generating and detecting electromagnetic waves (1888). Generation of gravitational waves strong enough to detect inside a laboratory is not a feasible option as terrestrial masses simply aren't big enough. In any case, typical wavelengths are hundreds or thousands of kilometers, which is the minimum distance apart that the transmitter and receiver would have to be placed at to qualify for a gravitational wave experiment. The only gravitational waves humans will ever detect will come from the heavens.

The flip side of this is that, once detection is achieved (now only a matter of time, provided Einstein was right!), gravitational waves will become an entirely new branch of astronomy. By its nature, gravitational astronomy will focus on supermassive compact objects such as neutron stars and black holes; what astronomers couldn't "see" in space, they may be able to "hear" via gravitational waves. Earth-based gravitational wave detectors are likely to see signals in the frequency range of a few Hz to a few kHz. This is roughly the same range of frequencies of sound waves that the human ear can hear, and so gravitational wave detection has an acoustic flavour. Some gravitational wave scientists "listen" to the output of their detectors.

Potential sources generally fall into one of two categories: either they emit gravitational waves for a short burst (milliseconds or, at the most, seconds), or else they emit for a long time, or at least much longer than gravitational wave experiments run continuously for.

The archetypal gravitational wave source is the neutron star binary. The steady orbit of such a binary emits gravitational waves, but typically at very low frequencies (PSR1913+16 emits at a period of about 4 hours). The frequency increases very slowly as the system radiates gravitational energy, and the interesting part comes after millions of years when the two stars in such a binary "spin up" and coalesce into a single star. The resulting chirp waveform sweeps right through the frequency range of terrestrial detectors, albeit only for a few milliseconds. These waveforms are easy to *spot* due to their distinctly recognisable waveform[2], but the drawback is that binary coalescences are few and far between, with much less than one event per year occurring within the seeing distance of first generation detectors.

Another potential "burst" source is the collapse of the iron core of an exploding supernova of the Type II kind. History shows that one or two visible supernovae occur each century, the most recent being Supernova 1987A. This source has the advantage that data can be correlated with other astronomical detections (optical and neutrino for example), to affirm detection confidence. However, core collapse inside a supernova is not well modelled, and it is not clear how strong the emitted gravitational waves will be. Again,

---

[2] A technique called *matched filtering* can improve sensitivity by orders of magnitude, provided one knows the waveform ahead of time.

the low event rates complicate matters; the nearest galaxy cluster, the *Virgo cluster*, may sport a few detectable supernovae per year if scientists are lucky.

Continuous sources include single pulsars, where asymmetries in the mass distribution of the rapidly spinning compact objects can radiate gravitational waves. Asymmetries caused by a pulsar's dipole magnetic field will produce weak gravitational waves, perhaps ten orders of magnitude below first generation detector noise floors, but an instability mechanism proposed by Wagoner may cause pulsars to produce waves stronger by a few orders of magnitude. The continuousness of the waves, and the fact that rotating pulsars generate highly stable frequencies, enables us to push down the noise floor by averaging data over long periods. Some future interferometric detectors will be operated in *narrow-band* mode with corresponding resonant enhancement of sensitivity. These factors, and the fact that many such pulsars may be found in our own galaxy, make this a promising source.

Also worth a mention is the possibility that spacetime anisotropies shortly after the big bang caused a burst of gravitational waves that still resonate today as a *stochastic background*. Analogous with the cosmic microwave background, this could give us new information about the big bang, but effects like cosmic inflation are required in order for a gravitational wave stochastic background to be loud enough to observe today.

Perhaps most exciting is the prospect of detecting objects not before seen by other methods of astronomy. Studies of dark matter show that the matter of the universe that astronomers can see makes up only a small fraction. There are also several parts of the sky, such as the galactic centre, that are obscured from vision by dust clouds, whereas gravitational waves will pass right through more or less any matter.

The first gravitational wave detections should occur before the end of the decade. Indeed detection evidence presented recently by two of the bar detector groups [5] is a sign of the anticipation held by the international gravitational wave research community that the first bankable detections are imminent. On the other hand, true gravitational wave astronomy, involving many and frequent detections, is still perhaps a generation away.

Gravitational wave astronomy will involve a world-wide collaboration. With low signal-to-noise ratios and infrequent events, the random chance of noise impersonating a signal becomes significant. Comparison of data from several continentally-separated detectors will significantly improve detection confidence, as well as providing a measure of immunity to noise sources local to a specific detector.

Gravitational wave detectors are equally sensitive to waves approaching from virtually every direction, bar a couple of zeros. Also, incoming waves cannot be focused as with a radio telescope dish. Detectors therefore "see everywhere at once", and determining the source and polarisation of the incoming wave requires signal amplitude and timing comparisons between multiple detectors. As with radio telescope arrays, the longer the *baseline* between detectors, the better the ability to triangulate the source direction, as well as the ability to triangulate and exclude terrestrial noise sources. However, since it takes a matter of milliseconds for a gravitational wave to propagate from one side of the Earth to the other, a very high precision of time stamping and synchronisation between detectors is essential.

1.4 Experimental Gravitational Wave Detection

Although Einstein's General Relativistic theory predicted the existence of gravitational waves, he was skeptical that they would ever be detected. The first attempts were not until the 1960s by Joseph Weber, who attempted to measure vibrations of massive aluminium cylinders. While he claimed success in 1969, attempts to repeat the measurement failed, and it was agreed that the technology of the day was not sensitive enough.

The descendants of Weber's original experiment are called resonant mass detectors or *bar* detectors, and today's technology is greatly refined. Moreover, interferometric detectors, similar to Michelson's and Morley's apparatus, are reaching science-quality sensitivity and hold great if not certain promise. The following is a discussion of how these detectors work, including an outline of the associated noise suppression challenges, and a catalogue of current detector technology.

1.4.1 Principle of Detection: Interferometry versus Resonant Mass

The stretching and squashing of spacetime caused by a gravitational wave quite literally means that the proper length between two points oscillates, with a similar effect on the proper time elapsing in the affected area. The image shown in Fig. 1.1 is an exaggeration in that one would not "see" the room stretching and shrinking (and not only because the effect is very very small), since the distortion of space-time changes how the light photons would travel to the viewer's eyes[3]. However, measuring the effect these spacetime distortions have on photons yields information about the time varying space-time curvature caused by a gravitational wave. The photon's travel time, rather than its trajectory deviation, is easiest to measure precisely. This is the basis of light interferometric detectors, which compare the relative travel times of two orthogonally propagating light beams down to the phase precision of a partial wave cycle.

To make a valid measurement of the proper length of a section of space using photons, each component of one's measuring equipment must be free to "stay put" at its respective location. To stay put in this case means to be in free-fall, since the metric distortion $h_{\mu\nu}$ discussed above is valid with respect to coordinates defined by freely-falling objects. The free-fall requirement means that no forces other than gravitational are acting on the object in question. The simplest such scenario is an experiment consisting of two space probes, one with a laser and photodetector, and the other with a mirror, as in Fig. 1.2. The return trip time of a photon travelling from the laser to the mirror and back to the detector then gives a measure of the oscillating proper length between the probes. Interferometric detectors are usually arranged such that the time one photon spends travelling is much less than the period of the wave, so that the accumulated phase lag is all of one sign and does not cancel with itself.

The other approach to gravitational wave detection is to observe its effect on extended objects. If a metal bar is present in a part of space through which a gravitational wave passes, then each section of the bar will initially follow the (free-falling) coordinates of

---

[3] An observer in a room with a passing gravitational wave that is sufficiently strong would see *something*. Therein lies a challenge for numerical relativity to determine exactly what.

(a)

(b)
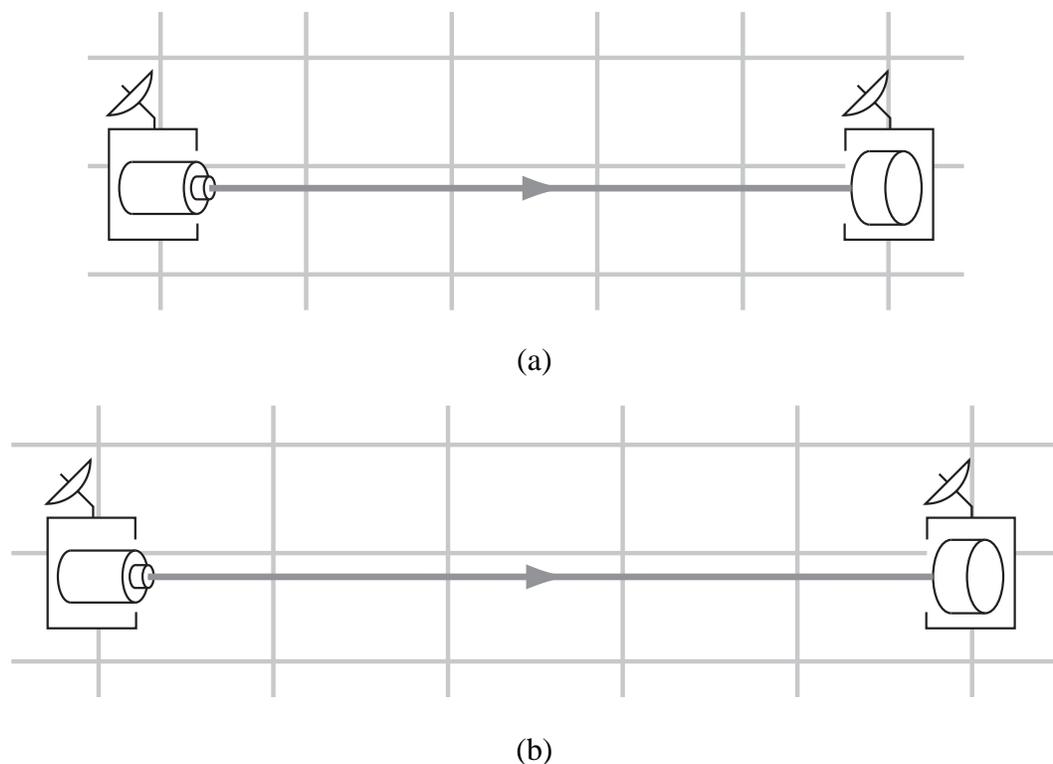
Fig. 1.2: (a) Two space probes measure their intervening distance with a laser and mirror system. (b) The two space probes under the influence of a gravitational wave (horizontal-dimension only for simplicity) follow their free-fall coordinates. The probes will now measure a longer travel time of the light photons that move between them.

spacetime as it stretches, as in Fig. 1.3. In other words, the proper length of the bar will increase with the proper length of the spacetime interval it occupies. However, the crystal lattice bonds holding the bar's atoms together are based on the same electromagnetic laws that were previously used to show that the speed of light was constant in all reference frames. As such, the lattice bonds respond with a restoring force that acts against the increase in proper length of the bar. An observer standing at one end of the bar would observe that end to shrink away from its default position[4]. As the phase of the gravitational wave changes from stretching to shrinking space-time, so the bar's response reverses, albeit with a phase lag since the crystal lattice forces take time to respond. If the bar has a mechanical resonance at the frequency in question, it will resonate in response to the gravitational wave and amplify the signal. Bar detectors work on this principle, and involve ultra-sensitive transducers being placed at the ends of high purity bars. Bar detectors measure at a single resonant frequency, usually of several hundred Hertz.

Ground-based interferometric detectors work on the same principle as space-based interferometric detectors, but arranging for the equipment to be in free-fall is much more difficult. The Earth itself can be thought of as a giant bar detector, except for being far too complicated to model usefully as a bar detector. Mirrors, which are given the label

---

[4] It is interesting to note that a stretch in space-time would cause objects to shrink rather than stretch in response.

(a)

(b)
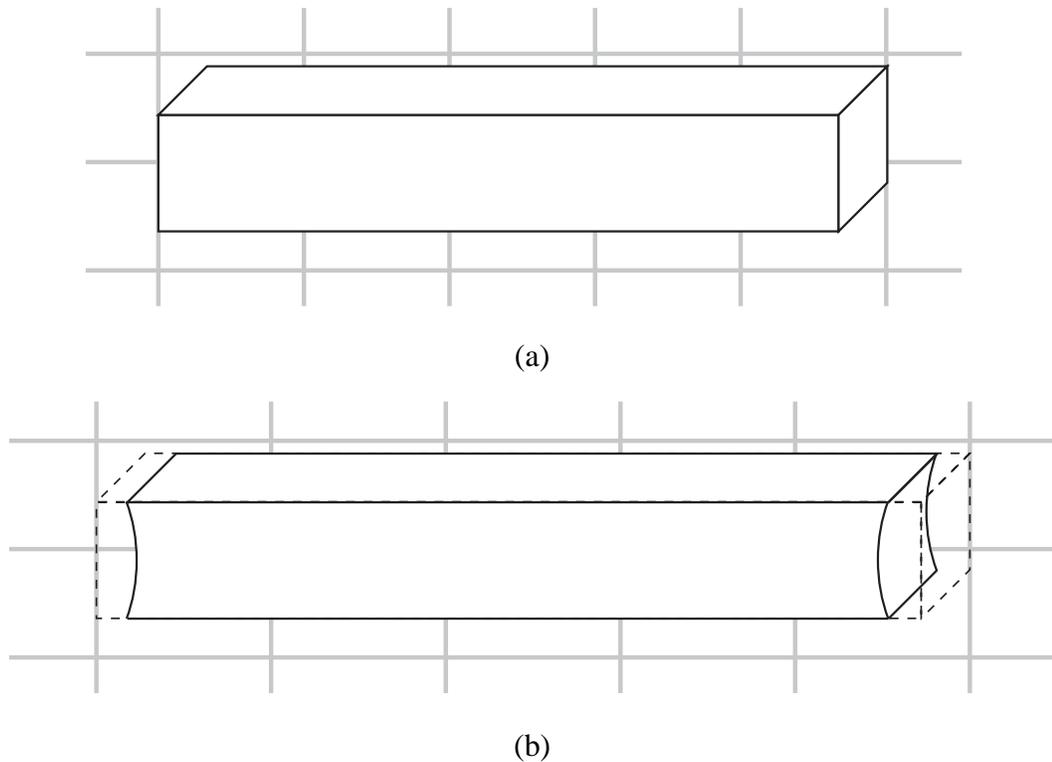
Fig. 1.3: (a) An extended object or bar shown against a background grid of coordinates defined by world lines of freely-falling objects. (b) The same object under the influence of a gravitational wave (horizontal-dimension only for simplicity) recoils from the metric stretch. The internal forces of the bar "correcting" for the metric change causes a measurable length contraction, and an oscillating metric strain hence causes an oscillating length change of the bar.

test masses due to the general relativistic context, must be suspended and isolated from the ground. While this is not technically free-fall, since the test masses are still prevented from falling to the Earth, it can be arranged so that the masses have no forces on them in the *horizontal* direction that the probe light travels in. The result then of a passing gravitational wave on a ground based detector is shown in Fig. 1.4. The test masses are observed to literally swing, though it is really the ground and laboratory, not the mirrors, that are being accelerated by non-gravitational forces[5].

This analysis has ignored the other half of the interferometric detector, dealing with only one arm of a Michelson interferometer. Actually, a second beam path is set up perpendicularly to the first, using a beam-splitter to divide the light into equal halves (classically speaking), and also to recombine the beams so that they interfere on a photodetector. Hence, the change in *relative* time travel is reported by the changing fringe condition of the Michelson. The absolute time of flight, or even the difference in time of flight between one arm and the other isn't available. However, as we're looking for such a small, time-varying fluctuation, this hardly matters. The fringe condition of the Michelson tells us the relative phase shift between the two interfering beams, which is a little like having a photo finish on the sub-wavelength scale. Interferometers, unlike bar

---

[5] ...if indeed one can make such *absolute* statements in the context of relativity!

(a)

(b)
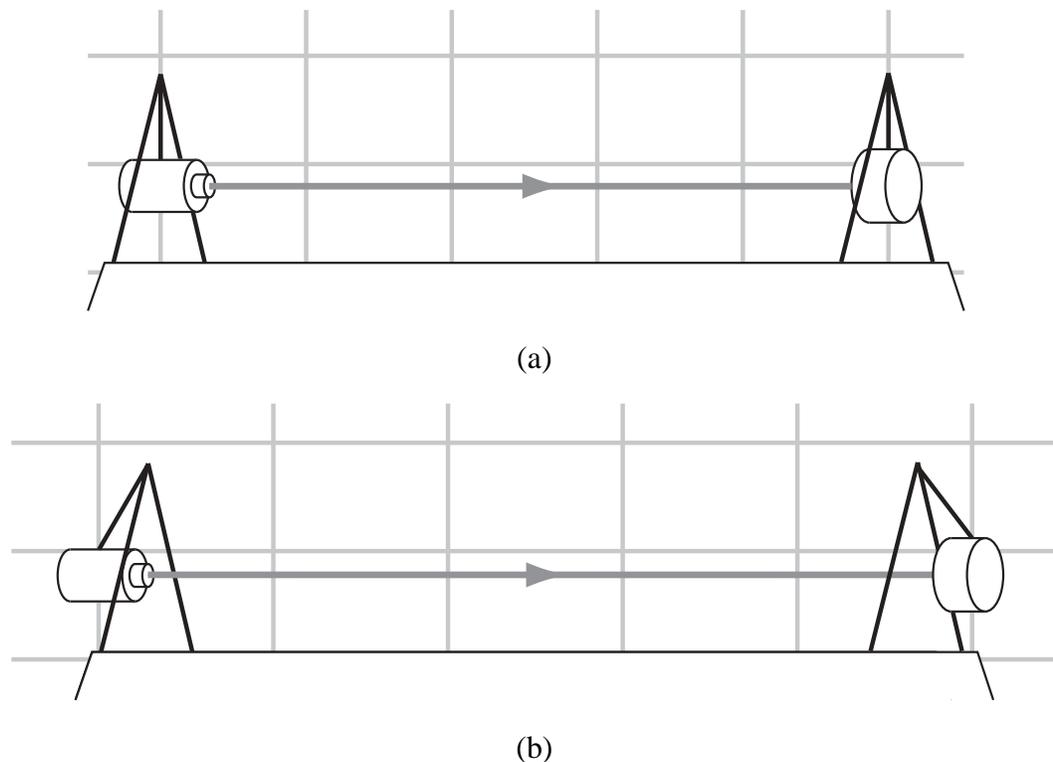
Fig. 1.4: (a) Optical equipment suspended from the ground in a ground-based detector. (b) As the gravitational wave passes through, the optical equipment will follow free-fall geodesics if it is properly suspended and isolated from ground motions, which will be induced in a manner similar to those of a bar detector. Hence, the principle of detection here is the same as for space-borne detectors. The tilt effect of a displaced pendulum has been ignored, but this is negligible for the minute strains produced by gravitational waves.

detectors, can detect gravitational waves at all frequencies where noise can be overcome; a frequency range of two or three orders of magnitude is typical.

A stripped-down gravitational Michelson interferometer responding to a gravitational wave incident from above is shown in Fig. 1.5. The quadrupole nature of gravitational waves couple perfectly into the differential mode of a Michelson interferometer. In other words, the quadrupole distortion lengthens one arm's optical path as it shortens the other, and so the measured change in fringe condition is twice what it would be from only one arm.

Of course, a gravitational interferometer such as this is maximally sensitive to waves incident perpendicular to the two arms, but still has significant sensitivity from most angles. There are a few specific angle/polarisation combinations where the path length change in the two arms are identical (or zero) and hence the detector is insensitive at those angles.

The above has dealt with the principle of how the detectors, but not the problem of how to make them sensitive enough to detect strains of better than one part in 10^{21}. The obvious place to start is to make the interferometers as long as possible, so that the absolute path length change is maximal. The absolute upper limit on length is set by the wavelength of the target gravitational waves since, if the light spends any longer than half
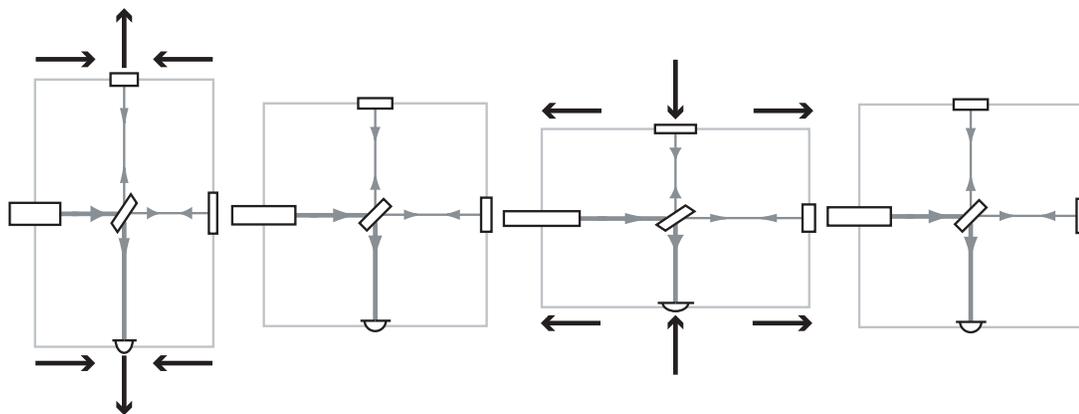
Fig. 1.5: A Michelson interferometer whose optical components are in free-fall responds with full effect to this gravitational wave that approaches from above (out of the page) with polarisation lined up with the axes of the Michelson.

a period in the device, the gravitational phase shift accrued will start to cancel itself out. At gravitational wave frequencies of 100 Hz, the optimum length of interferometer arms is of order 1000 km.

Obviously, engineers are limited in practical terms by the "3 mile limit" across the Earth's spherical surface when the land begins to fall away significantly. Current-day terrestrial detectors are no longer in arm length than 4 km. The light will therefore spend a time inside the interferometer much shorter than the gravitational wave period, accruing a correspondingly smaller time shift, and so the device is less sensitive than it might otherwise be. As such, there is the need to make up the many orders of magnitude in sensitivity in other ways. The subject of the next section is the ingenious technological advances of the past three decades that have led to the advanced Michelson interferometers of the present day.

1.4.2 Advanced Interferometer Configurations

Enhancements to interferometer sensitivity have come by adding in extra components to the basic Michelson. Fig. 1.6 shows a gravitational wave detection interferometer schematic with all its optical "inclusions". Three of these inclusions are established as part of the best case technology, and the fourth is likely to be a future inclusion pending further research. Each is examined briefly in turn.

The oldest improvement came in the 1970s as a result of the fact that orders of magnitude of sensitivity were being lost because Michelson arms could not be built longer than a few kilometers. The first idea was to fold the optical path over on itself many times, using multiple mirrors to reflect the beam back and forth before finally sending it back along its path. A refinement on the idea, the *Herriott delay line*, used a pair of single mirrors with a hole cut on one side of the entry mirror. The two mirrors were arranged such that the beam bounced multiple times on each mirror, with spots arranged in a circle, until the beam returned to exit the hole through which it entered. Ultimately, the delay line idea was abandoned as being infeasible: multiple mirrors involve too many extra degrees of
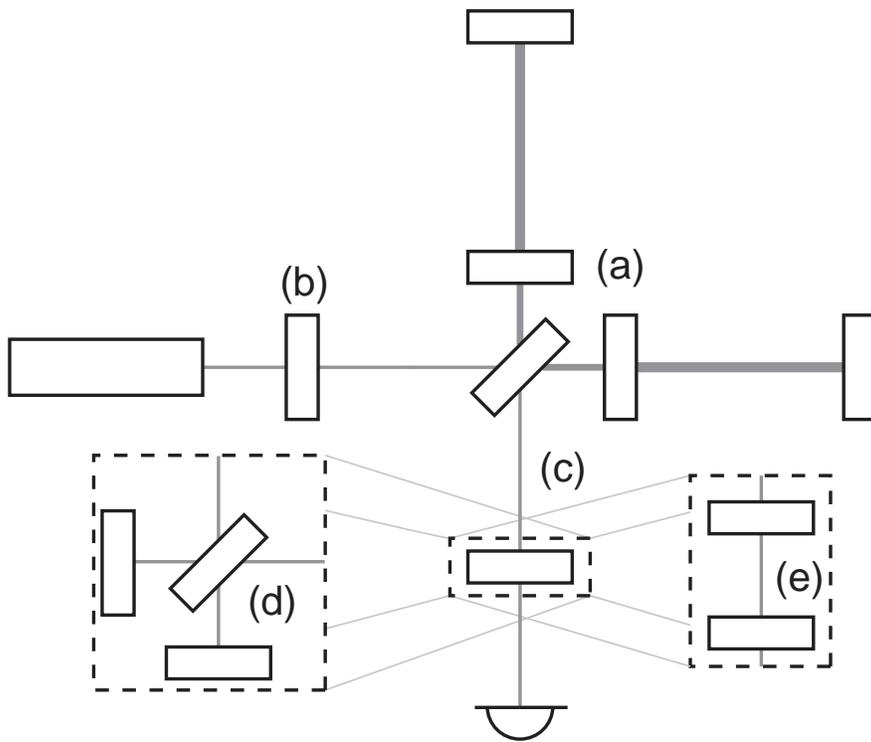
Fig. 1.6: Advanced gravitational wave detector longitudinal configuration. Shown are: (a) arm cavity input mirrors; (b) input power-cavity mirror; (c) output signal-cavity mirror; (d) Michelson-type variable reflectivity mirror to replace signal mirror; (e) cavity-type variable reflectivity mirror to replace signal mirror.

alignment, and the Herriott delay line requires too big a mirror surface in order to keep beam spots separate for avoiding crosstalk and diffraction problems.

The currently accepted alternative is to reflect the beam in each arm multiple times along an exactly coincident path between two mirrors. The resulting configuration is called a *Fabry-Perot cavity*, and the incident light *resonates* between the mirrors for much longer than the time of a single round trip. There is an extra requirement that an exact number of half-wavelengths of the probe light be equal to the mirror spacing, in order that the light constructively interferes, but since virtually monochromatic lasers are used, this is no barrier. Rather than crosstalk being a problem, the constructive and destructive interference of the beam meeting itself after one or more round trips, in and out of the cavity, is part and parcel of the operation of this improvement, known as Fabry-Perot *arm cavities*. Fabry-Perot cavities are discussed further in Chapter 4, albeit not in this context.

The next improvement came in 1981, suggested independently by Drever and Schilling [6]. A mirror placed between the laser and beamsplitter creates another cavity-like configuration, this time to "recycle" the light escaping back to the laser which would otherwise be wasted. So called *power recycling* is possible because interferometers are currently operated with virtually all of the light in the Michelson returning to the laser. The effect of the *power cavity* is to increase the intensity of light into the interferometer, well beyond the power that the laser can supply, and the improvement to gravitational wave sensitivity comes because the same gravitational-wave-induced phase shift produces

a bigger intensity change at the Michelson output.

Along similar lines again, Meers proposed in 1988 [7, 8] to add a mirror at the output port of the Michelson to reflect any light exiting the Michelson (carrying the signal) back into the Michelson for further Fabry-Perot-like resonance and constructive interference. So called *signal recycling* can once again enhance the signal, but over a select range of frequencies which constructively interfere inside the *signal cavity*. Typical configurations allow many gravitational wave signal frequencies centered around 0 Hz to resonate in *broadband* operation, or a narrower selection at a shifted centre frequency to resonate strongly in *narrowband* mode.

A variation on signal recycling called *resonant sideband extraction* (Mizuno et al 1993 [9]) uses the same configuration, but different operational parameters such as lengths and mirror reflectivities. The arm cavities are configured to over-resonate the light such that it spends "too long" accruing gravitational wave induced phase shift to the point where the effect will start to cancel. The signal mirror then plays a slightly different role, actually extracting the signal rather than returning it to the arm cavities. This is done by setting the signal mirror position so that it forms a resonance with the arm cavity input mirrors for a given gravitational wave signal frequency. The combination of the input arm cavity mirrors and signal mirror then become transparent at that frequency, and the signal is transmitted straight out the dark port. This option allows for higher storage times of the probe light in the arm cavities, and hence higher resonating powers which would otherwise require more power from input lasers.

The last improvement is an extension of signal recycling, but where more control is granted over the frequency response of the interferometer. To coin the term used by de Vine in 2001 [10], a Variable Reflectivity Signal Mirror (VRSM) can substitute for the mirror at the dark port of a signal recycling configuration. The VRSM is really an optical subsystem; either another Fabry-Perot cavity (as demonstrated by Strain and Hough in 1994 [11]), or another Michelson inteferometer (from de Vine and Shaddock et al 2002 [12]). Both options afford an extra degree of freedom that changes the proportion of light reflected back into the interferometer. This improvement therefore does not increase the sensitivity per se, but does provide more flexibility for tuning the frequency response of the gravitational wave detector, including the potential for real-time tuning. The VRSM idea forms a large part of the motivation of this thesis.

1.4.3   The Engineering Challenge of Noise Suppression

If one could build a hypothetical noise-free instrument, then the improvements in sensitivity outlined in the previous section would be irrelevant; one would just turn up the gain of the output signal until they saw something. But, as with all experimental instruments, there are a myriad of uncontrolled processes that produce unwanted signals which compete with the target signal. It is customary in gravitational waves to express the magnitude of these noise sources in terms of their equivalent strain h, ie the amount of gravitational-wave-induced strain that would produce a signal at the output of equal strength to that produced, through whatever mechanism, by the noise source. The following is a catalogue of the noise sources relevant to ground-based detectors, and the technology being used or developed in order to overcome each noise source.

At low frequencies (below about 50 Hz), seismic vibrations dominate the output channels of a ground-based detector. Test mass mirrors, and indeed all of the optics, are suspended by stacks (piles of alternating layers of material) and elaborate multi-stage pendulums to attenuate the transmitted ground motion by many orders of magnitude. The choice of site, with respect to both natural and man-made seismic disturbances, comes into play significantly here. Strong seismic bursts or earthquakes can knock a detector away from its operating point, but ought not be mistaken for gravitational waves as their much slower propagation speed gives them away when multiple detectors are in use.

The test masses themselves, being at temperatures well above zero Kelvin, are never completely at rest due to internal vibrations. So called *thermal noise* causes a small random fluctuations in the position of the test mass mirrors and, given the extreme sensitivity required, this is a significant noise source. Choosing high quality materials (mirrors of sapphire or fused silica, and similarly for suspension wires) can simplify the signature of noise source to an extent. Thermal noise of the suspension systems is most significant at middle frequencies (low hundreds Hertz).

A major fundamental noise source of quantum origin is the discreteness of light photons, which manifests in two ways. Photon *shot noise* is the phenomenon of power fluctuations at the photodetector due to the Poisson-statistical nature of photon arrivals. Shot noise scales inversely with beam power; a higher number of photons arriving per time interval can better mimic a continuous classical light wave. On the other hand, if the light power is too high, *radiation pressure* on the test masses can transfer power fluctuations in the arms into physical path length changes. This is particularly important due to the antisymmetric nature of the Michelson's beamsplitter, which sends a photon to one arm or the other but not both[6], and hence this effect is magnified when light powers are high enough to push the test masses significantly. Radiation pressure noise is more significant at low frequencies but is typically less significant than seismic or thermal noise at those frequencies. Shot noise dominates a detector's output at frequencies above a few hundred Hertz.

Detector beam paths and all optics are enclosed in vacuum tubes and tanks, and so the noise from residual gas is below noise from other competing sources. High power lasers in second generation detectors will be strong enough to heat optical substrates so that thermal lensing effects will become significant; annular compensation heating coils is among the ideas being developed to counter this. The lasers themselves are passed through "mode-cleaner" Fabry-Perot cavities to remove amplitude and frequency noise, as well as to mask unwanted higher order modes.

Elaborate feedback control systems are employed for active damping of oscillations in suspension stacks and pendulums. The interferometer itself is the subject of a multi-dimensional longitudinal feedback control system to keep cavities on resonance and to keep the main Michelson fixed at the correct operating point. Lateral feedback control systems, known as *autoalignment* systems, keep the mirror-beam system in optical alignment. Controlling optical configurations through feedback loops is one of the main

---

[6] This isn't quite true as, without a measurement, the photon isn't definitely in either arm, but a quantum analysis does show that the noise introduced by the beamsplitter is antisymmetric with respect to the two arms
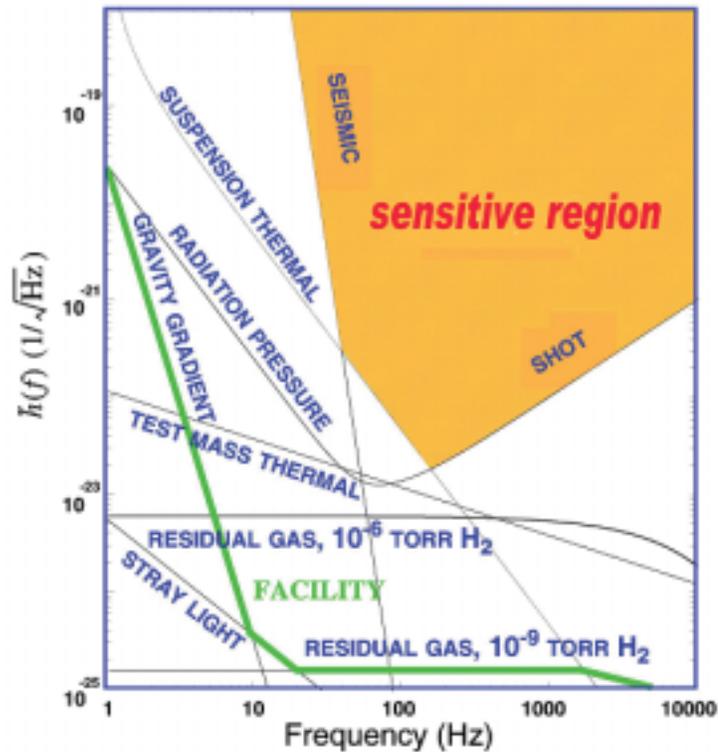
Fig. 1.7: Design sensitivity curve of the initial LIGO project, with noise floor contributions individually marked as shown. The vertical scale gives the amplitude spectral density of each effect's noise-equivalent gravitational strain. From [13].

themes of this thesis, so an entire section is devoted to basic concepts in Chapter 2.

Cables and circuit boards are electrically shielded to the extent that practicality allows. In any event, arrays of physical-environment-monitoring sensors are operated to keep track of any terrestrial disturbances (natural or man-made) that could couple in to the output signal via vibration, electrical inductance, or any other means. These channels are compared to the main interferometer channel, and sections of data are either conditioned (a knowledge of the noise source is used to remove it from the data) or vetoed as appropriate.

This list is not exhaustive, but gives an idea of the lengths scientists and engineers must go to in order to give their instrument a reasonable chance of detection.

1.4.4   Gravitational Wave Detectors of the World

Different research groups pay different amounts of attention to each engineering challenge, and are at different stages of development. The following is a snapshot of detector technology as it stands today in 2004.

The interferometric detector to be commissioned earliest[7] was TAMA300 in Tokyo,

---

[7] One can't entirely disqualify the suspended prototype interferometers from the status of "detector", since the 10 m Glasgow, 40 m CalTech, 30 m Garching, and 100 m Tenko prototypes have all been used to

Japan, which is a power-recycled Michelson with arm cavities and 300 m long arms. The TAMA group have recorded over 2000 hours worth of data since 1999, with a peak sensitivity of $h \sim 3 \times 10^{-21}$ Hz $^{-1/2}$ at 1.3 kHz [14], and a seeing distance of tens of kiloParsecs (kPc) for standard candle binary inspiral sources. TAMA succeeded in the face of the extra challenge of overcoming local seismic disturbances from the high population density of Tokyo, but were somewhat underprepared when it came time to analyse their data. The TAMA collaboration did not record a successful detection during their data taking, but did exchanged data with the American LIGO group (see below) in 2003. The group has forward plans for a long interferometer, the Laser Cryogenic Gravitational Telescope (LCGT), which will use cryogenic cooling of test masses to overcome thermal noise.

The Laser Interferometric Gravitational Observatory (LIGO) group was formed in the US by the National Science Foundation around 1980, and has developed three detectors on two sites at Hanford, Washington state, and Livingstone, Louisiana. The detectors are power recycled Michelsons with arm cavities, with both sites installed with 4 km instruments, and Hanford having a nested 2 km instrument. LIGO data taking runs, called Science runs, began in 2002, and continue to the present day with the Hanford 4 km instrument sensing at $h \sim 8 \times 10^{-23}$ Hz $^{-1/2}$ at 100 Hz, corresponding to a seeing distance of 6.5 MPc for standard candle binary inspirals (recalling that the Virgo cluster is a little over twice this distance away) [13]. The LIGO detectors are currently within an order of magnitude of their design sensitivity, and great hopes are held that these instruments will make a confirmed detection within the next few years. One of LIGO's strengths is in its attention to the problem of making sense of the output data, by developing an extensive, automated, conditioning, vetoing, filtering and data mining software package called the LIGO Data Analysis System (LDAS). LIGO's Achilles' heel has been its basic single-stage suspensions which have made it difficult to overcome seismic noise, although recent "fixes" seem to be helping [15]. LIGO plan to decommission and begin upgrading their detectors in 2007 in preparation for Advanced LIGO (AdvLIGO), which will include signal recycling as well as higher power lasers [16]. The technology presented in this thesis is aimed at incorporation into Advanced LIGO.

GEO is a cooperation between research groups at the Max Planck Institute in Garching, Germany, and at the University of Glasgow, Scotland, who have built a detector south of Hannover, Germany. The GEO 600 detector has 600 m folded arms (for 1200 m effective length) and power and signal recycling, together called *dual recycling* [17]. During their first science-quality data taking runs in 2003, GEO shared data with LIGO, and the instrument is currently sensitive at $h \sim 2 \times 10^{-21}$ Hz $^{-1/2}$ at 1000 Hz [18]. It is doubtful whether or not GEO can be made sensitive enough to detect gravitational waves with certainty, though its importance as the first detector-scale device with signal recycling should not be overlooked.

Another European collaboration of French and Italian groups have built a detector named Virgo, after the star cluster it aims to observe, at Cascina, Italy. The Virgo detector

---

set "upper limits" on gravitational wave events. The practise of setting an upper limit, where one calculates the magnitude of a gravitational wave that is just barely strong enough to be detected by the instrument in question, is all that any of the large scale interferometers have achieved so far.

is a power recycled Michelson with arm cavities of length 3 km, with recent test-phase sensitivity measurements being recorded at $h \sim 3 \times 10^{-20}$ Hz $^{-1/2}$ at 1000 Hz [19, 20]. The Virgo detector sports highly advanced multi-stage suspension systems, and holds promise as it nears science quality sensitivity.

The frontier of detection technology in the 1990s was the domain of the resonant mass (bar) detector groups. Allegro (Baton Rouge, Louisiana, US), Auriga (Padova, Italy), Explorer (CERN, Geneva, Switzerland), Nautilus (Rome, Italy) and Niobe (Perth, Western Australia) reached their peak sensitivities within the last ten years [21]. Coincidence studies stimulated much discussion on the best way to combine results from multiple detectors. Recent interest in frequencies between 1 kHz and 10 kHz has been generated by, amongst other factors, a proposed Brazilian spherical resonant mass detector called Minigrail [22].

Last but not least, and perhaps the biggest hope for certain detection, is the space-borne detector LISA (Laser Interferometric Space Antenna), which consists of three space probes in a heliocentric orbit exchanging laser signals in an equilateral Michelson configuration at a separation of 16 light seconds. The system is designed to be sensitive to low frequencies of between 10^{-1} and 10^{-3} Hz, with the lower limit nominally being set by a "confusion noise" generated by multiple binary star gravitational wave signals! If design sensitivity is met, signals are expected to be orders of magnitude above noise floors [23]. Research and development by NASA and ESA continues, with scheduled launch dates early in the next decade.

1.5   Thesis Summary

In this chapter, a broad perspective of the gravitational wave detection field has been maintained, but from here on the focus will narrow to the optical subsystems of a detector specific to the results of this thesis.

In Chapter 2, the literature review is continued with an introduction to feedback control theory, and a discussion of the purpose of locking in a gravitational wave detector.

Chapter 3 is a treatise on electro-optic modulation, covering simple and combined modulations, and devices to generate and measure modulations, and is presented in preparation for the modulation locking schemes that follow in the thesis. A new visual and mathematical representation for modulation, the *Modulation Sphere* formulation, is presented, and the formalism is demonstrated in a description of Michelson interferometer modulation locking via the Schnupp asymmetry.

Chapter 4 covers Fabry-Perot cavity optics and locking techniques, ranging from the standard Pound-Drever-Hall technique through to the newest spatial mode technique called *flip locking*. The techniques in this chapter are discussed particularly with respect to mode-cleaner cavities placed before or after a gravitational wave interferometer configuration.

A device called the Quadrature Variable Modulator (QVM) is introduced in Chapter 5 and shown to be capable of producing all possible modulation states of a given frequency, both in the standard modulation picture and in the modulation sphere description. Properties of the device are discussed, and conclusions are drawn about how best to operate

the device.

The QVM would be essential for the locking technique variations described in Chapter 6, which require a tunable modulation source for tuning lock points of cavity- or Michelson-like degrees of freedom in a gravitational wave detector. The aim, to tune the frequency response of a gravitational wave detector, is described. A theoretical demonstration of tune-locking of simple one-degree-of-freedom systems is presented as a first step towards realising this goal.

Finally, a QVM prototype is experimentally scrutinised in Chapter 7. The prototype is an in-house modification of an existing commercial modulator and exhibits high accuracy, predictability and stability in selecting modulation states. Technical difficulties with the device are scouted and dealt with.

The main outcomes of the research in this thesis are summarised in Chapter 8 along with a list of extensions to the work that could be carried out in future.

When Noah sailed the oceans blue,
He had his troubles, just like you.
For forty days he sailed his Ark,
Before he found a place to park.

— Anon (on the wall in the ANU Parking Office)

Feedback Control of Gravitational Wave Interferometers

A requirement of conducting interferometric gravitational wave detection is that the detection instrument be kept fixed at a specific operating point. This is of vital importance to the success of the instrument: the signal presents in the form of linear deviations from the operating point, and avoids the prohibitively-complicated devolvement which would be required for interpreting a drifting detector output, amongst other fatal problems. This chapter explains the details of the design choice to *lock* using feedback control, continues with a discussion of two significant operating points for a gravitational wave interferometer, and concludes with the frequency response of the full gravitational wave detector configuration.

Before proceeding, a short introduction follows as to the means whereby a detector is fixed to an operating point. A comprehensive treatment of *feedback control theory*, particularly the design of electronic filters as part of servomechanisms, is beyond the scope of this thesis. The reader is advised to consult the excellent text by Abramovici et al [24], from which the following section (2.1) is summarised. [2] is used as a reference for the entire chapter.

2.1 Feedback Control Theory: Outline

The task arises frequently in engineering applications to have a system "track" a signal input by the user. A simple example is the cruise control system of a modern car; the driver sets the desired road speed, and the car's cruise control system applies the brake or accelerator as required to maintain that speed. Apart from the car itself, the key elements of this system are: a component to measure the current speed of the car, a component that has direct access to the brake and accelerator functions, and a component that knows how to apply the brake or accelerator to best effect given the speed measurement. This is an example of a *feedback control system*, a term which encapsulates the process of taking a measurement of the target system's state, and using this information in a feedback loop to force the system toward a particular state.

The key elements of a feedback control system are generically named as follows (with cruise control equivalent): the *plant* is the system to be controlled (the car itself), the *sensor* measures some property of the system's state (the speedometer), the *filter* interprets the sensor measurement and converts it into an appropriate feedback signal (cruise con-
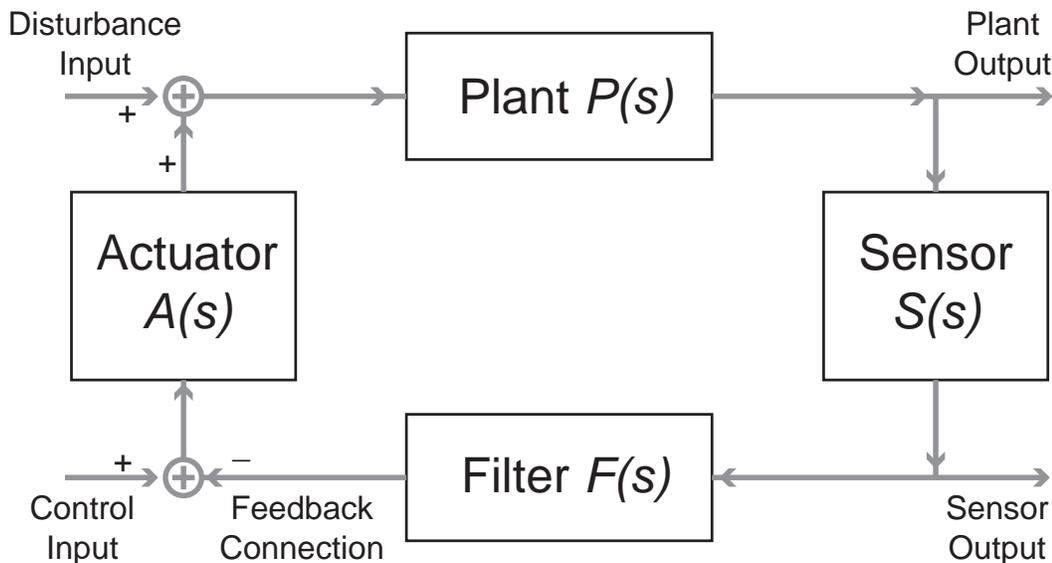
Fig. 2.1: Block schematic of a feedback control loop.

trol electronics that communicate between the speedometer and the accelerator/brake), and the *actuator* applies the correcting influence on the plant in accordance with the feedback signal (the accelerator/brake).

Fig. 2.1 is a feedback control schematic showing these components, where each component has an associated Laplace frequency transfer function[1] that represents the change applied to the incoming information to produce the outgoing information. Also labelled are key points in the information flow. The *plant output* is simply the state achieved by the plant (the car travels at some speed). The *sensor output* is the value of a state parameter as measured by the sensor (the car's speed). The *control input* is the entry point for the user's command (the cruise control speed selected by the driver). The subtraction between the control input and the filter output produces what is called an *error signal*, as it gives a measure of how far the target system's state is from the desired operating point. The *disturbance input* is where all the trouble is caused: some uncontrolled influence forces the plant away from the desired operating point (the car goes up or down a hill and tends to change speed).

The error signal is often defined instead as being the same as the sensor output, before the filter rather than after. This is common in applications where there is zero control input so that there is no distinction between the cases, bar the filter's frequency response.

When designing a feedback control system, part of the challenge is to find a measurable that produces a "good" error signal, as in Fig. 2.2. A good error signal should be antisymmetric and should have a zero crossing. The zero crossing should occur when the plant is at the desired state, and hence no feedback influence will be exerted at that

---

[1] Laplace transforms are more general than Fourier transforms for describing transfer functions of linear systems. The Laplace variable $s = \sigma + i\omega$ contains the frequency component ω which corresponds with the steady state response of a system, as functions like $exp(i\omega)$ are oscillatory. The additional σ corresponds to the transient response which, for a linear system, comprises decaying (or growing) exponentials $exp(\sigma)$.
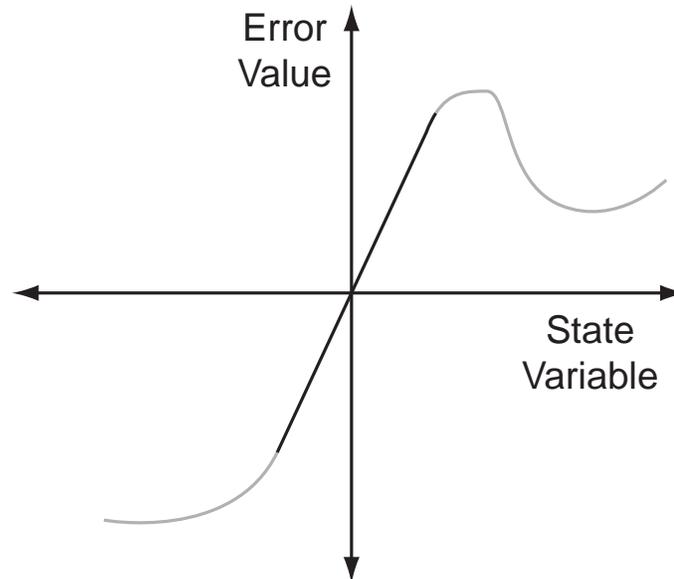
Fig. 2.2: However else an error signal behaves, it should accurately report the error of the state parameter being locked within a reasonable range about the lock point.

point. The error signal should have a strong linear slope (ie antisymmetry about the zero crossing) with respect to the state variable being locked, since then the feedback force will, when connected with the appropriate polarity, push the system back toward the lock state from either side.

Error signals typically have multiple zero crossings, often in repeating sequences as seen when the locking parameter is scanned. The *polarity* of a zero crossing, ie whether the error signal has a positive or negative slope, determines whether a feedback loop will force the system toward the zero crossing or away from it. In practise, a simple change of the polarity of the feedback connection turns a repulsive point into an attractive point or vice versa. Some systems allow locking to two different types of point, where a change in feedback polarity can be used to switch between one and the other.

A system is within "range" of a locking point (that is, the feedback force pushes the system closer to the locking point) up to the point of the next zero crossing, since this corresponds to a sign change of the feedback force. A common mistake is to think that a locking point's range instead only extends to the next turning point (zero-slope point) in the error signal, but this is not the case. So long as the error signal's value is of the one sign, the feedback force maintains a constant direction.

There are practical limitations to the effectiveness of a feedback loop, ie its ability to keep the system's operating point precisely fixed in the face of all disturbances. Actuators usually have some limit on the magnitude of force that can be applied (the hill may be too steep for the car to keep speed, even with the accelerator "floored"). More importantly, some disturbances happen "too quickly", and the feedback loop responds only after a short delay. It isn't critical if a cruise-controlled car slows down a little for a second or

two while the automatic transmission changes the car's gear. However, for gravitational wave detection applications the subject is important.

Mathematical analysis of the problem is possible using the Laplace transfer functions defined above. The first step is to define the *loop transfer function* by:

$$G(s) \equiv P(s)S(s)F(s)A(s), \tag{2.1}$$

which is the dimensionless factor by which the magnitude of a disturbance is scaled in one round trip through the feedback loop (the 'G' stands for *gain*, a term which usually describes an amplification factor). It can be shown that, if $P(s)$ is the plant's transfer function when feedback is *not* connected, then:

$$P_{cl}(s) = \frac{P(s)}{1 + G(s)} \tag{2.2}$$

is the *closed loop transfer function* of the plant. What this tells us is that, when feedback is connected, the plant's response to disturbances is suppressed by a factor of $1 + G(s)$, or $\approx G(s)$ for large $|G|$.

The frequencies at which the gain G is large are important. At high frequencies, the feedback loop's response will typically begin to lag, if for no other reason than that it takes time for electrical signals to travel along wires. This manifests mathematically as a phase lag in the loop gain $G(s)$ at high Laplace frequencies s. In addition, the magnitude of G usually decreases as a function of frequency because it becomes harder and harder to force a mechanical actuator to oscillate at higher and higher frequencies. There is a possibility that, if the magnitude of G falls to unity around the same frequency that the phase lag of G reaches π (so $G \approx -1$), then the value of $1 + G(s)$ becomes very small rather than very large, and disturbances are *amplified* rather than suppressed. In this case, the system provides a feedback influence about the same magnitude as the disturbance, but instead of cancelling the disturbance through negative feedback, the π phase lag produces positive feedback, and the influence is amplified over and over again.

The design challenge of avoiding such instabilities is the bread and butter of feedback control engineers. The engineer has the opportunity of "tweaking" the frequency response of the system by appropriate choice of the filter stage, which in practise means using an electronic priming circuit. For us here, it is sufficient to know that this challenge is usually solvable, and the resulting loop gain functions are large at low frequencies and then fall below unity before reaching frequencies at which the phase lag is anywhere near π.

For mechanical systems using feedback control to suppress vibration, it is sufficient to use feedback with strong gain at low frequencies, and more or less ignore the problem at high frequencies. This is particularly the case for gravitational wave detectors, where seismic noise causes a relatively large swing range in suspended optics, and the device simply will not work without feedback control at these frequencies. High frequency disturbances are less of a problem as they usually don't have anything like the same amplitude: compare the displacement of a grandfather clock pendulum at 1 Hz with the displacement of a tuning fork after striking a note at several hundred Hertz. Hence, while high frequency noise may prevent measurement at said frequency, the noise typically isn't strong enough to stop the device working altogether.

2.2   The Interferometer as a Null Instrument

The interferometric gravitational wave detector, as presented in the previous chapter, is designed to produce a signal in the form of a fringe shift. In literal terms, a passing gravitational wave causes a slight change in the relative optical path length of the interferometer arms. This changes the interference between the light beams as they recombine, and a change in power level is observed at the output port.

One might originally decide to let the Michelson swing across many fringes on the grounds that the more the Michelson can move the better the sensitivity of the device. The initial difficulty with this is that, when the output power reaches its maximum or minimum and then leaves again, it is impossible to tell "which way" the path length mismatch went after the turning point[2].

Bigger problems arise when one considers an advanced configuration such as that in Fig. 1.6, where the Michelson degree of freedom is supplemented by multiple cavity degrees of freedom. Optical Fabry-Perot cavities are only useful in a gravitational wave detector if they resonate the probe light, that is if successive internal reflections all add up in phase. Cavities require feedback control to do this. It is also important to lock the main Michelson itself to a single operating point so that the light being supplied to the cavities can reach a steady state. This accounts for all the degrees of freedom of the advanced interferometer, which are hence locked with an array of feedback loops, and so the system is fixed to a single collective operating point. This simplifies the entire gravitational wave interferometer into a linear system as long as the system remains locked, since the feedback loop ensures it only executes small linear excursions from the selected operating point.

There remains the question: "How does the system give a signal readout of gravitational disturbance information?" because oscillations are now suppressed in the same degree of freedom that the signal will appear in. The answer is that the *error signal* associated with feedback control of the main Michelson degree of freedom gives exactly the required information. Admittedly, the strength of the gravitational wave signal manifesting in the error signal is scaled down by the value of the feedback loop gain, but this suppression is applied uniformly to signal and noise alike at any given frequency. Hence, the signal is no harder to see on the noise background, at least relative to noise sources that can be suppressed by feedback.

In addition to this, other feedback loops are used elsewhere in the interferometer to independently reduce the level of some particular noise source. For example, the multi-stage pendula suspending the Michelson mirrors are fitted with active damping systems aimed at reducing vibrations at the system's resonant frequencies.

Upon locking the main Michelson degree of freedom, the gravitational wave detector becomes what is called a *null instrument*. This describes the fact that the output is actively forced to zero, or to whatever the default value is. The output of course still reports small fluctuations from the lock point from which the error signal is derived. In fact, the error

---

[2] This isn't fatal. It is actually possible to extract full information about the trajectory of the Michelson's state over several fringes, as shown in [25]. However, this particular technique still requires the use of equipment to generate an error signal in exactly the same way described in the text that follows, even though the error signal isn't actually used in a feedback loop.

signal now becomes the main detection channel of the interferometer, since it reports the desired quantity of how far from the operating point the interferometer has been displaced.

2.3 Interferometer Operating Points

The next step is to determine at what state the Michelson should be locked. The choices are to lock to a bright fringe (perfect constructive interference at the output), a dark fringe (perfect destructive interference), or somewhere in between. The following is a brief outline of the associated noise considerations.

With a simple Michelson (ignoring cavities for the moment) the obvious choice is to lock "half way up a fringe", where the slope of the output power curve is maximal. The output power curve itself becomes the error signal, where the user subtracts some constant voltage to generate an artificial zero crossing at the desired lock point. This method is referred to as *direct detection*, and sometimes as *offset locking*.

There are technical problems with direct detection. Fluctuations in the power output from the Michelson can come from a wide variety of sources other than a gravitational-wave-induced fringe shift. An important example is power fluctuations of the laser source input into the Michelson. In addition, there are a range of noise sources collectively referred to as "$1/f$ noise", which roughly describes the characteristic amplitude roll-off with increasing frequency, and includes seismic noise, acoustic vibrations, and electrical noise. This is a generic problem of any physics experiments, not just gravitational wave detectors, but in this case, the desired signal is in the acoustic band where $1/f$ noise is problematic.

While it is possible to do a certain amount to suppress and dampen these noise sources, there exists a way to circumvent low frequency noise to a large extent. It is possible to operate a gravitational wave interferometer at a dark fringe, using a trick that will be described fully at the end of Chapter 3. While a Michelson's power response is quadratic about a dark fringe (ie no first order response), its *electric field* response is linear, and the trick takes advantage of this fact. Meanwhile, the bulk of the light at the laser frequency is stopped by the Michelson from reaching the photodetector, along with any noise it carries.

The trick is traditionally referred to as *chopping*: a high frequency signature is applied to the signal to be measured. This enables the user to distinguish the signal from noise, which does not carry the signature, at the measurement stage. This evades the low frequency noise problem, at least for sources upstream of the Michelson, but the cost is introducing noise associated with the higher frequency band around the chopping frequency [26]. Usually this is a favourable trade off because the chopping can be done at a frequency high enough, beyond where $1/f$ noise is significant. A chopped gravitational wave detector is said to use *rf detection* (radio frequency detection) as the signature's frequency is Megahertz or higher.

Recently however, direct detection is being reconsidered as a possibility for gravitational wave interferometers. The impetus comes from the ultra-high-finesse cavities planned for second generation detectors such as Advanced LIGO which, if practise meets design expectations, will significantly attenuate input-noise at frequencies higher than about 1 Hz [16]. Hence $1/f$ noise will no longer be a barrier inside the detection band,

at least for sources upstream of the cavity system. The choice between direct detection and rf detection then comes down to more fundamental noise sources at the respective frequencies. Moreover, a recent paper by Chen et al [27] showed that quantum noise is greater in an rf detection scheme than in a direct detection scheme for parameter regimes such as those proposed for the Advanced LIGO detector.

This thesis does not deal with the problem of locking an entire advanced Michelson configuration, but instead presents newer additional optical subsystems in isolation, leaving the assimilation of these components into a full control system and configuration as future work. However, it is still possible to comment on how this new work may fit. This is deferred until Chapter 6 once the requisite components have been discussed in detail.

2.4   Frequency Response to Gravitational Waves

Interferometers with signal recycling are configurable to be sensitive over a broad or narrow band. Presented here are pertinent details about the frequency response of detectors, with a view to improving freedom to tune a detector's frequency response.

The frequency response to gravitational waves of a power and signal recycled interferometer with arm cavities can be shown to be:

$$\frac{E_{out}(\omega_s)}{E_{cav}(0)} = A(\omega_s)M(\omega_s)E(\omega_s) \tag{2.3}$$

where the transfer function ratio is between the gravitational wave output light signal $E_{out}(\omega_s)$, shifted by ω_s from the main laser frequency, and the light circulating in the arm cavities $E_{cav}(0)$. Shaddock showed in [28] that this transfer function can be decomposed into contributions from parts of the interferometer as follows:

- $A(\omega_s)$ is the arm cavity response, which responds to gravitational waves with a high-bandwith cavity-like resonance centered at zero frequency:

$$A(\omega_s) = \frac{-it_1^2 r_2 e^{i2\omega_s L_c/c}(\delta/2)}{(1-r_1 r_2)(1-r_1 r_2 e^{i2\omega_s L_c/c})}; \tag{2.4}$$

- $M(\omega_s)$ is imposed by the Michelson, whose differential mode oscillation converts the arm cavity's phase shift into an amplitude shift via an extra complex factor i:

$$M(\omega_s) = 2it_{bs}r_{bs}; \tag{2.5}$$

- $E(\omega_s)$ is the resonant enhancement of the signal due to the presence of the signal cavity, and can resonate with a broadband or narrowband response depending on the signal cavity detuning ϕ_s:

$$E(\omega_s)\frac{it_s e^{i\omega_s L_s/c}e^{i\phi_s/2}}{1-r_s(r_{bs}^2+t_{bs}^2)r_{cav}(\omega_s)e^{i2\omega_s L_s/c}e^{i\phi_s}}. \tag{2.6}$$
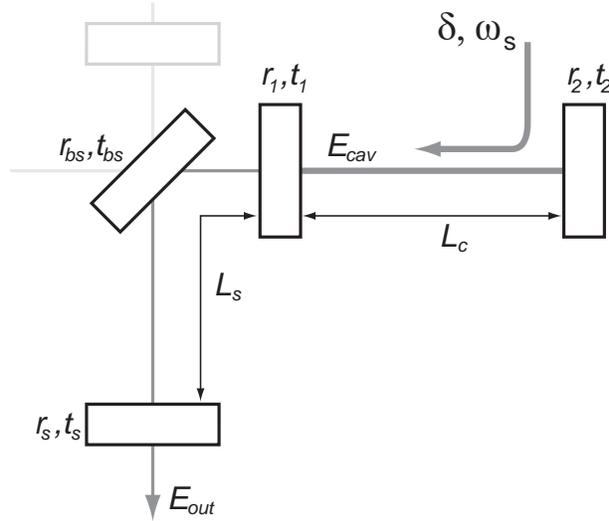
Fig. 2.3: Definitions for calculation of the gravitational wave detector's transfer function. A gravitational wave imposes, on the resonating probe light field E_{cav}, a modulation of amplitude δ and frequency ω_s. The transfer relations showing how this signal is converted and enhanced on the way to the output E_{out} are outlined in the text.

Parameters are defined as shown in Fig. 2.3. Specifically, the t_j and r_j are transmissivities and reflectivities of the input arm cavity mirrors ($j = 1$), end arm cavity mirrors ($j = 2$), the signal recycling mirror ($j = s$), and the beamsplitter in the centre of the Michelson ($j = bs$). L_c and L_s are the lengths of the arm cavities and the signal cavity respectively. The gravitational wave modulates the light in the arm cavities with amplitude δ.

As the gravitational wave signal is reflected by the signal mirror back into the arm cavities, the effective reflectivity of the arm cavities r_{cav} enters into the equation for $E(\omega_s)$. This factor is given by:

$$r_{cav} = \frac{r_1 - (r_1^2 + t_1^2)r_2 e^{i2\omega_s L_c/c}}{1 - r_1 r_2 e^{i2\omega_s L_c/c}} \tag{2.7}$$

The signal cavity detuning ϕ_s is defined as the optical phase accrued in the extra round trip distance travelled when the signal mirror's position is shifted back slightly. This parameter gives the user the means to alter the gravitational wave detector's frequency response from broadband to narrowband. Fig. 2.4 shows the shape of this response for several choices of ϕ_s, with the peak sensitivity and resonance width changing significantly.

One major issue is the fact that the resonant frequency and bandwith cannot be changed independently in this configuration. While it is possible to gain access to another set of frequency response curves by substituting the signal mirror with another of different reflectivity, it takes weeks to make this substitution in a large scale interferometer due to the requisite vacuum enclosures.

de Vine showed in [10] how independent control over the resonant frequency and bandwidth can be achieved by substituting either a cavity or a Michelson interferometer for the signal mirror. This enables future detector designs which incorporate a signal
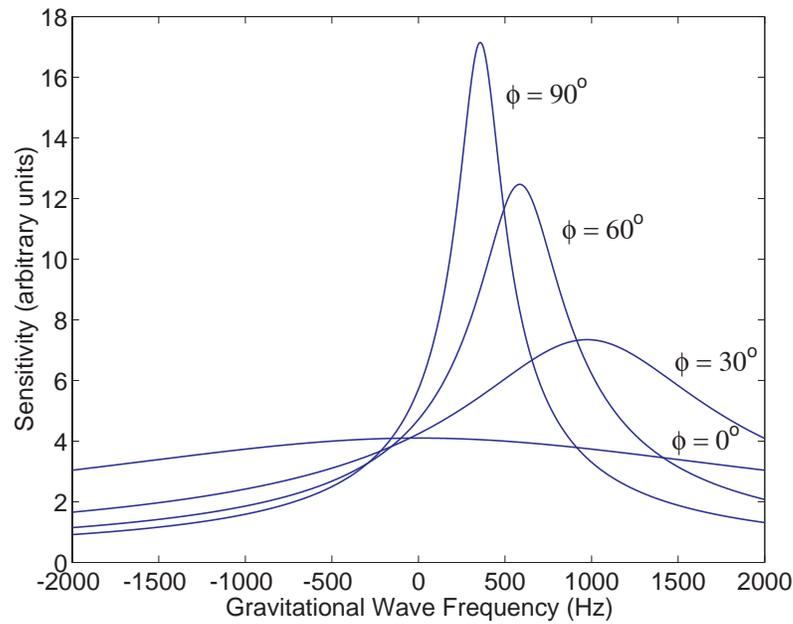
Fig. 2.4: Gravitational wave detector frequency response for several signal cavity detunings ϕ_s, ranging from broadband to narrowband, as computed from Eq. 2.3. Values used are $r_1^2 = 98\%$, $r_2^2 = 100\%$, $r_s^2 = r_{bs}^2 = 50\%$, $L_c = 4$ km, $L_s = 3$ m, which correspond to the signal recycling (not resonant sideband extraction) regime.

VRM (variable reflectivity mirror) to have their responses tuned more flexibly, with the possibility of real-time tuning to "zoom-in" on the frequency of a specific gravitational wave event as it occurs.

This thesis aims to supplement this discovery by investigating locking schemes that take full advantage of this tunability. Commonly in texts, especially theses, descriptions of optical locking methods are grouped together. However, the emphasis of this thesis is on the components of such locking systems, and so the locking techniques are introduced on the fly. Specifically, Michelson interferometer locking is described in Section 3.4, Fabry-Perot locking in Section 4.3, and Chapter 6 describes a variation on these that includes the ability to shift the lock point of the device. The central component of these locking schemes is electro-optic modulation, and hence this is reviewed in the following chapter.

"The only way to succeed in this business is to be so bloody good
at it that it doesn't matter what anyone else thinks of you!"

— from the BBC television drama *Red Cap*

Electro-Optic Modulation and the Modulation Sphere

The modulation of electromagnetic light beams is a significant topic here due to its prevalence in gravitational wave detector control schemes. This chapter digresses from gravitational wave detector technology to take a closer look at modulation in its own right.

In addition to the standard modulation description, given in Section 3.1 and 3.2, a new geometric phase formulation of modulation is presented in Section 3.3, called the *modulation sphere* formulation in analogy to the Poincare sphere for polarisation states. The strength of the modulation sphere representation is in its inherently visual nature, facilitating quick mental calculations that would otherwise take time, thought and possibly computer modelling to comprehend. Examples using the modulation sphere formulation follow in Section 3.4, and the formalism is used in subsequent chapters as a visualisation and analysis tool.

3.1 Standard Description of Modulation

3.1.1 Preliminaries

As a preliminary step, the basic properties of laser beams are iterated here, in order to be specific about the subset of properties that will be described in this chapter.

A laser beam is a field of electromagnetic light that propagates in some direction in three-dimensional space. In most applications, beams are well-described under the paraxial approximation. This means that, amongst other things, electric field components in the direction of propagation are negligibly small, and therefore beams have transverse field oscillations only. By default, the following treatment assumes[1] that beams are monochromatic prior to modulation, of Gaussian profile, of a single linear polarisation, stationary, and classical, except where otherwise stated.

Mathematically, details of transverse profile, longitudinal profile, time varying profile and polarisation will be implicitly suppressed:

$$E = |\vec{E}_0(x,y,z,t)|e^{i\omega t} = E_0 e^{i\omega t}, \tag{3.1}$$

---

[1] The author has observed that many time-consuming mistakes occur in the laboratory when implicit assumptions such as these are overlooked, hence it is better to be clear from the start..

and will be made explicit as required. The remaining explicit terms are the electric field amplitude E_0 (which could be a scale factor for the entire beam, or the amplitude at some specific point), and the time-varying field oscillation $e^{i\omega t}$ which rotates at ω, the light's angular frequency[2].

The intensity and power of the beam are given by:

$$I = E^*E, \qquad P = \int I dA, \tag{3.2}$$

where dA is a unit of area perpendicular to the propagation axis, and an asterisk denotes a complex conjugate. So long as any treatment to the beam E is done uniformly across its profile, one is free to simply write $P = E^*E$. The power is usually the important quantity as standard photodetectors do not retain spatial information.

3.1.2   Modulation Mathematics

A light beam is modulated when one of its properties is varied at some frequency, ω_m, which is much less than the frequency of the electromagnetic oscillation, ω. Modulation needn't occur only at a single frequency, but this treatment is restricted to single frequency modulations for simplicity. In any case, a Fourier treatment is effective for analysing multiple frequencies[29].

Two physically distinct kinds of modulation exist: amplitude modulation (AM) and phase modulation (PM). A beam can be modulated with one or the other or a combination of both; in the latter case the relative phase of modulations is of significant physical importance. Modulation methods are discussed later in more detail, but it is a fact that all modulations can be written as a combination of AM and PM at a selection (discrete or continuous) of frequencies ω_m.

A sinusoidally amplitude-modulated beam of light can be written:

$$E_{AM} = E_0 \left[1 + A\cos(\omega_m t + \phi_A) \right] e^{i\omega t} = E_0 \left[1 + \Re\{\tilde{A}e^{i\omega_m t}\} \right] e^{i\omega t}, \tag{3.3}$$

where A is the *modulation depth* and is typically, but not always, much less than one, and ϕ_A is a phase angle. It is sometimes convenient to write the amplitude modulation in complex phasor notation, $\tilde{A} = Ae^{i\phi_A}$, as in the second expression. From here on, $\Re$ and $\Im$ will denote real and imaginary components of complex quantities.

The intensity of an amplitude-modulated beam oscillates as one would expect:

$$I_{AM} = E_{AM}^*E_{AM} = E_0^2 \left[1 + 2A\cos(\omega_m t + \phi_A) + A^2\cos^2(\omega_m t + \phi_A) \right]. \tag{3.4}$$

This demonstrates that AM produces a physical effect that is directly detectable. The effect is an oscillation in light power at the modulation frequency, and is called a *beat*.

The last term represents a small shift to the DC signal plus a small oscillation at twice the modulation frequency. For small modulations, these can be neglected. For larger

---

[2] The physical electric field strength isn't actually *complex*, but complex notation simplifies the algebra enormously. The actual electric field strength requires taking the real component of these terms.

modulations, the story is much more complicated than this, as will be seen in Section 3.2.5.

A sinusoidally phase-modulated beam of light can be written:

$$E_{PM} = E_0 e^{i[\omega t + P\cos(\omega_m t + \phi_P)]} = E_0 e^{i[\omega t + \Re\{\tilde{P}e^{i\omega_m t}\}]}, \tag{3.5}$$

where once again $\tilde{P} = Pe^{i\phi_P}$ represents the modulation depth and phase angle. Phase modulation is not directly detectable:

$$I_{PM} = E_{PM}^* E_{PM} = 0, \tag{3.6}$$

and requires more complex techniques to measure. For small modulation depth, $|\tilde{P}| \ll 1$, the complex exponential can be expanded[3] and neglect terms of higher order than one:

$$E_{PM} = E_0 \left[1 + i\Re\{\tilde{P}e^{i\omega_m t}\}\right] e^{i\omega t}. \tag{3.7}$$

In this thesis, Eq. 3.7 is treated as the *definition* of phase modulation since it mirrors Eq. 3.7 as the definition of amplitude modulation, and the two form a simple, consistent basis for modulations. There is a strong resemblance between Eq. 3.7 and Eq. 3.3; the difference is in a factor of i in the oscillating term, which represents an optical phase difference of 90°.

3.1.3   Modulation Visualisations

Fig. 3.1 shows four different visual representations of optical modulation: a simple graph of the electric field as a function of time; an optical phasor and its time-varying behaviour; a three-dimensional frequency vector plot; and a modulation phasor diagram. The first three are fairly common, and the last is relatively rare.

In Fig. 3.1(a) the electric field of modulated beams are plotted versus time. The AM figure is clear in its (exaggerated) sinusoidally varying amplitude. PM appears as a compression wave, which is result of an alternating advancement and retardation of the wave's phase. This is the least abstract visualisation of modulation, and is really only useful as a teaching aid.

In Fig. 3.1(b) the light is represented by a rotating optical phasor. Phasors mathematically corresponds to complex quantities with terms like $e^{i\omega t}$, and phasor diagrams are produced by plotting these in an Argand diagram (the complex plane), as shown. The phasors of course don't actually rotate around the origin on the printed page; instead the viewer envisages his/her own point of reference rotating with the optical phasor at frequency ω, or else that the view of the phasor is "strobed" at the same frequency. The double-ended arrows represent the range over which the optical phasors oscillate due to modulation. The AM double-ended arrow is parallel to the main phasor, and the PM double-ended arrow is perpendicular to it; the latter was represented mathematically in Eq. 3.7 by the extra factor of i.

---

[3] Expansions of this type use the so called *Bessel functions*, and are described in Section 3.2.5.
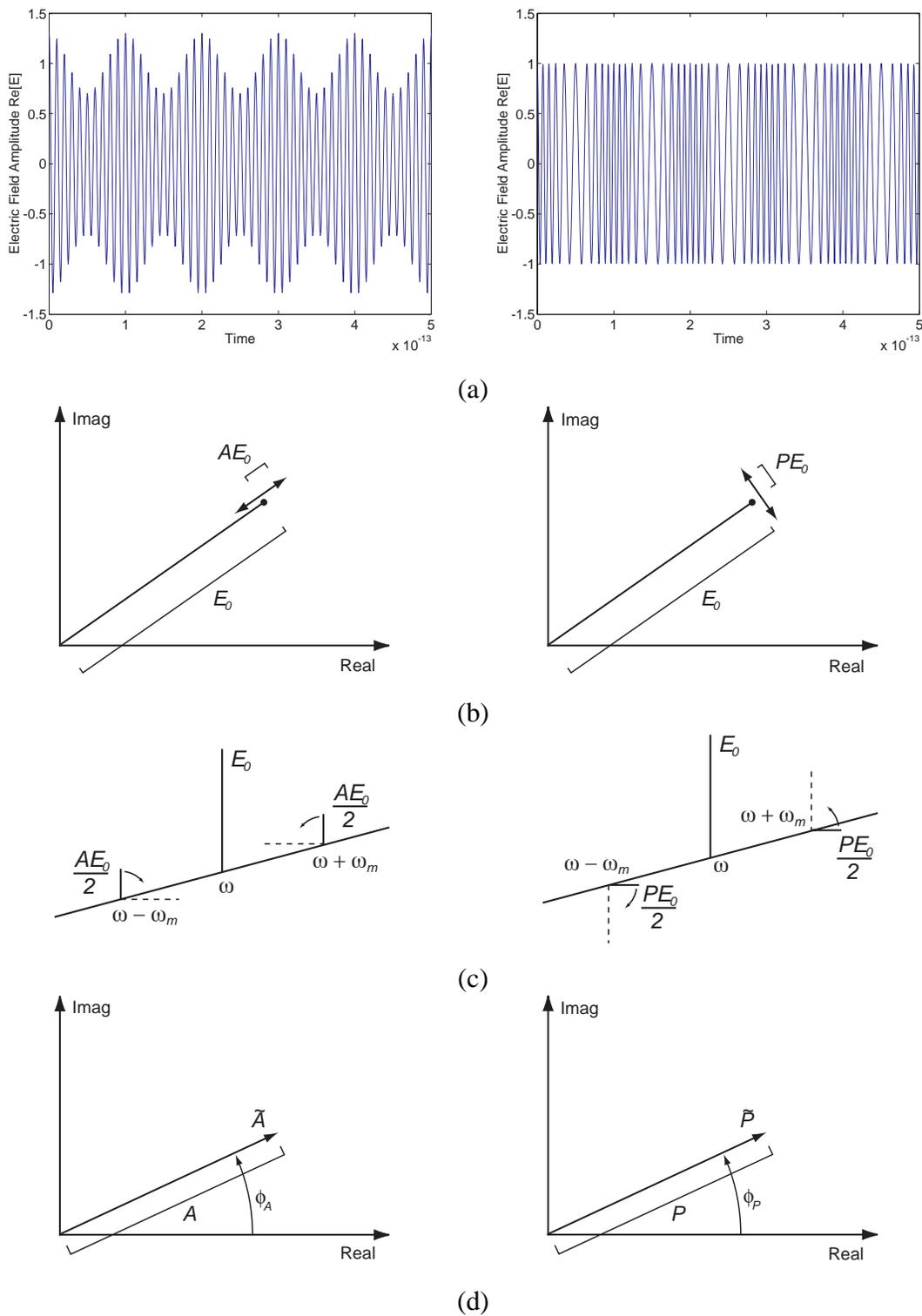
Fig. 3.1: Amplitude modulation (left) and phase modulation (right) represented as: (a) time varying electric fields; (b) optical phasors; (c) frequency vectors; and (d) modulation phasors.

True PM (Eq. 3.5) requires a curved double-ended arrow, in order that the magnitude of the phasor remains constant. PM as defined in Eq. 3.7 corresponds to a straight line, which is accurate enough for sufficiently small modulation depth, P.

Fig. 3.1(c) shows a schematic of a Fourier spectrum of the modulated light, with real and imaginary components of frequency lines plotted against their frequencies. It is seen that both PM and AM produce *sidebands* above and below the optical *carrier* frequency, ω, where the sidebands are displaced by an amount equal to the modulation frequency, ω_m. Once again, these frequency lines have the properties of phasors, this time rotating around the frequency axis. As with the previous phasor diagrams, one can visualise the upper and lower sidebands ($\omega + \omega_m$ and $\omega - \omega_m$ respectively) as rotating faster and slower, respectively, than the carrier. At any instant during this rotation, the three vectors can be added, in the normal sense of vector addition, to give the resultant time-varying electric field. It can be seen that, in the AM case, the sideband vectors reinforce in the direction of the carrier vector, and cancel in the perpendicular direction, so that the magnitude of the carrier field changes without the phase changing. Similarly for PM, where the sideband vectors only ever contribute perpendicularly to the carrier vector. The difference between PM and AM is therefore as simple as a different phase relation between the carrier and sidebands.

Fig. 3.1(d) is a less-physical representation of modulation, but directly represents the phasors $\tilde{P}$ and $\tilde{A}$ introduced in Eq. 3.3 and Eq. 3.7. Here, PM and AM are distinguished by labelling, not by form, and the carrier is not represented at all. However, the phase (ϕ_A or ϕ_P) of a modulation is immediately clear in this representation, whereas in Fig. 3.1(c) phase information is more obscure, and modulation phase is not represented at all in Fig. 3.1(b). This representation is more useful for dealing with complicated modulation states, as only one frequency component is represented at a time.

3.2 Advanced Modulation

The above description of modulation is complete in that it is possible to analyse complicated modulations from these basic foundations. However, this thesis deals extensively with modulation, and the following thorough review is a necessary building block for understanding and interpreting the developments to come.

3.2.1 Modulation Quadratures

It is worth taking a second look at the phase of the modulation (ϕ_P or ϕ_A), which is better known as the modulation *quadrature*. Physically, this is the phase of the wave imposed as modulation on a light beam, and is of significance when comparing two beams, or a beam and an electrical signal. If two modulations have the same phase, they are said to be *in-phase*. If two modulations are different in phase by $90°$, they are in *quadrature*. The term "quadrature" often carries an assumption of rectangularity, so that in-phase and quadrature vector-wise *components* of a modulation are discussed (mathematically, real and imaginary components), rather than magnitudes and phases of a modulation as used above.

It is common to speak of in-phase or quadrature components of a single modulation
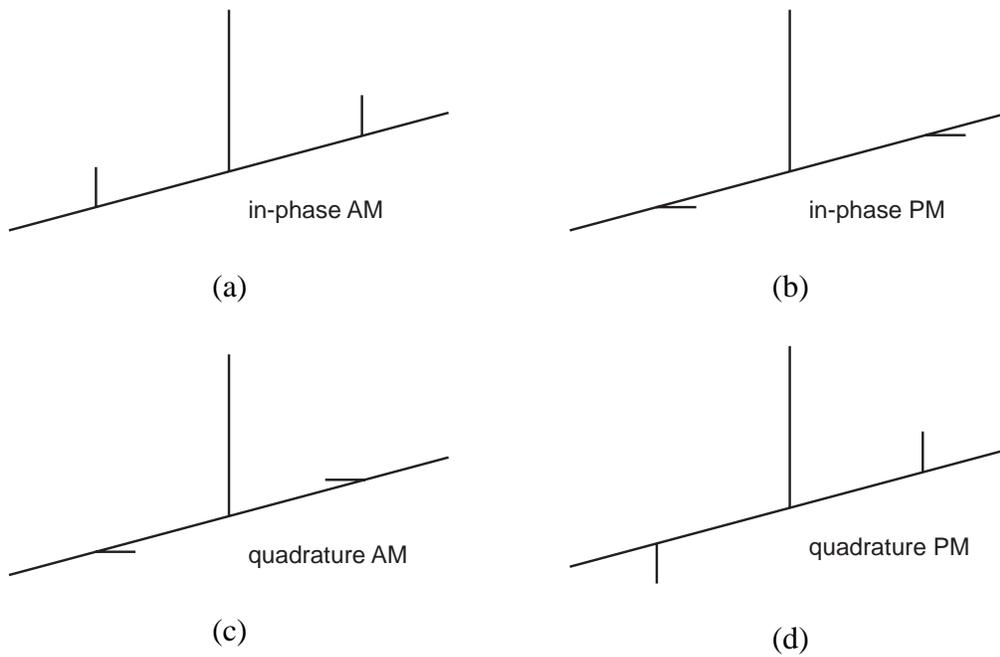
Fig. 3.2: Frequency vector diagrams of: (a) in-phase AM ($\phi_A = 0$); (b) in-phase PM ($\phi_P = 0$); (c) quadrature AM ($\phi_A = \pi/2$); and (d) quadrature PM ($\phi_P = \pi/2$).

even without specifying a phase reference by which these terms are defined. This is usually for concreteness, and is perfectly valid since deliberate modulations always have a well-defined phase. Typical experiments will include an electronic oscillator signal which was used in the original electro-optic modulation; a copy of this signal provides the phase reference, so that there is always something concrete to refer back to.

Mathematically, the significance of *relative* phases between modulations becomes clear with a change to the zero-point of the time parameter, by making the substitution $t \leftarrow (t - \tau)$. This effectively can change a modulation phase angle arbitrarily, but of course the time shift applies equally to all modulations and electrical signals in the system, so that relative phases are preserved.

Fig. 3.2 shows frequency vector diagrams of the same modulations with different quadratures. The left-hand-side diagrams both represent AM, but the lower diagram shows what the AM in the upper diagram will look like after one quarter of an oscillation. In particular, the $+\omega_M$ sideband has rotated anti-clockwise by $90°$ and the $-\omega_M$ sideband clockwise by $90°$. The two PM diagrams on the right similarly represent the same state with a time difference equal to one quarter of a cycle. It is conventional, in the absence of a specific phase reference, to refer to the upper diagrams as being "in-phase" modulations and the lower diagrams to be "quadrature" modulations.

3.2.2   Modulation Sidebands

The term *sidebands* was in the previous section, and the mathematical description can be manipulated to shed some light on the situation. Observing that:

$$\Re\{\tilde{X}\} = \frac{1}{2}\left[\tilde{X} + \tilde{X}^*\right], \tag{3.8}$$

Eq. 3.3 can be rewritten as:

$$E_{AM} = E_0 e^{i\omega t} + \frac{1}{2}\tilde{A}E_0 e^{i(\omega+\omega_m)t} + \frac{1}{2}\tilde{A}^* E_0 e^{i(\omega-\omega_m)t}, \tag{3.9}$$

and Eq. 3.7 as:

$$E_{PM} = E_0 e^{i\omega t} + \frac{1}{2}i\tilde{P}E_0 e^{i(\omega+\omega_m)t} + \frac{1}{2}i\tilde{P}^* E_0 e^{i(\omega-\omega_m)t}. \tag{3.10}$$

Now the individual frequency components are explicit, and one can see the correspondence with Fig. 3.1(c). Both PM and AM produce upper and lower sidebands of the same magnitude (for the same modulation depth) but with different phases.

In fact, it is possible to define upper and lower sideband phasors, $\tilde{S}_{+\omega_m}$ and $\tilde{S}_{-\omega_m}$, and write a modulation state in terms of these:

$$E = E_0 \left[1 + \tilde{S}_{+\omega_m} e^{i\omega_m t} + \tilde{S}_{-\omega_m} e^{-i\omega_m t} \right] e^{i\omega t} \tag{3.11}$$

where

$$\tilde{S}_{+\omega_m} = \tilde{A}/2 \quad \text{and} \quad \tilde{S}_{-\omega_m} = \tilde{A}^*/2 \tag{3.12}$$

for AM, and

$$\tilde{S}_{+\omega_m} = i\tilde{P}/2 \quad \text{and} \quad \tilde{S}_{-\omega_m} = i\tilde{P}^*/2 \tag{3.13}$$

for PM.

3.2.3   Combined Modulations

Texts describing modulation often stop short of reviewing the physics of modulations composed of both AM and PM. While the description up to this point provides adequate means to describe all modulations (since modulation phasors add with all the usual rules for vectors), combined modulations carry a degree of subtlety that is worth examining.

As mentioned earlier, when a modulation state has some component of both AM and PM, the phase relation between the two components is critical. There are two extreme cases: PM and AM components in-phase, and PM and AM components in quadrature. In Section 3.3, these states, along with states purely of either PM or AM, will take on cardinal significance in defining the modulation sphere representation.

Firstly, a beam of light with the most general single-frequency modulation can be written:

$$E = E_0 \left[1 + \Re\{\tilde{A}e^{i\omega_m t}\} + i\Re\{\tilde{P}e^{i\omega_m t}\} \right] e^{i\omega t}. \tag{3.14}$$

Equally, the same expression can be written in terms of sideband phasors as in Eq. 3.11 above, with the correspondence:

$$\begin{aligned}
\tilde{S}_{+\omega_m} &= \frac{1}{2}\left[\tilde{A} + i\tilde{P}\right] \\
\tilde{S}_{-\omega_m} &= \frac{1}{2}\left[\tilde{A} - i\tilde{P}\right]^*.
\end{aligned} \tag{3.15}$$

The first case examined here is a modulation state with equal-magnitude PM and AM
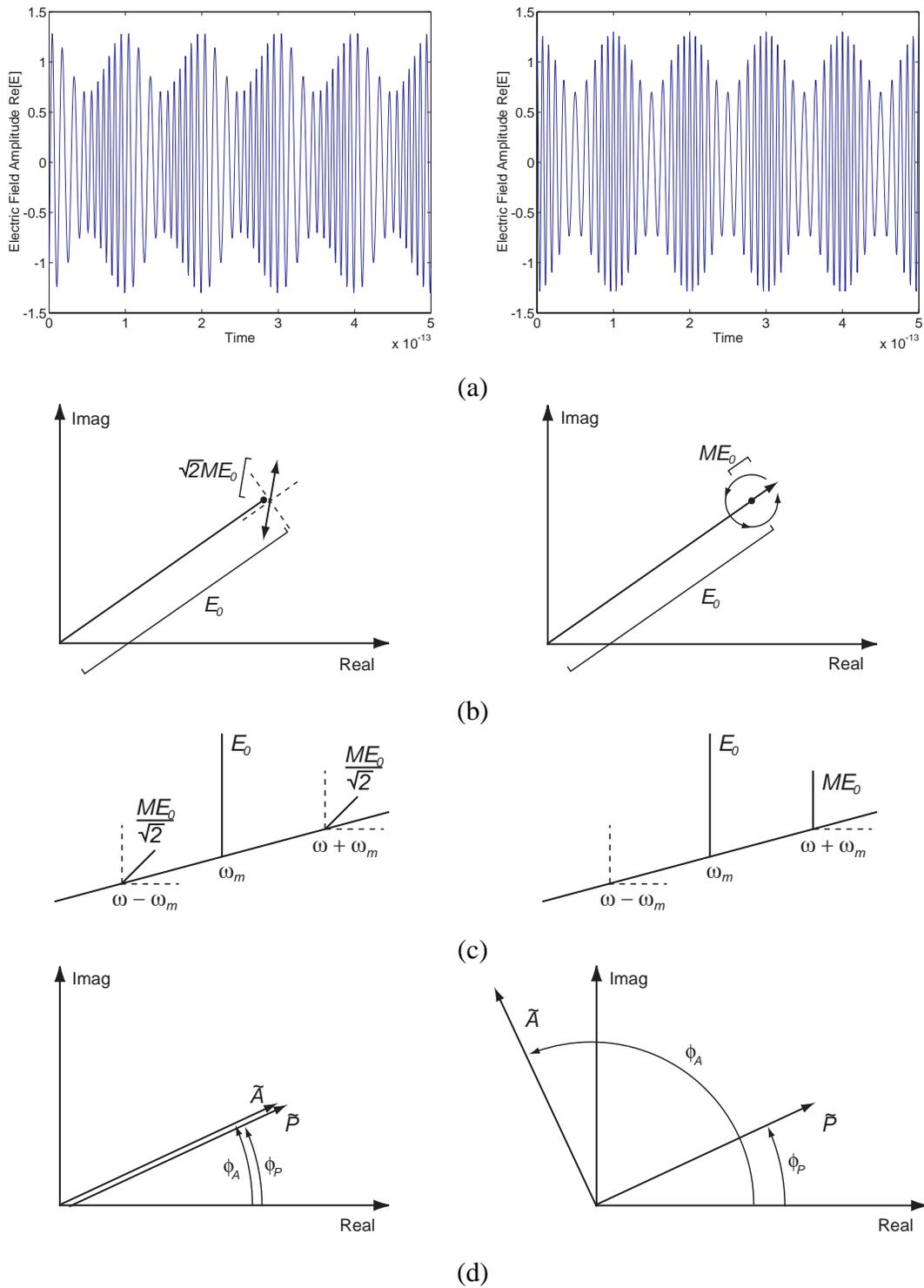
Fig. 3.3: In-phase AM/PM combination (left) and single sideband modulation (right) represented as: (a) time varying electric fields; (b) optical phasors; (c) frequency vectors; and (d) modulation phasors.

components that are in phase, ie $\tilde{A} = \tilde{P} = \tilde{M} = Me^{i\phi_M}$, where $\tilde{M}$ is a modulation phasor introduced for convenience. Then, Eq. 3.14 becomes:

$$\begin{aligned}
E &= E_0 \left[1 + \Re\{\tilde{M}e^{i\omega_m t}\} + i\Re\{\tilde{M}e^{i\omega_m t}\} \right] e^{i\omega t} \\
&= E_0 \left[1 + \sqrt{2}e^{i\pi/4}\Re\{\tilde{M}e^{i\omega_m t}\} \right] e^{i\omega t}.
\end{aligned} \tag{3.16}$$

This modulation has an optical phase at $45°$ relative to the carrier. Further, the sidebands also receive this phase shift:

$$\tilde{S}_{+\omega_m} = \frac{1}{\sqrt{2}} \left[\tilde{M}e^{i\pi/4} \right] \quad \text{and} \quad \tilde{S}_{-\omega_m} = \frac{1}{\sqrt{2}} \left[\tilde{M}^* e^{i\pi/4} \right], \tag{3.17}$$

The modulation state is shown in Fig. 3.3, and (b) and (c) in particular show clearly how the $45°$ shift is manifested.

The second case examined here is a modulation state with equal-magnitude PM and AM components that are in quadrature, ie $\tilde{A} = i\tilde{P} = \tilde{M}$:

$$\begin{aligned}
E &= E_0 \left[1 + \Re\{\tilde{M}e^{i\omega_m t}\} + i\Re\{-i\tilde{M}e^{i\omega_m t}\} \right] e^{i\omega t} \\
&= E_0 \left[1 + \Re\{\tilde{M}e^{i\omega_m t}\} + i\Im\{\tilde{M}e^{i\omega_m t}\} \right] e^{i\omega t} \\
&= E_0 \left[1 + \tilde{M}e^{i\omega_m t} \right] e^{i\omega t}.
\end{aligned} \tag{3.18}$$

The final expression represents a modulation state with only one sideband, called a *single sideband* (SSB) state, since a comparison with Eq. 3.11 gives:

$$\tilde{S}_{+\omega_m} = \tilde{M} \quad \text{and} \quad \tilde{S}_{-\omega_m} = 0 \ ! \tag{3.19}$$

Again Fig. 3.3 shows this single sideband modulation in a number of different pictures, of which (c) shows the absence of the lower sideband most clearly. The result in (b) is quite instructive; the carrier phasor traces out a circle in the Argand diagram. One way to understand this is to compare the situation to an electronic oscilloscope set in X-Y mode, where two signals are mapped to the X and Y axes of the display. When the two signals are $90°$ different in phase, ie in quadrature, the resulting figure is a circle[4]. In the modulation case, the AM and PM components define the X and Y axes, and the fact that the two components are in quadrature phase produces this circular oscillation. One might hence use the term *circular modulation* to describe this state.

3.2.4 Modulation Devices

The discussion so far has dealt with modulation in the abstract, and now turns to the means of taking a laser beam and adding a modulation. Phase modulation devices are described first, followed by their more-complicated amplitude modulation counterparts.

Fig. 3.4(a) shows a schematic of an electro-optic *phase modulator*, or *Pockel's cell*

---

[4] Such a shape is called a *Lissajous figure*, which is the trace resulting when coordinates in orthogonal axes oscillate at frequencies related by a ratio of two (small) integers. A circle, or diagonally aligned ellipse, is the simplest such Lissajous figure, when the ratio between X and Y frequencies is unity.
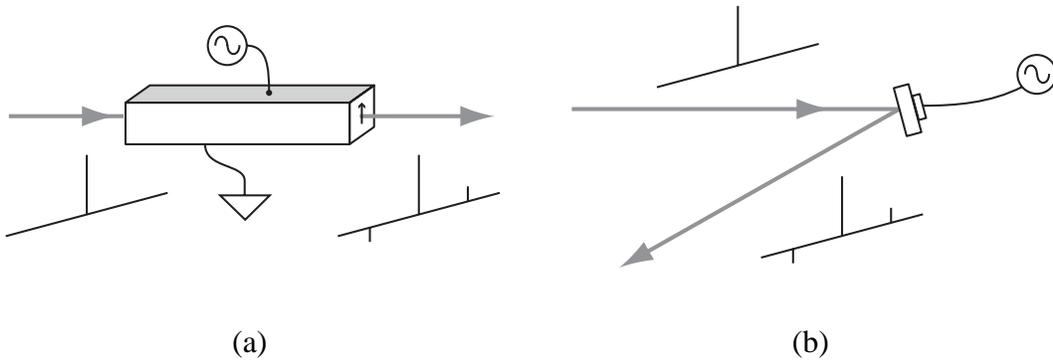
Fig. 3.4: Optical devices that mount phase modulation on a beam using: (a) a Pockel's cell; (b) a mirror mounted on a piezo-electric device.

as the devices were better known as a couple of decades ago. This device exploits a nonlinearity possessed by certain crystals called the *Pockel's effect*: an electric field E applied to the crystal causes a first-order change in refractive index n, via:

$$n(E) = n - \frac{1}{2}r_P n^3 E,\tag{3.20}$$

where r_P is the Pockel's coefficient and depends on the material of the cell (typical values are 10^{-12} to 10^{-10} m/V). Fig. 3.4(a) shows a transverse modulator, where the crystal is such that the refractive index change occurs perpendicularly to the applied field. The direction of the electric field determines whether the refractive index increases or decreases, as does the optical path length travelled by a beam going through the device. Hence, a crystal with a Pockel's (first order) nonlinearity is not symmetrical, but has a preferred direction along the axis where the electric field is applied.

The two electrodes of a Pockel's cell form a capacitor around the crystal and so, to a good approximation, the applied electric field is constant within the crystal's volume. For this geometry of modulator, the optical phase acquired by a laser beam passing through is:

$$\phi = \phi_0 - \pi\frac{V}{V_\pi}\tag{3.21}$$

where ϕ_0 is the zero-voltage phase shift, V is the applied voltage. V_π, the *half-wave voltage*, is the applied voltage at which the phase shift changes by π, and depends on the material properties, the light's wavelength, and the aspect ratio (length to height) of the modulator. When one makes the requirement that the modulation depth is small, this means that the maximum applied voltage should be well short of the half wave voltage.

It is clear that a sinusoidally varying voltage applied to the cell will produce a sinusoidally varying optical path length, and so the laser beam will acquire a phase modulation as in Eq. 3.5.

Pockel's cells are made with a variety of crystals, with the cells tested in this thesis being of $LiNbO_3$. Cells are available that can produce modulations at frequencies anywhere between DC and GigaHertz. Cells can be designed for broadband operation, with half wave voltages of kiloVolts, or at single frequencies, when a built-in resonant circuit amplifies the voltage input such that the half wave voltage is of order 10 Volts.

(a)   (b)
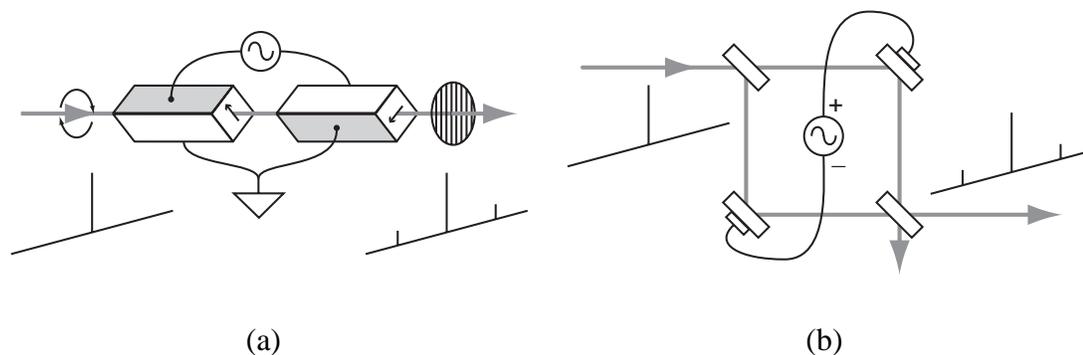
Fig. 3.5: Optical devices that mount amplitude modulation on a beam using: (a) two Pockel's cells; (b) a Mach-Zehnder interferometer.

Fig. 3.4(b) shows a far simpler, but less flexible method of phase-modulating a beam. The beam shown reflects off a mirror attached to a piezo-electric material that causes the mirror to move forward or backward in response to an applied voltage. Hence the optical path length that the beam travels in modulated in the simplest way possible, by a mechanical path length change. The obvious drawback is the introduction of a lateral shift of the beam which, while small, may be problematic for some applications. Instead, the beam can be retroreflected (ie perpendicular incidence on the mirror) to avoid this, but then other optics are required to handle the coincidence of the input and output beams. Also, piezo-electric devices have diminishing responses for high frequencies, so this method is rarely used above acoustic frequencies.

Many other phase-modulation devices are possible, and usually involve some sort of optical path length oscillation. In practise, the choice of modulator is based on cost, frequency response, and availability.

An electro-optic *amplitude modulator* shown in Fig. 3.5(a). This consists of two Pockel's cells aligned with their modulating axes at right angles to one another, followed by a polar filter or other polarising element. A voltage electrode is connected in parallel to both crystals with opposite polarity, such that an applied voltage will increase the refractive index of one crystal as it decreases the refractive index of the other. This ensures that the net phase modulation on the output beam is zero, so that only amplitude modulation is possible.

Understanding how the device actually produces amplitude modulation is tricky; the mathematical details are postponed until Chapter 5 and a qualitative description will suffice at this stage.

This type of amplitude modulator needs to operate on circularly-polarised light[5]. When a sinusoidal voltage is applied, the two Pockel's cells operate on the light to form a *polarisation modulation* state, shown in Fig. 3.6(a). Lastly, a polarising element separates the vertical and horizontal polarisation components, each of which separately is an amplitude modulated beam, and has half the intensity of the input light.

---

[5] Technically, the light needs only some component of circular polarisation in order for the device to produce amplitude modulation. Very often the inherent birefringence of the crystals will produce this at the output, so that just about any beam will produce *some* amplitude modulation. Chapter 5 demonstrates in

(a)

(b)

(c)
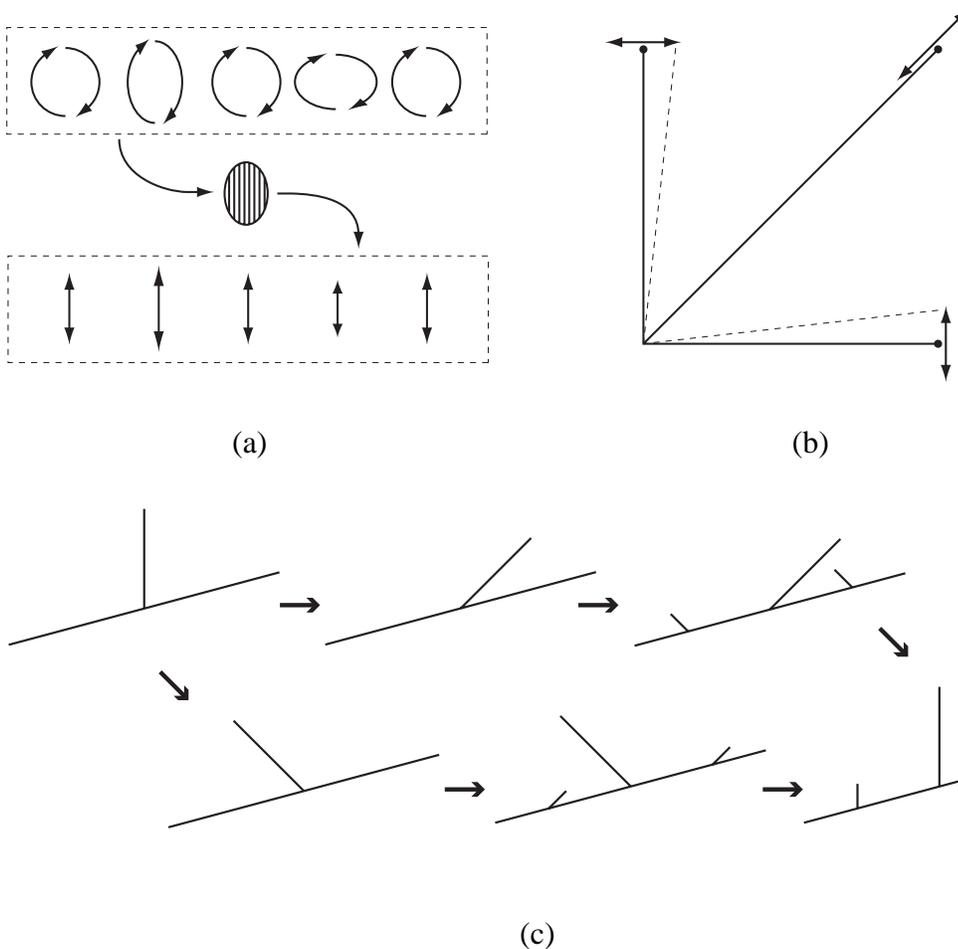
Fig. 3.6: Generation of AM. (a) The device in Fig. 3.5(a) produces a polarisation modulation state, that becomes an AM state after one polarisation component is extracted. (b) Phasor diagram, and (c) frequency vector diagram, both showing how two PM beams can be interfered to produce an AM beam.

Fig. 3.6(b) and (c) show two alternative diagrammatic descriptions of amplitude modulation production, using phasors and frequency vectors, respectively. In Fig. 3.6(b), two phasors, representing the light's polarisation components aligned with each of the two Pockel's cells, are separately phase modulated exactly out-of-phase, so that the phase of one beam is accelerated as the phase of the other beam is decelerated. Since circular polarisation is equivalent to two equal rectangular light components with a 90° phase difference, the two phasors are at right angles. Then, the polarising element has the effect of interfering these two components and producing a phasor which oscillates in length but not direction, ie amplitude modulation.

Fig. 3.6(c) tells the same story as Fig. 3.6(b) but using frequency vector diagrams. Two separate polarisation components acquire phase modulations that are exactly out-of-phase. One of these light components has a 90° optical phase shift relative to the other, since the beam is circularly polarised. Lastly, the vector sum of these two beams, representing the interference enacted by the polarising element, produces a modulation

more detail how this works, and how one gets the best out of this type of modulator.

state of the AM type.

Fig. 3.5(b) shows an alternative optical configuration for generating amplitude modulation based on a *Mach-Zehnder interferometer*. The interferometer is biased so that equal power components exit the two output ports (ie "half way up a fringe"). The mirrors in both interferometer paths are mechanically oscillated with piezo-electric devices as before, but out-of-phase so that the interferometer fringe condition oscillates without the net phase changing. Either output then has an amplitude modulated beam of half the intensity of the input beam.

In a variation on Fig. 3.5(b), a Pockel's cell can be placed in each interferometer arm to form a device called a *Mach-Zehnder modulator*.

The two AM devices here, Fig. 3.5(a) and (b), are in fact topologically equivalent. They differ only in the fact that one uses polarisation, and the other uses spatial separation, to arrange for the two light-components to receive out-of-phase PM. As such, the above descriptions are very similar, and Fig. 3.6(b) and (c) could just as easily apply to either modulator setup.

In general, AM devices all have some means whereby light power is dumped or diverted away from the output beam. The AM is ultimately achieved via a time-variation in the proportion of light allowed to pass through the device. As such, all passive AM devices must decrease the average power of a light beam by some factor, however small.

3.2.5 Higher-order Frequency Components

Up to this point, only modulations at the fundamental harmonic frequency have been considered. In fact, most modulation devices, including all the devices described in the previous section, produce frequency components at multiples of the base modulation frequency. These are important when the modulation depth is large, but may be important even for relatively low modulation depths because they appear at different frequencies.

The details of the higher-order harmonics for true phase modulation is obtained by using the *Jacobi expansion*:

$$e^{iz\cos y} = J_0(z) + 2\sum_{n=1}^{\infty} i^n J_n(z)\cos ny \tag{3.22}$$

to expand the exponential-cosine in the original true phase modulation equation Eq. 3.5, after which one obtains:

$$E_{PM} = E_0 \left[J_0(P) + 2\sum_{n=1}^{\infty} i^n J_n(P)\cos\left(n\omega_m t + n\phi_P\right) \right] e^{i\omega t}. \tag{3.23}$$

The $J_n(z)$ are the *Bessel functions* (or *Bessel coefficients*) *of the first kind*, and are given by:

$$J_n(z) = \sum_{m=0}^{\infty} \frac{(-)^m (z/2)^{n+2m}}{m!(n+m)!}. \tag{3.24}$$

The first few Bessel functions are iterated in Table 3.1. From this it is clear that each Bessel function is a power series where the lowest order term is of order n and

m	0	1	2
$J_0(z)$	1	$-\dfrac{z^2}{2^2}$	$+\dfrac{z^4}{2^6}$
$J_1(z)$	$\dfrac{z}{2}$	$-\dfrac{z^3}{2^4}$	$+\dfrac{z^5}{2^7\cdot 3}$
$J_2(z)$	$\dfrac{z^2}{2^3}$	$-\dfrac{z^4}{2^5\cdot 3}$	$+\dfrac{z^6}{2^{10}\cdot 3}$
$J_3(z)$	$\dfrac{z^3}{2^4\cdot 3}$	$-\dfrac{z^5}{2^8\cdot 3}$	$+\dfrac{z^7}{2^{11}\cdot 3\cdot 5}$

Tab. 3.1: The first few terms of the first few Bessel functions of the first kind.

(a)   (b)
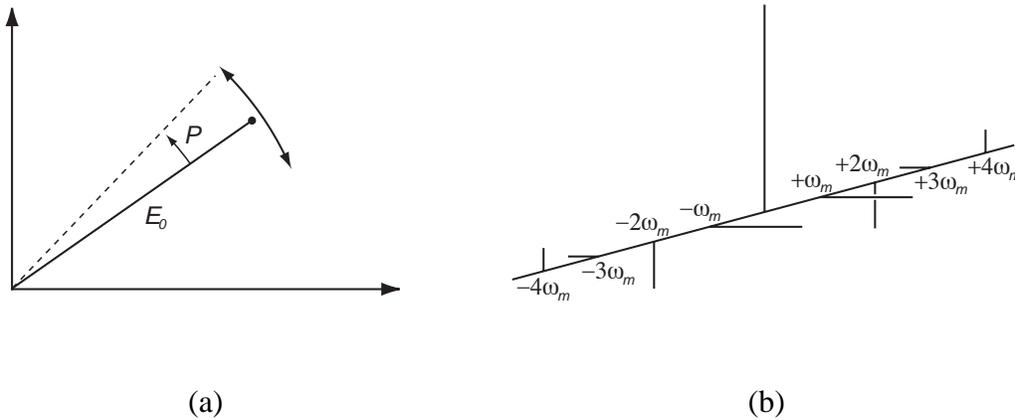
Fig. 3.7: True phase modulation as represented by (a) a phasor diagram and (b) a frequency vector diagram.

each subsequent term is two orders higher than the last. The latter fact means that the small-modulation-depth requirement is not as restrictive as it first seemed, and that the weaker mathematical constraint $|\tilde{P}|^2 \ll 1$ is adequate for neglecting all but the first terms. For example, the lowest order Bessel function, $J_0(z)$, has no linear term, so $J_0(z) \approx 1$ is accurate so long as the quadratic term is not too big.

Each harmonic at frequency $n\omega_m$ is represented in the summation of cosines in Eq. 3.23, each with a Bessel coefficient of order n, so that higher harmonics scale as higher powers of the modulation depth. The factor of i^n implies that at the second harmonic, fourth and sixth and so on, there is actually modulation of the *amplitude* kind, not the *phase* kind (see Fig. 3.7(b))! These higher order amplitude modulations are part and parcel of true phase modulation; they actually *prevent* the overall light power from oscillating. If one considers only the carrier and first-order sidebands (refer back to Fig. 3.2(b)) they will observe that, when the sidebands are perpendicular to the carrier, the resultant phasor is slightly longer than the carrier by itself, especially when the first-order sidebands are relatively large. The resulting oscillation in amplitude of the carrier phasor occurs at $2\omega_m$, and so true PM has amplitude modulation sidebands at $2\omega_m$ that exactly cancel this amplitude oscillation. There is a flow-on effect to higher order sidebands, each of which adds a finer correction to the shape of the modulation. Also, comparing Fig. 3.1(b) with

(a)   (b)
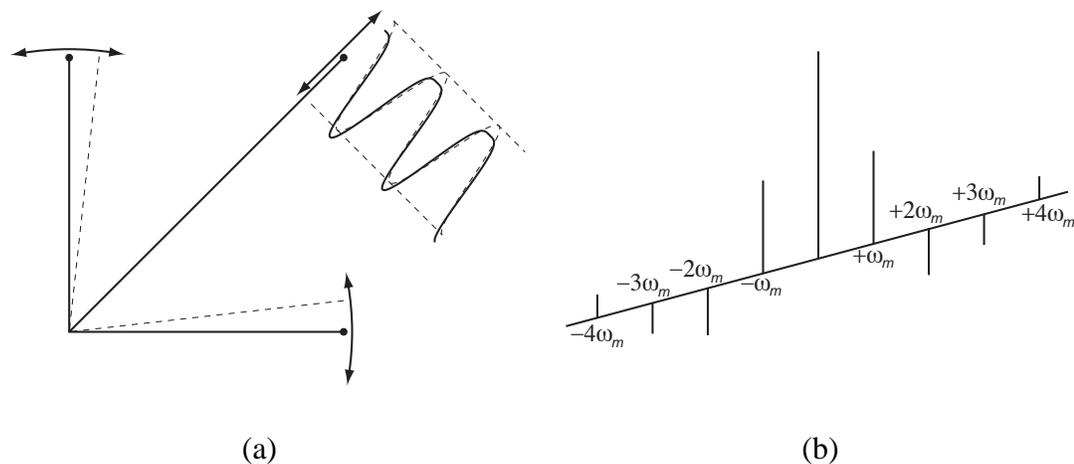
Fig. 3.8: True amplitude modulation as represented by (a) phasors, and by (b) frequency vectors.

Fig. 3.7(a), one sees that the optical phasor of first-order-approximated PM traces a line, whereas the optical phasor of true PM traces a circular arc.

A useful approximation is to take only the first term of each Bessel function:

$$J_n(z) \approx \frac{z^n}{2^n.n!}. \tag{3.25}$$

which gives the order of magnitude of each harmonic, again for small z. One has to be careful with this since, for example, the z^2 term in $J_2(z)$ is kept, but the z^2 term in $J_0(z)$ is discarded.

Since the AM devices introduced in the previous section make use of PM cells, the AM from these devices must also come with higher-order harmonics. Fig. 3.8(a) shows AM being generated from two orthogonal PM beams but, unlike Fig. 3.6, the modulation depth is stronger so that the PM phasors trace out circular arcs. While the resultant phasor oscillates purely in the amplitude direction (the phases of the PM beams still exactly cancel over the entire oscillation, so that there is no PM resultant), the shape of the AM waveform isn't quite sinusoidal. This is equivalent to having higher-order components (all of the AM type), as seen in Fig. 3.8(b).

Hopefully, it is now clear why the distinction was made between *true* and *defined* PM or AM, since most devices produce modulations that are quite complicated. The PM and AM definitions here are restricted to single frequencies, and so make natural building blocks with which to describe the multi-frequency modulations produced by real devices.

3.2.6 Modulation Measurements

The subject of modulation phenomenology was touched on in Section 3.1.2, where the comment was made that AM produces a measurable power variation but PM does not. This thread is continued in this section with descriptions of a number of readout systems for measuring the type, amount and quadrature of a modulation mounted on a laser beam.

Measurement of lasers invariably involves the use of a *photodetector*, a single photo-sensitive element that produces an electric current in proportion to the power of the light that lands on it. The current is converted to a voltage signal by op-amp filter circuits, and the output can then be analysed with other equipment, such as a cathode ray oscillo-
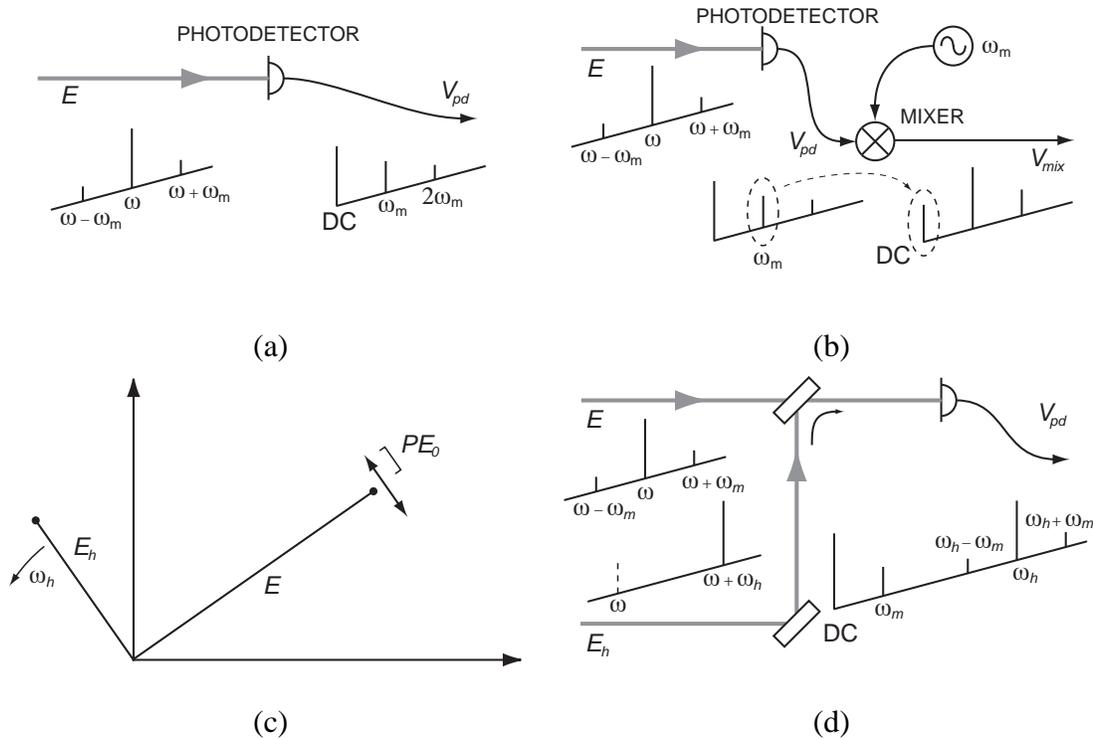
Fig. 3.9: Modulation measurement techniques. (a) Direct detection. (b) Detection and demodulation. (c) Phasor diagram and (d) frequency vector diagram depicting heterodyne detection.

scope or spectrum analyser. Without any extra elements being added, this direct-detection method is sensitive the presence of AM and not PM, since AM produces a measurable beat at the modulation frequency. The photodetector voltage, V_{pd}, from the detector is proportional to the beam's intensity I_{AM} from Eq. 3.4. Without a phase reference, this method is not very good for identifying the quadrature of the AM: a spectrum analyser only measures the magnitude of a frequency signal.

The direct detection method is illustrated schematically in Fig. 3.9(a), which shows explicitly a schematic frequency spectrum of the photodetector voltage, with DC, ω_m and $2\omega_m$ components as in Eq. 3.4. This representation is consistent with plots produced by a spectrum analyser, which gives the magnitude of frequency components averaged over a few milliseconds.

A slightly more complicated measurement called *demodulation* can give more information about the quadrature of a directly-detected AM component. The photodetector output is *mixed* with another oscillating signal of the same frequency (typically the same electronic signal used to generate the modulation is used); an electronic mixer component takes two signals and produces an output proportional to their product[6] (see Fig. 3.9(b)). A simple result from Fourier theory shows that this produces new frequency components at the sum and difference of the two input frequencies, ie at $2\omega_m$ and at DC. The DC component is of particular interest, and one says that a signal gets *mixed down* or *demodulated*

---

[6] Mixers consist of an arrangement of diodes that fire when the two input signals oscillate in-phase. Due to the nonlinear output of diodes, the output of a mixer is nearly always followed by a low-pass-filter. The reader can consult electronics catalogues for details on the workings of mixers, or [30] for a lucid description.

from ω_m to DC by this mixing process.

Mathematically, if the photodetector output is mixed with a signal proportional to $\cos(\omega_m t + \phi_{mix})$, then the output from the mixer is:

$$
\begin{aligned}
V_{mix} \;\; &\propto \;\; V_{pd} \times \cos(\omega_m t + \phi_{mix}) \\
&\propto \;\; E_0^2 A \cos(\phi_A - \phi_{mix}) + \text{h.o.}, \quad\quad\quad (3.26)
\end{aligned}
$$

where the constant DC only is shown for the sake of brevity. The mixer output hence gives a readout of the strength of AM in the beam; specifically it reads out the amplitude of AM that is *in-phase* with the mixing signal, as the above equation takes the form of a dot product of the two inputs. In principle, by varying the phase of the mixing signal, the magnitude and phase of the AM can be determined[7].

It is clear that, after a beam hits a photodetector, any PM information is lost, so PM detection must start before the photodetector. One general approach to remedy this is to interfere the phase-modulated beam (which will be called the *target beam*, as it is the target of the measurement) with another beam (called the *optical local oscillator*, a common term for a phase reference), prior to photodetection. The general principle of this approach is that, when the optical local oscillator is at 90° optical phase to the target beam, then the optical local oscillator will be in-phase with the target beam's PM oscillation (see Fig. 3.9(c)). Then, the optical local oscillator and target beam's PM combine like an amplitude modulation to produce a detectable oscillation, and hence the details of the PM can be measured.

If the optical local oscillator is of the same light frequency ω as the target beam, the method is called *homodyne detection*. Homodyne detection is beyond the scope of this thesis[8]. If the optical local oscillator and target beam have different frequencies, then the measurement is a *heterodyne detection*, which is discussed here in detail.

An optical heterodyne system makes use of a frequency-shifted local oscillator beam (of frequency $\omega + \omega_h$ for argument's sake), which beats with each individual sideband in the target beam with a different beat frequency. Fig. 3.9(d) shows that, at photodetection, these beats produce an electronic copy of the target beam's carrier and sidebands, centred around the heterodyne frequency ω_h. The photodetector signal also retains sufficient phase information to completely describe the original modulation state of the target beam. The experiment in Chapter 7 includes a demonstration of how a combination of heterodyne detection and demodulation can decode the modulation information from the photodetector signal.

3.3   The Modulation Sphere

In this section, a new theoretical representation for modulation is presented. The *modulation sphere* representation is a variation on the theme of geometric phase sphere represen-

---

[7] Electronic phase shifter circuits can be employed to vary the phase of an electronic oscillation, but more often one simply changes the length of coaxial cable until the relative phases are about right. For a signal at say 30 MHz, a few metres of cable is enough to delay the signal by a whole cycle.

[8] See for example [31].

tations, a theme which includes the Poincaré sphere representation for optical polarisation states [32], the Bloch sphere representation for the quantum state of coherently coupled two-level atoms [33], and the three space representation of optical evolution in directional couplers demonstrated by Korotky [34]. All these systems have a common algebra: a physical state is described by two complex quantities whose magnitudes and relative phases completely describe the physics of the system. The geometric phase representation of such systems involves transforming to three real-valued parameters which carry the same information as the two complex quantities, except for the overall phase of said quantities which is lost.

To move to the modulation sphere representation, one transforms a given modulation state from the complex quantities $\tilde{P}$ and $\tilde{A}$ to a set of real-valued coordinates (M_1, M_2, M_3) in *M-space* (see Fig. 3.10) via the transformation:

$$
\begin{aligned}
M_1 &= P^2 - A^2 \\
M_2 &= 2PA\cos(\phi_P - \phi_A) = 2\Re\{\tilde{P}\tilde{A}^*\} \\
M_3 &= 2PA\sin(\phi_P - \phi_A) = 2\Im\{\tilde{P}\tilde{A}^*\}
\end{aligned}
\tag{3.27}
$$

A single point in M-space represents a distinct modulation state. In particular, the M-space representation suppresses the common (beat) phase of the modulation state; the equations are defined only in terms of the relative phase difference between AM and PM contributions. This is a useful simplification in that now *every physically distinct modulation corresponds to one point and one only* in M-space. The downside is that one cannot use this representation for demodulation-phase-tracking calculations, since the common phase has been discarded. In demodulation situations, the physical significance of the modulation's overall phase is renewed by the presence of an external (electrical) phase reference.

The three M-space parameters can be interpreted as follows: M_1 is a measure of the extent to which one kind of modulation (AM or PM) dominates over the other; M_2 is a measure of the degree to which the present PM and AM components are correlated or anti-correlated in phase; and M_3 is a measure of the degree to which the present PM and AM components are in quadrature phase to each other, and thus also of the extent to which one frequency sideband has more power than the other.

The *modulation power* is defined by $M_0 = P^2 + A^2$, from whence $M_0^2 = M_1^2 + M_2^2 + M_3^2$ follows. Therefore, a surface of constant M_0 is a sphere in M-space. This is why the representation is called "modulation sphere", because if a modulation changes in such a way that it maintains a constant power M_0, then the trajectory of the modulation state in M-space will lie on the surface of a sphere.

Particular significance is attached to modulation states whose M-space coordinates lie on one cardinal axis. The point $(+M_0, 0, 0)$ represents pure PM, and the point $(-M_0, 0, 0)$ represents pure AM. The points $(0, \pm M_0, 0)$ represent correlated $(+)$ and anti-correlated $(-)$ PM and AM, and $(0, 0, \pm M_0)$ represent upper and lower single sideband states. These last two modulation states were encountered in Section 3.2.3, where they were referred to as "equal AM and PM in-phase" and "equal AM and PM in quadrature" respectively, but they appear in a new light, as part of a set of three basis vectors for modulation. Each *physically distinct* modulation state is comprised of some combination of the three types
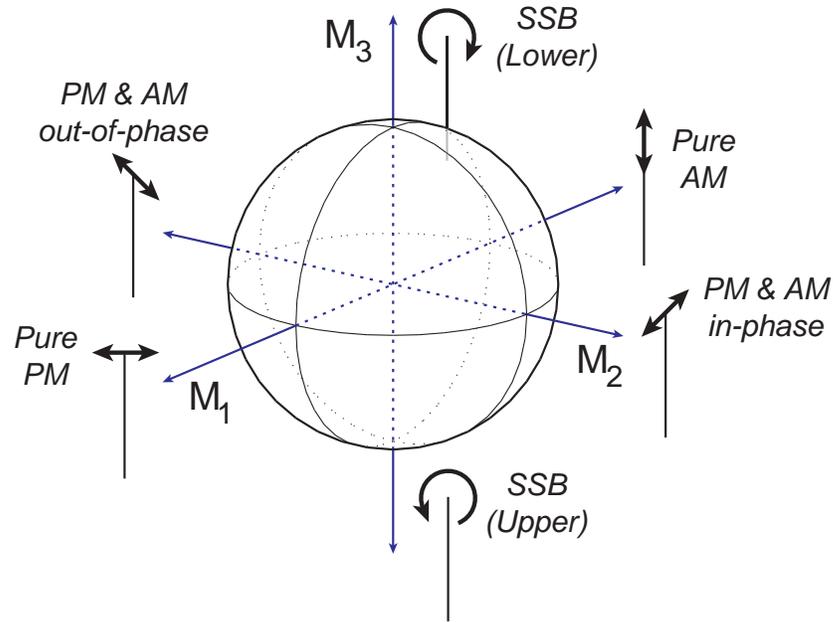
Fig. 3.10: The modulation sphere. Modulated phasors are shown as a visual reminder of the physical significance of the three M-space axes.

of modulation described by the real-valued parameters M_1, M_2 and M_3.

The occasion may arise when a reverse transform is required. This is not unique since the M-space representation does not include information about the overall modulation phase, but if such a phase is supplied (for example, the phase of the PM component, ϕ_P) then the original modulation phasors can be recovered via:

$$\tilde{P} = \sqrt{\frac{M_0 + M_1}{2}}\, e^{i\phi_P}$$
$$\tilde{A} = \frac{M_2 - iM_3}{\sqrt{2(M_0 + M_1)}}\, e^{i\phi_P}. \tag{3.28}$$

It is also possible to transform to the M-space basis from the basis of upper and lower sideband phasors:

$$M_1 = -4\Re\{\tilde{S}_{-\omega_m}\tilde{S}_{+\omega_m}\}$$
$$M_2 = +4\Im\{\tilde{S}_{-\omega_m}\tilde{S}_{+\omega_m}\}$$
$$M_3 = 2(S^2_{-\omega_m} - S^2_{+\omega_m}) \tag{3.29}$$

with $M_0 = 2(S^2_{-\omega_m} + S^2_{+\omega_m})$. There is a strong similarity between these transfer relations and those above in Eq. 3.27; the difference is that the M-space axes have been cyclically permuted[9]. Now M_3 represents the extent to which one sideband dominates over the

---

[9] There must logically be a third pair of basis phasors, with corresponding M-space transfer relations cyclically permutated once more, and whose physical significance correspond with the two ends of the M_2-axis. This third set of phasors do have their own significance as will be seen in Chapter 5.

(a)

(b)
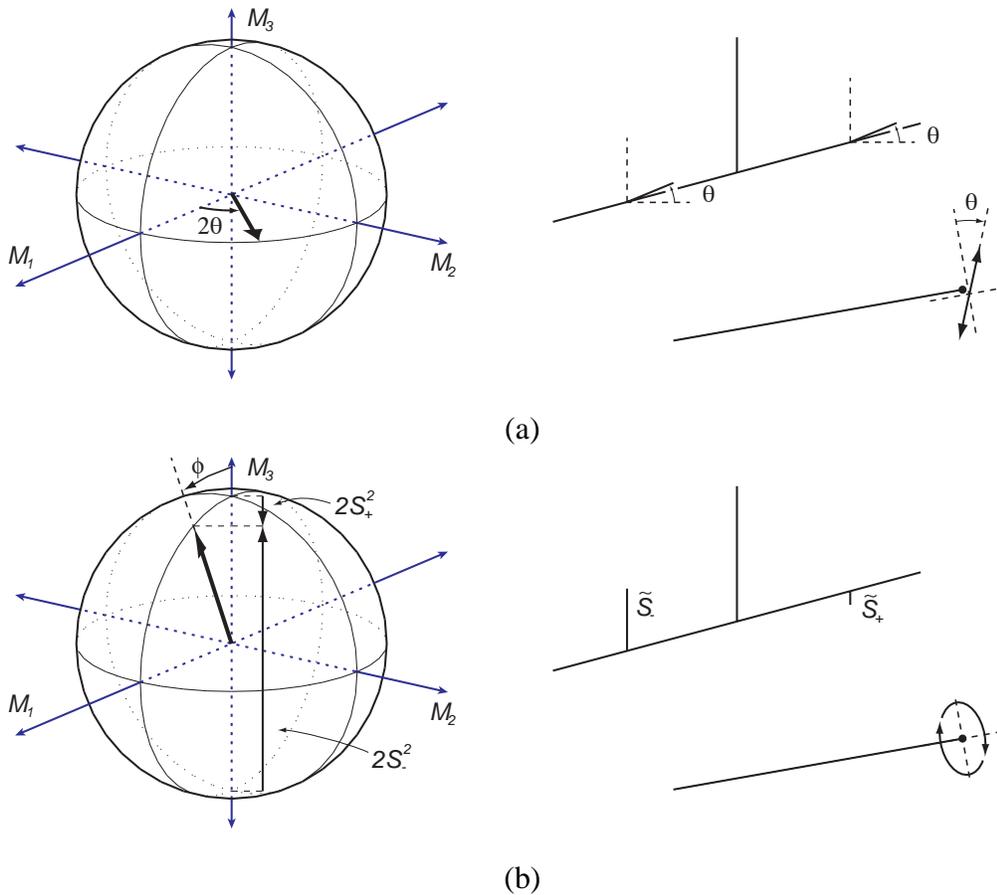
Fig. 3.11: Graphical links between the modulation sphere diagram and the phasor and frequency vector diagrams. (a) The angle of the M-space vector subtended in the M_1-M_2 plane is twice the angle that the phasor oscillation is set at. (b) The M_3 coordinate is related to the asymmetry in sideband power as shown: intervals from the M_3-intercept to the modulation sphere poles are in direct proportion to the sideband powers.

other; and M_1 and M_2 measure the amount of projected modulation where the averaged phase of the two sidebands is $0°$ and $45°$ respectively.

To illustrate the use of the modulation sphere, two examples follow with vectors in M-space that represent modulation states. The first example is a state composed of an unequal combination of PM and AM, combined in-phase. Mathematically, the state is written as $(M_1, M_2, 0)$, where $M_3 = 0$ implies that there is no quadrature combination of PM and AM, and also that the sidebands are balanced in power. A useful observation is that the angle that the modulation vector subtends with the M_1-axis varies as twice the angle that the sidebands make relative to the carrier, as shown in Fig. 3.11(a). This result holds for modulations vectors that don't lie in the M_1-M_2 plane; in this case, the carrier phasor traces out an ellipse, and the major axis of this ellipse is set at the angle in question.

The second example is a state composed of an unequal combination of PM and AM, combined in quadrature, represented $(M_1, 0, M_3)$. In this case, the nonzero M_3 component indicates that the modulation has unbalanced sidebands. In this case, it is possible to relate the polar angle of the modulation vector, ϕ, to the relative proportions of sideband

power, by:

$$\frac{|\tilde{S}_{-\omega m}|^2 - |\tilde{S}_{+\omega m}|^2}{|\tilde{S}_{-\omega m}|^2 + |\tilde{S}_{+\omega m}|^2} = \cos\phi = \frac{M_3}{M_0}. \tag{3.30}$$

which is obtained by manipulating the relations in Eq. 3.29. This relation has particular geometric significance: the vectors drawn from the two poles of the modulation sphere to the M_3 intercept of the modulation vector are proportional in magnitude to the powers of the two sidebands, as shown in Fig. 3.11(b). This result is true also for modulation vectors not confined to the M_1-M_3 plane.

To summarise the conclusions drawn from the above two examples, the azimuthal angle of a modulation vector in M-space is equal to twice the "major axis" of the modulation, and the polar angle gives a measure of the extent to which the modulation sidebands are unbalanced in power.

The modulation sphere by default only applies for one given frequency of modulation, and only describes the lowest order sidebands at that frequency. It is still possible to use multiple modulations spheres for different frequencies, or to use the same sphere with careful labelling when dealing with higher order harmonics. The latter option comes in handy in Chapter 5.

Lastly, the formalism here has a pitfall that is worth mentioning. The modulation mathematics given so far in this chapter is set up to deal with situations where the optical carrier is present. Things get complicated when the optical carrier goes to zero because the modulation parameters are defined as fractions of the carrier amplitude (refer back to Eq. 3.14, which has the form: *constant* $\times (1 + sidebands)$). In order to avoid having the modulations terms blow up (which is the consequence of defining them by their ratio to something that goes to zero), one has to bring some part of the *constant* back inside the parentheses, and define the *sidebands* as amplitudes in their own right.

However, this introduces another dilemma as the *constant* term often has a complex phase associated with it, which can cause confusion as to which oscillating terms represent AM and which represent PM. There is no natural way to deal with this once and for all, and it is best to just work around the problem as it arises, and bear in mind that scaling factors may need to be included before the parameters $\tilde{A}$ and $\tilde{P}$ correspond directly to measurables.

3.4   Modulation locking

In this section, the modulation formalism developed in this chapter is applied by demonstrating a method of locking a Michelson interferometer to a specific operating point. First, the mathematics of a Michelson interferometer are given, and then the method of modulation locking using the Schnupp asymmetry is described.

3.4.1   Michelson Interferometer

The optical transfer function of a Michelson interferometer is given here. The arm length L and arm length mismatch ΔL are defined as shown in the Michelson schematic in Fig. 3.12.
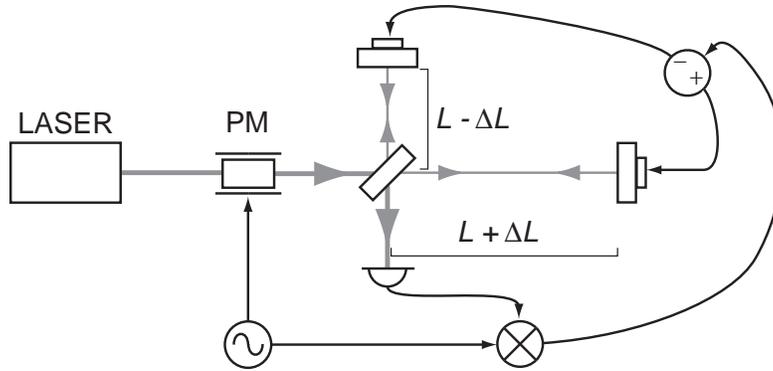
Fig. 3.12: Michelson interferometer with the Schnupp asymmetry, set up in a modulation locking configuration.

Only the longitudinal aspects of this device are considered here, ie it is implicitly assumed that spatial mode matching and focusing optics are taken care of. The exception is the following comment: the terminology of Michelson *fringes* arises from the optical patterns exhibited by interferometers that are misaligned or mode-mismatched, which appear as alternating bright and dark lines, or bright and dark concentric circles. On the other hand, the output of a well-aligned and modematched Michelson is entirely uniform and shows no spatial features. Still, the terminology is carried over so that when one speaks of the *fringe condition* of a Michelson, this refers to whether the entire beam output is bright or dark or in between.

The optical transfer function of a simple Michelson interferometer is[10]:

$$\frac{E_{\text{OUT}}}{E_{\text{IN}}} = 2t_{\text{BS}}r_{\text{BS}}e^{i\omega 2L/c}sin(\omega 2\Delta L/c). \qquad (3.31)$$

Variations of L and ΔL are variations in the so-called *common mode* and *differential mode*, respectively, of a Michelson, and provide a natural analysis basis for the device. The common mode and differential mode transmission responses of a Michelson are plotted in Fig. 3.13. Clearly, a change in L changes only the phase of the output, and a change in ΔL causes a sinusoidal change in the amplitude of the output with no net phase change.

The plots in Fig. 3.13 are also a function of optical frequency ω, so that different frequencies of light will encounter different transfer functions.

3.4.2 Fringe Lock of Michelson

As discussed in Section 2.3, a Michelson in a gravitational wave detector is typically locked to a dark fringe, so that the main light beam is totally reflected along its input path.

---

[10] I've cheated a little bit here by setting the interferometer to a dark fringe by default (ie the interferometer transmits zero light when $\Delta L = 0$), but actually if the Michelson arms are exactly equal, the interferometer is on a bright fringe. Mathematically, one arm needs to be a distance of $\lambda/4$ longer than the other to default to a dark fringe, ie to produce the transfer function here.
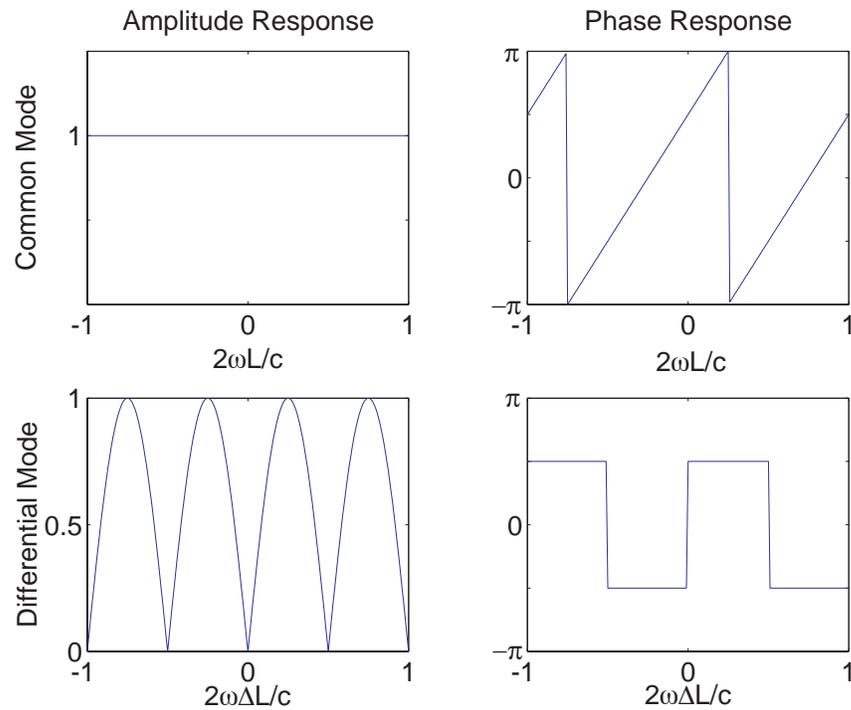
Fig. 3.13: Michelson interferometer response. A change in the common mode produces only a phase change. A change in the differential mode produces only a sinusoidal amplitude change.

Referring to Fig. 3.12, the input beam is phase modulated at some frequency ω_m, and a feedback signal is obtained from the beat signal in the detected light at the output port (also known as the *dark port*). The beat signal at the output port is demodulated down to baseband as described in Section 3.2.6 to produce an error signal. The demodulation is done in *quadrature*; that is the electrical signal used in the mixdown is shifted by 90° relative to the signal powering the input PM device.

In particular, for the input modulation to appear at the dark port, a large arm length mismatch ΔL must be introduced. This removes the degeneracy of the Michelson transfer function with respect to frequency, so that the modulation sidebands are partially transmitted when the carrier is at a dark fringe. The idea for this asymmetry is due to Schnupp[11] [36].

To mathematically describe the error $\phi(t)$, ie the phase offset of how far the Michelson is away from a dark fringe, the arm length mismatch is split into fixed and variable components[12]:

$$\Delta L(t) \leftarrow \Delta L + \frac{\phi(t)c}{2\omega}. \tag{3.32}$$

For an input phase modulation of $\tilde{P}_{\text{IN}}$, the electronic error signal after demodulation

---

[11] Previous systems did not need an asymmetry as they used a pair of phase modulators *inside* the Michelson, an idea attributed to Weiss [35].

[12] The equation here isn't strictly valid as the error ϕ should have units of length, being a mechanical offset; problems arise if the light frequencies ω in the system differ considerably. However, so long as $\omega_m/\omega \ll 1$, the above is a valid approximation. The ratio ω_m/ω in most experiments is of order 10^{-6} to 10^{-9}.
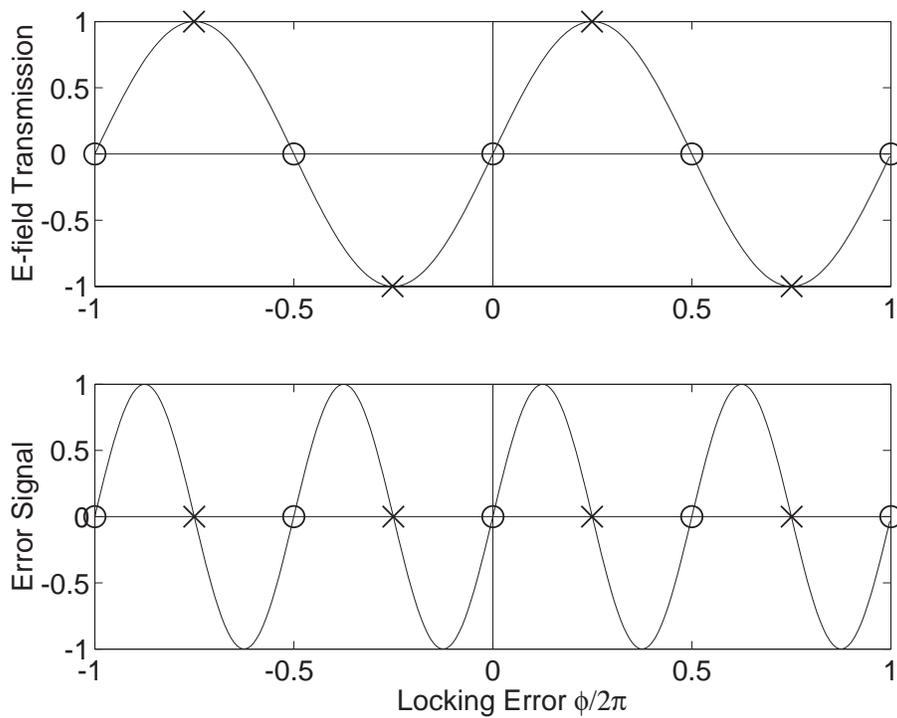
Fig. 3.14: Michelson electric field transfer function, with modulation-induced error signal (both arbitrary units). Positive sloping zero-crossings in the error signal can be used to lock to dark fringes (circles); negative sloping zero-crossings will lock to bright fringes (crosses).

is given by:

$$V_{\text{mix}} \propto 4t_{\text{BS}}^2 r_{\text{BS}}^2 E_0^2 \left[\sin\left(2\omega_m \Delta L/c\right) \sin\left(2\phi\right) \Re\{ie^{-\phi_{\text{mix}}} e^{i\omega_m 2L/c} \tilde{P}_{\text{IN}}\} \right] \qquad (3.33)$$

and is plotted in Fig. 3.14. The derivation of this equation is deferred until Chapter 6 where the Michelson's response to arbitrary modulation is calculated. The following is a qualitative explanation of the origin of this error signal.

An important point that is often missed is that the input PM is converted by the (dark fringe) Michelson *entirely* into AM. This can be seen in Fig. 3.15(a), where the transfer function of the Michelson, with respect to input frequency ω, is antisymmetric about a dark fringe, and so one of the sidebands receives a $180°$ phase shift relative to the other. The shape of the error signal near the dark fringe is then determined by the amount of carrier transmitted, since this contributes linearly to the strength of the beat measured at the output port.

It is also possible to lock to a bright fringe using this system; the error signal has a corresponding zero-crossing at bright fringes, with opposite polarity to those of dark fringes. However, near the bright fringe, the PM sidebands are preserved as PM on transmission through the Michelson (the response is symmetrical about a bright fringe). If the carrier moves to one side of the bright fringe, one of the sidebands will be transmitted with greater amplitude than the other (see Fig. 3.15(b)). This introduces a differential in the PM sidebands, which equates to the introduction of a component of quadrature AM. The carrier (which stays more or less at 100% near the bright fringe) beats with this AM
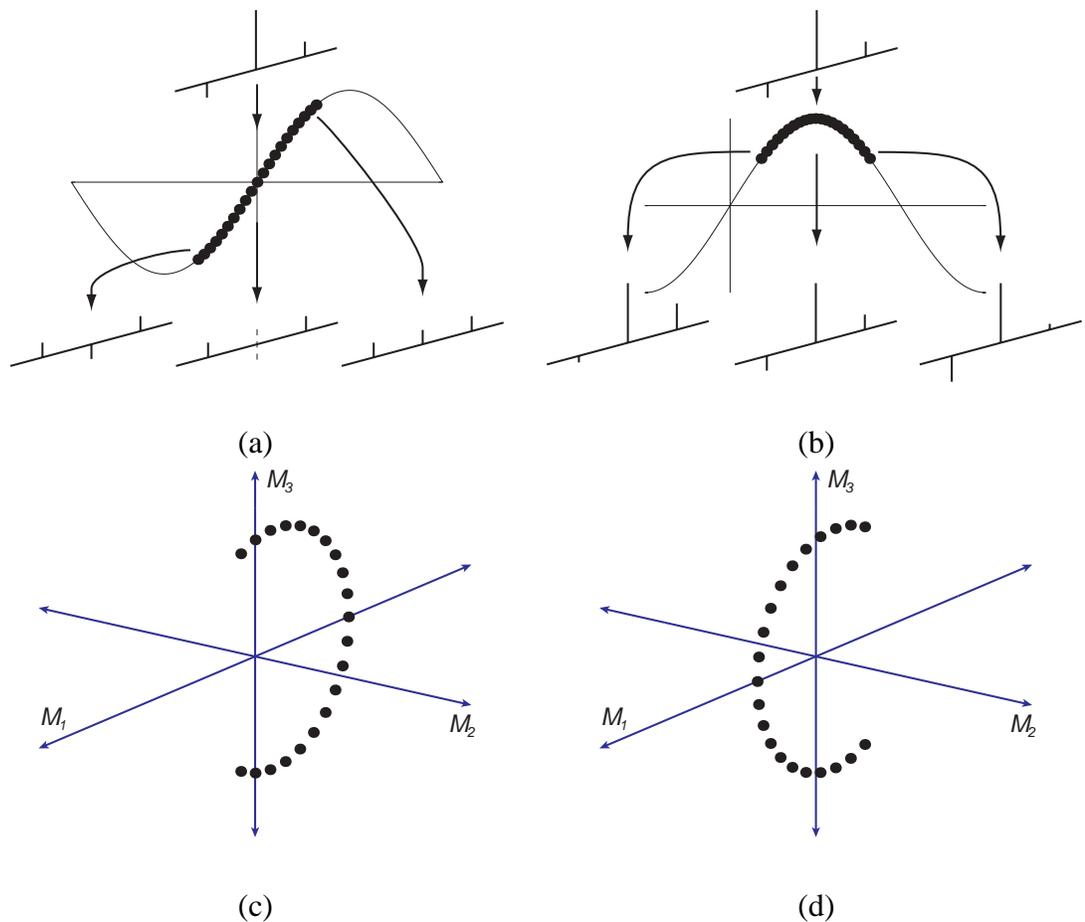
Fig. 3.15: Frequency vector diagrams showing the Michelson's response to a modulated beam (a) near a dark fringe, and (b) near a bright fringe, overlayed on the Michelson's electric field transfer function. (c) and (d) show the M-space trajectories of the output modulations near dark and bright fringes, with similarly spaced data points.

component to produce the linear error signal shape in this regime.

Fig. 3.15(c) and (d) show the M-space trajectories of the modulation states produced respectively near dark and bright fringe locking points. As the input PM sidebands are only differentially attenuated, and never phase-shifted relative to each other, so the trajectories stay in the M_1-M_3 plane. This fact will become important later, as the unused M_2 axis gives us an independent degree of freedom with which to control and lock other parts of an optical system. Also, the points in M-space are evenly spaced: in fact the phase angle of the M-space coordinates goes as twice the phase offset ϕ of the Michelson's operating point.

While the above treatment of Michelson locking has served as a good opportunity to exercise the modulation formalism developed in this chapter, it is also prerequisite information for what follows. In particular, Chapter 6 shows how a modification to the input modulation state can be used to shift the lock point of a Michelson. The intervening chapters cover the equivalent requisite knowledge for cavities, and also describe a modulation device capable of providing the input modulation required for such a tune-locking set up.

3.5   Summary

In this chapter, modulation of a light beam was introduced and two physically distinct types, amplitude modulation (AM) and phase modulation (PM), were identified. A number of graphical visualisation tools were presented to aid the user's understanding. The significance of modulation quadratures and modulation sidebands were discussed and then demonstrated via combinations of AM and PM. Devices that mount modulation on light beams were described, and it was shown that these devices produce higher order sidebands in addition to their fundamental harmonic modulations. Several ways of detecting modulations on a light beam were described.

The *modulation sphere* formulation, a geometric phase representation for modulation, was introduced and demonstrated. The representation involves transforming to three *M-space* parameters which represent *three* physically distinct modulation types possible in an arbitrary combination of PM and AM. The representation suppresses the common phase of the modulation, so that modulations that differ only by a time shift appear degenerate.

The transfer function of a Michelson interferometer was given, and a method was described of locking the Michelson to a dark or bright fringe using modulation and the Schnupp asymmetry. The locking technique was analysed with both the standard and geometric phase representations of modulation.

*"You must be able to talk fluently about the subject, or else you
don't know anything about it."*

— Professor Julius Sumner Miller

Fabry-Perot Cavity Locking Techniques

4.1 Context

A gravitational wave interferometer requires a laser input that is as close to being stationary as possible. Amongst other things, this requires the minimisation of intensity and frequency fluctuations, of directional fluctuations or "jitter", and the removal of higher order spatial modes. One approach is to include one or more *mode cleaner* ring cavities [37] in the laser preparation stage; the output of a passive ring cavity is typically much more stable in intensity and frequency, and spatially "clean", than the output of an active lasing cavity. The mode cleaner cavity is employed by research and detection groups in various combinations: the LIGO [15] and GEO [17] groups use two mode cleaners in series, the Virgo [20] group use one mode cleaner in parallel with a reference cavity, the GEO and Virgo groups use output mode cleaners, and various groups use mode cleaners as part of detector research and development.

High laser input powers improve the sensitivity of a gravitational wave interferometer with respect to quantum noise [2]. As much as 200W of laser input power is planned for Advanced LIGO, with kilowatts of power resonating within power cavities and arm cavities [38, 39]. Such high powers can cause difficulties: transmissive elements such as substrates, modulators and isolators, become heated causing thermal lensing and distortion [40]. Careful testing and design are required to compensate for these effects, especially considering that detector alignment and lock acquisition phases are carried out with varying power levels. High laser powers can also potentially damage detectors used in cavity locking schemes.

In this chapter, a variety of techniques are discussed for locking optical cavities that avoid these problems. These techniques are intended for use in input and output mode cleaner cavities, and reference cavities, for gravitational wave detectors and associated research experiments. First, in Section 4.2, Fabry-Perot cavities and Gaussian beam optics are reviewed, and the Pound-Drever-Hall method of Fabry-Perot locking is described in Section 4.3. In Section 4.4, a rearrangement of the optical components is proposed, known as *double pass locking*, where the detector is put safely beyond the reach of the high power input beam. Section 4.5 discusses two locking techniques based on spatial mode interference which do not require a modulator, thus removing one potentially problematic transmissive element. Specifically, *tilt locking* is discussed in Section 4.5.1, and a new variation on the theme, *flip locking*, is discussed in Section 4.5.2.
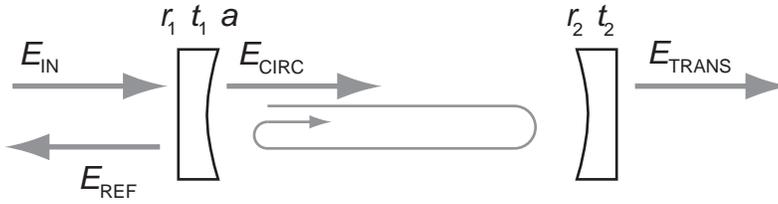
Fig. 4.1: Fabry-Perot cavity schematic. An input laser beam E_{IN} enters the cavity and produces a resonating field E_{CIRC}, which couples out to a transmitted beam E_{TRANS} and a reflected beam E_{REF}. Mirror transmissivities t and reflectivities r are defined as shown, with losses a occuring at the cavity entrance for the sake of convenience in the model. The relationships between these parameters are derived in the text. The self-consistent convention of assigning a factor of r to reflections and it to transmissions is used.

4.2 Fabry-Perot Optics

The following is a rigorous description of Fabry-Perot optical resonator cavities, and their utility as mode cleaners, including the standard mathematical description of cavities and Gaussian beams.

A Fabry-Perot cavity, shown in Fig. 4.1, consists of two partially-transmitting mirrors aligned such that a beam of laser light entering through one mirror will rebound from the second mirror back to its original entry point. The essence of a Fabry-Perot cavity is that, when the distance between the two mirrors equals an exact multiple or half-multiple of the wavelength of the light, a bright standing wave forms inside the cavity. Under these circumstances, one says that the light-and-cavity system is *on resonance*, or that the light is *resonant* inside the cavity.

4.2.1 Longitudinal modes

Provided certain geometric requirements are met, a cavity is capable of resonating light in a number of *modes*. For example, there is an infinite array of possible wavelengths λ (and frequency $\nu = c/\lambda = \omega/2\pi$) of light that could form standing waves inside a cavity of length L, the requirement being that $L = q\lambda/2$ where q is an integer. Each frequency of light capable of resonating has one corresponding *longitudinal mode* in the cavity. Consecutive longitudinal modes are separated by a frequency gap of $\nu_F = c/2L$, a parameter called the *free spectral range* (FSR).

A brief derivation of a Fabry-Perot cavity's transfer equations follows. The parameter conventions indicated in Fig. 4.1 are used, and the approach is taken of solving for the steady state laser field amplitudes, in and around the cavity, after any initial transients have died out [41]. First, the electric field just inside the cavity must be equal to the fraction of light entering the cavity plus the component still present from the previous round trip:

$$E_{CIRC} = E_{IN} i t_1 + E_{CIRC} r_2 r_1 e^{i2\omega L/c} \tag{4.1}$$

or, upon rearrangement,

$$\frac{E_{CIRC}}{E_{IN}} = \frac{it_1}{1 - r_1 r_2 e^{i2\omega L/c}}. \tag{4.2}$$
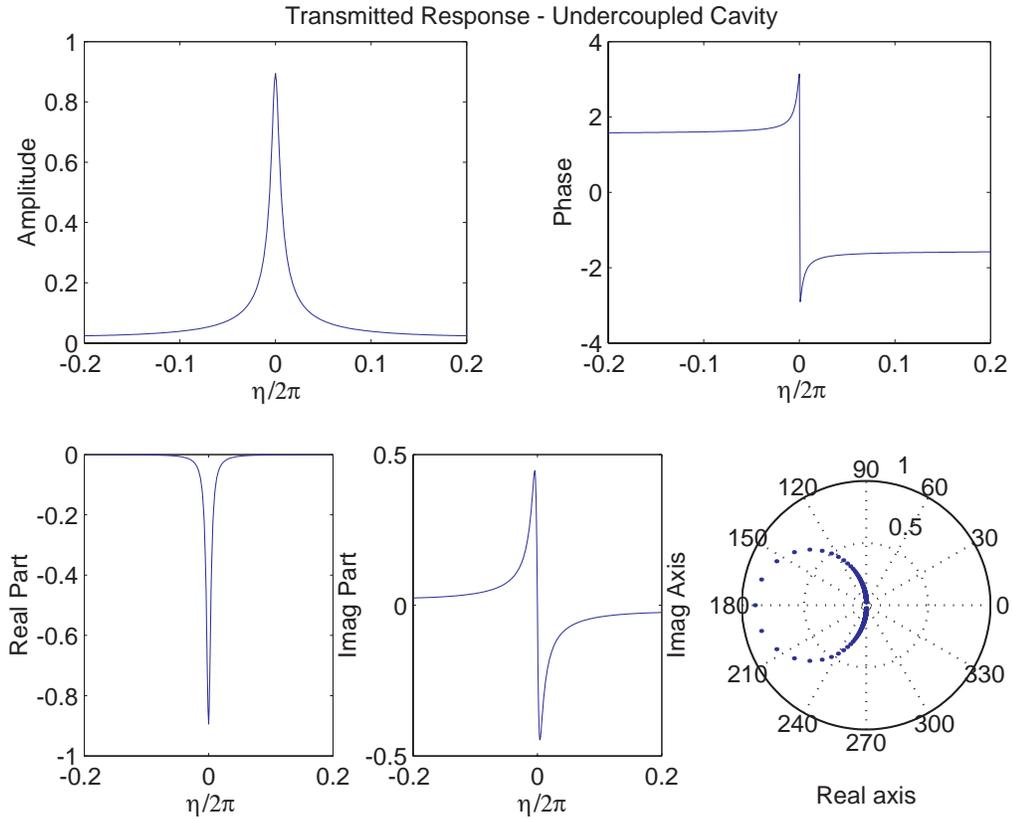
Fig. 4.2: Cavity response in transmission $\frac{E_{\text{TRANS}}}{E_{\text{IN}}}$ as a function of the parameter $\eta \equiv 2\omega L/c$, for $r_1 = r_2 = \sqrt{0.95}$ and $a = \sqrt{0.01}$ (an undercoupled cavity). The transmitted amplitude is negligibly small except when the parameter η is near a multiple of 2π. The bottom right plot is a polar diagram showing how the electric field phasor traces out a circular arc in the complex plane on sweeping through resonance.

Similarly, the transmitted field is equal to the fraction of the circulating field that leaks out of the second mirror:

$$E_{\text{TRANS}} = E_{\text{CIRC}} it_2 e^{i\omega L/c} \tag{4.3}$$

$$\frac{E_{\text{TRANS}}}{E_{\text{IN}}} = \frac{-t_1 t_2 e^{i\omega L/c}}{1 - r_1 r_2 e^{i2\omega L/c}}. \tag{4.4}$$

Finally, the reflected field is the field directly reflected off the first mirror added to the component of circulating field leaking out of the first mirror:

$$E_{\text{REF}} = E_{\text{IN}} r_1 + E_{\text{CIRC}} it_1 r_2 e^{i2\omega L/c} \tag{4.5}$$

$$\frac{E_{\text{REF}}}{E_{\text{IN}}} = \frac{r_1 - r_2(r_1^2 + t_1^2)e^{i2\omega L/c}}{1 - r_1 r_2 e^{i2\omega L/c}}. \tag{4.6}$$

In Fig. 4.2 and Fig. 4.3, the transmitted and reflected responses are plotted near resonance as a function of η (which varies linearly with both L and ω), for a cavity with reflectivities $r_1 = r_2 = \sqrt{0.95}$ and loss coefficient $a = \sqrt{0.01}$ (where power conservation requires that $r_1^2 + t_1^2 + a^2 = 1$). This cavity is an *undercoupled* cavity and is so called because the amount of light that reflects directly off the front mirror is more than the amount
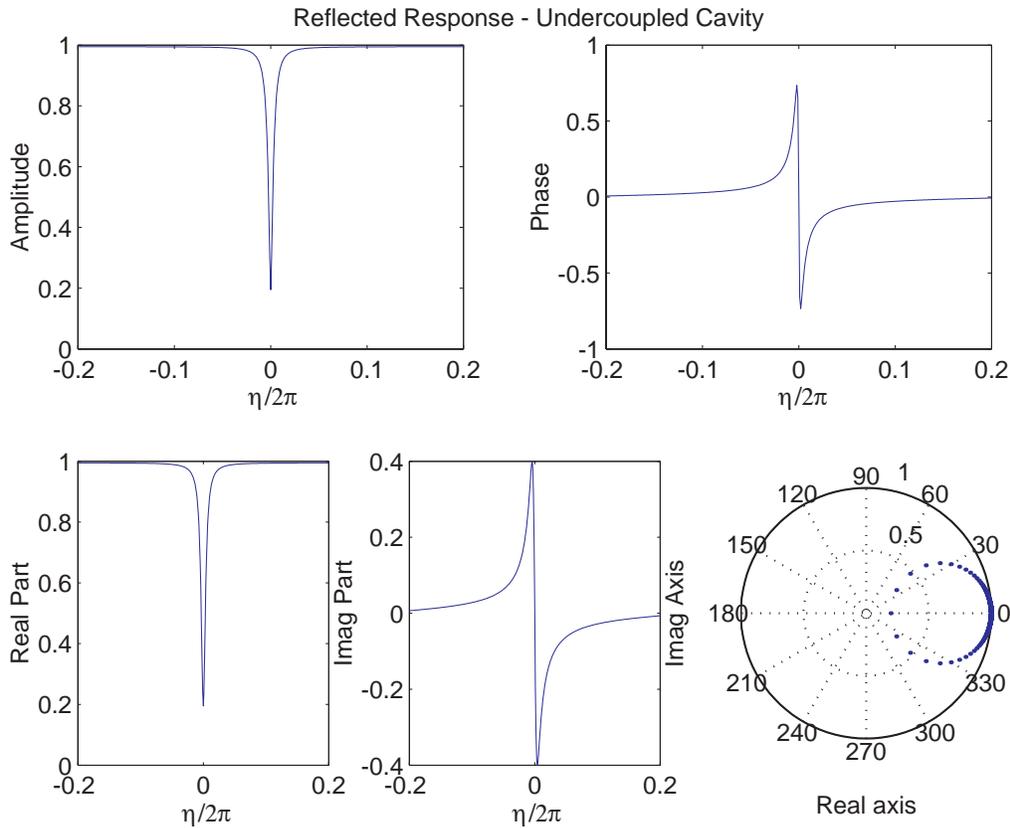
Fig. 4.3: Cavity response in reflection $\frac{E_{REF}}{E_{IN}}$ as a function of the parameter $\eta \equiv 2\omega L/c$, for $r_1 = r_2 = \sqrt{0.95}$ and $a = \sqrt{0.01}$ (an undercoupled cavity). The phase dispersion in reflection is exploited later to produce a locking signal (see Section 4.3).

of light that couples into the cavity, circulates and couples out the front mirror again. This corresponds mathematically to Eq. 4.6: the first term in the numerator, r_1, is bigger in magnitude than the second term, $r_2(r_1^2 + t_1^2) = r_2(1 - a^2)$.

If the situation were to be reversed, that is if the amount of light reflecting directly off the front mirror were less than the amount coupling out after circulating in the cavity, the cavity would be called *overcoupled*; the front mirror has sufficiently low reflectivity r_1 so that "too much" light is coupled into the cavity. The reflected response near resonance for a cavity with $r_1 = \sqrt{0.93}$, $r_2 = \sqrt{0.97}$ and $a = \sqrt{0.01}$ is plotted in Fig. 4.4.

Both undercoupled and overcoupled cavities never quite transmit all of the light from one side of the cavity to the other, even on resonance, because some small amount of light is still reflected. If the reflectivities are chosen such that the reflected components destructively interfere and cancel exactly, then the cavity is *critically coupled* or *impedance matched* (a term borrowed from electronics). Theoretically, equal reflectivities $r_1 = r_2$ and no losses $a = 0$ achieves impedance matching, but real cavities always have some component of loss that must be compensated for by a slight imbalance in reflectivities.

4.2.2 Transverse modes

Thus far, the discussion has centred on the length-wise characteristics of cavities. The transverse properties are equally important. Fabry-Perot cavities permit the resonance of
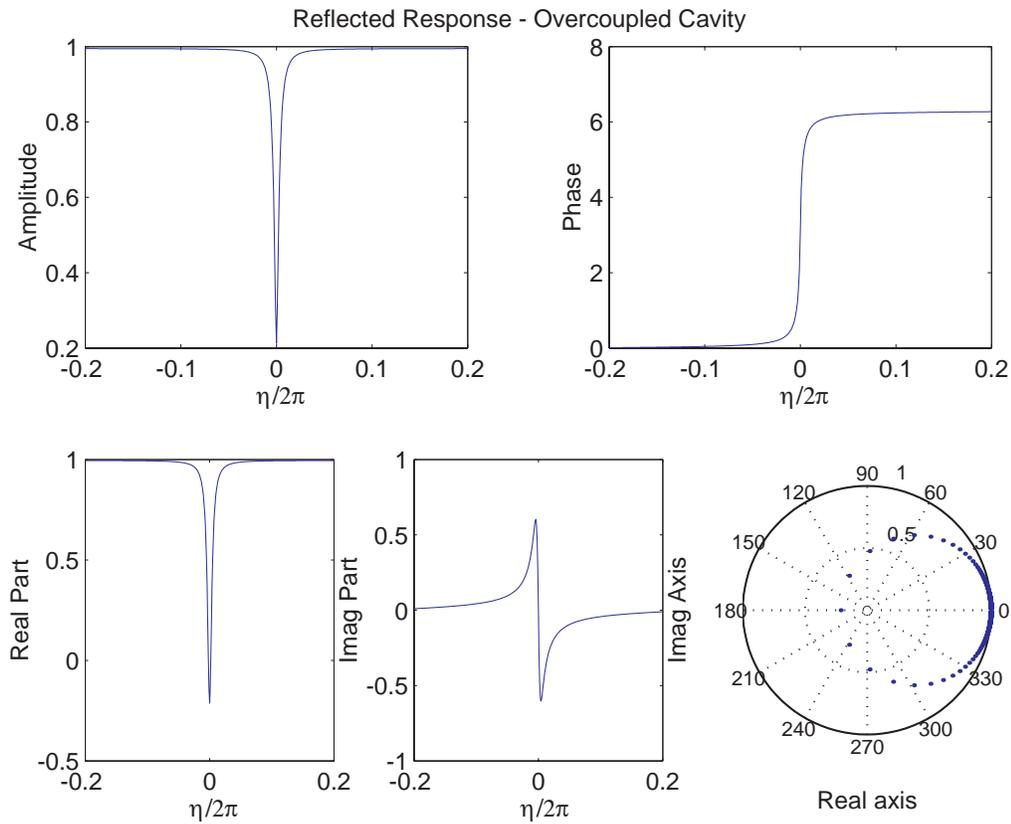
Fig. 4.4: Cavity response in reflection $\frac{E_{\text{REF}}}{E_{\text{IN}}}$ as a function of the parameter $\eta \equiv 2\omega L/c$, for $r_1 = \sqrt{0.93}$, $r_2 = \sqrt{0.97}$ and $a = \sqrt{0.01}$ (an overcoupled cavity). Here, the phase executes a complete $360°$ rotation on sweeping through resonance. The imaginary component, however, maintains the same shape as in the undercoupled case.

particular *spatial modes*, a term which describes a resonating laser beam whose transverse profile remains the same after a complete circuit of the Fabry-Perot. The family of spatial modes permitted by a Fabry-Perot are called the *Hermite-Gaussian modes*, whose transverse profile is described by the product of a Gaussian and a Hermite polynomial:

$$G_l(u) = H_l(u) exp\left(\frac{-u^2}{2}\right), \qquad l = 0, 1, 2, \dots \qquad (4.7)$$

(see Fig. 4.5(a)) where the first few Hermite polynomials are

$$H_0(u) = 1, \quad H_1(u) = 2u, \quad H_2(u) = 4u^2 - 2, \quad H_3(u) = 8u^3 - 12u, \qquad (4.8)$$

and higher order Hermite polynomials can be found using the recurrence relation:

$$H_{l+1}(u) = 2uH_l(u) - 2lH_{l-1}(u). \qquad (4.9)$$

Hermite-Gaussian modes have special significance even without consideration of Fabry-Perot cavities: they maintain the same transverse profile along their entire length when propagating in free space. The dimensions of the beam, however are not constant;

(a)

TEM$_{00}$ TEM$_{01}$ TEM$_{10}$ TEM$_{11}$

(b)
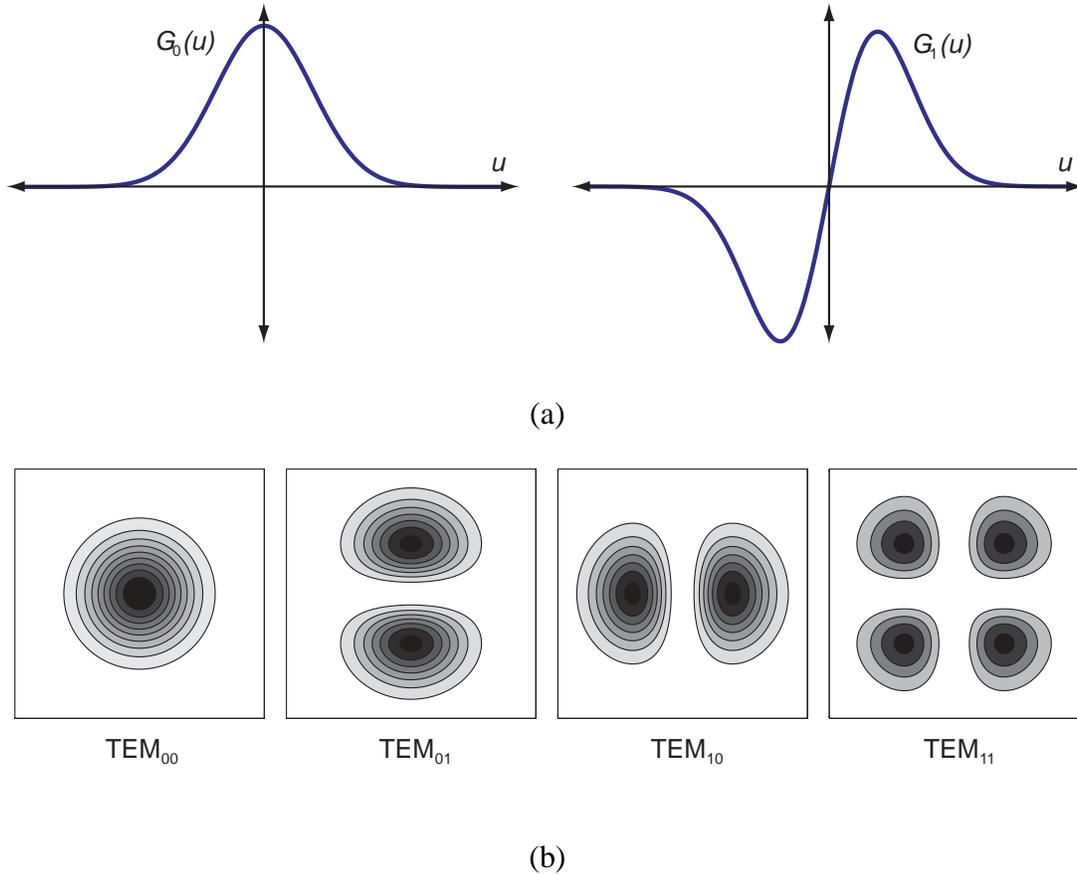
Fig. 4.5: (a) The first two Hermite-Gaussian functions. (b) Two-dimensional intensity profiles of some low order spatial modes.

while the transverse shape of the beam remains the same, the size of the beam diverges (or converges) steadily. The beam width parameter, $W(z)$, represents the transverse radius within which most of the beam's power is contained (for the lowest order Gaussian beam, this represents $1 - 1/e^2 \approx 86\%$):

$$W(z) = W_0 \left[1 + \left(\frac{z}{z_0} \right)^2 \right]^{1/2},$$
(4.10)

(see Fig. 4.6(a)) where z represents the distance along the length of the beam, and $z = 0$ is the location of the narrowest part of the beam, called the *waist* which has half-width W_0. The Rayleigh length, z_0, represents the distance from the waist required for the beam to diverge by a factor of $\sqrt{2}$. The waist radius and the Rayleigh length are related by:

$$W_0 = \left(\frac{\lambda z_0}{\pi} \right)^{1/2}.$$
(4.11)

The radius of curvature of wavefronts also changes steadily along the length of a Hermite-Gaussian beam:

$$R(z) = z \left[1 + \left(\frac{z_0}{z} \right)^2 \right]$$
(4.12)
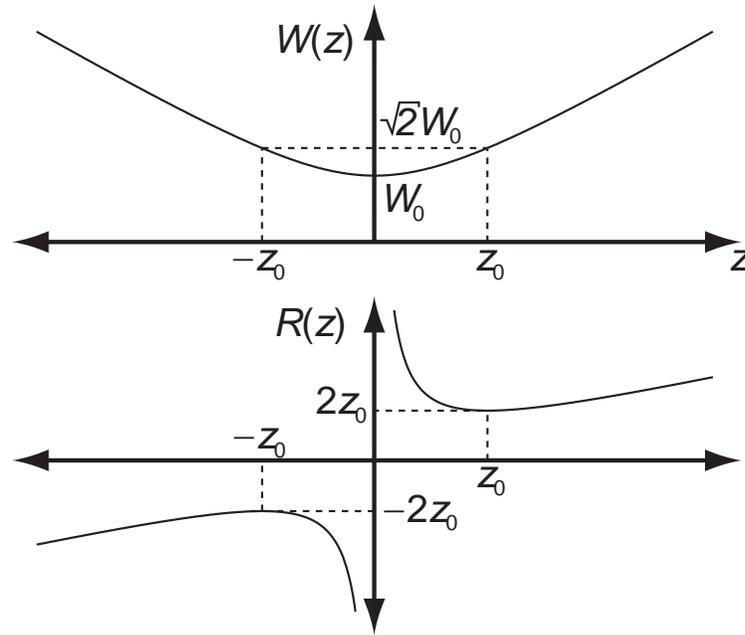
Fig. 4.6: Waist half-width $W(z)$ and radius of curvature $R(z)$ as a function of distance along a Gaussian beam's length.

(see Fig. 4.6(b)).

Armed with the above definitions, the full expression for the electric field of a Hermite-Gaussian beam can be written, and is:

$$U_{lm}(x,y,z) = A \left[\frac{W_0}{W(z)}\right] G_l \left[\frac{\sqrt{2}x}{W(z)}\right] G_m \left[\frac{\sqrt{2}y}{W(z)}\right]$$

$$\times exp \left[-ikz - ik\frac{x^2+y^2}{2R(z)} + i(l+m+1)\zeta(z)\right], \tag{4.13}$$

where $k = 1/\lambda$ is the wavenumber, and A is an arbitrary constant.

Mathematically speaking, the Hermite-Gaussian modes comprise a complete set of solutions to the paraxial Helmholtz equation. Any travelling wave that satisfies the paraxial approximation (beam wavefronts are approximately planar) can be written as a superposition of Hermite-Gaussian modes.

Note that the Hermite-Gaussian functions in Eq. 4.13 run in two directions, such that the indices l and m correspond to the x and y directions respectively. It is conventional to refer to spatial modes by these indices in the format TEM$_{lm}$, for example TEM$_{00}$ for the fundamental Gaussian beam. TEM stands for Transverse Electric-Magnetic, which is a term borrowed from fibre-optics. Fig. 4.5(b) shows the two-dimensional profile of the lower order spatial modes.

Other complete bases exist for describing localised travelling waves. One basis is the Laguerre-Gaussian modes, which also have profiles that are invariant with respect to propagation, and also self-replicate in a Fabry-Perot cavity, but have cylindrical symmetry as compared to the rectangular symmetry of the Hermite-Gaussian modes. The reason for
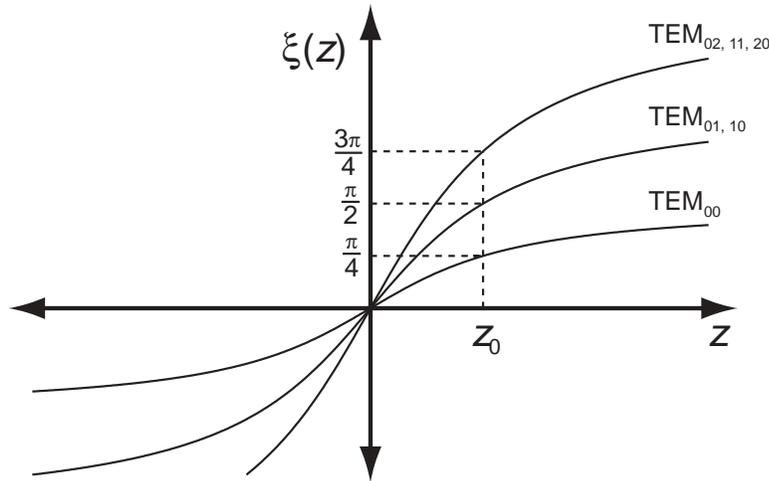
Fig. 4.7: Guoy phase shift $\zeta(z)$ as a function of distance along a Gaussian beam's length. The Guoy phase is different for different spatial modes, and can be exploited to ensure that only selected modes resonate inside a cavity at once.

the coexistence of these bases is that a certain degeneracy exists for groups of higher-order modes, and there are multiple ways to choose a set of basis modes to span each of these degenerate spaces. A more detailed explanation follows shortly.

4.2.3   A Fabry-Perot cavity as a mode cleaner

There remains one term in Eq. 4.13 still unexplained: the term $\zeta(z)$ is called the *Guoy phase* and represents a phase retardation compared with the phase of a plane wave:

$$\zeta(z) = \tan^{-1}\frac{z}{z_0} \qquad (4.14)$$

The rate that the Guoy phase accumulates is different for different spatial modes TEM_{lm}, as can be seen in Eq. 4.13 where the Guoy phase term is multiplied by the sum $(l+m+1)$ (see Fig. 4.7). This fact can be exploited to ensure that only one spatial mode is resonant inside a cavity at once. Due to the Guoy effect, all other modes will accumulate, during a round trip, a phase that is *not* equal to a multiple of π radians, and will hence destructively interfere[1]. Hence, the term *mode cleaner* describes a cavity that is employed to filter off higher-order spatial modes from a beam and pass only the TEM_{00} component.

Technically, modes can co-resonate if their indices, l and m, have the same sum, for example TEM_{01} and TEM_{10} , because their accumulated Guoy phase is the same. Such modes are degenerate with respect to propagation, ie a linear combination will also propagate with a constant profile. It is from this degeneracy that multiple bases arise, as discussed above.

---

[1] Linear cavities have this property provided that they are not *confocal*, meaning that both mirrors share a common focus.

Mode cleaner cavities also "clean" a beam in terms of reducing intensity and frequency fluctuations. Typically, a laser's frequency is locked to a mode cleaner cavity because the latter can be made more immune to temperature drift. Also, a high finesse cavity will filter out amplitude fluctuations of the input light, since the resonating light will take time to respond to the change in input. Moreover, the output of a laser can fluctuate in its emitted direction which, when the light is incident on an experiment, can produce an effect similar to the presence of higher-order spatial modes. The use of a passive Fabry-Perot removes this effect, only passing the light that matches its preferred resonating mode.

In some cases, a small tap-off of a laser is redirected and locked to a *reference cavity*, and the remainder of the beam, now frequency-stabilised, is used for the main experiment. Also, mode cleaners can be employed at the output of an experiment, if the experiment produces unwanted higher-order modes.

Gravitational wave detectors variously employ input mode cleaners, output mode cleaners and reference cavities. Mode cleaners usually consist of rigidly spaced mirrors. Alternatively, each mirror can be independently suspended and isolated under vacuum to form a *suspended mode cleaner*. Suspended mode cleaners, when properly engineered, can produce beams far cleaner than rigid, Earth-mounted mode cleaners, as they are isolated from seismic, and acoustic disturbances, and can achieve much higher finesses.

The locking techniques described in this chapter apply particularly to rigid mode cleaners, but in some cases may be employed with the suspended variety. Each example is discussed in turn.

4.3   Pound-Drever-Hall locking

Invariably, the desired operating point to lock a cavity to is exact resonance, whether the cavity is for mode-cleaning purposes or for resonating light in the arms of a gravitational wave detector[2]. The most commonly used method of locking a cavity to resonance is called *Pound-Drever-Hall* (PDH) locking after the authors that originally presented the idea [42]. A review is given here as a lead in to the alternative cavity locking techniques discussed later in this chapter. PDH locking is also extended to include a tuning feature in Chapter 6.

Most but not all cavity locking techniques exploit the dispersive phase shift given to a cavity's reflected output (refer back to Fig. 4.3) to derive an error signal giving the cavity's displacement from resonance. PDH locking does this by phase-modulating the input beam. A typical PDH locking system [42] for a ring mode cleaner cavity is shown in Fig. 4.8(a).

The modulation frequency is selected such that the modulation sidebands are well away from any cavity resonance when the carrier is near resonance. The modulation sidebands then act as a phase reference for comparison with the carrier, since the phase shift to the modulation sidebands themselves is negligible. In particular, changing the

---

[2] The gravitational wave detector signal cavity is the exception, where locking the cavity to near-resonance for the carrier frequency is desirable in order to resonate a frequency slightly shifted from the carrier frequency.
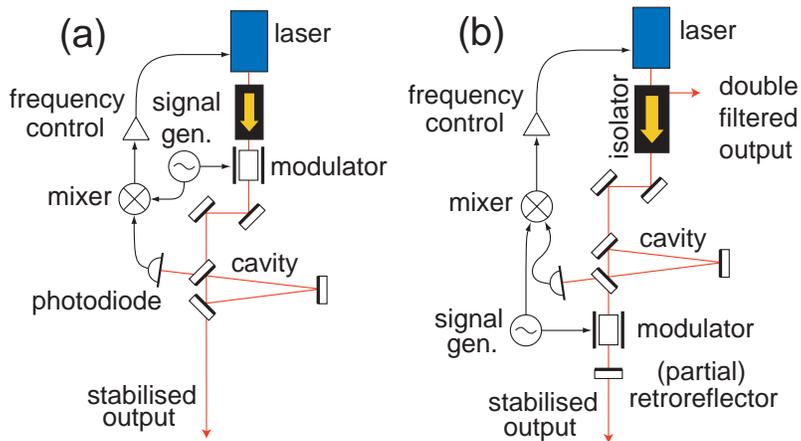
Fig. 4.8: Configurations to PDH lock a ring mode cleaner cavity in the (a) single pass and (b) double pass regimes.

(a)   (b)
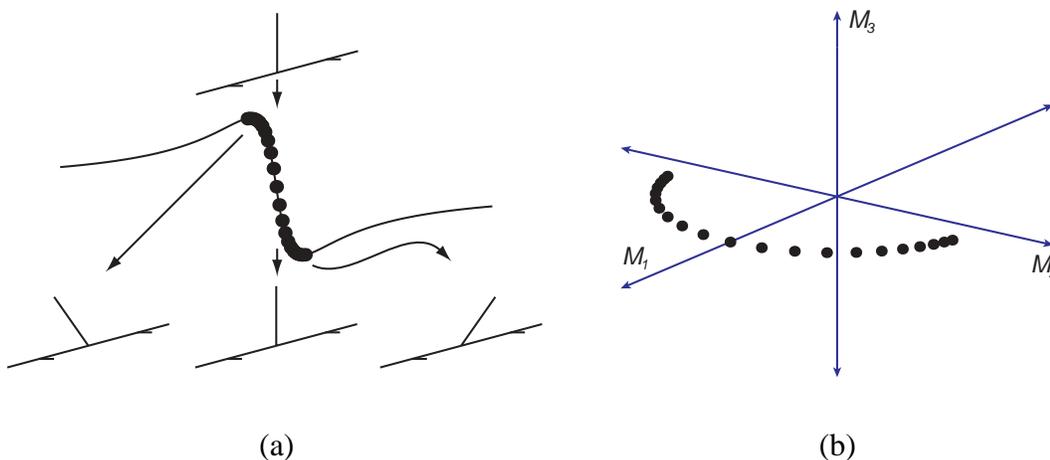
Fig. 4.9: (a) Frequency vector diagrams showing a cavity's response to a modulated beam near resonance, overlayed on the cavity's phase response in reflection. (b) The M-space trajectory of the output modulations near resonance, with similarly spaced data points.

phase of the carrier without changing the phase of modulation sidebands will convert a fraction of the modulation from PM into AM of the same beat phase (see Fig. 4.9(a)), or equivalently converting M_1 into M_2 (see Fig. 4.9(b)). The AM beat is then demodulated in-phase with the PM input to produce the required error signal. This can be fed back to the position of a cavity mirror, or to the input laser's frequency controls).

As with Michelson locking, a variable $\phi(t)$ is assigned to be the error (in radians) that the cavity-laser system is away from a resonance operating point. It can be shown that, for an input phase modulation of $\tilde{P}_{\text{IN}}$ at frequency ω_m, the electronic error signal after demodulation is given by:

$$V_{\text{mix}} \propto \frac{E_0^2}{2} \Re \left\{ e^{-\phi_{\text{mix}}} i \left(g(\phi,\eta) g(-\phi,0) - g(-\phi,\eta) g(\phi,0) \right) \tilde{P}_{\text{IN}} \right\}, \qquad (4.15)$$

where

$$g(\phi,\eta) = \frac{r_1 - r_2 e^{i\eta} e^{i\phi}}{1 - r_1 r_2 e^{i\eta} e^{i\phi}}, \qquad \eta \equiv 2\omega L/c. \qquad (4.16)$$
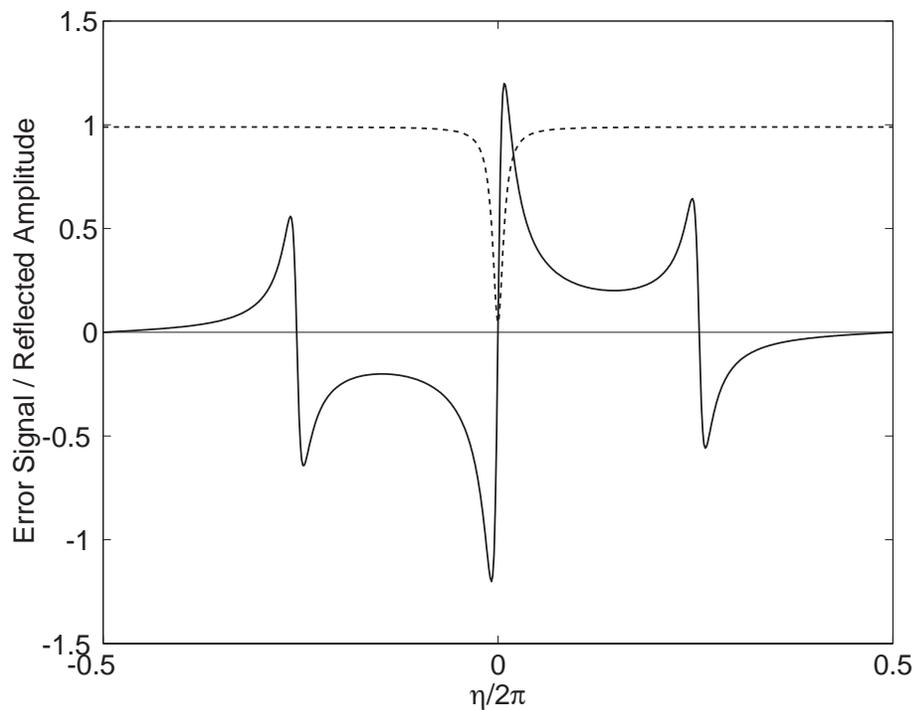
Fig. 4.10: PDH error signal (solid line) with cavity reflection amplitude (dashed line) shown for comparison. The error signal takes on the shape of the imaginary component of the cavity's reflected response near resonance, but also has similar dispersion curves (with opposite slope) where the sidebands are resonant.

This transfer function is derived fully in Chapter 6.

For now, a short qualitative description is provided of the origins of this PDH error signal, which is plotted in Fig. 4.10. The error signal shows a large-slope zero crossing on resonance as expected, which has a similar shape near resonance to the imaginary component of the cavity reflectivity function. However, the sidebands also create a dispersion curve in the error signal when they go through resonance, though the slope is of the opposite sign. Hence it is also possible to lock the sidebands to resonance in the cavity, by switching the polarity of the feedback loop.

Pound-Drever-Hall locking is well-established as the standard method for cavity locking, and has been so for two or three decades. It is used for locking both rigid and suspended cavities.

4.4 Double pass locking

The PDH scheme described in the previous section is a *single-pass* locking configuration, ie the laser travels through the cavity in the forward direction and continues on. In a single-pass configuration, when the laser is non-resonant in the cavity, the entire laser power is incident on the detector. For gravitational wave detectors, powers are high enough to cause damage to detectors. The modulation power is typically several orders of magnitude smaller than the carrier power, so simple attenuation of the reflected beam alone does not solve the problem. A common solution is to protect the detector with a
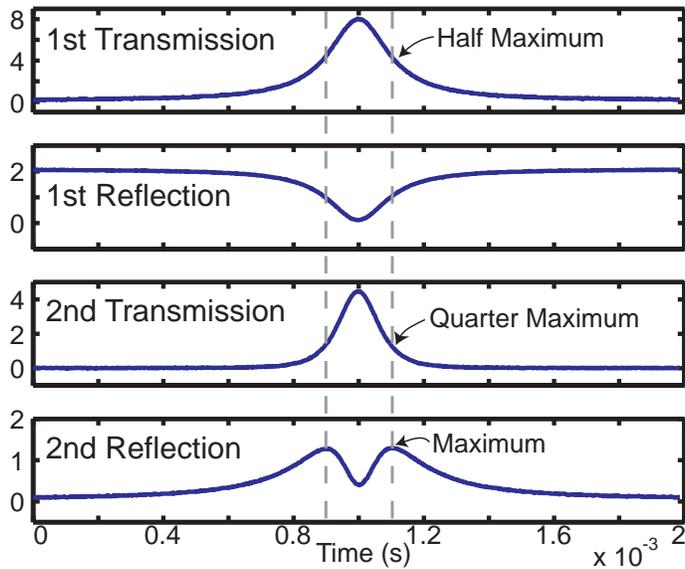
Fig. 4.11: Powers at the four cavity outputs while sweeping the laser frequency through the cavity resonance: transmission and reflection of the first pass, and transmission and reflection of the second (retroreflected) pass. The second pass transmission resonance is half as wide as the first pass transmission resonance as expected. The second pass reflection has two maxima when the first pass transmission is at half maximum, as explained in the text. A glass wedge was used for retroreflection; the latter two traces represent approximately 2 orders of magnitude less in power than the former two.

fast shutter which triggers when the cavity drops lock. The following is an alternative configuration that solves the problem more naturally.

A *double pass* Pound-Drever-Hall locking system involves putting the detector out of range of the input beam, as shown in Fig. 4.8(b). A retroreflector sends a small fraction of the output beam back into the cavity in the reverse direction. The resonance of this second pass is locked, which is enough to ensure that both the first and second passes are on resonance. The modulator and detector are placed accordingly, to modulate the retroreflected beam, and detect at the second pass reflected port, respectively. In a variation, 100% retroreflection can be employed, and the twice-mode-cleaned output is obtained from the isolator.

It is clear that very little power lands on the detector of a double pass locking system whether the system is locked or not. The maximum amount of power that the detector ever sees must be a quarter of the input power, when at the half maximum of resonance and hence when the cavity is transmitting and reflecting equally; the maximum is much less if only partial retroreflection is used. In any case, the detector sees this maximum power for a short time of order milliseconds during lock acquisition or lock drop, and hence the risk of damage to the photodetector is greatly reduced.

Fig. 4.11 shows photodetector voltages at the four possible ports of a cavity in a double pass arrangement during a sweep through resonance: the transmitted and reflected outputs associated with the first pass, and those associated with the second pass. The second pass

transmission profile is narrower than its first pass counterpart by a factor of 2 because the cavity's resonance curve is applied to the light twice over. The second pass reflected output has a minimum on resonance and two maxima at (first pass) half-resonance. This is the port where the photodetector is placed, and so the above prediction of a reduced maximum power is verified. In fact, the output of this port is simply the product of the transmissivity and reflectivity of the cavity.

Another advantage of using a double pass is that the detector sees virtually no parasitic mode mismatch, thus reducing shot noise contributed from non-resonant higher order modes. The error signal power is reduced if a partial retroreflector is used, as if an attenuator had been placed in front of the photodetector, with a corresponding drop in the shot noise limited sensitivity. However, this becomes less important if the main interferometer of a gravitational wave detector imposes tighter locking constraints on the laser frequency, which is typically the case [43].

The locking of a double pass system is insensitive to any jitter or mode mismatch of the input laser system; the first pass reflection is not used in the locking scheme, and the first pass output is fixed relative to the cavity. As such, double pass locking lends itself to suspended, as well as rigid, mode-cleaner cavities.

One apparent drawback of placing the modulator after the cavity is that the mode-cleaned output beam now has relatively strong modulation sidebands. This could be turned to advantage in a gravitational wave detection system: a subsequent mode-cleaner stage, or even part of the main interferometer, could be single pass locked using these "extraneous" modulation sidebands without the addition of a second modulator.

Double pass locking solves one difficulty of cavity locking with high laser powers, by placing the locking detector where it never sees the full power of the input beam. In the next two sections, schemes are presented that further simplify the design of a high power mode-cleaner cavity by removing one transmissive component, namely the modulator.

4.5   Spatial mode locking

This section describes two cavity locking techniques, based on spatial mode interference, that are suitable for rigid mode cleaner cavities in gravitational wave detectors. Both are modulation-free, and hence avoid the difficulties of modulator thermal lensing in high power systems [44]. *Tilt locking*, in the single and double pass regimes, has previously been presented elsewhere [1, 45–47]; here a brief overview is given in the context of mode cleaner cavities. For the first time, a relative of tilt locking called *flip locking* is introduced and demonstrated. Many unexplored variations of flip locking are possible, and some of these hold promise for locking suspended, as well as rigid, mode cleaners.

4.5.1   Double Pass Tilt locking

Tilt locking is similar in principle to PDH locking in that the dispersive phase shift of the near-resonant cavity is compared to some non-resonant local oscillator to produce an error signal. For PDH locking, modulation sidebands play the role of the local oscillator. For tilt locking, an added component of TEM_{01} spatial mode does the job. Here, demodu-
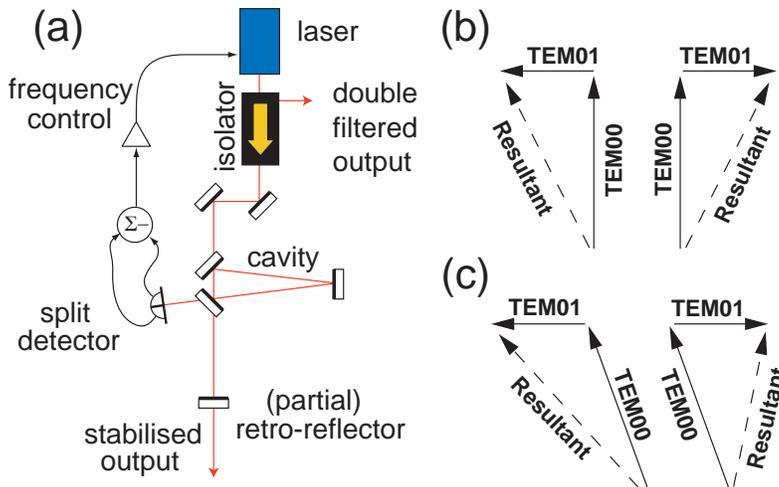
Fig. 4.12: (a) Double pass tilt locking configuration. The retroreflector is given a slight tilt. (b) & (c) Phasor diagrams for TEM_{00} and TEM_{01} modes (from [1]) showing interference on the two sides of the split detector, when (b) on resonance and (c) near resonance. Subtraction of the resultant powers provides an error signal.

lation is replaced by subtraction of the signals from the two halves of a two-element split detector (a detector with two separately-wired photosensitive cells). An error signal is obtained in this way since the near-resonant TEM_{00} mode interferes with the TEM_{01} mode constructively on one side of the split detector, and destructively on the other side (see Fig. 4.12(b) and (c)). The resulting error signal is equal to the imaginary component of the cavity reflectivity.

An experimental arrangement for the double pass tilt locking of a ring cavity is shown in Fig. 4.12(a). Compared with double pass PDH locking, the modulator is removed and the detector is replaced with a two-element split detector. In addition, a tilt is given to the retroreflector to misalign the beam's second pass slightly, such that a small amount of the light couples into the antisymmetric, non-resonant TEM_{01} spatial mode, and reflects off the cavity.

The component of TEM_{01} directly reflected from the cavity should be incident at the centre of the split detector, at an angle slightly tilted with respect to the normal. This ensures that the Guoy phase shift between the exactly resonant TEM_{00} mode (or any non-resonant mode-mismatch mode) and the TEM_{01} mode is 90 degrees. The interference of the two modes produces equal power in the two halves of the split detector, and the error signal zero crossing occurs precisely on resonance. Achieving this in the laboratory requires a small lateral shift to the split detector, since a tilt to the retroreflector will also produce a small lateral offset of the beam when it reaches the detection plane.

Single pass tilt locking is also possible. A tilt to the input beam is added, and the split detector is placed at the first pass reflected port. The retroreflector is not required. For high precision experiments, double pass tilt locking is preferred as fewer degrees of freedom susceptible to jitter will affect the locking performance. Aligning and double pass locking a cavity configuration in practise is made simpler by initially single-pass tilt locking the system to aid placement of the retroreflecter.
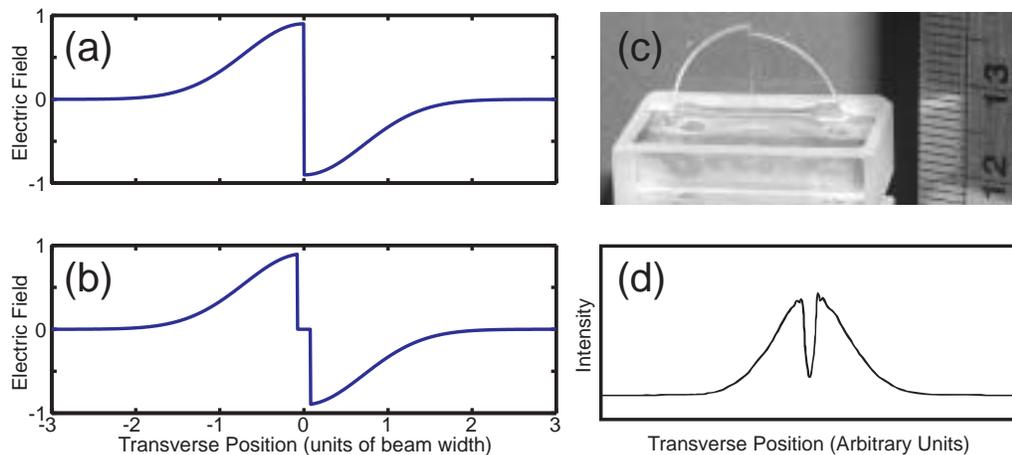
Fig. 4.13: Ideal electric field amplitudes of (a) flipped mode and (b) spatial jitter immunity flipped mode (see text). (c) The modified half wave plate "flipped mode generator", used to generate (d) a laboratory demonstrated flipped mode, from [48].

Previous demonstrations of tilt locking have shown that its noise suppression performance matches PDH locking [1]. One experiment involved tilt locking two lasers to a single cavity, and a fractional locking stability of 1.3×10^{-14} was demonstrated [47].

Tilt locking is more appropriate for rigid mode-cleaner cavities than the suspended equivalent, due to its first-order dependence on lateral displacement of the cavity mode with respect to the split detector.

4.5.2 Double Pass Flip locking

In tilt locking, the tilt added to the beam at the retroreflector is equivalent to a linear lateral phase shift across the beam profile. Here, the possibility of adding a stepped phase shift to a TEM$_{00}$ mode is considered. Such a mode should have improved efficiency of interference with the TEM$_{00}$ mode when split-detecting. The transverse electric field profile of a *flipped mode* is shown in Fig. 4.13(a).

One way to generate a flipped mode is by using a modified half wave plate, shown in Fig. 4.13(c). A regular half wave plate is cut into quarters, and two opposite quarters are joined such that the optical axis of one is at 90 degrees relative to the other. A TEM$_{00}$ gaussian beam with vertical or horizontal polarisation is directed exactly at the interface between the wave plate quarters. Just as a regular half wave plate gives a 180 degree phase shift between horizontal and vertical polarisations, this modified wave plate gives the required 180 degree phase shift between the left and the right side of the transverse beam profile. Fig. 4.13(d) shows the intensity profile of a laboratory flipped mode generated in this way.

Flip locking is very similar to tilt locking, bit instead a small amount of flipped mode replaces the small component of TEM$_{01}$ tilt mode generated by the tilted retroreflector. The configuration in Fig. 4.12(a) can be used where the modified half wave plate is placed in front of the retroreflector; the retroreflected beam passes through the wave plate twice, effecting a 360 degree phase shift between the two halves and hence returning the original
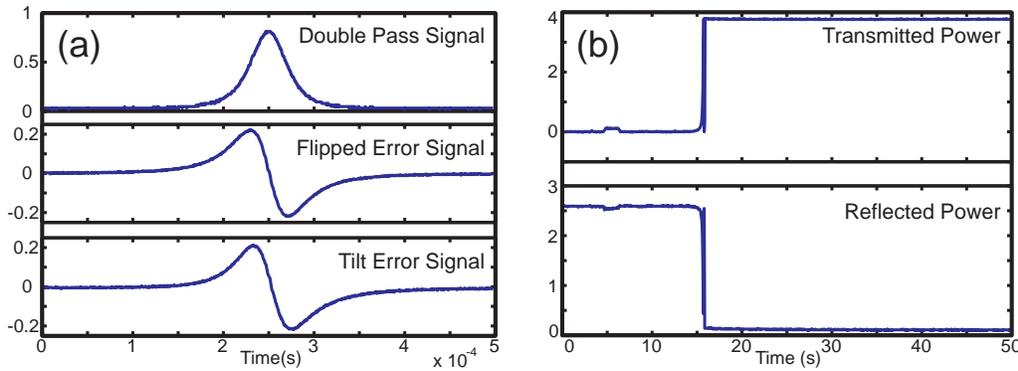
Fig. 4.14: (a) Resonance sweep signals from a double pass locking experiment simultaneously generating tilt and flip locking error signals (arbitrary units). (b) Cavity transmission and reflection power (arbitrary units) during lock acquisition.

TEM_{00} mode (neglecting diffraction effects). A slight tilt to the wave plate extends the path length and results in a slight phase mismatch between the two beam halves. This new beam can be expressed as a TEM_{00} mode plus a small fraction of flipped mode with a 90 degree phase shift. The component of flipped mode (comprised of non-resonant odd TEM modes) then reflects from the cavity and interferes to generate an error signal in the same manner as in tilt locking.

A ring cavity was successfully double pass flip locked using the configuration discussed above. Locking was stable and repeatable; the first pass transmission and reflection during lock acquisition are shown in Fig. 4.14(b). For comparison, the retroreflector was tilted in the vertical direction to generate a tilt locking error signal which was measured by the vertical subtraction of a quadrant detector; both error signals are shown in Fig. 4.14(a).

It is useful to note that the Rayleigh range of the cavity mode is typically metres so that, at the worst, flipped mode diffraction in the Fresnel regime applies. In practise, the resulting Guoy phase shifts between spatial modes are compensated for either by a slight lateral shift to the detector [49], or by using a focusing lens.

The above configuration is merely a demonstration; a realistic locking system would use a flipped mode generator that adds only a small phase shift between beam halves such that the single pass output is still approximately a (mode-cleaned) TEM_{00} mode. The flipped mode generator need not even be a distinct optical component; a stepped dielectric coating on either the retroreflector or the output coupler would achieve the same goal. Single pass flip locking is also possible and would require an optic element that adds a small component of flipped mode to the input beam, though the double pass regime is preferred due to the protection provided against jitter of the input optics.

A modified flipped mode with a central gap has also been generated (Fig. 4.13(b)) [48]. The gap width could be set greater than the spatial jitter noise at the split detector, in which case the error detection system would become completely immune to spatial jitter. Some variation on this theme may be appropriate for the locking of suspended mode cleaner cavities as well as rigid cavities, pending further research.

4.6   Summary

A review of Gaussian beam optics, Fabry-Perot cavities and Pound-Drever-Hall locking was presented in the context of using mode-cleaner cavities in gravitational wave detectors.

Solutions were offered to some of the difficulties arising from the use of high laser powers with mode-cleaner cavities. Double pass locking was suggested as a way to drastically reduce the maximum power that can ever reach the locking detector, thus simplifying the detector design. Double pass tilt locking removes the need for a modulator, hence avoiding the difficulties of high power optical transmissions. Furthermore, a demonstration experiment was detailed of double pass flip locking, an adaptable tilt locking variation, which has some potential for locking suspended mode cleaner cavities.

"See, I have learned that not everybody likes being told they're wrong."

— John McNabb

The Quadrature Variable Modulator

This chapter presents the theory of a device that mounts an arbitrary modulation state at a given frequency onto a laser beam. The device is termed a *Quadrature Variable Modulator* (QVM) in an attempt to encapsulate the ability of the device to tune around the parameter space of modulations. The input-output relations of the device are derived, using both the standard modulation formalism and the modulation sphere formalism. The properties and capabilities of the device are discussed, along with a prescription for getting the best performance out of the device, with gravitational wave detector applications in mind.

5.1 Quadrature Variable Modulator Schematic

Electro-optic modulators are ubiquitous in modern optics laboratories. Typical devices are designed specifically to produce either amplitude modulation or phase modulation, mounted on laser beams. Many applications use modulation as a means of information transmission, or as a means of measurement by exploiting a frequency-dependent dispersion in the subject optical system. In these applications, a single device capable of producing an arbitrary quadrature of modulation (amplitude or phase or a combination) is invaluable for simplifying the design of experiments, for improving the quality of results, and for diagnostics.

The QVM (see Fig. 5.1(a)) consists of two modulating crystals in series, with separate voltage sources, and with modulating axes at right angles to each other. It is used by sending a laser beam of elliptical polarisation through the device, and then through a linearly polarising filter angled at 45° to the crystals' modulating axes, which taps off or dumps a proportion of the optical power.

The choice of polarisation state of the incident beam has a crucial effect on the transfer characteristics of the system. The analysis for this chapter includes finding the transfer characteristics for all polarisation states, but the QVM works best with one of a subset of polarisation states. The subset is those with equal power in the polarisation components aligned with each modulating crystal axis. Equivalently, this requires that the polarisation is elliptical with major and minor axes aligned at 45° to both crystal axes. For concreteness, the choice is made here that the ellipse axes be horizontal and vertical, and the crystal axes be left and right diagonal (as in Fig. 5.1(a)).

The simplest polarisation state out of this subset is circular polarisation. However, by retaining a degree of freedom to choose vertically- or horizontally-aligned elliptical
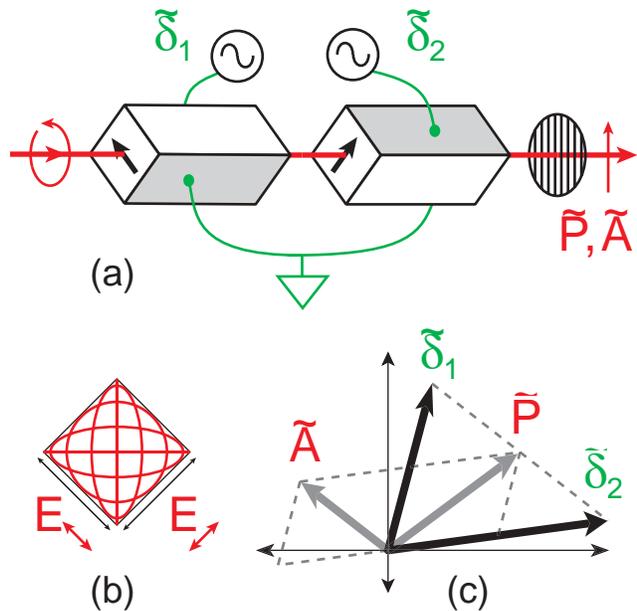
Fig. 5.1: (a) QVM schematic. Two modulating crystals are positioned in series with modulating axes at $90°$. A laser beam of elliptical polarisation passes through each crystal and then through a vertically polarising element. (b) The subset of input beam polarisation states considered here is all states such that there is an equal optical field amplitude in both the left- and right-diagonal components. σ equals the phase between these components: shown are $\sigma = 0$ (vertical), $\pi/4$, $\pi/2$ (circular), $3\pi/4$, and π (horizontal). (c) Phasor diagram representing the transfer function given by Eq. 5.4. If the input electrical signals are represented as rotating phasors $\tilde{\delta}_1$ and $\tilde{\delta}_2$, then the corresponding phasors that represent the output phase and amplitude modulations ($\tilde{P}$ and $\tilde{A}$) are proportional to $(\tilde{\delta}_1 + \tilde{\delta}_2)$ and $(\tilde{\delta}_1 - \tilde{\delta}_2)$ respectively.

polarisation, the option is retained of varying the amount of beam power tapped off by the vertical polariser at the exit of the QVM device. This is important for applications where beam power is at a premium, and gravitational wave detection is such an application.

Before proceeding, a simple analogy may help to clarify the inner workings of a QVM device, and lend physical insight. A Mach-Zehnder modulator (MZM) involves splitting a beam into two parts, separately phase-modulating each part, then recombining the two parts on a beamsplitter. The MZM device shown previously in Fig. 3.5(b) is hard-wired to the amplitude modulation operating point (a single electronic oscillator controls the modulations in both arms), but if two independent voltage sources were used, this device would be physically equivalent to the QVM.

The QVM is a spatially-degenerate interferometer that uses two orthogonal polarisations of light rather than two separate optical paths. Each QVM crystal modulates one and only one polarisation component, in the same way that modulators in a MZM act each on their own inteferometer arm. The birefringent waveplates used to create the initial polarisation for a QVM take the place of the input beamsplitter of a MZM, and likewise a polarising beamsplitter facilitates the output interferometric recombination. The phase between the electric fields in the two polarisation states is the interferometric recombination phase.

Through manipulation of the input parameters (polarisation state and electrical signals input to crystals), the user has complete control over the interference condition of the two carrier beams (in each of the two polarisation states), and separately over the interference of the lower and upper sidebands generated by the modulating crystals. This amounts to having complete freedom to choose any modulation state by appropriate choice of the input parameters.

As an example, consider two identically phase-modulated beams, where the optical phase of one beam is advanced by $90°$, and the modulation phase of the other beam is advanced by $90°$. The result, upon interfering the two beams, is that one resultant sideband is exactly cancelled out and the other is additively reinforced, producing a single sideband state. Indeed, the earlier description of amplitude modulation devices (Section 3.2.4, especially Fig. 3.6(c)) is another example of two phase modulated beams interfering to make a beam with a different modulation. One can intuitively arrive at all of the modulation states discussed in this paper by reasoning or drawing diagrams along these lines.

This reasoning will now be quantified into a mathematical theory of the QVM.

5.2 QVM Transfer Function

The first step is to quantify prescribed restriction on polarisation states as described in the previous section, ie elliptical polarisation with major axis aligned horizontally or vertically (see Fig. 5.1(b)). The angle σ is defined as the phase difference between left-diagonal and right-diagonal electric field components of the light exiting the crystals. Therefore, σ completely describes the input polarisation state for this case. The identity

$$\tan^2\left(\frac{\sigma}{2}\right) = \frac{P_h}{P_v} \tag{5.1}$$

follows, where $P_{h,v}$ are the powers in the horizontal and vertical polarisation components of the input laser beam.

Note that, in the ideal situation where both crystals have identical refractive indices and lengths, the phase angle σ (and hence the polarisation state) is the same before and after the modulator. However, a real device almost certainly does not have identical crystals so that the relative phase between left- and right-diagonal components may well change upon transmission. In this case σ should describe the polarisation state of the light *exiting* the modulator, just before it reaches the linearly polarising filter. The experimentalist ought to bear this in mind: if one were to place a beam tap-off to monitor the polarisation state, the tap-off should be placed after the modulator.

The inputs to the two crystal electrodes are sinusoidal voltages of frequency ω_m; these are written as complex phasors $\tilde{\delta}_j = \delta_j e^{i\phi_j}$ ($j = 1, 2$, and $i = \sqrt{-1}$) so that the single-crystal phase modulations are given by $\Re\{\tilde{\delta}_j e^{i\omega_m t}\} = \delta_j \cos(\omega_m t + \phi_j)$, where $\Re$ and $\Im$ respectively return the real and imaginary components of a complex variable.

The δ_j are modulation *depths*, and are related to the applied voltages V_j by:

$$\delta_j = \pi \times \frac{V_j}{V_{(\pi)j}} \tag{5.2}$$

where the π-voltages $V_{(\pi)j}$ may differ if the two QVM crystals are not identical.

The optical field amplitude exiting the QVM can be expressed as:

$$E_{\text{OUT}} = E_{\text{IN}} \left[\cos(\sigma/2) + i\Re\{\tilde{P}e^{i\omega_m t}\} + \Re\{\tilde{A}e^{i\omega_m t}\} \right] \tag{5.3}$$

where E_{IN} is the input optical field amplitude. In Eq. 5.3, small single-crystal modulation depths are assumed ($\delta_j \ll 1$, see Section 5.4 for the general case). This equation resembles Eq. 3.14, but the explicit factor of $\cos(\sigma/2)$ covers the possibility of the carrier amplitude going to zero[1].

The relationship between input and output parameters for the QVM device can be shown to be:

$$\begin{aligned} \tilde{P} &= \frac{1}{2}\cos\left(\frac{\sigma}{2}\right)\left(\tilde{\delta}_1 + \tilde{\delta}_2\right) \\ \tilde{A} &= \frac{1}{2}\sin\left(\frac{\sigma}{2}\right)\left(\tilde{\delta}_1 - \tilde{\delta}_2\right). \end{aligned} \tag{5.4}$$

5.2.1 Proof of Transfer Function

This is an outline of the derivation of Eq. 5.4. First, the (more general) transfer function is derived for the QVM when the input beam is allowed to have any polarisation. At the end, the equation is simplified to the subcase described in the main text.

The input beam's polarisation will be characterised by two electric field phasor components, $\tilde{L} = Le^{i\sigma_L}$ and $\tilde{R} = Re^{i\sigma_R}$, as defined in a set of left-diagonal and right-diagonal spatial coordinate axes, with unit vectors $\hat{L}$ and $\hat{R}$ respectively. The electric field exiting the QVM can be written as a vector (to include polarisation information) as:

$$\vec{E}_{\text{ExitingQVM}} = \left[\tilde{L}e^{i\Re\{\tilde{\delta}_1 e^{i\omega_m t}\}}\hat{L} + \tilde{R}e^{i\Re\{\tilde{\delta}_2 e^{i\omega_m t}\}}\hat{R} \right] e^{i\omega t}. \tag{5.5}$$

From hereon in, the $e^{i\omega t}$ term is suppressed for the sake of brevity. Upon passing through the vertically aligned linear polariser (equivalent to taking a dot product with the vector $(\hat{L} + \hat{R})/\sqrt{2}$), and expanding the complex exponentials to first order (hence assuming $|\tilde{\delta}_{1,2}| \ll 1$), the vertical electric field amplitude becomes:

$$E_{\text{OUT}} = \frac{1}{\sqrt{2}}\left[\tilde{L}(1 + i\Re\{\tilde{\delta}_1 e^{i\omega_m t}\}) + \tilde{R}(1 + i\Re\{\tilde{\delta}_2 e^{i\omega_m t}\}) \right]. \tag{5.6}$$

Next, DC terms and oscillating terms are collected, and the overall optical phase is

---

[1] The *carrier suppression* operating point, where the carrier component goes to zero, is discussed in Section 5.2.3.

factored out:

$$
\begin{aligned}
E_{\text{OUT}} &= \frac{\tilde{L}+\tilde{R}}{\sqrt{2}}\left[1+\frac{i(L^2+\tilde{L}\tilde{R}^*)}{|\tilde{L}+\tilde{R}|^2}\Re\{\tilde{\delta}_1 e^{i\omega_m t}\}\right.\\
&\quad+\left.\frac{i(\tilde{R}\tilde{L}^*+R^2)}{|\tilde{L}+\tilde{R}|^2}\Re\{\tilde{\delta}_2 e^{i\omega_m t}\}\right].
\end{aligned}
\tag{5.7}
$$

Real and imaginary components are now separated for the oscillating terms:

$$
\begin{aligned}
E_{\text{OUT}} &= \frac{\tilde{L}+\tilde{R}}{\sqrt{2}}\left[1+\Re\left(\frac{\Im\{\tilde{R}\tilde{L}^*\}(\tilde{\delta}_1-\tilde{\delta}_2)}{|\tilde{L}+\tilde{R}|^2}e^{i\omega_m t}\right)\right.\\
&\quad+\left.i\Re\left(\frac{\Re\{\tilde{R}\tilde{L}^*\}(\tilde{\delta}_1+\tilde{\delta}_2)+L^2\tilde{\delta}_1+R^2\tilde{\delta}_2}{|\tilde{L}+\tilde{R}|^2}e^{i\omega_m t}\right)\right].
\end{aligned}
\tag{5.8}
$$

These real and imaginary oscillating terms correspond to AM and PM respectively, so the following parametrisation is made:

$$
E_{\text{OUT}}=\frac{|\tilde{L}+\tilde{R}|}{\sqrt{2}}+\Re\{\tilde{A}e^{i\omega_m t}\}+i\Re\{\tilde{P}e^{i\omega_m t}\}
\tag{5.9}
$$

with

$$
\begin{aligned}
\tilde{P} &= \frac{\Re\{\tilde{R}\tilde{L}^*\}(\tilde{\delta}_1+\tilde{\delta}_2)+L^2\tilde{\delta}_1+R^2\tilde{\delta}_2)}{\sqrt{2}|\tilde{L}+\tilde{R}|}\\
\tilde{A} &= \frac{\Im\{\tilde{R}\tilde{L}^*\}(\tilde{\delta}_1-\tilde{\delta}_2)}{\sqrt{2}|\tilde{L}+\tilde{R}|}
\end{aligned}
\tag{5.10}
$$

where the net optical phase shift has been discarded. Eqs. 5.9 and 5.10 constitute the general QVM transfer function for arbitrary input polarisation.

Upon restricting to the prescribed subset of polarisation states, $L=R=E_{\text{IN}}/\sqrt{2}$ (equal power in the two polarisation axes), and defining $\sigma=\sigma_R-\sigma_L$ as the phase between the two polarisation components, the equations reduce to:

$$
E_{\text{OUT}}=E_{\text{IN}}\cos\left(\frac{\sigma}{2}\right)+\Re\{\tilde{A}e^{i\omega_m t}\}+i\Re\{\tilde{P}e^{i\omega_m t}\}
\tag{5.11}
$$

and

$$
\begin{aligned}
\tilde{P} &= \frac{E_{\text{IN}}}{2}\cos\left(\frac{\sigma}{2}\right)\left(\tilde{\delta}_1+\tilde{\delta}_2\right)\\
\tilde{A} &= \frac{E_{\text{IN}}}{2}\sin\left(\frac{\sigma}{2}\right)\left(\tilde{\delta}_1-\tilde{\delta}_2\right).
\end{aligned}
\tag{5.12}
$$

Eqs. 5.3 and 5.4 are the same as these, but with E_{IN} factored out of the definitions of $\tilde{P}$ and $\tilde{A}$.

5.2.2 Physical Insights

An important observation from Eq. 5.4 is that the phases of the AM and PM track the phases of the sum and difference of the electrical signals, respectively. Thus, an electronic oscillator for use in demodulation schemes is readily available, and can be calibrated once for all operating points. This phase-tracking property holds for any choice of σ; the effect of a change of polarisation ellipticity is to change the overall quantity of PM or AM available (as well as changing the overall beam power in Eq. 5.3). The property is depicted in Fig. 5.1(c) which shows the modulation phasors $\tilde{A}$ and $\tilde{P}$ in relation to the electrical signal phasors $\tilde{\delta}_j$.

Note that without a restriction on the polarisation state, the PM component in Eq. 5.10 does not have the property of tracking the phase of the sum of the input electrical signals, since it generally depends on the δ_j in different proportions. This is one of the reasons why the prescribed restriction on polarisation is emphasised here as being desirable.

As discussed earlier, the polarisation subset still gives the freedom to choose how much optical power to tap off from the beam. However, there is a trade off, since some proportion of power must be removed in order to achieve any appreciable amount of AM. The following is a useful rule-of-thumb for selecting an appropriate amount of power to tap off.

Consider Eq. 5.4 when setting $\tilde{\delta}_1 = i\tilde{\delta}_2$ such that $\tilde{\delta}_1 + \tilde{\delta}_2$ and $\tilde{\delta}_1 - \tilde{\delta}_2$ have the same magnitude (in order to focus on the *coefficients* of the complex $\tilde{\delta}_j$ phasors). The ratio of modulation power between the AM and PM components is $|\tilde{A}|^2/|\tilde{P}|^2 = \tan^2(\sigma/2)$ and, comparing this with Eq. 5.1:

$$\frac{|\tilde{A}|^2}{|\tilde{P}|^2} = \frac{P_h}{P_v} \tag{5.13}$$

So, the fraction of power tapped off is directly proportional to the fraction of modulation power available to be expressed as AM rather than PM. Typically then, if 10% of the output power is tapped off, and the user wishes to alternately produce a pure PM state and then a pure AM state, each of the same modulation power, then the AM state will require 9 times the input electrical power in compensation, compared to the PM state.

5.2.3 QVM States

Here, explicit instructions are given for preparing the basic modulation states using a QVM, as well as for a new modulation state not discussed earlier. Table 5.1 details six cardinal modulation states with the electrical signals required in order to "dial up" these states. These results follow directly from Eq. 5.4.

Pure PM and AM states require electrical signals equal in amplitude, with PM requiring the signals to be in-phase, and AM requiring that they are exactly out-of-phase[2]. Note that these two operating points have the same inputs no matter what the value of σ is.

Combinations of PM and AM that are correlated or anticorrelated require electrical signals with the same phase but different amplitudes. The entries into Table 5.1 give the

---

[2] The QVM shown in Fig. 5.1 has crystals hard-wired in-phase, unlike the amplitude modulator in Fig. 3.5(a) which was hard-wired out-of-phase.

Operating point	Modulations	Electrical Inputs
Pure PM	$\tilde{P} \neq 0, \tilde{A} = 0$	$\tilde{\delta}_1 = \tilde{\delta}_2$
Pure AM	$\tilde{P} = 0, \tilde{A} \neq 0$	$\tilde{\delta}_1 = \tilde{\delta}_2 e^{i\pi}$
Correlated PM & AM	$\tilde{P} = \tilde{A}$	$\tilde{\delta}_1 \cos(\sigma) = \tilde{\delta}_2 e^{i\pi}(1 + \sin(\sigma))$
Anticorrelated PM & AM	$\tilde{P} = \tilde{A} e^{i\pi}$	$\tilde{\delta}_1(1 + \sin(\sigma)) = \tilde{\delta}_2 e^{i\pi}\cos(\sigma)$
Single sideband $-\omega_m$	$\tilde{P} = \tilde{A} e^{i\pi/2}$	$\tilde{\delta}_1 = \tilde{\delta}_2 e^{i(\pi+\sigma)}$
Single sideband $+\omega_m$	$\tilde{P} = \tilde{A} e^{-i\pi/2}$	$\tilde{\delta}_1 = \tilde{\delta}_2 e^{i(\pi-\sigma)}$

Tab. 5.1: Details of significant operating points. The modulation parameters, expressed both in terms of phasors and M-space parameters, are given for each point. The electrical inputs required to produce these operating points are also given (both in terms of phasors and Q-space parameters), and these inputs vary with the polarisation parameter σ.

relative proportions on δ_1 and δ_2 required. In the case of circular polarisation, $\sigma = \pi/2$ and $\cos(\sigma) = 0$, these modulation states require one modulating crystal to be switched off altogether. This is an instructive point: correlated and anticorrelated PM & AM combinations are the natural basis of the QVM when used with circularly polarised light. One could choose to work with phasors of this type as an alternative to pure PM and pure AM phasors as is done here.

The reader will recall that single sideband modulations are composed of equal PM and AM combinations in quadrature. When using circular polarisation, these are achieved with equal electrical signals in quadrature. When away from circular polarisation, equal electrical signals are still required, but at a quadrature phase determined by σ.

It is apparent from Eqs. 5.3 and 5.4 that for $\sigma = \pi$ the carrier and the PM contribution vanish, leaving only the AM component. This is the *carrier suppression* operating point, where only two sidebands remain (see Fig. 5.2(a)). The modulation ceases to be of the phase or amplitude type due to the absence of a carrier as a phase reference, though the modulation retains its beat phase, which is expressed through the second harmonic in optical power. $\sigma = \pi$ corresponds to a horizontally polarised state, and Fig. 5.2(b) shows physically how sidebands but no carrier are transmitted through the vertical polariser.

5.3 QVM in the Modulation Sphere Description

The modulation sphere formalism was originally developed by the author as a visualisation tool for understanding the workings of the QVM. Here, the input-output relations of the QVM are transformed into the modulation sphere representation to gain further insight into the transfer characteristics of the device, especially regarding changes to the optical polarisation state used.

It is convenient to define a *Q-space* representation, (Q_1, Q_2, Q_3), for the electrical input parameters, analogous to the M-space representation for modulation, via:

$$
\begin{aligned}
Q_1 &= \delta_1^2 - \delta_2^2 \\
Q_2 &= 2\delta_1\delta_2\cos(\phi_1 - \phi_2) = 2\Re\{\tilde{\delta}_1\tilde{\delta}_2^*\} \\
Q_3 &= 2\delta_1\delta_2\sin(\phi_1 - \phi_2) = 2\Im\{\tilde{\delta}_1\tilde{\delta}_2^*\}
\end{aligned}
\tag{5.14}
$$

(a)   (b)
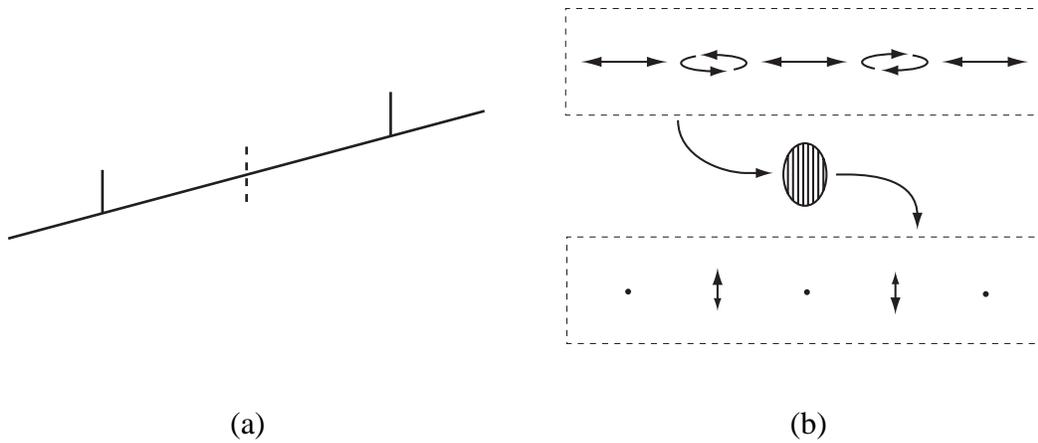
Fig. 5.2: (a) Frequency vector diagram of the *carrier suppression* modulation state. Without the carrier present, the modulation is neither AM nor PM. (b) The near-horizontally polarised state exiting the QVM is blocked by the vertical polariser, leaving only the oscillation component of a carrier-suppressed modulation state.
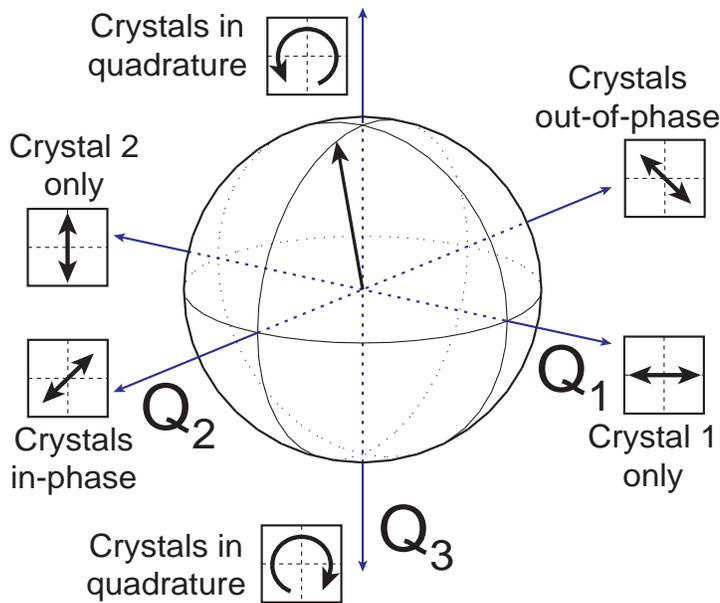
Fig. 5.3: The electrical signal sphere, in analogy with the modulation sphere. Axis labels are visual reminders of relative amplitudes and phases of input electrical signals (visualise the display of a cathode ray oscilloscope run in X-Y mode with δ_1 and δ_2 as the X and Y inputs).

where $Q_0 = \delta_1^2 + \delta_2^2 = \sqrt{Q_1^2 + Q_2^2 + Q_3^2}$ is the sum input electrical power. Physical interpretations for these parameters correspond with the M-space equivalents (see Fig. 5.3 for a diagrammatic description).

Now, the QVM transfer function, Eq. 5.4, can be rewrittenin terms of the M-space and

Q-space parameters:

$$
\begin{aligned}
M_0 &= \frac{1}{4}\left(Q_0 + Q_2\cos(\sigma)\right) \\
M_1 &= \frac{1}{4}\left(Q_0\cos(\sigma) + Q_2\right) \\
M_2 &= \frac{1}{4}\sin(\sigma)Q_1 \\
M_3 &= -\frac{1}{4}\sin(\sigma)Q_3.
\end{aligned}
\tag{5.15}
$$

5.3.1 Proof of Transfer Function

This is an outline of the derivation of Eq. 5.15. The Q-space to M-space transfer function of the QVM is derived for any input light polarisation state. At the end, the equations are reducedto the case where the allowed polarisations are restricted. The derivation consists of transforming parameters $\tilde{P}$ and $\tilde{A}$ to M-space parameters (M_1, M_2, M_3) using Eq. 3.27, and transforming parameters $\tilde{\delta}_1$ and $\tilde{\delta}_2$ to Q-space parameters (Q_1, Q_2, Q_3) using Eq. 5.14, and hence converting Eq. 5.10 from one set of coordinates to another.

Firstly, it is convenient to write down a few intermediate terms using Eq. 5.10 as a starting point. Conversion to Q-space parameters (via Eq. 5.14) is done on the fly.

$$
\begin{aligned}
P^2 = \tilde{P}^*\tilde{P} &= \frac{\Re\{\tilde{R}\tilde{L}^*\}^2[Q_0 + Q_2] + 2L^2R^2Q_2}{2|\tilde{L}+\tilde{R}|^2} \\
&\quad + \frac{\frac{1}{2}[(L^4+R^4)Q_0 + (L^4-R^4)Q_1]}{2|\tilde{L}+\tilde{R}|^2} \\
&\quad + \frac{\Re\{\tilde{R}\tilde{L}^*\}[(L^2+R^2)Q_0 + (L^2-R^2)Q_1 + (L^2+R^2)Q_2]}{2|\tilde{L}+\tilde{R}|^2}
\end{aligned}
\tag{5.16}
$$

$$
A^2 = \tilde{A}^*\tilde{A} = \frac{\Im\{\tilde{R}\tilde{L}^*\}^2[Q_0 - Q_2]}{2|\tilde{L}+\tilde{R}|^2}
\tag{5.17}
$$

$$
\begin{aligned}
\Re\{\tilde{P}\tilde{A}^*\} &= \frac{\Re\{\tilde{R}\tilde{L}^*\}\Im\{\tilde{R}\tilde{L}^*\}Q_1}{2|\tilde{L}+\tilde{R}|^2} \\
&\quad + \frac{\frac{1}{2}\Im\{\tilde{R}\tilde{L}^*\}\left[(L^2-R^2)Q_0 + (L^2+R^2)Q_1 - (L^2-R^2)Q_2\right]}{2|\tilde{L}+\tilde{R}|^2}
\end{aligned}
\tag{5.18}
$$

$$
\Im\{\tilde{P}\tilde{A}^*\} = \frac{-\Im\{\tilde{R}\tilde{L}^*\}\left[\Re\{\tilde{R}\tilde{L}^*\} + \frac{1}{2}(L^2+R^2)\right]Q_3}{2|\tilde{L}+\tilde{R}|^2}
\tag{5.19}
$$

Things can be simplified somewhat by using a Stokes parameter representation for the polarisation from this point[3]. Here, the Stokes parameters are (unusually) defined in

---

[3] The Stokes parameters define the Poincaré sphere representation for polarisation in a manner analogous

Operating point	M-space	Q-space
Pure PM	$(+M_0, 0, 0)$	$(0, \frac{4M_0}{\cos(\sigma)+1}, 0)$
Pure AM	$(-M_0, 0, 0)$	$(0, \frac{4M_0}{\cos(\sigma)-1}, 0)$
Correlated PM & AM	$(0, +M_0, 0)$	$(\frac{4M_0}{\sin(\sigma)}, -\frac{4M_0\cos(\sigma)}{\sin^2(\sigma)}, 0)$
Anticorrelated PM & AM	$(0, -M_0, 0)$	$(-\frac{4M_0}{\sin(\sigma)}, -\frac{4M_0\cos(\sigma)}{\sin^2(\sigma)}, 0)$
Single sideband $-\omega_m$	$(0, 0, +M_0)$	$(0, -\frac{4M_0\cos(\sigma)}{\sin^2(\sigma)}, -\frac{4M_0}{\sin(\sigma)})$
Single sideband $+\omega_m$	$(0, 0, -M_0)$	$(0, -\frac{4M_0\cos(\sigma)}{\sin^2(\sigma)}, \frac{4M_0}{\sin(\sigma)})$

Tab. 5.2: Details of significant operating points in M-space with corresponding Q-space coordinate requirements, which vary with the polarisation parameter σ.

terms of left-diagonal and right-diagonal electric field components (equivalent to rotating the usual X-Y axes anti-clockwise by 45°):

$$
\begin{aligned}
S_0 &= R^2 + L^2 \\
S_1 &= R^2 - L^2 \\
S_2 &= 2RL\cos(\sigma) = 2\Re\{\tilde{R}\tilde{L}^*\} \\
S_3 &= 2RL\sin(\sigma) = 2\Im\{\tilde{R}\tilde{L}^*\}
\end{aligned}
\tag{5.20}
$$

and so the previous four equations become:

$$
\begin{aligned}
P^2 &= \frac{1}{4}\left[S_0 Q_0 - S_1 Q_1 + S_2 Q_2 \right. \\
&\quad \left. - \frac{S_3^2}{2(S_0 + S_2)}(Q_0 - Q_2) \right] \\
A^2 &= \frac{1}{8}\frac{S_3^2}{S_0 + S_2}(Q_0 - Q_2) \\
\Re\{\tilde{P}\tilde{A}^*\} &= \frac{1}{8}\left[S_3 Q_1 - \frac{S_3 S_1}{S_0 + S_2}(Q_0 - Q_2) \right] \\
\Im\{\tilde{P}\tilde{A}^*\} &= -\frac{1}{8}S_3 Q_3
\end{aligned}
\tag{5.21}
$$

These intermediate quantities are then converted to M-space parameters (via Eq. 3.27). The result in matrix form is:

$$
\begin{bmatrix} M_0 \\ M_1 \\ M_2 \\ M_3 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} S_0 & -S_1 & S_2 & 0 \\ S_2 + \frac{S_1^2}{S_0+S_2} & -S_1 & S_0 - \frac{S_1^2}{S_0+S_2} & 0 \\ \frac{-S_1 S_3}{S_0+S_2} & S_3 & \frac{S_1 S_3}{S_0+S_2} & 0 \\ 0 & 0 & 0 & -S_3 \end{bmatrix} \begin{bmatrix} Q_0 \\ Q_1 \\ Q_2 \\ Q_3 \end{bmatrix}
\tag{5.22}
$$

The restriction on polarisation states discussed in the text is equivalent to setting S_1 to

to the M-space parameters defining the modulation sphere representation. A useful description of Stokes parameters can be found in [50].

zero, which in turn forces $S_0 = E_{IN}^2$, $S_2 = E_{IN}^2\cos(\sigma)$ and $S_3 = E_{IN}^2\sin(\sigma)$, when Eq. 5.22 reduces to:

$$\begin{bmatrix} M_0 \\ M_1 \\ M_2 \\ M_3 \end{bmatrix} = \frac{E_{IN}^2}{4} \begin{bmatrix} 1 & 0 & \cos(\sigma) & 0 \\ \cos(\sigma) & 0 & 1 & 0 \\ 0 & \sin(\sigma) & 0 & 0 \\ 0 & 0 & 0 & -\sin(\sigma) \end{bmatrix} \begin{bmatrix} Q_0 \\ Q_1 \\ Q_2 \\ Q_3 \end{bmatrix} \tag{5.23}$$

Once again, Eq. 5.15 stated originally is the dimensionless equivalent of Eq. 5.23, since the former factored out E_{IN}^2 from the definition of the M-space parameters.

5.3.2   Physical Insights

A closer look at the placement of terms or zeros in the transfer matrix of Eq. 5.23 gives a lot of information. Since the modulation power M_0 depends partly on Q_2 (and hence depends on the *detail* of the input signals, not just the overall input power Q_0), a sphere of constant electrical input power in Q-space does *not*, in general, map to a sphere of constant modulation power in M-space. In fact, Eq. 5.15 describes a transfer function from a sphere in Q-space to an *ellipsoid* in M-space. It can be shown that the ellipsoid is centered at $(M_1, M_2, M_3) = (Q_0\cos(\sigma)/4, 0, 0)$, has radii $[Q_0/4, Q_0\sin(\sigma)/4, Q_0\sin(\sigma)/4]$, and has an eccentricity of $\varepsilon = |\cos(\sigma)|$. The ellipse is always prolate with long axis aligned with the M_1-axis, and always has one focus at the M-space origin. The orthogonality of axes is preserved in the transformation, which consists only of a translation (the ellipsoid is not origin-centered) and a dilation (the ellipsoid is squashed in the M_2 and M_3 directions). The centrepoint and proportions of the ellipsoid are parametrised by the input light's polarisation parameter, σ (see Fig. 5.4(a)).

Similarly, reviewing Eq. 5.22, the general Q-space to M-space transfer function for the QVM with any input polarisation, gives further insight into the manner in which choosing $L \neq R$ affects the properties of the system. In this case, the parameters M_1 and M_2 depend on both Q_1 and Q_2, which means that a sphere in Q-space will be distorted in the M_1-M_2 plane by the transformation. The resulting surface is an ellipsoid whose major axis is no longer collinear with the M_1-axis, but is at an angle subtended in the M_1-M_2 plane (see Fig. 5.4(b)). This distortion certainly interferes with most of the favourable properties of the system. For example, none of the poles of the ellipse coincide with a coordinate axis, making any of the six operating points discussed more complicated to find.

On the other hand, since the only nonzero coefficient in the bottom row of the transfer matrix in Eq. 5.22 is $-S_3$, there is therefore a requirement to have *some* component of circularly polarised light in order to have access all types of modulation. Specifically, the coefficients for both M_2 and M_3 vanish if $S_3 = 0$.

As well as being necessary, some component of S_3 light is also sufficient to access all modulation states. If the choice is pure circularly polarised light via $S_3 = S_0$ and $S_1 = S_2 = 0$ (equivalent to choosing $\sigma = \pi/2$ for the restricted polarisation case) Eq. 5.23

(a)

(b)

(c)
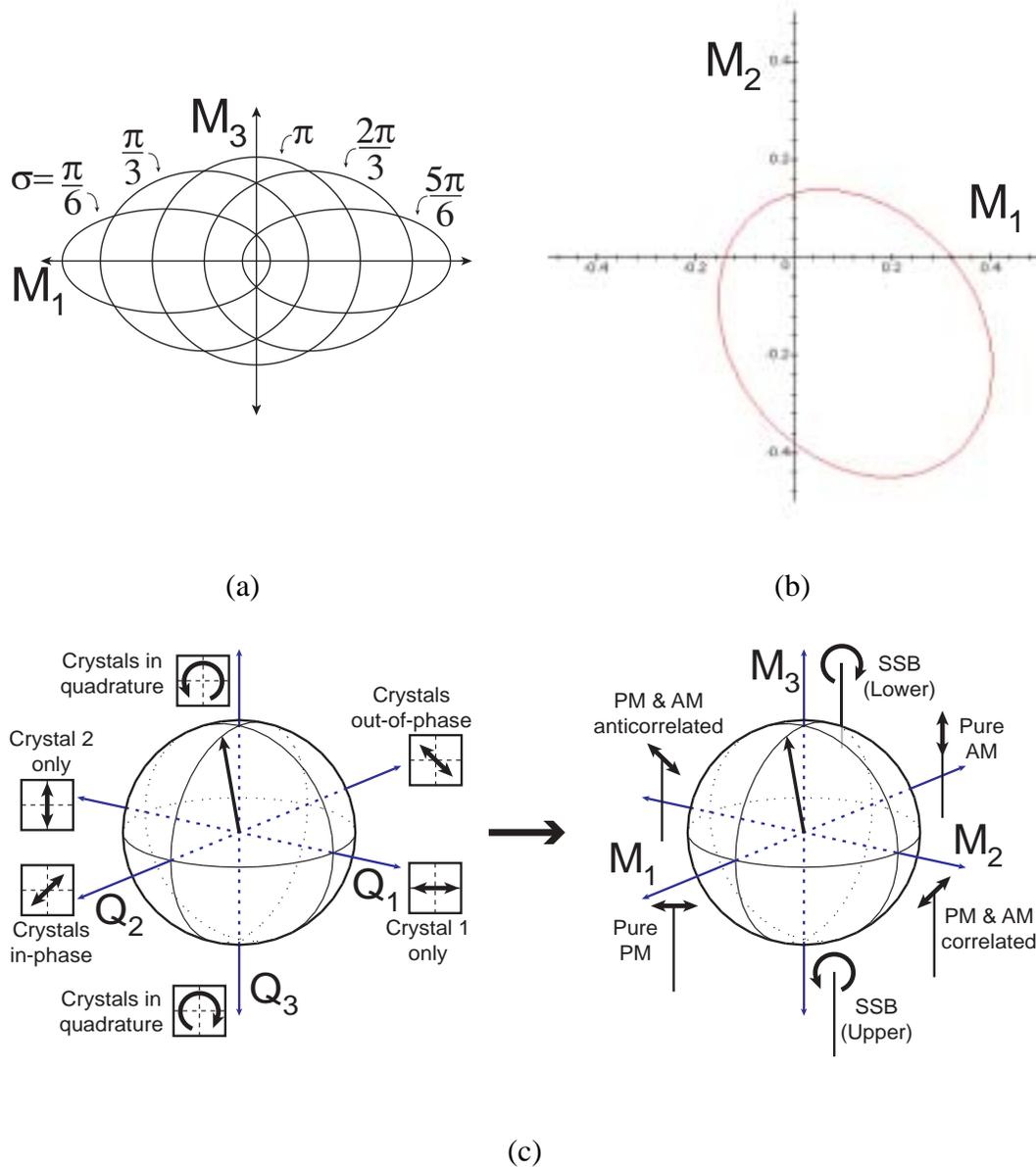
Fig. 5.4: (a) A sphere in Q-space maps to an ellipsoid in M-space in general. With restricted polarisation, the ellipse is aligned with the axes as shown. The ellipsoid dimensions in the M_2 and M_3 axes are the same. (b) With unrestricted polarisation, the ellipsoid's major axis lies at an angle in the M_1-M_2 plane. It appears (but remains unproven) that the ellipsoid retains two focii (possible only if two out of the three ellipsoid dimensions are equal), and that one of these focii remains at the origin for all possible polarisations. In this numerical Maple-generated plot: $(S_1, S_2, S_3) = (0.8, 0.0, 1.0)$. (c) In the special case where circular polarisation is used with the QVM, there is a direct graphical transfer correspondence between the Q-space sphere and the M-space sphere.

reduces to:

$$M_0 = \frac{1}{4}Q_0 \qquad M_2 = \frac{1}{4}Q_1$$

$$M_1 = \frac{1}{4}Q_2 \qquad M_3 = -\frac{1}{4}Q_3 \qquad\qquad (5.24)$$

Here, a sphere in Q-space maps to a sphere in M-space. This case is particularly instructive as the modulation and electrical signal spheres take on a new significance: there is now a direct graphical correspondence between the two spheres, in accordance with Eq. 5.24 (see Fig. 5.4(c)). So, as an example of using these spheres as a visual tool, one can see "at a glance" that the PM operating point is obtained by running both crystals with equal in-phase electrical inputs, or that a single sideband is obtained by running both crystals with equal inputs in quadrature.

Finally, the six cardinal modulations, corresponding to the six coordinate axis directions in M-space, are summarised in Table 5.2. Interestingly, it is not possible with this device to achieve either M_2 or M_3 type modulations when there is no carrier, as when σ is a multiple of π, the corresponding coefficients in the transfer function disappear (and the entries into Table 5.2 blow up). The inability of the device to generate M_2 with no carrier is immaterial since, without a carrier, there is no difference between M_1 and M_2 modulation. However, M_3 modulation, ie single sideband modulation, is physically distinct even without a carrier. The conclusion is that a QVM *cannot* simultaneously suppress the carrier and one sideband. One way to understand this is to note that single-sideband modulation is composed of as much AM as PM, and AM is not possible without power discrimination at the output polarising filter.

5.4   QVM Higher-Order Frequencies

In this section, the full QVM output electric field is derived for cases where the modulation depth is large enough that higher order sidebands cannot be neglected. Many applications particularly in fibre optics operate with high modulation depth, and hence a knowledge of the scaling of power in higher frequencies is essential. Some feedback control systems specifically rely on the presence or absence of second and third order sidebands [51].

Also, the possibility of removing problematic higher-order sidebands has been raised [52], for situations where the presence of these sidebands somehow interferes with an experiment or measurement. The QVM would be an ideal device to cancel higher order sidebands via careful application of higher order electrical inputs, since the QVM can emulate any quadrature and type of modulation at the required frequencies.

The derivation departs from the algebra at Eq. 5.5, instead using the Bessel function expansions for the phase modulations as per Eq. 3.23. After several stages of algebra, one

obtains the lengthy but complete transfer function:

$$
\begin{aligned}
E_{\text{OUT}} \;=\; & \frac{\tilde{L}J_0(\delta_1)+\tilde{R}J_0(\delta_2)}{\sqrt{2}} \left[1 + 2\sum_{n=1}^{\infty} i^n \times \right. \\
& \left(\frac{\left[|L|^2 J_0(\delta_1) + \Re\{\tilde{R}\tilde{L}^*\}J_0(\delta_2)\right]\Re\{J_n(\delta_1)e^{in\omega_m t}e^{in\phi_1}\}}{|\tilde{L}J_0(\delta_1)+\tilde{R}J_0(\delta_2)|^2} \right. \\
& + \frac{\left[\Re\{\tilde{R}\tilde{L}^*\}J_0(\delta_1) + |R|^2 J_0(\delta_2)\right]\Re\{J_n(\delta_2)e^{in\omega_m t}e^{in\phi_2}\}}{|\tilde{L}J_0(\delta_1)+\tilde{R}J_0(\delta_2)|^2} \\
& + i\frac{\Im\{\tilde{R}\tilde{L}^*\}J_0(\delta_1)\Re\{J_n(\delta_2)e^{in\omega_m t}e^{in\phi_2}\}}{|\tilde{L}J_0(\delta_1)+\tilde{R}J_0(\delta_2)|^2} \\
& \left. \left. - i\frac{\Im\{\tilde{R}\tilde{L}^*\}J_0(\delta_2)\Re\{J_n(\delta_1)e^{in\omega_m t}e^{in\phi_1}\}}{|\tilde{L}J_0(\delta_1)+\tilde{R}J_0(\delta_2)|^2} \right) \right] e^{i\omega t}.
\end{aligned}
\tag{5.25}
$$

Here, real and imaginary components are sorted as far as possible, so that when the i^n term is evaluated for each harmonic, it is relatively straightforward to determine which components are PM-like and which are AM-like.

Extracting interesting physics from the above equation is hard without being more specific in defining an area of interest in the parameter space. However, an idea of where to look can be gained by taking only the first terms of each Bessel function, via the approximation given in Eq. 3.25. The above equation then reduces to:

$$
\begin{aligned}
E_{\text{OUT}} \;=\; & \frac{\tilde{L}+\tilde{R}}{\sqrt{2}} \left[1 + 2\sum_{n=1}^{\infty} \frac{i^n}{2^n n!} \times \right. \\
& \left(\frac{\Re\left\{ \left[\Re\{\tilde{R}\tilde{L}^*\}(\tilde{\delta}_1^n + \tilde{\delta}_2^n) + |L|^2\tilde{\delta}_1^n + |R|^2\tilde{\delta}_2^n\right] e^{in\omega_m t} \right\}}{|\tilde{L}+\tilde{R}|^2} \right. \\
& \left. \left. + i\frac{\Re\left\{ \left[\Im\{\tilde{R}\tilde{L}^*\}(\tilde{\delta}_2^n - \tilde{\delta}_1^n)\right] e^{in\omega_m t} \right\}}{|\tilde{L}+\tilde{R}|^2} \right) \right] e^{i\omega t}.
\end{aligned}
\tag{5.26}
$$

This is starting to look more like the previous equations, and hence one might wonder, for example, when first order sidebands are suppressed if a similar effect is possible with the higher harmonics. The answer is yes. This is demonstrated here in the simplest case of circularly polarised light, leaving the general case open as a future direction for this research.

When selecting circular polarisation, Eq. 5.26 becomes:

$$
\begin{aligned}
E_{\text{OUT}} \;=\; & \frac{E_{\text{IN}}}{\sqrt{2}} \left[1 + \sum_{n=1}^{\infty} \frac{i^n}{2^n n!} \times \right. \\
& \left. \left(\Re\left\{ (\tilde{\delta}_1^n + \tilde{\delta}_2^n)e^{in\omega_m t} \right\} + i\Re\left\{ (\tilde{\delta}_2^n - \tilde{\delta}_1^n)e^{in\omega_m t} \right\} \right) \right] e^{i\omega t}.
\end{aligned}
\tag{5.27}
$$

From this, it is relatively straightforward to determine the types of modulations occurring at each frequency. Table 5.3 gives this information for the first four harmonics, for

Operating point	ω_m	$2\omega_m$	$3\omega_m$	$4\omega_m$
Pure PM	PM	AM	PM	AM
Pure AM	AM	AM	AM	AM
Corr'd PM & AM	PM + AM	PM − AM	PM + AM	PM − AM
Anticorr'd PM & AM	PM − AM	PM + AM	PM − AM	PM + AM
Single sideband $-\omega_m$	SSB $-\omega_m$	AM	SSB $+3\omega_m$	PM
Single sideband $+\omega_m$	SSB $+\omega_m$	AM	SSB $-3\omega_m$	PM

Tab. 5.3: Types of modulations at higher order harmonics produced by the QVM used with circular polarisation.

each of the main operating points listed earlier in Table 5.1. Higher harmonics repeat this pattern when circular polarisation is used.

The higher harmonics of PM and AM have already been seen in Chapter 3. Here, the higher order harmonics of correlated or anticorrelated PM with AM are also of the same M_2-type (see Fig. 5.5(a) and (b)). The single sideband operating points come with a suppressed sideband at each odd harmonic, and either pure AM or pure PM at the even harmonics (see Fig. 5.5(c) and (d)), and hence rotate around in the M_1-M_3 plane.

Perhaps the most crucial observation for these expansions is that all higher-order sidebands are scaled by the same polarisation-dependent factor corresponding to their first order counterparts; this factor does not increase in powers with the same index as that of each sideband but remains the same. As such, this puts a limitation on extent to which one can compensate for a reduction in AM power available, due to a reduction in the proportion of power being tapped off, by increasing the modulation depth in the two crystals. While a constant power of first order AM sidebands can be maintained in this way, higher-order sidebands increase more quickly with the change in modulation depth. In each application, there will be a threshold where the advantage of an increase in transmitted power will be outweighed by problems caused by increasingly strong higher-order sidebands.

5.5   Summary

In this chapter, an electro-optic device called the quadrature variable modulator was presented. Mathematical calculations showed that the QVM can produce arbitrary combinations of PM and AM, provided the input light has some component of circular polarisation. A restriction by design on input polarisation, maintaining equal power components aligned with the respective axis of each crystal, was shown to simplify the electrical-to-optical transfer characteristics of the device, while still allowing enough flexibility to retain the majority of the input light power. The beat phase of the modulation was shown to track the phases of the input electrical signals in a simple way, so that a reliable phase reference for demodulation would be easy to come by. The QVM was shown to be capable of a carrier-suppressed modulation state, consisting of only the sideband pair. The QVM's higher order sideband response was calculated, and exhibited similar patterns of behaviour to that of the first order sidebands. The modulation sphere representation, which was developed with the QVM device in mind, provided a quick visual "look-up"

(a)

(b)

(c)

(d)
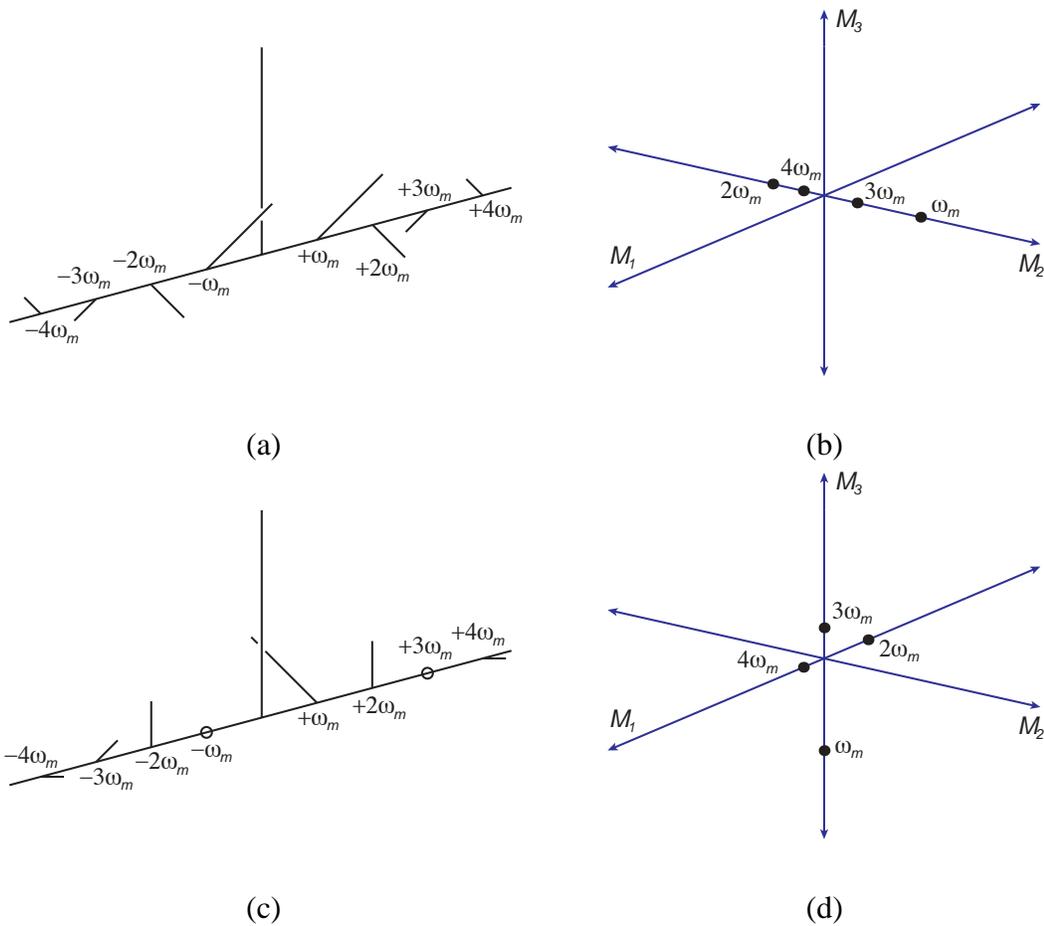
Fig. 5.5: (a) Frequency vector diagram showing higher order sidebands for a correlated PM & AM combination state, generated by a QVM operated with a circularly polarised light state. (b) M-space diagram showing that all higher order components of this state are of the M_2 type. (c) Frequency vector diagram and (d) M-space diagram for a single sideband modulation state (again using circular polarisation through a QVM), which includes both M_1 and M_3 type modulations at higher harmonics.

tool for identifying the electrical signals required to achieve a particular modulation state.

"Fakts still exist, even if they are ignored."

— from *Harvie Krumpet*'s Fakt Notebook

Tune-Locking by Modulation

6.1 Context

One anticipated use for the QVM device is in optical feedback control of the resonance condition of a Fabry-Perot cavity, or the fringe condition of a Michelson interferometer, as discussed in the previous chapter. The rf modulation techniques used to lock these devices are Pound-Drever-Hall locking (discussed in Chapter 4) and Schnupp modulation locking (discussed in Chapter 3), respectively, and are both based on properties of the devices that convert injected PM to AM, which is then demodulated to give an error signal readout. In both cases, the default locking point is at a turning point in the transmitted or reflected optical field intensity.

In this chapter, it will be shown that injecting additional AM along with the usual PM into either a Michelson or a Fabry-Perot cavity causes the device to lock with an offset relative to the default. Moreover, the quadrature of AM required for Pound-Drever-Hall locking is very nearly orthogonal to that required for Schnupp modulation locking, introducing the possibility of independently tuning two lock points at once in a coupled system. This so called *tune-locking* technique may be useful for future large-scale gravitational wave detector configurations [16], with the feature of facilitating real-time tuning of detector frequency responses [9, 11, 12].

Development of a full 6-degree-of-freedom control system for a power recycled Michelson with arm cavities and signal VRM (variable reflectivity mirror) is beyond the scope of this thesis. However, the author has endeavoured to keep existing configurations and control systems in mind, especially those developed in [53–56] which cater for up to 5 degrees of freedom but do not include a signal VRM. It should be possible to augment such systems with the techniques discussed here, and suggestions are given as to how this might be done in each case.

On the other hand, the reintroduction of direct detection into the list of possibilities for gravitational wave interferometer configurations complicates matters. There is currently little research toward locking configurations catering for this possibility; the author is not aware of any publications of the calibre of the corresponding rf detection papers. While it is not clear how carrier laser outcoupling will mesh with the presence of the signal recycling mirror, it seems sensible that if this problem can be solved then the extra degree of freedom of replacing the signal mirror with a VRM should not cause any new complications.

The main Michelson degree of freedom is *not* intended to be tune-locked with these
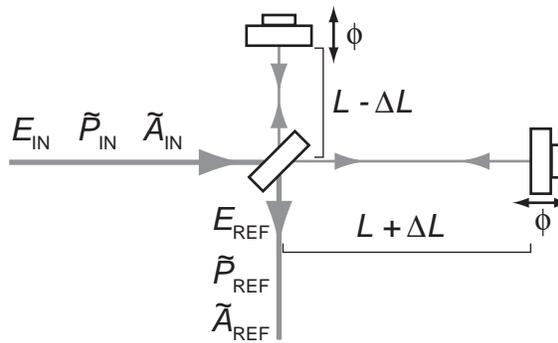
Fig. 6.1: Michelson interferometer with the Schnupp asymmetry. In this section, the Michelson's transmission transfer function for modulation is calculated by relating the parameters shown.

techniques. The tuning methods are based on rf locking; using them to offset lock the main Michelson degree of freedom as part of a direct detection control scheme defeats the purpose of the idea, since the quantum noise associated with the rf detection would still apply. The locking techniques presented here are specifically aimed at detuning the signal cavity and tuning the reflectivity of the signal VRM, in such a way as to change the gravitational wave detector's frequency response.

6.2   Tuning the Lock Point of a Michelson

Here, the modulation transfer function of a Michelson interferometer is derived. In other words, the modulation state input parameters of a Michelson are related to the parameters of the modulation state that is transmitted through to the Michelson dark port.

The transfer relations will of course depend on the Michelson's operating point, which is parametrised in Eq. 3.32, with the variable fringe condition phase ϕ, by:

$$\Delta L(t) \leftarrow \Delta L + \frac{\phi(t)c}{2\omega}. \tag{6.1}$$

Upon substituting this into the Michelson's transfer function, from Eq. 3.31, the result is:

$$\frac{E_{\text{OUT}}}{E_{\text{IN}}}(\omega + \omega_n) = 2t_{\text{BS}}r_{\text{BS}}e^{i\omega_n 2L/c}\sin(\omega_n 2\Delta L/c + \phi), \tag{6.2}$$

where the transfer function's dependence on frequency is written explicitly, and the laser's frequency, ω, and Michelson dimensions, L and ΔL, are calibrated such that both $2\omega L/c = 2\pi j$ and $2\omega\Delta L/c = 2\pi k$ for integers j and k,ie such that the Michelson defaults to a dark fringe.

The input electric field is written using the sideband decomposition as per Eq. 3.11, with a generic single-frequency modulation at ω_m:

$$E_{\text{IN}} = E_0 \left[1 + \tilde{S}_{\text{IN}}^+ e^{i\omega_m t} + \tilde{S}_{\text{IN}}^- e^{-i\omega_m t}\right] e^{i\omega t} \tag{6.3}$$

in order that each frequency component (carrier and sidebands) can be treated separately with individual applications of Eq. 6.2. The output electric field at the Michelson dark

port is then:

$$
\begin{aligned}
E_{\text{OUT}} \;=\; & 2t_{\text{BS}}r_{\text{BS}}E_0 \Big[\sin\phi \\
& + e^{i\omega_m 2L/c}\sin(\omega_m 2\Delta L/c + \phi)\tilde{S}^+_{\text{IN}}e^{i\omega_m t} \\
& - e^{-i\omega_m 2L/c}\sin(\omega_m 2\Delta L/c - \phi)\tilde{S}^-_{\text{IN}}e^{-i\omega_m t} \Big] e^{i\omega t}.
\end{aligned}
\tag{6.4}
$$

The aim is to see how PM and AM components transmit through a Michelson, so the conversion is made from sideband notation via:

$$
\begin{aligned}
\tilde{P} \;&=\; \frac{\left(\tilde{S}_+ - \tilde{S}_-^*\right)}{i} \\
\tilde{A} \;&=\; \left(\tilde{S}_+ + \tilde{S}_-^*\right),
\end{aligned}
\tag{6.5}
$$

which gives

$$
E_{\text{OUT}} = 2t_{\text{BS}}r_{\text{BS}}E_0 \left[\sin\phi + \Re\{\tilde{A}_{\text{OUT}}e^{i\omega_m t}\} + i\Re\{\tilde{P}_{\text{OUT}}e^{i\omega_m t}\} \right] e^{i\omega t},
\tag{6.6}
$$

with

$$
\tilde{A}_{\text{OUT}} = e^{i\omega_m 2L/c} \left[i\tilde{P}_{\text{IN}}\sin(2\omega_m \Delta L/c)\cos\phi + \tilde{A}_{\text{IN}}\cos(2\omega_m \Delta L/c)\sin\phi \right],
\tag{6.7}
$$

and

$$
\tilde{P}_{\text{OUT}} = e^{i\omega_m 2L/c} \left[\tilde{P}_{\text{IN}}\cos(2\omega_m \Delta L/c)\sin\phi - i\tilde{A}_{\text{IN}}\sin(2\omega_m \Delta L/c)\cos\phi \right].
\tag{6.8}
$$

This can be abbreviated into matrix form via:

$$
\begin{bmatrix} \tilde{P}_{\text{OUT}} \\ \tilde{A}_{\text{OUT}} \end{bmatrix} = e^{i\omega_m 2L/c} \begin{bmatrix} \cos(2\omega_m \Delta L/c)\sin\phi & -i\sin(2\omega_m \Delta L/c)\cos\phi \\ i\sin(2\omega_m \Delta L/c)\cos\phi & \cos(2\omega_m \Delta L/c)\sin\phi \end{bmatrix} \begin{bmatrix} \tilde{P}_{\text{IN}} \\ \tilde{A}_{\text{IN}} \end{bmatrix}
\tag{6.9}
$$

The transfer matrix has pure real components on the diagonal, and pure imaginary components off-diagonal. If AM or PM are transmitted as is then they are transmitted with their beat phase unmodified, though the amount of modulation may have been attenuated[1]. By contrast, if PM is converted to AM or vice versa, it is converted with a 90° shift in beat phase. This is characteristic of the Michelson's differential mode property of changing the amplitude of a frequency component but not changing its phase.

To get a meaningful error signal, one is required to demodulate in quadrature relative to the injected PM. Eq. 6.7 shows that, if one wishes to move the locking point, they will need to inject AM that is in quadrature with the injected PM. Any component of injected AM that is in-phase with the PM will show up in the wrong quadrature, and hence not be detected in the mixdown.

To solidify the above observations, the beam intensity and demodulated voltage should

---

[1] There is, of course, an overall beat phase shift factored out of the matrix, but this is simply a phase shift accumulated when traversing the length of the Michelson arms, which any beam accrues when travelling any distance. It is safe to remove it, since it gets experimentally "factored out" by adjusting the demodulation phase as described in Chapter 3.
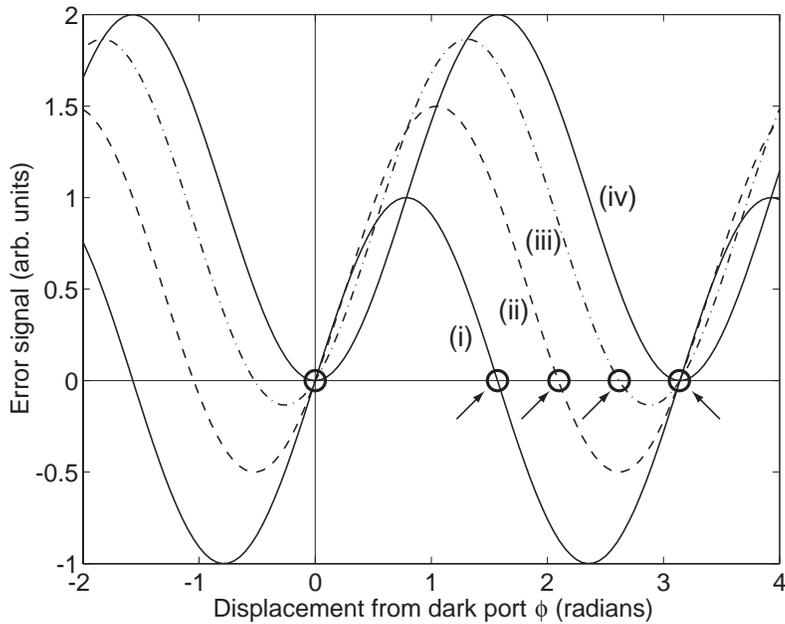
Fig. 6.2: Series of Michelson tune-locking error signals. The lock point corresponding to the dark port does not shift at all, but the lock point at the bright port shifts as indicated by arrows. Functions plotted are described in the text, with: (i) PM injected only; (ii) PM and AM injected ($p/a = \sqrt{3}$); (iii) PM and AM injected ($a/p = \sqrt{3}$); and (iv) AM injected only. In the last case, the function only touches zero and hence is not a legitimate error signal.

be examined. They are, respectively:

$$
\begin{aligned}
I_{\text{OUT}} = \ & 4t_{\text{BS}}^2 r_{\text{BS}}^2 E_0^2 \Big[\sin(2\omega_m \Delta L/c)\sin 2\phi \Re\{\tilde{P}_{\text{IN}} i e^{i\omega_m 2L/c} e^{i\omega_m t}\} \\
& + \cos(2\omega_m \Delta L/c)(1 - \cos 2\phi)\Re\{\tilde{A}_{\text{IN}} e^{i\omega_m 2L/c} e^{i\omega_m t}\} \Big]
\end{aligned}
\tag{6.10}
$$

and

$$
\begin{aligned}
V_{\text{mix}} \propto \ & 4t_{\text{BS}}^2 r_{\text{BS}}^2 E_0^2 \Big[\sin(2\omega_m \Delta L/c)\sin 2\phi \Re\{\tilde{P}_{\text{IN}} i e^{-\phi_{\text{mix}}} e^{i\omega_m 2L/c}\} \\
& + \cos(2\omega_m \Delta L/c)(1 - \cos 2\phi)\Re\{\tilde{A}_{\text{IN}} e^{-\phi_{\text{mix}}} e^{i\omega_m 2L/c}\} \Big],
\end{aligned}
\tag{6.11}
$$

where ϕ_{mix} is the phase angle of the electronic oscillator used in the demodulation[2].

In a nutshell, the demodulated voltage signal looks like a weighted sum of two sinusoidal functions of the error ϕ, ie $p\sin 2\phi + a(1 - \cos 2\phi)$. The first term appears due to the presence of injected PM, and the second from injected AM. These functions are plotted in Fig. 6.2 to show the shape of the error signal.

A number of observations are apparent:

- the error signal produced by injecting *only* AM (ie the function $1 - \cos 2\phi$) does not have a zero crossing, and hence does not function as an error signal;

---

[2] The special case of this equation where $\tilde{A}_{\text{IN}} = 0$ was given previously as Eq. 3.33

- the lock point corresponding to a dark fringe *does not* shift in response to injecting AM in addition to PM;

- the lock point corresponding to a bright fringe *does* shift depending on the PM/AM balance;

The second observation is critical because it means that tune-locking the error signal crossing point away from the dark fringe is actually impossible using modulation injection techniques[3]. However, this does not prohibit the idea altogether, since it is possible to instead lock to the error signal crossing corresponding to the bright fringe which *can* be tuned. The tuning range should be quite broad, though the lock becomes more and more likely to jump a fringe (or several) the closer one tunes to the dark fringe, since the error signal has decreasing swing range in one direction.

It is worth pausing to examine where the factors of 2ϕ have come from in the arguments of the sinusoids, considering that a Michelson's electric field response includes sinusoids with arguments of just ϕ. The answer is in the trigonometric identities that relate the product of two sinusoids to another sinusoid with twice the frequency. Two sinusoids enter the equation when the power of the modulation beat frequency is detected. One sinusoid comes from the Michelson's response to the carrier field, and the other comes from the Michelson's response to the sideband field. If these two sinusoids are in-phase, they multiply together to produce an average level away from zero (eg $2\sin^2 x = 1 - \cos 2x$ where the 1 represents the shift to the average level of the sinusoid). This shift is just enough to prevent the sinusoid crossing zero, though it still touches zero. If the two sinusoids are in quadrature, the product still averages out to zero (eg $2\sin x \cos x = \sin 2x$), so that zero crossings result. A glance at Eqs. 6.6 and 6.7 shows that the carrier and AM components are transmitted with a factor of $\sin\phi$ but the PM component transmission carries the factor $\cos\phi$, and hence some component of PM must be input in order to produce a zero-crossing error signal.

6.2.1   The Tuned Error Signal's Locking Range

As stated above, when the lock point is tuned away from the bright fringe, the range of swing that the system has before dropping lock is reduced. An accurate model of lock dropping is well beyond the scope of this thesis, but some useful information can be obtained by calculating the distance that the interferometer can be displaced before dropping lock or jumping a fringe. One might imagine a situation where the interferometer gets a jolt that shifts the operating point by some displacement in a time faster than the feedback system can respond. For the interferometer to jump a fringe, the jolt would have to carry the interferometer *beyond the next zero crossing* of the error signal. The next zero crossing will of course be of opposite polarity to that of the current lock point, and so signifies the point where the feedback force changes sign and pushes the interferometer on to the next fringe. Therefore, calculating the lock range simply means calculating the displacement

---

[3] On reflection, this is self-evident. There must always be a zero crossing in a modulation-based error signal when the Michelson is at a dark port, because there is no carrier to beat with.

to the next nearest zero crossing of the error signal. This is straightforward, and it can be shown that the zero crossings occur at:

$$\phi_{\text{dark}} = 0 \pm \pi n, \qquad\qquad n = 0, 1, 2, ... \qquad\qquad (6.12)$$

(the fixed dark-port zero crossings) and at

$$\phi_{\text{bright}} = \tan^{-1}\left(-\frac{p}{a}\tan\eta\right) \pm \pi n, \qquad n = 0, 1, 2, ... \qquad (6.13)$$

(the variable bright-port zero crossings), where p and a are the projections of PM and AM in the detection quadrature (ie the result of evaluating the $\mathfrak{R}$ operator on the contents of the corresponding brackets in Eq. 6.11).

The interferometer can hence be displaced by up to ϕ_{bright} radians or by $\pi - \phi_{\text{bright}}$ radians, whichever is smaller. Ordinarily, the maximum displacement is $\pi/2$. The ratio of ϕ_{bright} to $\pi/2$ ought to give an indication of the fraction that the locking "tightness" is reduced by as a result of selecting an offset-locked operating point.

The corresponding transmissivity of the Michelson, taken as a unit, with respect to the carrier is:

$$\frac{E_{\text{OUT}}}{E_{\text{IN}}} = 2t_{\text{BS}}r_{\text{BS}}\sin\phi_{\text{bright}}. \qquad\qquad (6.14)$$

For an example of using these equations, if $p = a$ and $\eta = \pi/4$ (sidebands placed halfway up a fringe by default), the result is $\phi_{\text{bright}} = -\pi/4$ or $3\pi/4$, which means that the Michelson will lock the carrier to halfway up a fringe, for a transmissivity of $1/\sqrt{2}$, and the Michelson will be locked half as tightly as usual.

A few quick calculations then show that a Michelson VRM can be locked to 50% reflectivity with half the locking tightness of a dark-port-locked Michelson, or 75% reflectivity with one third the locking tightness. Achieving very high VRM reflectivities of say 99% would mean locking very close to a dark port (approximately $8°$, or one tenth the locking tightness) and would be unreasonable. However, in most gravitational wave detector schemes, the reflectivity of the signal recycling mirror is likely to be in the mid range values. Hence this method is feasible for tuning the reflectivity of a signal VRM.

6.2.2   M-space Transfer Function

It is also possible to transform the above modulation transfer functions into the modulation sphere representation, by applying Eq. 3.27 to both input and output modulations. The result in matrix form is:

$$\begin{bmatrix} M_0 \\ M_1 \\ M_2 \\ M_3 \end{bmatrix}^{\text{OUT}} = \frac{1}{2}\begin{bmatrix} 1 - \cos2\xi\cos2\phi & 0 & 0 & -\sin2\xi\sin2\phi \\ 0 & \cos2\xi - \cos2\phi & 0 & 0 \\ 0 & 0 & \cos2\xi - \cos2\phi & 0 \\ -\sin2\xi\sin2\phi & 0 & 0 & 1 - \cos2\xi\cos2\phi \end{bmatrix}\begin{bmatrix} M_0 \\ M_1 \\ M_2 \\ M_3 \end{bmatrix}^{\text{IN}}$$
$$(6.15)$$

where $\xi \equiv 2\omega_m\Delta L/c$. The proof is straightforward and is omitted for the sake of brevity.

The above *M-space to M-space* transfer function is interesting from an academic perspective, as it is an example of the transfer characteristics of a system that can change
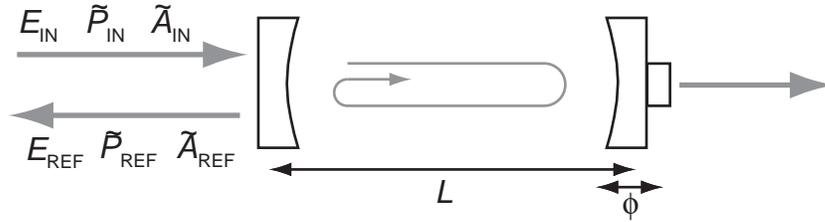
Fig. 6.3: Fabry-Perot cavity schematic. In this section, the cavity's reflection transfer function for modulation is calculated by relating the parameters shown.

the amplitude, but not the phase, of frequencies that it transmits. All such systems will have zeros in their transfer matrices in the same pattern as here, ie only two off-diagonal components will be nonzero. These off-diagonal components show that the system is capable of mixing M_3 type modulation with M_0 type modulation, but otherwise does not mix modulation types. The physical interpretation of this is that a Michelson can unbalance (or re-balance) a set of sidebands, but is otherwise insensitive to the details of the modulation state.

These observations do not in themselves tell anything new about the Michelson's transfer characteristics; they merely provide a new language with which to express and discuss the Michelson's properties. It may however be of interest to consider the type of device which has non-zero elements in the transfer matrix in place of the zeros in Eq. 6.15. This theme is discussed further in the next section.

6.3 Tuning the Lock Point of a Cavity

The modulation transfer function of a Fabry-Perot cavity is now derived. As with the Michelson treatment above, a variable resonance condition phase $\phi(t)$ is defined by[4]:

$$L(t) \leftarrow L + \frac{\phi(t)c}{2\omega}, \tag{6.16}$$

for cavity length L and laser frequency ω.

From Eq. 4.6, the reflection transfer function of a Fabry-Perot is:

$$\frac{E_{\text{REF}}}{E_{\text{IN}}}(\omega + \omega_n) = \frac{r_1 - r_2(1 - a^2)e^{i2\omega_n L/c}e^{i\phi}}{1 - r_1 r_2 e^{i2\omega_n L/c}e^{i\phi}}, \tag{6.17}$$

where the transfer function's dependence on frequency is highlighted, and the laser-cavity system is set by default to resonance at $\phi = 0$ (mathematically $2\omega L/c = 2\pi j$ for integer j).

---

[4] Again the equation here isn't strictly valid as the error ϕ should have units of length, being a mechanical offset. However, the light frequencies dealt with in practice are so close as to make this irrelevant, and the equation above is very nearly exact.

The input modulated field, again explicitly writing individual sidebands, is:

$$E_{\text{IN}} = E_0 \left[1 + \tilde{S}_{\text{IN}}^+ e^{i\omega_m t} + \tilde{S}_{\text{IN}}^- e^{-i\omega_m t} \right] e^{i\omega t}, \tag{6.18}$$

and Eq. 6.17 is applied for each frequency component to obtain the reflected field:

$$E_{\text{REF}} = E_0 \left[f(\phi) + g(\phi,\eta)\tilde{S}_{\text{IN}}^+ e^{i\omega_m t} + g(\phi,-\eta)\tilde{S}_{\text{IN}}^- e^{-i\omega_m t} \right] e^{i\omega t}, \tag{6.19}$$

where the following parameterisations have been made:

$$g(\phi,\eta) = \frac{r_1 - r_2(1-a^2)e^{i\eta}e^{i\phi}}{1 - r_1 r_2 e^{i\eta}e^{i\phi}}, \tag{6.20}$$

$$f(\phi) = g(\phi,0) = \frac{r_1 - r_2(1-a^2)e^{i\phi}}{1 - r_1 r_2 e^{i\phi}}, \tag{6.21}$$

$$\eta = \frac{2\omega_m L}{c}. \tag{6.22}$$

Since the term $f(\phi)$ is complex, it will need to be factored out in order to determine PM and AM components. $f(\phi)$, which is identically the same as the cavity's reflection for the carrier, goes to zero for an impedance matched cavity which is exactly on resonance, and so there will be a division by zero problem in that case, but otherwise it is safe to factor it out. The reflected field, in terms of PM and AM components is then:

$$E_{\text{REF}} = E_0 f(\phi) \left[1 + \Re\{\tilde{A}_{\text{OUT}} e^{i\omega_m t}\} + i\Re\{\tilde{P}_{\text{OUT}} e^{i\omega_m t}\} \right] e^{i\omega t}, \tag{6.23}$$

$$\begin{bmatrix} \tilde{P}_{\text{OUT}} \\ \tilde{A}_{\text{OUT}} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \left(\dfrac{g(\phi,\eta)}{f(\phi)} + \dfrac{g(-\phi,\eta)}{f(-\phi)} \right) & -i\left(\dfrac{g(\phi,\eta)}{f(\phi)} - \dfrac{g(-\phi,\eta)}{f(-\phi)} \right) \\ i\left(\dfrac{g(\phi,\eta)}{f(\phi)} - \dfrac{g(-\phi,\eta)}{f(-\phi)} \right) & \left(\dfrac{g(\phi,\eta)}{f(\phi)} + \dfrac{g(-\phi,\eta)}{f(-\phi)} \right) \end{bmatrix} \begin{bmatrix} \tilde{P}_{\text{IN}} \\ \tilde{A}_{\text{IN}} \end{bmatrix}. \tag{6.24}$$

The keen observer will note that these equations form a template for *any* optical device that has a well-defined, linear, frequency-dependent transfer function $g(\phi,\eta)$, and so wone can also make general pronouncements about the properties of said subset of optical systems.

At first glance, Eq. 6.24 might seem to indicate that, if there is conversion between PM and AM, that it is done in quadrature as with the Michelson case above, due to the presence of the i on the diagonal term. However, this time the matrix elements are themselves complex, so it isn't as intuitively obvious what quadratures the modulations will appear in.

Figs. 6.4 and 6.5 show the functional form of these matrix elements and their component functions. In essence, the important and interesting features of the functions occur when either the carrier or a sideband goes through resonance (mathematically when $\phi = 0$ or when $\phi = \pm\eta$). Fig. 6.5 also shows that the four matrix elements from Eq. 6.24 have much smaller imaginary components than real components, even in spite of the explicit factors of i on the diagonal elements. The imaginary components only become large when a sideband imbalance is caused, which happens when one sideband passes through the resonance of the cavity, and is hence reflected at a much smaller amplitude. Near car-
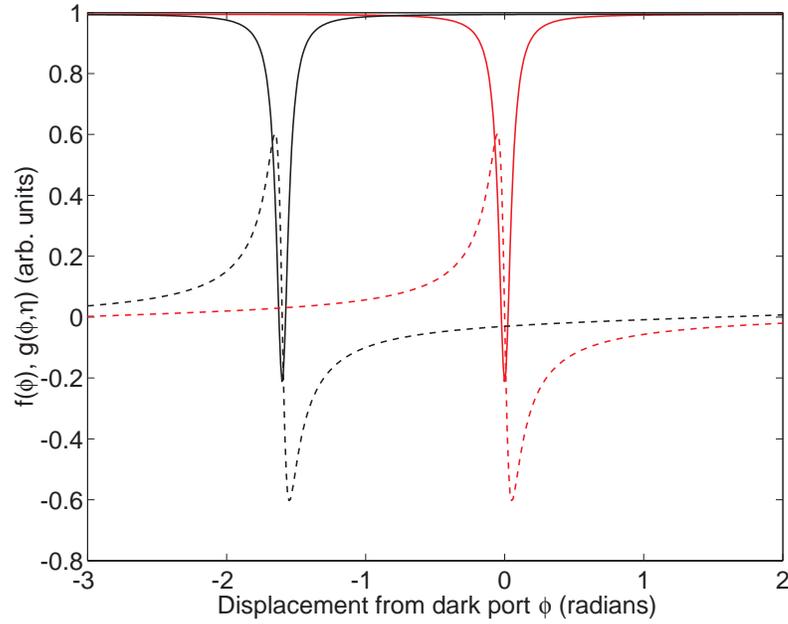
Fig. 6.4: Real and imaginary parts (solid and dashed lines respectively) of functions $f(\phi)$ (centered at $\phi = 0$) and $g(\phi, \eta)$ (centered at $\phi = -\eta$) for $\eta = 1.6$ radians. These functions are simply the cavity's dispersive reflection response, centered either at zero or at η. The ratio of these functions gives the transfer matrix elements in Eq. 6.24.

(a)   (b)
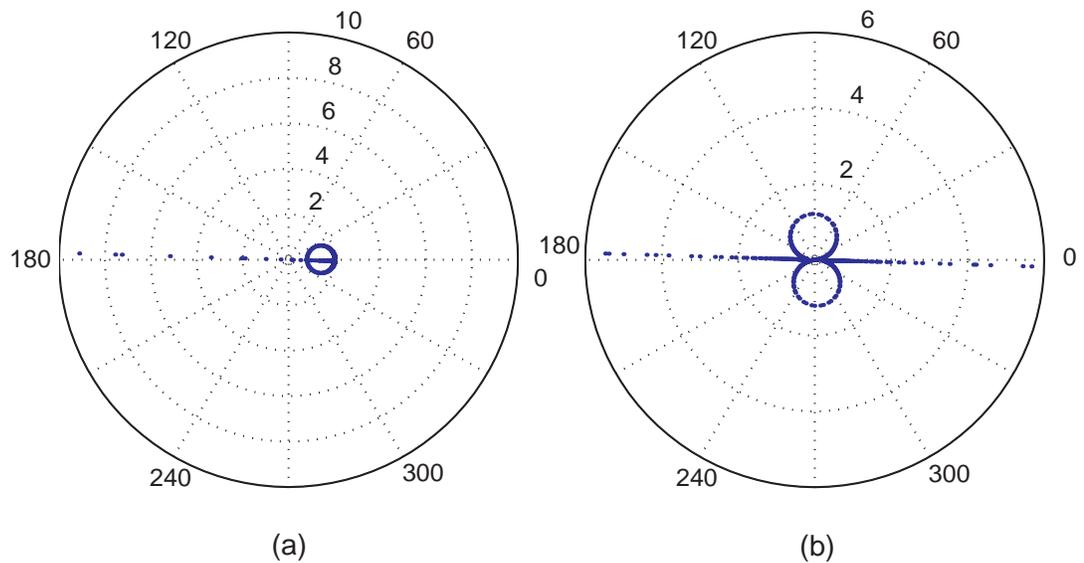
Fig. 6.5: Complex diagrams of transfer coefficients from Eq. 6.24. (a) shows the function $\frac{g(\phi,\eta)}{f(\phi)} + \frac{g(-\phi,\eta)}{f(-\phi)}$ and (b) shows the function $i\left(\frac{g(\phi,\eta)}{f(\phi)} - \frac{g(-\phi,\eta)}{f(-\phi)} \right)$. Both functions stay mainly on the real axis. The exception is the loops which occur when a sideband passes through resonance, thus producing a large sideband mismatch in the reflected field. There is also a gentle slope away from the real axis, which represents the slight sideband imbalance when the carrier is only partially resonant.

rier resonance at least, PM is converted to AM in-phase and vice versa, rather than being converted to quadrature as in the Michelson case. This is consistent with the fact that a Fabry-Perot has a strong effect on the *angle* that modulations make with their carrier, but generally causes only a small degree of unbalancing in sideband power, particularly for a high finesse cavity, where non-resonant frequencies are almost entirely reflected. In fact, if the sidebands are set such that they are exactly non-resonant when the carrier is resonant ($\eta = \pi$), the imaginary component of the matrix transfer coefficients disappears altogether.

The consequence of this analysis is that, if the experimenter wishes to inject AM to shift the lock point, they should inject AM in-phase with the present PM, as expected. This is then compatible with Michelson tune-locking; there is no fundamental reason why both a Michelson and a Fabry-Perot in the same optical system cannot both be tune-locked using the same modulation frequency.

Continuing on, the error signal voltage after photodetection and mixdown can be shown to be:

$$V_{\text{mix}} \propto \frac{E_0^2}{2}|f(\phi)|^2 \Re \left\{ \tilde{A}_{\text{IN}} \left(\frac{g(\phi,\eta)}{f(\phi)} + \frac{g(-\phi,\eta)}{f(-\phi)} \right) e^{-\phi_{\text{mix}}} \right.$$
$$\left. + \tilde{P}_{\text{IN}} i \left(\frac{g(\phi,\eta)}{f(\phi)} - \frac{g(-\phi,\eta)}{f(-\phi)} \right) e^{-\phi_{\text{mix}}} \right\}, \qquad (6.25)$$

or more compactly

$$V_{\text{mix}} \propto \frac{E_0^2}{2}|f(\phi)|^2 \Re \left\{ \frac{g(\phi,\eta)}{f(\phi)} \tilde{S}_{\text{IN}}^+ e^{-\phi_{\text{mix}}} + \frac{g(\phi,-\eta)}{f(\phi)} \tilde{S}_{\text{IN}}^- e^{+\phi_{\text{mix}}} \right\}. \qquad (6.26)$$

This error signal, for a selection of AM/PM balances, has been plotted in Fig. 6.6, and shows clearly that the zero crossing that is normally used to lock can be tuned to the side of resonance. There is a point where this zero crossing disappears, which occurs outside the cavity resonance.

6.3.1 Locking Range and Feasibility

Determining the locking range of a tune-locked cavity is difficult to do in general; the mathematics have already become quite cumbersome. Numerical modelling ought to be employed on a case-by-case basis. A qualitative analysis, however, is still enlightening here.

It seems, from Fig. 6.6, that the system can only be reliably asked to tune the cavity resonance to within half a linewidth of the default, since the error signal crossing disappears shortly. Clearly then, a Fabry-Perot offset locking system, based on modulation such as this, requires for some significant laser frequency component to be on or near resonance in order to produce a tune-locking signal. This restricts the possibilities for a system to be used in tuning a gravitational wave detector's frequency response. For example, a narrowband detuned frequency response would be impossible with this system if the centre frequency was well away from any carrier or modulation sideband frequencies in the system. One would have to include an extra frequency sideband near the desired
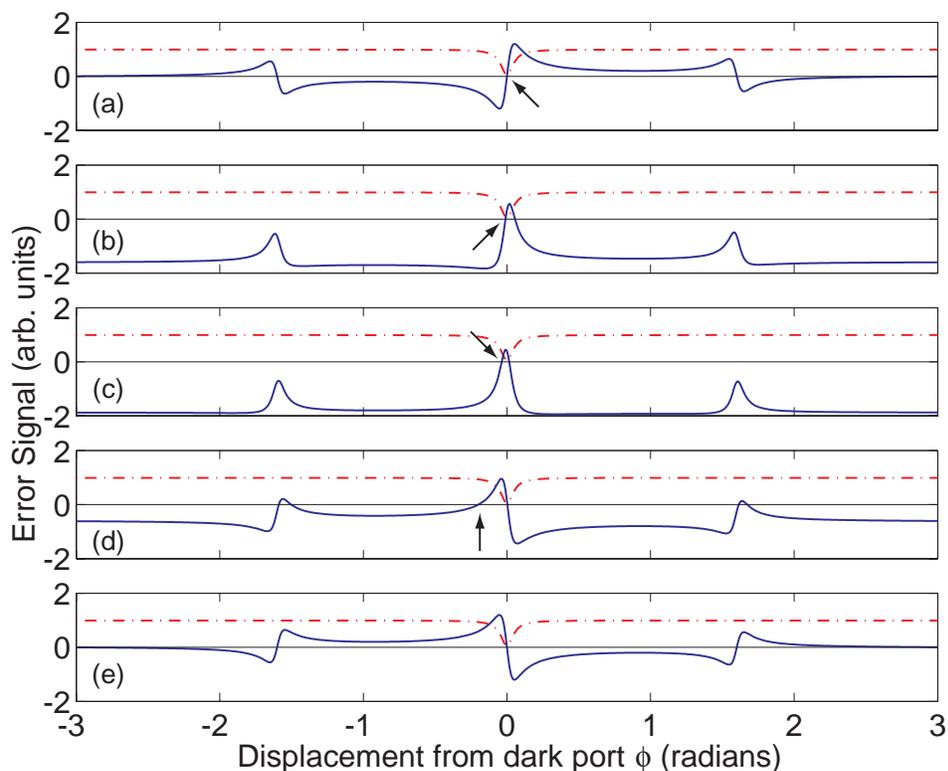
Fig. 6.6: Series of Fabry-Perot cavity tune-locking signals (solid lines). The lock point (indicated with arrows) shifts all the way to the foot of the cavity reflectivity resonance (shown with a dashed line for comparison) before the error signal finally loses the corresponding zero crossing. Diagrams shown are: (a) pure PM ($\theta = 0$); (b) PM and AM ($\theta = 3\pi/10$); (c) near pure AM ($\theta = 6\pi/10$); (d) PM and AM ($\theta = 9\pi/10$); and (e) pure PM of opposite polarity to the original ($\theta = \pi$). θ dictates the relative proportions of in-phase PM and AM combinations (M_1 and M_2 type modulations), and is defined in Fig. 3.11(a).

signal-resonance frequency, and since systems designed in this way are already capable of tuning the signal cavity resonance (for example [28]), this modulation offset-locking scheme would be redundant.

However, the system in [28] tunes the signal cavity by tuning a laser frequency that must take discrete jumps in frequency in order to pass through the input mode-cleaners (the jumps are equal to the mode-cleaner's free spectral range), which may prevent tuning while maintaining detector lock. In this case, the modulation offset locking system presented here may be useful as an augmentation that allows small continuous detunings while maintaining lock.

The system seems more feasible for broadband frequency responses to gravitational waves, when the cavity resonance is wide enough to include the carrier inside its frequency band. There is a secondary issue in this case: the carrier is not generally supposed to resonate inside the signal cavity, if the main Michelson is locked to a dark port, and hence the carrier is not available as a locking reference. However, as the possibility of transmitting some part of the carrier through the interferometer has recently been reintroduced [27], this variation should not be overlooked.

6.3.2   M-space Transfer Function

In the modulation sphere representation, Eq. 6.24 becomes:

$$
\begin{bmatrix} M_0 \\ M_1 \\ M_2 \\ M_3 \end{bmatrix}^{\text{OUT}} = \frac{1}{2} \begin{bmatrix} A & 0 & 0 & B \\ 0 & C & D & 0 \\ 0 & D & C & 0 \\ B & 0 & 0 & A \end{bmatrix} \begin{bmatrix} M_0 \\ M_1 \\ M_2 \\ M_3 \end{bmatrix}^{\text{IN}}
\tag{6.27}
$$

$$
\begin{aligned}
A &= \left| \frac{g(\phi,-\eta)}{f(\phi)} \right|^2 + \left| \frac{g(\phi,\eta)}{f(\phi)} \right|^2 \\
B &= \left| \frac{g(\phi,-\eta)}{f(\phi)} \right|^2 - \left| \frac{g(\phi,\eta)}{f(\phi)} \right|^2 \\
C &= \Re \left\{ \frac{g(\phi,\eta)g(\phi,-\eta)}{f(\phi)^2} \right\} \\
D &= -\Im \left\{ \frac{g(\phi,\eta)g(\phi,-\eta)}{f(\phi)^2} \right\}
\end{aligned}
$$

The M-space transfer matrix has zeros for half of the entries. The matrix shows the system is capable of mixing between M_0 and M_3 types of modulation, and now also between M_1 and M_2 types of modulation. The latter mixing represents the ability of the Fabry Perot to give a phase shift to the carrier that is different from the phase shift given to the sidebands. The Fabry-Perot also has the ability to attenuate the sidebands to different degrees as the Michelson did in the previous section, but mostly does this to a very small degree, the exception being when one sideband passes through resonance. Earlier, it was pointed out that this ability disappears when $\eta = \pi$ (sidebands exactly non-resonant) since no matter what the error ϕ from resonance for the carrier, the sidebands are reflected with the same amplitude. Mathematically, $\eta = \pi$ means that $g(\phi,\eta) = g(\phi,-\eta)$ and so the parameter B in the above matrix disappears, confirming the above physical interpretation.

As noted before, the above transfer matrix Eq. 6.27 must be the general form of transfer matrices for optical systems with a well-defined linear frequency-dependent transfer function (ie for linear optical systems). This begs the question: "What are the properties of a system that has nonzero entries in place of the zeros seen here?" An example is a system that has non-diagonal terms mixing M_0 and M_1 components. Just as $M_0 - M_3$ mixing represents the ability of a system to pass a lower sideband and an upper sideband unequally, so $M_0 - M_1$ mixing must represent the ability of a system to pass either AM or PM preferentially. Such a property seems to correspond to some class of nonlinear optical devices. There are four such mixing effects representable here (corresponding to the eight remaining zeros) in addition to the two linear mixing mechanisms already discussed.

It is not clear what nonlinear devices this transfer matrix is capable of representing, though it is clear that nonlinear systems involving frequency-shifting *cannot* be represented here. This topic represent a future direction in this field of research.

6.4 Summary

A technique of tune-locking optical systems by varying the injected modulation state was theoretically investigated. Optical locking schemes using modulation usually require that PM is injected into the system, whose dispersive response converts some component to AM. The technique involves injecting some component of AM to tune the system's lock point away from its default.

It was found that the lock point corresponding to a Michelson's dark fringe could not be tuned, but that the bright fringe lock point could be tuned over a significant range. Low to mid-range reflectivities (up to say 75%) could be achieved for a Michelson VRM, which makes the technique a possible option for use with a Michelson VRM in place of the signal mirror in a gravitational wave detector configuration.

It was shown that a Fabry-Perot cavity could be tuned over a range where the laser carrier was resonant or near-resonant. From this, it was inferred that complicated systems would need some laser frequency component to be on or near resonance of a cavity that one is attempting to tune-lock. The technique could be used in gravitational wave configurations for detuning the signal cavity between broadband operating points, but narrowband detuning would require an extra laser frequency component near the required narrowband centre frequency.

An M-space transfer function analysis demonstrated two types of modulation mixing possible in linear optical systems: those that introduce a differential between sideband magnitudes, and those that introduce a phase shift between sidebands and the carrier. The former effect is the primary effect producing an error signal for a Michelson, and the latter effect is dominant for a Fabry-Perot error signal. The two effects are orthogonal, in that the AM required to tune-lock a cavity is of the opposite quadrature to the AM required to tune-lock a Michelson. Therefore, the possibility exists of a system with both Michelson-like and a cavity-like degrees of freedom that are tune-locked using one modulation frequency.

"The large print giveth; the small print taketh away."

— Tom Waits

# Experimental Characterisation of a Prototype Qaudrature Variable Modulator

In this chapter, the results of an experiment to test and characterise a prototype quadrature variable modulator (QVM) device are reported. The device is tested with the application of tune-locking for gravitational wave detectors in mind, but other applications are also considered.

7.1   QVM in Experimental Context

The QVM is by no means the first such device of its kind, with other optical configurations capable of producing arbitrary combinations of amplitude and phase modulations already in use. These include an arrangement of a phase modulator and amplitude modulator in series [57], and an arrangement of two phase modulators used in a Mach-Zehnder interferometric setup. The latter is known as a Mach-Zehnder modulator (MZM), and was described earlier in Section 3.2.4. A fibre-optic version of the MZM is commercially available as a one-piece unit [58]. MZM devices are well-characterised elsewhere, particularly for fibre-optic applications [59]. The QVM is presented as an alternative that may be more suitable in certain circumstances, and it is arguably more accurate, more versatile, and less vulnerable to drift.

One anticipated use for the QVM device is in optical feedback control of the resonance condition of a Fabry-Perot cavity, or the fringe condition of a Michelson interferometer, as discussed in the previous chapter. This *rf tune-locking* application is used as a case study example to demonstrate the variability of a prototype QVM.

The QVM device is capable of single sideband (SSB) modulation, which requires an equal combination of AM and PM in quadrature. The need for this has arisen a number of times with optic fibre technology where, for example, a SSB-modulated signal is immune to fibre dispersion penalties [59]. The technique has also been suggested for subcarrier-multiplexing systems [60]. SSB modulation has been achieved by other means: optically filtering out one sideband [61]; cascaded amplitude and phase modulators [57]; Mach-Zehnder modulators [59]; and more complex arrangements [62, 63]. Here, the QVM is demonstrated to be at least comparable regarding suppression capabilities, achieving sideband suppression of 39 dB.

(a)　　　　　　　　　　　　　　　　　　(b)
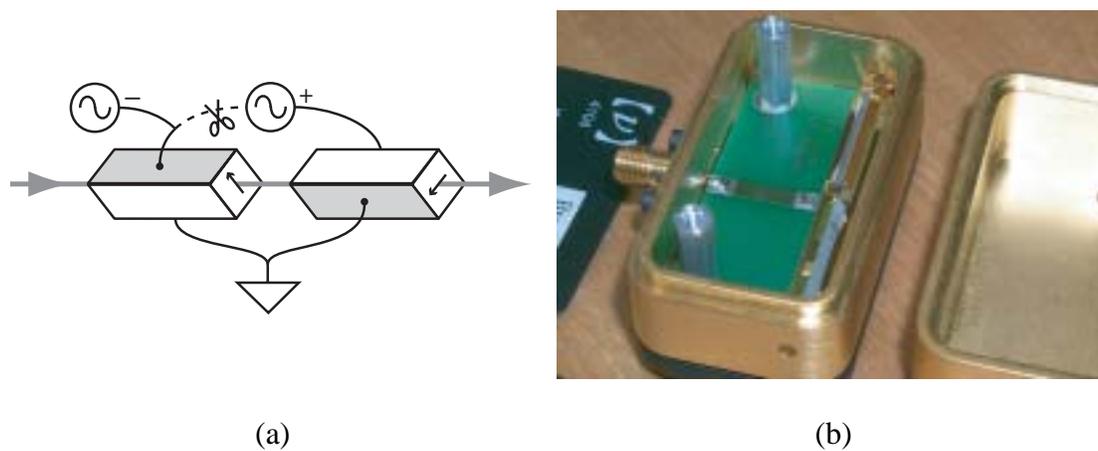
Fig. 7.1: (a) Schematic showing the operation converting an amplitude modulator into a QVM. Since an AM device is hard-wired with opposite polarity for the two crystals, one must take this into account and include a polarity reversal in one of the applied electrical signals, in order to achieve the QVM behaviour described in this thesis. (b) A photograph of the QVM schematic before the operation. Two gold wires connect the electrode to the crystals. One wire was cut and connected to a second circuit board which was placed over the top of the existing wiring, and connected to a second external electrode.

The QVM can produce PM or AM states, with purity limited by the accuracy of po-larisation optics including the birefringent effects of the device itself. While these states are obtainable using off-the-shelf amplitude or phase modulators, applications in coherent state quantum cryptography require fast-switching between AM and PM [64], for which the QVM is ideal. Other applications, including some quantum communication proto-cols [65, 66], also benefit from having easy access to a tuned, stationary combination of AM and PM. In these situations, it is essential that the modulation device maintains a high level of accuracy when holding a selected modulation state. The QVM was able to produce a pure PM state, with the residual AM beat suppressed by 38 dB compared to a similar-input-electrical-power pure AM state.

Additionally, the QVM (or indeed any two-crystal amplitude modulator) can produce a *carrier suppression* state, described in Section 5.2.3, where the output consists only of two modulation sidebands. The QVM was used to generate 30 dB of carrier suppression as measured by a heterodyne beat signal.

7.2 The Experiment

The following characterisation experiment was designed to measure the amplitude and phase response of a prototype QVM device with respect to both the AM and PM output states. For the purposes of this experiment, the multitude of applications for which device may be used are grouped into two categories of measurement: those that measure the *variability* of the QVM device, and those that measure the *purity* of the modulations that the QVM device can produce. Variability refers to the capability of the QVM prototype to tune around the parameter space of modulations, or "dial up" a particular modulation, in a predictable, theory-matching fashion and with a reasonable level of precision. Purity

means the QVM prototype's ability to accurately attain a particular modulation, especially those that involve suppression of a frequency line, and to hold that modulation indefinitely without drifting.

7.2.1 The QVM Prototype

The QVM prototype itself was built by modifying a commercial electro-optic amplitude modulator. The type of amplitude modulator used (a New Focus broadband amplitude modulator, Model 4104) is almost identical to a QVM, except it has only one electrode connected to both crystals. The connections are hard-wired to have antisymmetric polarity, as symmetric polarity would produce PM instead of AM. The prototype QVM was built by separating these two electrodes such that independent voltages could be applied to the two crystals, as shown in Fig. 7.1.

The original New Focus amplitude modulator (Model 4104) carried the specification: max $V_\pi = 300$ V @ $1.06\,\mu$m. While V_π was not explicitly measured for the prototype QVM after the modification from the original AM device, it is reasonable to assume it to be at worst of the same order of magnitude. Therefore, since the input voltages used did not exceed 10 V, the following results are in the small modulation depth limit as described in Section 3.1.2.

7.2.2 Measurement Layout

The experiment is shown in Fig. 7.2. The polarisation of a laser source was prepared with a half- and quarter-wave plate in series, and the laser source was passed through the QVM, which was operated at $\omega_m/2\pi = 5$ MHz. A polarising beam splitter completed the process; the vertically polarized output was the subject of for subsequent measurements.

The horizontally polarised output was used to keep track of the polarisation state, and hence to measure the value of σ. For experimental convenience, an optical state of near-circular polarisation with the vertical component stronger was used, corresponding to $\sigma = 75°$.

A heterodyne detection scheme formed part of the measurement apparatus, where a shunted beam was frequency-shifted by $\omega_h/2\pi = 80$ MHz with an acousto-optic modulator (AOM). Following on from Section 3.2.6, the result of this heterodyne interference is now calculated. If the heterodyne oscillator beam is represented by $E_{\mathrm{HET}} = \gamma E_{\mathrm{IN}} e^{i\omega_h t}$ with $\gamma \ll 1$ and interferes with the modulated beam from Eq. 3.14, then the detected power is given by:

$$
\begin{aligned}
P_{\mathrm{DET}} &= E_{\mathrm{IN}}^2 \left[P_{\mathrm{DC}} + P_{\omega_m} + P_{\omega_h} + P_{\omega_h - \omega_m} + P_{\omega_h + \omega_m} \right]; \\
P_{\mathrm{DC}} &= \cos^2\left(\frac{\sigma}{2}\right); \\
P_{\omega_m} &= 2\cos\left(\frac{\sigma}{2}\right) \Re\{\tilde{A} e^{i\omega_m t}\}; \\
P_{\omega_h} &= 2\cos\left(\frac{\sigma}{2}\right) \Re\{\gamma e^{i\omega_h t}\}; \\
P_{\omega_h - \omega_m} &= 2\Re\left[\gamma(\tilde{A} + i\tilde{P})^* e^{i(\omega_h - \omega_m)t}\right]; \\
P_{\omega_h + \omega_m} &= 2\Re\left[\gamma(\tilde{A} - i\tilde{P}) e^{i(\omega_h + \omega_m)t}\right]
\end{aligned}
\tag{7.1}
$$
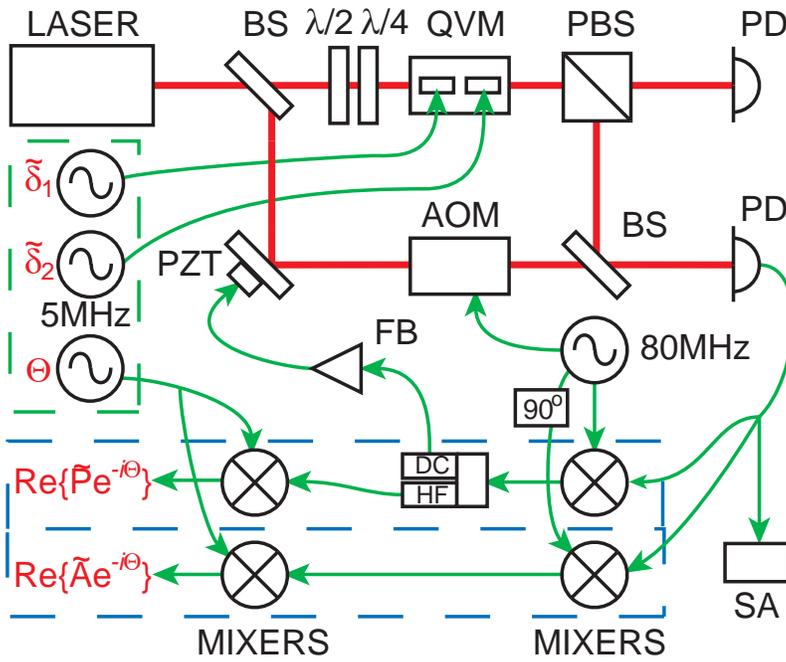
Fig. 7.2: QVM heterodyne characterisation experiment layout. A spectrum analyser provides frequency line data of the direct AM beat, and heterodyned copies of sidebands for establishing the purity of the QVM. Two parallel double-demodulation circuits provided the complete set of modulation measurables for testing the variability of the QVM. BS: dielectric beam splitter; $\lambda/2$ half-wave plate; $\lambda/4$ quarter-wave plate; PBS: polarising beam splitter; PZT: piezo-electric actuator; AOM: acousto-optic modulator; FB: feedback servo and HV amp; 90°: electronic phase-shifter; DC/HF: bias-T; SA: spectrum analyser. The 5 MHz signal generators are electronically phase-locked.

where the power components are split up according to their respective frequencies. P_{ω_m} represents the measurable beat due to the presence of AM, P_{ω_h} is the beat between the carrier and the heterodyne oscillator, and $P_{\omega_h-\omega_m}$ and $P_{\omega_h+\omega_m}$ represent heterodyned copies of the two modulation sidebands. A spectrum analyser was used to monitor the strength of these frequency lines. The use of the spectrum analyser facilitated the assessment of the *purity* of states produced by the QVM prototype, by observing operating points that involved suppression of one of these frequency lines.

An alternative set of measurables was obtained by electronically mixing the heterodyne frequencies down to baseband via a *double-demodulation* scheme. This gave a DC readout of the PM and AM amplitudes at a selected beat phase. A complete description of the modulation strengths and quadratures was obtained in this way, and was hence useful to characterise the *variability* of the QVM prototype.

The method of signal extraction via double demodulation is best understood by re-working the last two components of Eq. 7.1 to:

$$P_{\omega_h-\omega_m} + P_{\omega_h+\omega_m} = 2\cos\left(\frac{\sigma}{2}\right)\left[\Re\{\tilde{P}e^{i\omega_m t}\}\Im\{\gamma e^{i\omega_h t}\}\right.$$
$$\left. + \Re\{\tilde{A}e^{i\omega_m t}\}\Re\{\gamma e^{i\omega_h t}\}\right] \tag{7.2}$$

This signal was mixed down to base band by demodulating at $\omega_h/2\pi = 80$ MHz and then $\omega_m/2\pi = 5$ MHz in series. The output of a demodulation is sensitive to the relative phases of the signals being mixed: if they are exactly in-phase or out-of-phase, the output will be maximally positive or negative; if they are in quadrature, the output will be zero. Eq. 7.2 shows that the ω_h oscillator components of the AM and PM portions are orthogonal (one is the real component, the other is the imaginary component) so that the appropriate choice of electrical oscillator phase can force the readout of PM only, or AM only, or some linear combination of the two. Similarly, the demodulation quadrature of the ω_m stage determines the beat phase onto which the output signal is projected.

The output DC component of the *first* stage of demodulation was extracted with a bias-T component and used as an error signal for locking the optical recombination phase of the heterodyne. This error signal varies sinusoidally with respect to the optical recombination phase, with a zero crossing where the optical heterodyne beat (P_{ω_h}) and the electronic demodulation oscillator are in quadrature. In other words, the feedback loop "locks out" the heterodyne beat. A comparison between Eq. 7.1 and Eq. 7.2 shows that the AM term has the same phase as the P_{ω_h} term; they both contain the terms $\Re\{\gamma e^{i\omega_h t}\}$). Therefore, the feedback loop also locks out the AM component of modulation, leaving only the PM component, which has a $\Im\{\gamma e^{i\omega_h t}\}$ term. This is an important part of the process as, without a locking loop, the demodulation phase of the circuit would be uncharacterised, and would be free to drift.

In the experiment, two separate double demodulation schemes were used in order to measure PM and AM simultaneously; the ω_h signal was split with a 90° electronic splitter to ensure that the two double demodulation circuits scanned orthogonal modulations. Hence, when the heterodyne locking loop was connected using a signal from one double-demodulation circuit, this ensured that that particular circuit was sensitive to PM and that the other was sensitive to AM.

In the second stage of demodulation, electronic phase locking of the signal generators was sufficient to time-stabilise the demodulation quadrature, and a known modulation was used for calibration.

7.2.3   Results

Fig. 7.3 shows spectrum analyser traces for four operating points, demonstrating the suppression capabilities of the device. The suppression factors are 34.9 dB and 39.5 dB for the left- and right-hand sidebands (Fig. 7.3(a) and Fig. 7.3(b)). The figures shown suppress the frequency lines down to the electronic noise floor; higher suppression factors may be possible. A more detailed trace exhibiting sideband suppression is given in Fig. 7.4, with 35.2 dB relative suppression recorded.

Around 38 dB suppression of the AM beat was measured compared with a similar-input-power pure-AM state (Fig. 7.3(c)). In this and the other diagrams in Fig. 7.3, "max-hold" data is shown to demonstrate that the device is highly stable, maintaining these operating points without significant drift on the timescale of hours.

In Fig. 7.3(d), the carrier is suppressed by selecting the AM operating point and setting $\sigma = \pi$ (horizontal polarisation). The heterodyne measurement of the carrier is down by around 30 dB, from which it was inferred that the carrier power (which goes as the square
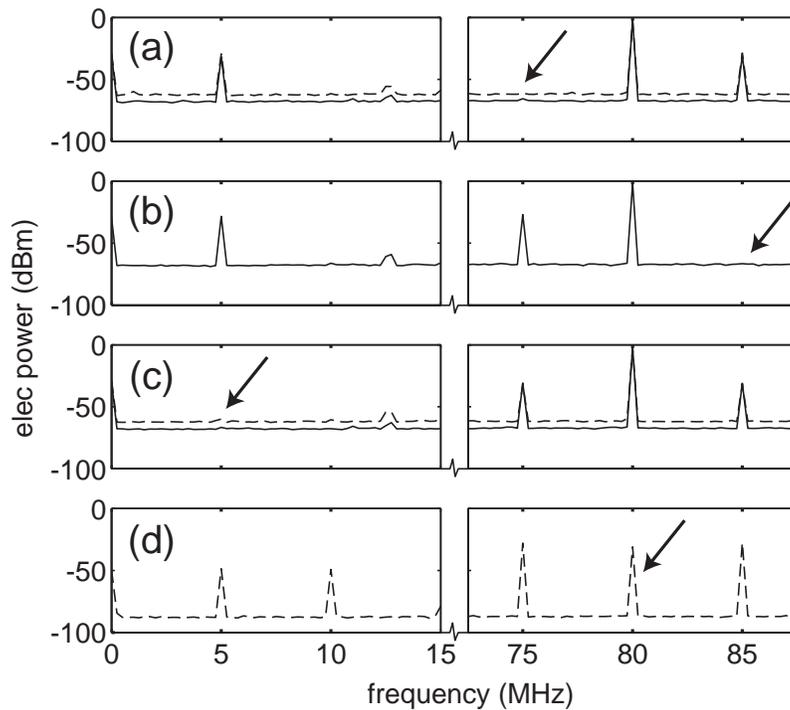
Fig. 7.3: Spectrum analyser traces of heterodyne measurements of four operating points. Frequencies of interest are 5 MHz (direct AM beat), 80 MHz (heterodyne-carrier beat), and 75 MHz and 85 MHz (heterodyne-sideband beats). A solid line indicates data averaged over a few seconds, a dotted line indicates max-hold data acquired over approximately an hour, and arrows indicate lines that have been suppressed. The plots are: (a) upper single-sideband (lower sideband suppressed by 34.9 dB); (b) lower single-sideband (upper sideband suppressed by 39.5 dB); (c) pure PM (AM beat suppressed by around 38 dB); (d) carrier suppression (heterodyne beat suppressed by around 30 dB).

of the heterodyne measurement) was down by 60 dB. This heterodyne measurement is further supported by observing that the first and second harmonics of direct AM beat are approximately equal, which is consistent with the fact that the carrier (as measured by the heterodyne) is approximately 6 dB weaker than the sidebands.

The experimenters were especially careful to validate this heterodyne measurement of the carrier, because the directly detected power only dropped by 40 dB. The reason for this is that the majority of the residual carrier power was now in a higher-order, antisymmetric spatial mode which did not interfere efficiently with the heterodyne oscillator beam. The higher-order modes were observed directly by viewing the lateral beam profile. They were thought to result from a spatially non-uniform polarisation of the light exiting the QVM which, when subsequently passed through a polarising beam splitter, produced modes reflecting the symmetry of the modulator. As such, for applications where spatial mode interference is important, the employment of a mode-cleaner cavity with free spectral range equal to the modulation frequency, or one of its integer divisors, should solve the problem.

In an attempt to confirm the variability of the QVM prototype, significant paths were mapped in the modulation parameter space. Fig. 7.5 shows the result of sweeping through M-space in two cardinal directions: varying the relative phase between the two constant-
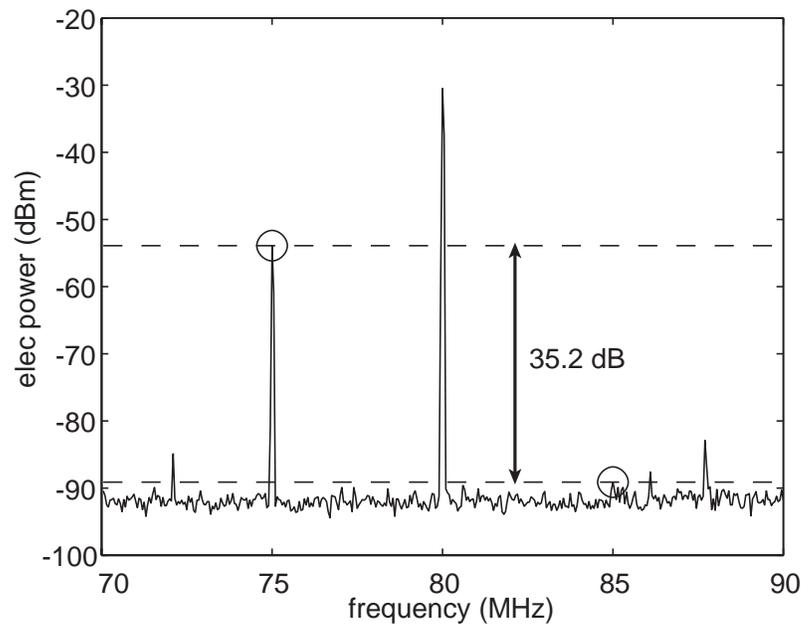
Fig. 7.4: A more detailed spectrum analyser trace of sideband suppression, with 35.2 dB difference between sidebands. Frequency lines at 72.1 MHz, 86.1 MHz and 87.7 MHz are from radio interference with the electronic equipment.

amplitude electrical signals; and varying the amplitude of one electrical signal while keeping the other signal amplitude, and the relative phase, constant. For matters of calibration, the QVM was initially set to the PM operating point. This initial state was used as the reference in labelling "in-phase" and "quadrature" components. The relative gains and phases of the two double demodulation circuits were measured and factored out via correspondence with spectrum analyser data. The input polarisation was set to be elliptical with a vertical component slightly larger than the horizontal component. The results showed that changing the phase between the two electrical signals produced AM in quadrature to the original PM, whereas a change to the amplitude of one electrical signal generated AM in-phase with the original PM. The data points are generally in good agreement with the theoretical predictions. Some small systematic errors are apparent, and are thought to be associated with unmatched electrical and optical impedances between the two crystals, and possibly to do with slightly unmatched optical power levels probing each crystal. Overall, the device is shown to be highly predictable, and that one can "dial-up" a particular modulation state on call, having calibrated the device initially.

It is instructive to notice that the cardinal operating points are not evenly spaced when sweeping across the parameter space. In particular, the two single-sideband operating points are shifted toward the AM operating point in Fig. 7.5(c), and the correlated PM and AM operating point is similarly shifted in Fig. 7.5(d), relative to the perpendicular. This effect is a product of the non-circular polarisation of light input into the modulator device, and can be understood by reviewing the modulation ellipses in the two figures: the M_2 and M_3 axes intercept the ellipse at points that are closer to the AM operating point than the PM. In fact, this provides a means of calibration of the relative transfer-function-amplitudes of the PM and AM double-demodulation circuits: they are set using the knowledge that (in Fig. 7.5(c)) the PM and AM components are equal at the phase ϕ
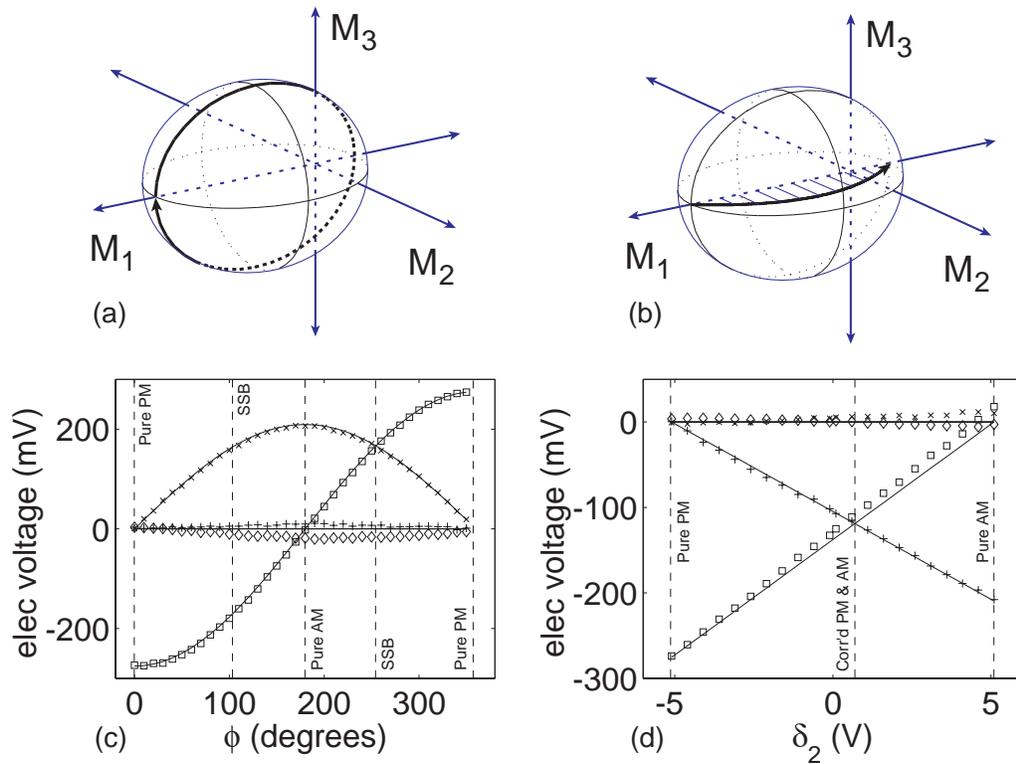
Fig. 7.5: (a) Modulation ellipse showing trajectory of measurement sweep through parameter space, where the phase difference between electrical signals, $\phi = \phi_1 - \phi_2$, is varied through $360°$. The ellipse shown has an exaggerated σ of $\approx 53°$. (b) Modulation ellipse showing trajectory of measurement sweep through parameter space, where the strength of one electrical signal, δ_2, is varied from $+5.1$ V to -5.1 V. The sweep involves varying the overall input electrical power, so the modulation trajectory does not stay on the ellipse surface, and instead traces out a parabolic curve. (c) and (d) Double-demodulation measurements of PM (squares and diamonds for in-phase and quadrature components respectively) and AM (pluses and crosses for in-phase and quadrature components respectively), with corresponding theoretical predictions (solid lines). (c) shows measurements corresponding to the trajectory in (a), and the majority of modulation present is either PM or quadrature AM as predicted. This plot corresponds to $\sigma \approx 75°$, a value derived by measuring, from corresponding spectrum analyser data, the phase ϕ at which the single sideband operating points occur. A result from Table 5.1 is that the phase ϕ between the pure AM operating point and either of the single sideband operating points, is precisely equal to σ. (d) corresponds to the trajectory in (b), and there is diverging agreement in the PM which may be caused by imperfectly matched electrical impedances for the two crystals. As in (c), the location of the correlated PM and AM point is shifted toward the pure AM point by about 0.7 V out of 5.1 V which is consistent with a value of $\sigma \approx 75°$.

where the single sideband is found in spectrum analyser data. Also, by using the identities in Table 5.1, a value of the polarisation parameter can be derived: $\sigma \approx 75°$.

Fig. 7.6 shows scans of the local region near the PM operating point, with deviations in four directions. These data were taken with a tune-locking application in mind, as discussed in Chapter 6. The in-phase PM component is by far the dominant signal, and has been scaled down by a factor of 10 to fit in the graphs. As expected, the QVM device can produce AM that is in-phase with the present PM, or AM that is in quadrature with the present PM, by changing the relative electrical signal amplitudes or phases respectively. In addition, both of these parameters can be varied together or oppositely, to give independent control over the two quadratures of AM. In the case of producing both in-phase and quadrature AM, a small component of quadrature PM appears. In other words, the phase of the PM changes, relative to that of the pure PM point used for the initial calibration. In actual fact, the quadrature PM data have a noticeably larger systematic error than the other three signals[1]. This is most likely due to pollution from its in-phase counterpart, caused by the demodulation oscillator's phase drifting marginally. This was observed to happen even in spite of the electronic phase locking between signal generators.

7.2.4   Experimental Difficulties

A number of experimental difficulties deserve mentioning. Perhaps the most frustrating inconsistency observed during the experiment was during early calibrations. The heterodyne measurements produced copies of the two modulation sidebands which were viewable on a spectrum analyser, and these sidebands were observed to have a height mismatch in circumstances where this ought to have been impossible. The most striking example was when only one QVM crystal was being operated which, armed with the theoretical knowledge presented earlier, should only ever produce modulations of the M_1 or M_2 type. M_3 type (sideband height differential) modulation can only result when two electrical signals are applied in quadrature (or at least not exactly in or out of phase), and this is clearly impossible with only one electrical signal present. Since the crystals have a common ground connection, the possibility of cross-coupling between crystals, with some RC-circuit-like phase shift, was investigated. However, the transfer coefficient from one electrode to the other was measured at less than -80 dB, so this possibility was ruled out. A mechanism for this inconsistency has never been identified.

The sideband height mismatch for single crystal operation was typically 1 or 2 dB, but was measured as being as high as 7 dB on occasion. The differential was observed to change dramatically with alignment of the heterodyne recombination and of the QVM device, and also with modulation frequency, although the range was not generally large enough to cancel out the problem altogether. Ultimately, the problem was attributed to inefficiency in the heterodyne interference, or impurities in the crystals, or a combination of both, since fluctuations of this sort were not observed in the direct AM beat signal. Calibrating the QVM to the PM operating point, by suppressing the AM beat, was hence

---

[1] The data points were taken in groups of three working outward from the central PM operating point, over a period of about twenty minutes. The signature of the errors seen here reflects this pattern, and hence the experimenters identified the error as being a time-dependent drift, sampled in blocks of three.
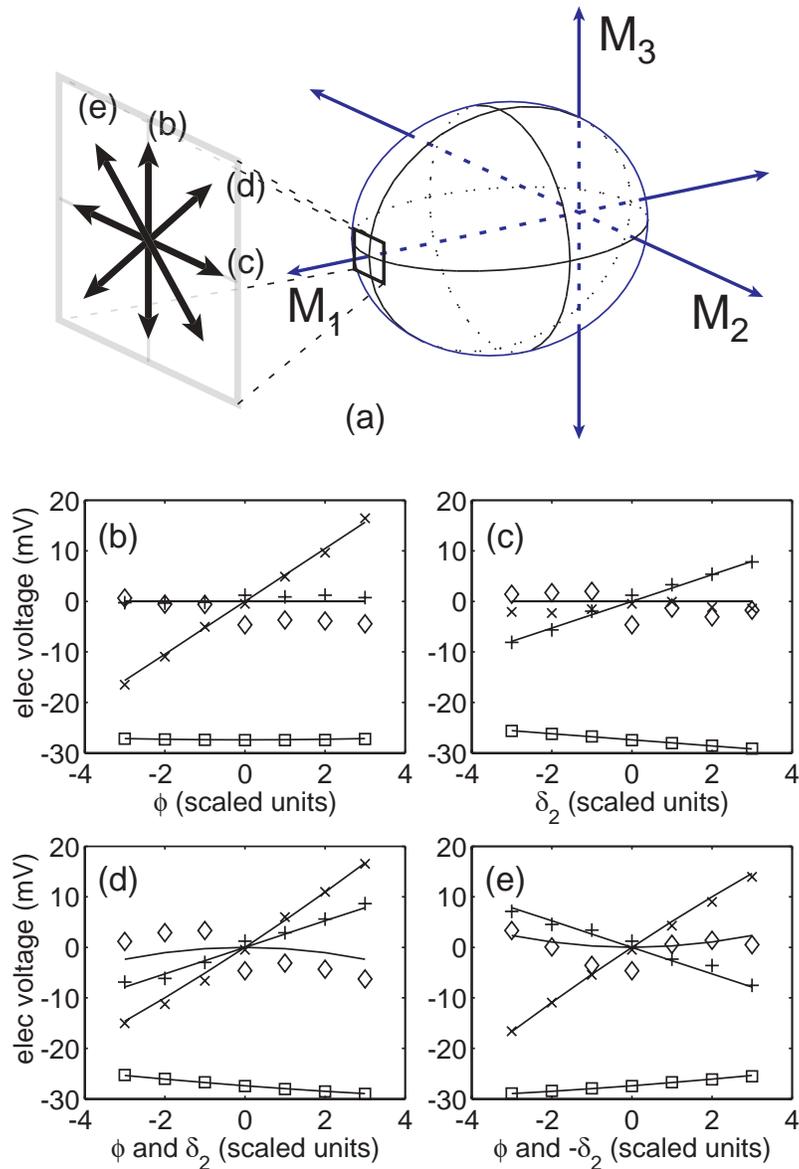
Fig. 7.6: (a) A modulation ellipse showing short-range trajectories designed to map the local region near the PM operating point. The electrical parameters $\phi = \phi_1 - \phi_2$ and δ_2 were varied to achieve these results, with units worth $5°$ and 0.44 V respectively. (b), (c), (d) and (e) show double-demodulation measurements of PM (squares and diamonds for in-phase and quadrature components respectively) and AM (pluses and crosses for in-phase and quadrature components respectively), with corresponding theoretical predictions (solid lines). Parameters varied were (b) ϕ, (c) δ_2, (d) ϕ and δ_2 together with the same polarity, and (e) ϕ and δ_2 together with opposite polarities. *Note:* the in-phase PM data points (squares) have been scaled down by a factor of 10 to fit in the diagram. Also, the systematic error in the quadrature PM data is probably due to an electronic phase drift between signal generators, causing a small amount of in-phase PM data (which is an order of magnitude stronger) to couple across.

identified as the most reliable calibration method. Without knowing the mechanism behind this effect, it is hard to dismiss, but the overall success of the results convinced us that the device was behaving at least *close* to a sensible manner.

As described in the context of the carrier suppression results, the QVM generated a spatially varying polarisation state, which produced a small percentage of higher-order spatial modes (order 1% of overall power) upon selecting out a polarisation component. This obstructed direct measurements of the polarisation state, which was done by measuring the overall detected power while rotating a diagnostic half-wave plate placed before the PBS. In this way, a value of $\sigma \approx 70°$ was measured for the results shown in Fig. 7.5, as compared to $\sigma \approx 75°$ measured by inference from spectrum analyser data. A reasonably large error was assigned to the direct measurement value due to the presence of the higher-order spatial modes. Measuring σ by inference from the data was favoured as being more reliable.

The overall optical path lengths of the modulating crystals were found to change significantly as they warmed up after starting the laser. The dual crystal design of the device goes a long way toward minimising this problem, and the device is stable once it has warmed up. However, long term drift of the crystal lengths is possible, having a direct effect on the polarisation state leaving the device. In circumstances where this becomes a problem, a feedback loop may be employed to lock the polarisation state. A possible scheme would see the power level out of one PBS port monitored and used (minus an offset equal to the desired power level) as a feedback signal to the DC optical path length of one of the crystals, thus compensating for a mechanical path length change by feeding back to the refractive index.

Finally, the electrical impedances of the two crystals were not well matched for two reasons: the electronics were not identical, and the crystals themselves had "good" and "bad" spots which generated varying levels of modulation. In general terms, careful alignment can largely overcome this problem and return the two crystals to an equal footing. The differing electrical impedances complicate the issue of generating a local oscillator with the phase-tracking property discussed earlier, since the voltages $V_{1,2}$ and modulation depths $\delta_{1,2}$ are not then related by the same factor. The device tested here is merely a prototype, and careful management should suffice to deal with these issues as the need arises.

7.3   Summary

The results of an experiment to characterise a prototype QVM were reported. The experiment involved using dual double-demodulation circuits to measure both AM and PM components simultaneously. Data sets were obtained and analysed to illustrate the variability and purity characteristics of the device, which was shown to be highly predictable and capable of highly pure states. Applications for the QVM were discussed.

"He tried to collect himself but could not succeed; for it is especially in the hours where men have the most need of thought that all the threads are broken in the brain."

— Victor Hugo, from *Les Miserables*

Conclusions

The primary result of this thesis is the successful development of the *quadrature variable modulator*, an electro-optic modulation device. Chapters 5 and 7 presented the theory and experimental characterisation of the device, and demonstrated its wide range of capabilities for applications which require precise modulations or a tunable source of modulation.

The QVM was shown to be capable of both amplitude modulation and phase modulation, and of combining the two with an arbitrary phase relation. A prescription for use was developed, requiring that the input light be of a certain polarisation, in order that the transfer characteristics of the device were simplified. In this case, it was shown that an electronic local oscillator for use in demodulation schemes was readily available, and the oscillator's phase relation to the modulation was preserved when tuning the QVM output. The prescription also allows for the majority of the input light beam's power to be conserved; total power conservation is impossible as some percentage of loss always results when generating amplitude modulation. The higher harmonics produced by the QVM were also theoretically examined.

The subject of Chapter 6 was an anticipated application for the QVM in gravitational wave detector longitudinal control schemes as part of a *tune-locking* arrangement, where the detector's frequency response is tunable in real-time. The technique is intended to work in conjunction with an enhancement to signal recycling where the signal mirror is replaced by an optical subsystem whereby the effective reflectivity can be varied. A theoretical investigation of tune-locking was conducted via modelling of two single-degree of freedom systems: a Michelson interferometer, and a Fabry-Perot cavity. It was found that the technique will work within limits, and has potential to be incorporated into future gravitational wave detector configurations.

In Chapter 3, electro-optic modulation was reviewed in depth, and a new theory of modulation, called the *modulation sphere* formulation, was presented. The purpose of the modulation sphere is for use as a visual tool for aiding mental or intuitive calculations and for comprehending complicated modulation situations, but it also shows promise as a mathematical analysis tool. The modulation sphere representation involves transforming the usual amplitude and phase modulation parameters into three real-valued parameters called *M-space* parameters, suppressing the beat phase of the modulation. The M-space parameters have the significance of defining physically distinct modulation states, and encapsulating the physics of all single-frequency modulation states. Modulation sphere diagrams were used throughout the thesis to supplement explanations of systems involving modulation.

Chapter 4 discussed a range of methods for locking optical Fabry-Perot cavities, specifically for use with mode-cleaners in gravitational wave detectors. Alternatives to the commonly used Pound-Drever-Hall method were presented, including a new spatial mode interference technique called *flip locking*. Flip locking, along with its older cousin *tilt locking* (originally presented elsewhere), were experimentally demonstrated as part of a *double pass* locking arrangement. Spatial mode interference locking, along with a double pass arrangement, were suggested as methods that avoid problems with high laser powers in gravitational wave detectors.

Future work stemming from this research could include:

- Conducting a proof-of-principle experiment demonstrating that a Michelson interferometer or a Fabry-Perot cavity can be tuned away from the default operating point via injection of amplitude modulation as theoretically described in Chapter 6;

- Conducting a further proof-of-principle experiment to show that the two offset locking techniques are compatible, by attmpting to tune-lock systems of 2 degrees of freedom or more, and also demonstrating that one modulation frequency can be used to tune both the Michelson and Fabry-Perot degrees of freedom simultaneously and orthogonally;

- Developing gravitational wave configurations and corresponding locking schemes that incorporate full tunability of frequency response to gravitational waves, with rf detection or with direct detection;

- Further investigate the discrepancy in results produced by the prototype QVM regarding the precise details of the polarisation state used (thought to be due to inaccurate polarisation optics), and develop a precise, repeatable calibration sequence for the QVM to circumvent these difficulties;

- Continue the analysis of M-space transfer functions of optical systems and investigate the possible utility of the formalism for analysing certain classes of nonlinear optical systems.

When asked for his reaction to being denied, due to his colour, the right to stay in the same hotel he had performed at that night, Duke Ellington responded:

"I took the energy it takes to pout, and wrote some blues."

Bibliography

[1] D A Shaddock, M B Gray, and D E McClelland. Frequency locking a laser to an optical cavity by use of spatial mode interference. *Optics Letters*, 24(21):1499–1501, 1999.

[2] P R Saulson. *Fundamentals of Interferometric Gravitational Wave Detectors*. Singapore: World Scientific, 1994.

[3] W Rindler. *Introduction to Special Relativity*. Oxford Science Publications, New York, 2nd edition, 1991.

[4] D Halliday, R Resnick, and J Walker. *Fundamental of Physics*. John Wiley & Sons, Inc., USA, 4th edition, 1993.

[5] P Astone et al. Comments on the 2001 run of the Explorer/Nautilus gravitational wave experiment. *Classical & Quantum Gravity*, 20(17):S785–S788, 2003.

[6] R W P Drever. *Gravitationao Radiation*, page 321. North Holland, Amsterdam, 1983.

[7] B J Meers. Recycling in laser-interferometric gravitational-wave detectors. *Physical Review D*, 38(8):2317–2326, 1988.

[8] K A Strain and B J Meers. Experimental demonstration of dual recycling for inteferometric gravitational-wave detectors. *Physical Review Letters*, 66(11):1391–1394, 1991.

[9] J Mizuno, K A Strain, P G Nelson, J M Chen, R Schilling, A Rudiger, W Winkler, and K Danzmann. Resonant sideband extraction: a new configuration for interferometric gravitational wave detectors. *Physics Letters A*, 175(5):273–276, 1993.

[10] G de Vine. Variable reflectivity mirrors for signal recycled michelson interferometer gravitational wave detectors. Honours thesis, The Australian National University, 2001.

[11] K A Strain and J Hough. Experimental demonstration of the use of a Fabry-Perot cavity as a mirror of variable reflectivity. *Review of Scientific Instruments*, 65(4): 799–802, 1994.

[12] G de Vine, D A Shaddock, and D E McClelland. Experimental demonstration of variable reflectivity signal recycling for interferometric gravitational wave detectors. *Optics Letters*, 27(17):1507–1509, 2002.

[13] LIGO Scientific Collaboration. http://www.ligo.caltech.edu/ lazz/distribution/LSC_Data/, 2004.

[14] R Takahashi and the TAMA Collaboration. Status of TAMA300. *Classical & Quantum Gravity*, 21(5):S403–S408, 2004.

[15] D Sigg. Commissioning of LIGO detectors. *Classical & Quantum Gravity*, 21(5): S409–S416, 2004.

[16] P Fritschel. Advanced LIGO systems design. Technical Note T010075-00-D, LIGO Scientific Collaboration, http://www.ligo.caltech.edu/, 2001.

[17] B Willke et al. Status of GEO 600. *Classical & Quantum Gravity*, 21(5):S417–S424, 2004.

[18] Private communcation with Benno Willke.

[19] Private communcation with Adalberto Giazotto.

[20] F Acernese, P Amico, and N Arnaud et al. Status of VIRGO. *Classical & Quantum Gravity*, 21(5):S385–S394, 2004.

[21] P Astone. Resonant mass detectors: present status. *Classical & Quantum Gravity*, 19(7):1227–1236, 2002.

[22] O D Aguiar, L A Andrade, and J J Barroso et al. The Brazilian spherical detector: progress and plans. *Classical & Quantum Gravity*, 21(5):S457–S464, 2004.

[23] LISA Study Team. Laser interferometer space antenna for the detection and observation of gravitational waves. Pre Phase A Report MPQ233, Max-Planck-Institüt für Quantenoptik, 1998.

[24] A Abramovici and J Chapsky. *Feedback control systems: a fast-track guide for scientists and engineers*. Kluwer Academic, Boston, 2000.

[25] B J J Slagmolen, X Brokmann, D A Shaddock, M B Gray, and D E McClelland. High dynamic range flexure transfer function measurement. *Classical & Quantum Gravity*, 19(7):1683–1688, 2002.

[26] J 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. *Journal of the Optical Society of America A*, 17(1):120–128, 2000.

[27] A Buonanno, Y Chen, and N Mavalvala. Quantum noise in laser-inteferometer gravitational-wave detectors with a heterdyne readout scheme. *Physical Review D*, 67(12):2005, 2003.

[28] D A Shaddock. *Advance Interferometry for Gravitational Wave Detection*. Phd thesis, The Australian National University, 2000.

[29] J F James. *A student's guide to Fourier transforms*. Cambridge University Press, 1995.

[30] M B Gray. *Quantum Noise Limited Interferometry*. Phd thesis, The Australian National University, 1995.

[31] H -A Bachor. *A guide to experiments in quantum optics*. Wiley VCH, Weinheim, New York, 1998.

[32] G G Stokes. *Trans. Camb. Phil. Soc.*, 9:399, 1852.

[33] M D Levenson and S Kano. *Introduction to nonlinear laser spectroscopy*. Boston Academic Press, 1988.

[34] S K Korotky. Three-space rpresentation of phase-mismatch switching in coupled two-state optical systems. *IEEE Journal of Quantum Electronics*, QE-22(6):952–958, 1986.

[35] R Weiss. Quarterly progress report. *MIT Research Lab of Electronics*, 105:54, 1972.

[36] L Schnupp. Conference Presentation in *European Collaboration Meeting on Interferometric Detection of Gravitational Waves*, 1988.

[37] B Willke, N Uehara, E K Gustafson, R L Byer, P J King, S U Seel, and R L Savage, Jr. Spatial and temporal filtering of a 10-W Nd:YAG laser with a Fabry-Perot ring-cavity premode cleaner. *Optics Letters*, 23(21):1704–1706, 1998.

[38] E Gustafson, D Shoemaker, K A Strain, and R Weiss. LSC white paper on detector research and development. Internal Report T990080-00-D, LIGO, http://www.ligo.caltech.edu/, 1999.

[39] G Sanders. LIGO II conceptual project book. Internal Report M990288-A-M, LIGO, http://www.ligo.caltech.edu/, 1999.

[40] J D Mansell, J Hennawi, E K Gustafson, M M Fejer, R L Byer, D Chubley, S Yoshida, and D H Reitze. Evaluating the effect of transmissive optic thermal lensing on lase beam quality with a Shack-Hartmann wave-front sensor. *Applied Optics*, 40(3):366–374, 2001.

[41] A E Siegman. *Lasers*. University Science Books, Mill Valley, California, 1986.

[42] R W P Drever, J L Hall, F V Kowalski, J Hough, G M Ford, A J Munley, and H Ward. Laser phase and frequency stabilisation using an optical resonator. *Applied Physics B - Photophysics & Laser Chemistry*, 31(2):97–105, 1983.

[43] P Fritschel, G Gonzalez, B Lantz, P Saha, and M Zucker. High power interferometric phase measurement limited by quantum noise and application to detection of gravitational waves. *Physical Review Letters*, 80(15):3181–3184, 1998.

[44] G Mueller, R S Amin, D Guagliardo, D McFeron, R Lundock, D H Reitze, and D B Tanner. Method for compensation of thermally induced modal distortions in the input optical components of gravitational wave interferometers. *Classicla & Quantum Gravity*, 19(7):1793–1801, 2002.

[45] D J Ottaway, M B Gray, D A Shaddock, C Hollitt, P J Veitch, J Munch, and D E McClelland. Stabilisation of injection-locked lasers using spatial mode interference. *IEEE Journal of Quantum Electronics*, 37(5):653–657, 2001.

[46] D A Shaddock, B C Buchler, W P Bowen, M B Gray, and P K Lam. Modulation-free control of a continuous-wave second-harmonic generator. *Journal of Optics A: Pure and Applied Optics*, 2(5):400–404, 2000.

[47] B J J Slagmolen, D A Shaddock, M B Gray, and D E McClelland. Frequency stability of spatial mode inteference (tilt) locking. *IEEE Journal of Quantum Electronics*, 38(11):1521–1528, 2002.

[48] V Delaubert, D A Shaddock, P K Lam, B C Buchler, H-A Bachor, and D E McClelland. Generation of a phase flipped gaussian mode for optical measurements. *Journal of Optics A: Pure and Applied Optics*, 4(4):393–399, 2002.

[49] G Heinzel, A Rudiger, R Schilling, K Strain, W Winkler, J Mizuno, and K Danzmann. Automatic beam alignment in the Garching 30-m prototype of a laser-interferometric gravitational wave detector. *Optics Communications*, 160(4-6):321–334, 1999.

[50] E Hecht. *Optics*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1987.

[51] O Miyakawa, K Somiya, G Heinzel, and S Kawamura. Development of a suspended-mass RSE interferometer using third harmonic demodulation. *Classical & Quantum Gravity*, 19(7):1555–1560, 2002.

[52] Private communcation with David Reitze, Peter Bender and Ken Strain.

[53] D A Shaddock. A power recycled Michelson interferometer with resonant sideband extraction. *Applied Optics*, 42(7):1283–1295, 2003.

[54] K A Strain, G Muller, T Delker, D H Reitze, D B Tanner, J E Mason, P A Willems, D A Shaddock, M B Gray, C Mow-Lowry, and D E McCElland. Sensing and control in dual recycling laser interferometer gravitational-wave detectors. *Applied Optics*, 42(7):1244–1256, 2003.

[55] J E Mason and P A Willems. Siugnal extraction and optical design for an advanced gravitational wave interferometer. *Applied Optics*, 42(7):1269–1282, 2003.

[56] G Mueller, T Delker, D B Tanner, and D Reitze. Dual recycled cavity enhanced Michelson interferometer for gravitational wave detection. *Applied Optics*, 42(7):1257–1268, 2003.

[57] B Davies and J Conradi. Hybrid modulator structures for subcarrier and harmonic subcarrier optical single sideband. *IEEE Photonics Technology Letters*, 10(4):600–602, 1998.

[58] JDS uniphase, www.jdsu.com.

[59] G Smith, D Novak, and Z Ahmed. Technique for optical SSB generation to overcome dispersion penalties in fibre-radio systems. *Electronics Letters*, 33(1):74–75, 1997.

[60] A Loayssa, C Lim, A Nirmalathas, and D Benito. Simple optical single-sideband modulator for fibre-radio applications. *Electronics Letters*, 39(1):97–99, 2003.

[61] K Yonenaga and N Takachio. A fiber chromatic dispersion compensation technique with an optical SSB transmission in optical homodyne detection systems. *IEEE Photonics Technology Letters*, 5(8):949–951, 1993.

[62] A Loayssa, D Benito, and M J Garde. Single-sideband sppressed-carrier modulation using a single-electrode electro-optic modulator. *IEEE Photonics Technology Letters*, 13(8):869–871, 2001.

[63] S Shimotsu, S Oikawa, T Saitou, N Mitsugi, K Kubodera, T Kawanishi, and M Izutsu. Single side-band nodulation performance of a *LiNbO$_3$* integrated modulator consisting of four-phase modulator waveguides. *IEEE Photonics Technology Letters*, 13(4):364–366, 2001.

[64] F Grosshans, G Van Assche, J Wenger, R Brouri, N J Cerf, and P Grangier. Quantum key distribution using Gaussian-modulated coherent states. *Nature*, 421(6920):238–241, 2003.

[65] W P Bowen, N Treps, B C Buchler, R Schnabel, T C Ralph, H-A Bachor, T Symul, and P K Lam. Experimental investigation of continuous-variable quantum teleportation. *Physical Review A*, 67(3):32302–1–4, 2003.

[66] A Furusawa, J L Sorensen, S L Braunstein, C A Fuchs, H J Kimble, and E S Polzik. Unconditional quantum teleportation. *Science*, 282(5389):706–709, 1998.