

**Comparison of optical configurations for
laser-interferometric gravitational-wave detectors**

Von dem Fachbereich Physik
der Universität Hannover
zur Erlangung des Grades eines
Doktor der Naturwissenschaften

— Dr. rer. nat. —

genehmigte Dissertation von

Jun Mizuno

Master of Science in Physics

geboren am 5. Februar 1966 in Saitama, Japan

1995

Referent:	Prof. K. Danzmann
Korreferent:	Prof. W. Ertmer
Tag der Promotion:	3. Juli 1995
Datum der Veröffentlichung:	17. Juli 1995

Vergleich optischer Anordnungen in laser-interferometrischen Gravitationswellen-Detektoren

Zusammenfassung

Verschiedene optische Konfigurationsmöglichkeiten von laser-interferometrischen Gravitationswellen-Detektoren werden vergleichend diskutiert. Es wird gezeigt, daß die erreichbare Empfindlichkeit wesentlich bestimmt ist durch die Größe der in dem optischen System gespeicherten Lichtenergie, mit nur schwacher Abhängigkeit von der gewählten Konfiguration. Um die in zukünftigen Gravitationswellen-Observatorien (wie LIGO und VIRGO) angestrebten Empfindlichkeiten zu erreichen, müssen 20 J bei $1\ \mu\text{m}$ Wellenlänge gespeichert werden.

Eine neue Konfiguration, *Resonante Seitenband-Entnahme* (RSE), wird vorgeschlagen und unter Benutzung eines linearen Modells ausführlich untersucht. Damit können die Parameter für eine gewünschte Empfangs-Charakteristik passend gewählt werden. Die zu erwartende Empfindlichkeit ist gleichwertig oder überlegen der anderer Anordnungen, insbesondere dann wenn die Detektor-Bandbreite groß ist.

Die Empfindlichkeiten der verschiedenen Anordnungen bei Anwesenheit von thermisch verursachten Verformungen des Lichtstrahls werden abgeschätzt. Es wird gezeigt, daß bei Auftreten solcher thermischer Probleme die neue Anordnung (RSE) vorteilhaft ist, weil weniger Lichtleistung durch die Substrate geschickt werden muß, bei denen die stärksten Verformungen des Lichtstrahles zu erwarten sind. Wenn man Quarzglas als Substratmaterial für die optischen Komponenten annimmt und die geringsten derzeit möglichen Verluste in den optischen Beschichtungen, dann werden die thermisch verursachten Verformungen die speicherbare Lichtenergie begrenzen. Dies kann durch geringfügig komplexere optische Anordnungen umgangen werden. In solchen Fällen kann die Anordnung von RSE besonders nützlich sein.

Comparison of optical configurations for laser-interferometric gravitational-wave detectors

Abstract

Various optical configurations for laser-interferometric gravitational-wave detectors are discussed. It is found that the amount of light energy stored in the optical system determines the achievable sensitivity, with only weak dependence on the configuration employed. In order to achieve the sensitivities expected in future advanced gravitational-wave observatories currently proposed (such as LIGO and VIRGO), 20 joule of $1\ \mu\text{m}$ wavelength light must be stored.

A newly proposed configuration, *resonant sideband extraction*, is fully analyzed assuming a linear model. This makes it possible to choose appropriate parameters for the desired response. The expected sensitivity using this configuration is similar to or better than those of other configurations. The new configuration is favorable especially when the detector bandwidth is broad.

The sensitivity of each configuration in the presence of thermally induced distortions of the beams is estimated. It is shown that when thermal problems become significant, the new configuration can be advantageous because less light power is transmitted through the optical substrates where the strongest distortion of the beam is expected. Assuming fused silica as the substrate material of optical components and coating losses of state-of-art level, thermally induced distortions are likely to limit the amount of light energy that can be stored. This may be overcome with the use of slightly more complicated configurations. In such cases, an application of resonant sideband extraction will be useful.

Contents

Zusammenfassung / Abstract	i/ii
Contents	iii
List of figures	vi
List of notations	viii
Acknowledgments	x
Chapter 1: Introduction	1
1.1: Gravitational waves and their detection	1
1.2: Expected signal	4
1.3: Possible noise	9
Chapter 2: Principle of gravitational wave detection	13
2.0: Introduction	13
2.1: Effect of gravitational waves	13
2.1.1: Gravitational waves in General Relativity	13
2.1.2: The effect on free masses	15
2.1.3: Optical measurement of gravitational wave effects	19
2.2: Modulation, sidebands, and their detection	20
2.2.1: Amplitude/phase modulation	20
2.2.2: Suppressed carrier and single sideband	25
2.2.3: Heterodyne and homodyne techniques	27
2.3: Optical phase detection	29
2.3.1: Photo-detection	29
2.3.2: Local oscillator light	32
2.3.3: Optical heterodyne	34
2.3.4: General double sidebands	37
2.4: A Michelson interferometer as a phase detector	40
2.4.1: Differential method—homodyne	40
2.4.2: Dark fringe operation—balanced modulator	44
2.4.3: External modulation	45
2.4.4: Internal modulation	48
2.4.5: Pre-modulation	50
Chapter 3: Various configurations	53
3.0: Introduction	53
3.1: The sensitivity theorem	53
3.1.1: Sensitivity	53
3.1.2: Normalized response	55
3.1.3: Simplified theorem	56

Contents

3.2: Delay lines and Fabry-Perot cavities	58
3.2.1: Multi-reflection schemes (optical delay-lines)	58
3.2.2: Fabry-Perot cavities	61
3.2.3: Frequency responses	64
3.2.4: Comparison of delay-lines and Fabry-Perot cavities	68
3.3: Power recycling	71
3.3.1: Principle of power recycling	71
3.3.2: Power recycling with delay lines	74
3.3.3: Power recycling with Fabry-Perot cavities	77
3.4: Other schemes	80
3.4.1: Signal recycling: Broad-band operation	80
3.4.2: Signal recycling: Narrow-band operation	85
3.4.3: Detuned recycling	89
3.4.4: Synchronous recycling	90
3.5: More about the theorem	92
3.5.1: Complete theorem	92
3.5.2: Shot Noise	94
3.5.3: Energy, or number of photons	95
3.5.4: Cavity effect (boundary condition)	96
3.5.5: The arm length effect	98
Chapter 4: Resonant sideband extraction	101
4.0: Introduction	101
4.1: The principle of resonant sideband extraction	101
4.1.1: Basic concept	101
4.1.2: Optical configuration	103
4.1.3: Brief discussion	105
4.2: Three-mirror coupled cavities	107
4.2.1: Optics matrices	107
4.2.2: Fabry-Perot cavities	109
4.2.3: Under- and over-coupling	112
4.2.4: Three-mirror coupled cavities	115
4.3: Frequency response of resonant sideband extraction	119
4.3.1: The effect of the signal cavity	119
4.3.2: Frequency response	121
4.3.3: Coupling of the signal extraction cavity	123
4.3.4: Choosing the length of the signal extraction cavity	126
4.4: Application as gravitational wave detectors	129
4.4.1: Parameters in resonant sideband extraction	129
4.4.2: Broad-band detector	130
4.4.3: Narrow-banding	134
4.4.4: Low frequency detector	138

Chapter 5: Thermal distortion	143
5.0: Introduction	143
5.1: Non-linear problems	143
5.1.1: Non-linearity	143
5.1.2: Thermal distortion	145
5.1.3: The beam size effect	146
5.2: Mode description of scattering	147
5.2.1: Loss mechanisms	147
5.2.2: Reflection at a non-ideal surface	148
5.2.3: Mode stability of a cavity	150
5.2.4: Scattering as mode conversion	154
5.2.5: Cavity effect (mode cleaning effect)	158
5.3: Thermal distortion	163
5.3.1: Thermal expansion	163
5.3.2: Thermal lensing	165
5.3.3: Thermal properties of substrate materials	166
5.3.4: The effect of thermal distortion on the interference	167
5.3.5: Thermal distortion in a cavity	169
5.4: Sensitivity with thermal distortion	171
5.4.1: Sample parameters of an interferometer	171
5.4.2: Delay lines in the arms	172
5.4.3: Fabry-Perot cavities in the arms	174
5.4.4: Resonant sideband extraction	177
5.4.5: Possible remedies	178
Summary and prospects	182
Appendices	184
Appendix A: An additional factor of two	184
Appendix B: Basic assumptions	186
Appendix C: Cancellation of phase delay	189
Appendix D: Derivation of narrow-band parameters	192
Appendix E: Et cetera	196
Appendix F: Frequency response and the local oscillator	199
Appendix G: Gaussian beam	203
G.1: Paraxial equation	203
G.2: Fundamental Gaussian mode	204
G.3: Hermite-Gaussian mode	207
G.4: Laguerre-Gaussian mode	211
G.5: Gaussian beam with different parameters	214
Appendix H: Higher order local oscillators	217
References	219

List of figures

1.1	The quadrupole force field of a gravitational wave	2
1.2	The expected signal strength	7
1.3	The typical noise spectrum in an advanced detector	10
2.1	An illustration of a gravitational wave	17
2.2	The displacement caused by a gravitational wave	18
2.3	The phasor diagram of amplitude modulation	21
2.4	A swinging vector	22
2.5	The phasor diagram of phase modulation	23
2.6	A swinging vector with a different phase	24
2.7	Suppressed-carrier double sideband	26
2.8	A beamsplitter	33
2.9	General double sidebands	38
2.10	Differential method	42
2.11	External modulation scheme	46
2.12	Internal modulation scheme	49
2.13	Pre-modulation scheme	51
3.1	An interferometer with Herriott delay lines	59
3.2	An interferometer with Fabry-Perot cavities	62
3.3	Response of interferometers with delay lines	66
3.4	Response of interferometers with Fabry-Perot cavities	67
3.5	A comparison of delay lines and Fabry-Perot cavities	69
3.6	Power-recycled interferometers (delay line/Fabry-Perot)	72/73
3.7	Response of power-recycled interferometers with delay lines	76
3.8	The input efficiency	78
3.9	Response of power-recycled interferometers with Fabry-Perot cavities	79
3.10	Dual-recycled interferometers (delay line/Fabry-Perot)	82/83
3.11	The effect of the signal cavity	86
3.12	Response of dual-recycled interferometers with delay lines	88
3.13	Interferometers using synchronous recycling	91

4.1	The optical layout of resonant sideband extraction	104
4.2	The signal cavity in resonant sideband extraction	105
4.3	Optics as a matrix	108
4.4	A Fabry-Perot cavity	109
4.5	The power transmittance of a Fabry-Perot cavity	110
4.6	Fabry-Perot cavities at resonance and at anti-resonance	111
4.7	The phasor diagram of the light reflected from a cavity	113
4.8	The phase shifts of the light reflected from cavities	114
4.9	A three-mirror coupled cavity	116
4.10	A tile pattern (under-coupled)	117
4.11	Tile filling the x - y plain	118
4.12	The power transmittance (under-coupled)	122
4.13	A tile pattern (over-coupled)	124
4.14	The power transmittance (over-coupled)	125
4.15	The dependence on the signal extraction cavity lengths	127
4.16	Response of broad-band interferometers	133
4.17	The normalized sensitivity and the bandwidth	136
4.18	Response of narrow-band interferometers	137
4.19	Response of interferometers for low-frequencies	139
5.1	A non-ideal surface of a mirror	149
5.2	The beam radius in a Fabry-Perot cavity	151
5.3	The stability of a Fabry-Perot cavity	153
5.4	Scattering due to sinusoidal fluctuation of a surface	156
5.5	Suppression or enhancement factor in a cavity	159
5.6	The scattering with the mode cleaning effect of a cavity	161
5.7	The sagitta and the local curvature of a mirror	164
5.8	Problems associated with thermal distortions	175
5.9	Possible remedies for thermal distortion problems	179
D.1	The loci of the denominator and of its approximation	194
F.1	A single sideband in a dual-recycled interferometer	200
F.2	Dual recycling with different local oscillator phases	200
F.3	A single sideband in resonant sideband extraction	201
F.4	Resonant sideband extraction with different local oscillator phases	201

List of notations

Relations	Operators
\sim : order estimation	dx ($\llbracket x \rrbracket$) : differentiation of x
$=$: exact relation, identity	∂x ($\llbracket x \rrbracket$) : partial differentiation of x
\equiv : definition	Δx ($\llbracket x \rrbracket$) : difference of two x s
\simeq : approximate relation	δx ($\llbracket x \rrbracket$) : time variant change in x
\approx : numerical result	\tilde{x} ($\llbracket x \rrbracket / \sqrt{\text{Hz}}$) : linearized spectral density of x
\cong : numerical definition	S_x ($\llbracket x \rrbracket^2 / \text{Hz}$) : double sided spectrum of x^2

Mathematical/physical constants

$i \equiv \sqrt{-1}$ (1)	*	: imaginary unit
$\pi \cong 3.141592653$ (1)		: circular constant
$e \cong 2.718281828$ (1)		: base of natural logarithm
$\hbar \cong 1.055 \cdot 10^{-34}$ (J · s)		: reduced Planck's constant
$c \cong 2.998 \cdot 10^8$ (m/s)		: speed of light
$e \cong 1.602 \cdot 10^{-19}$ (C)		: elementary charge
$G_N \cong 6.672 \cdot 10^{-11}$ (N m ² /kg ²)		: gravitational constant
$g \cong 9.806$ (m/s ²)		: gravitational acceleration
$N_A \cong 6.022 \cdot 10^{23}$ (1/mol)		: Avogadro number
$k_B \cong 1.381 \cdot 10^{-23}$ (J/K)		: Boltzmann constant
$\mu_0 \cong 4\pi \cdot 10^{-7}$ (N/A ²)		: magnetic permeability of vacuum
$\epsilon_0 \cong \frac{10^7/4\pi}{(2.998 \cdot 10^8)^2}$ (C ² /N m ²)		: electric permittivity of vacuum
$Z_0 \cong 376.7$ (Ω)	$= \sqrt{\mu_0/\epsilon_0}$: characteristic impedance of vacuum
$\alpha_{EM} \cong 7.297 \cdot 10^{-3}$ (1)	$= e^2/4\pi\epsilon_0\hbar c$: fine structure constant
$\alpha_G \cong 5.905 \cdot 10^{-39}$ (1)	$= G_N M_p^2/\hbar c$: characteristic strength of gravitation
$\mathcal{H}_0 \cong 50 \sim 100$ (km s ⁻¹ /Mpc)		: Hubble's constant

Mathematical symbols

H (1) : Hermite polynomials	\mathbf{C} (1) *: Complex number set
J (1) : Bessel functions	\mathbf{N} (1) : Natural number set
L (1) : Laguerre polynomials	\mathbf{N}_0 (1) : Natural numbers and zero
O (1) : same order	\mathbf{Q} (1) : Rational number set
o (1) : higher order	\mathbf{R} (1) : Real number set
δ (1) : Kronecker delta/delta function	\mathbf{Z} (1) : Integer number set
σ (-) : standard variance	Γ (1) *: gamma function
\mathcal{P} (1) : probability	\prod (1) : product
\wp (-) : probability distribution	\sum (1) : summation

Symbols for physical variables

F (N)	: force	h (1)	: amplitude of gravitational wave
G (1)	*: transfer function (gain)	k (1/m)	: wave number
I (A)	: current	ℓ (m)	: arm length
M (kg)	: mass	m (1)	: modulation index
N (1)	: number	n (1)	: refractive index
P (W)	: power	p (N·s)	: momentum
Q (1)	: quality factor	q (C)	: charge
R (1/m ²)	: Riemann or Ricci tensor, scalar curvature	r (m)	: radius
S (1)	: scattering matrix	s (m)	: proper time
T (J/m ³)	: energy momentum tensor	t (s)	: time
V (V)	: voltage	u (1)	*: normalized eigen function
Y (1/Ω)	*: admittance	v (m/s)	: velocity
Z (Ω)	*: impedance	w (m)	: beam radius
\mathcal{A} (1)	: relative loss	x (m)	: position, coordinate, axis
\mathcal{C} (1)	: contrast	y (m)	: position, coordinate, axis
\mathcal{E} (J)	: energy	z (m)	: position, coordinate, axis
\mathcal{F} (1)	: finesse	α (1/K)	: expansion coefficient
\mathcal{G} (1)	: equivalent power gain	β (1/K)	: temperature dependence of refraction index
\mathcal{R} (m)	: curvature radius	γ (1/m)	: absorption coefficient
\mathcal{T} (K)	: absolute temperature	ε (C/Vm)	: electric permittivity
\mathcal{V} (m ³)	: volume	ζ (1)	: geometric parameter
\mathbf{B} (Vs/m ²)	: magnetic displacement	η (1)	: efficiency/Minkowski metric
\mathbf{D} (C/m ²)	: electric displacement	θ (rad)	: angular coordinate
\mathbf{E} (N/C)	: electric field	ϑ (rad)	: angle
\mathbf{H} (A/m)	: magnetic field	κ (W/Km)	: thermal conductivity
Θ (1)	: spatial phase	λ (m)	: optical wavelength
Λ (m)	: spatial wavelength	μ (N/A ²)	: magnetic permeability
Φ (rad)	: normalized amplitude of spatial fluctuation	ρ (1)	: amplitude reflectivity
Ψ (1/m)	*: wave function	ς (m)	: sagitta
Ω (1)	: normalized spatial frequency	τ (1)	: amplitude transmittance
Γ (1/m)	: Christoffel's symbol	φ (rad)	: rotational coordinate
Υ (1)	: matching error	ϕ (rad)	: tuning (phase)
a (V)	*: normalized field amplitude	χ (1)	: ratio
f (Hz)	: Fourier frequency	ψ (rad)	: phase
g (1)	: metric tensor	ω (rad/s)	: angular frequency

$A \sim E$ are used for temporary purposes; j, l, m and $\kappa, \lambda, \mu, \nu, \xi$ are also used for dummy indices. Those with * can have a complex value. $\llbracket x \rrbracket$ and $\llbracket x \rrbracket$ means the numeric value and the unit of a physical variable $x = \llbracket x \rrbracket \cdot \llbracket x \rrbracket$.

Acknowledgments

First of all, I must declare that most of the results here are natural extensions of the late Dr. B.J. Meers' works during 1988 to 1991. His death in early 1992 was a great loss to this field. I am sure that this thesis would have been totally different if he had not pursued his pioneering works on signal recycling and relevant topics. Though some of my results may conflict with his, most of the contents—including the idea of resonant sideband extraction—might have been established before this thesis if he were alive. Again, I really deplore his too early death, and I would pray for the repose of his soul.

The idea of resonant sideband extraction was strongly motivated by the original analyses of thermal distortion effects investigated by Dr. W. Winkler and Dr. K.A. Strain. Their results revealed the importance of thermal effects, which led me to seek a way to avoid this problem. Furthermore, their advice was indispensable in analyzing thermal effects in resonant sideband extraction. I appreciate their many contributions in doing this work.

I am owing greatly to R. Schilling for his valuable answers to many questions, from simple ones to the most complicated ones that I found in writing throughout this thesis. It is his enrichment that made this thesis as comprehensive as it is.

I must acknowledge the help from A. Rüdiger in many aspects—physics of course, mathematics, English, German, writing style, and lots more. I regret that I could not follow all of his suggestions just because they spread so widely in so many fields.

I would like to thank Professor K. Danzmann for giving me a chance to continue this work in Germany. Without his offer, this thesis might not have appeared. I should also mention his kind permission to use the experimental facilities at Garching and to 'consume' optical components including many supermirrors, for the experiment of resonant sideband extraction that is now under way. The experiment is going slowly but smoothly with the assistance from G. Heinzl, thanks also to the environment of the Max-Planck-Institut für Quantenoptik.

There are many other people with whom I have discussed one or the other item in this thesis; these include Dr. P.G. Nelson now at Technical Manufacturing Corporation (TMC), Dr. J.M. Chen now at Siemens, Professor J. Hough, Dr. E. Morrison, Dr. D.I. Robertson, and Dr. H. Ward at the University of Glasgow, Dr. D. Shoemaker at Massachusetts Institute of Technology (MIT), and Dr. S. Kawamura at California Institute of Technology (Caltech). With all of them I had instructive as well as enjoyable discussions.

Last but not least, I am grateful to Dr. J.E. Logan who read the original manuscript and made lots of corrections of my English. Her suggestions helped me in writing this thesis and made it much easier to read. I am, of course, responsible for any remaining errors.

I also want to thank the Japanese Society for Promotion of Science (JSPS) and the Max-Planck-Gesellschaft (MPG) for their support during this work.

Chapter 1

Introduction

1.1: Gravitational waves and their detection

The detection of gravitational waves remains as one of the most challenging aims in experimental physics today. In spite of the efforts in the past decades, there is no successful result so far. These efforts were, however, not in vain. The experiences, together with recent developments of various technologies, made a consensus that interferometers with several kilometer arm lengths will be capable of reaching this very attractive target.

This motivated a few groups to build such interferometers; among them are the LIGO project [1] of the U.S.A. and the VIRGO project [2], a French-Italian collaboration. Both projects officially started in 1994 and are planning to operate their initial interferometers around the end of this century. In addition, several countries have similar plans but on a somewhat smaller scale, i.e. interferometers with several hundred meters. These include the GEO-600 [3] of the German-British collaboration, TAMA-300 [4] in Japan, and possibly AIGO-400 in Australia [5]. These groups may also extend their interferometers to a larger scale in the future (though it may not necessarily be at the same site).

The basic principle of these interferometric gravitational-wave detectors is described in Chapter 2. The actual interferometers will, however, employ more complicated designs in order to improve the shot-noise-limited sensitivity. The optical designs previously proposed are comprehensively reviewed in Chapter 3. A detailed analysis of a new optical design, *resonant sideband extraction*, can be found in Chapter 4. As is shown in Chapter 5, this new configuration has an advantage when thermally induced distortions of the beam become significant.

A gravitational wave can be understood in close analogy to an electromagnetic wave [6]. Both are transverse waves which propagate at the speed of light through space. What is different between the two cases is that there is only positive ‘charge’ of gravitation whereas both positive and negative ones in electro-magnetic interactions. This results in a difference on the lowest mode of

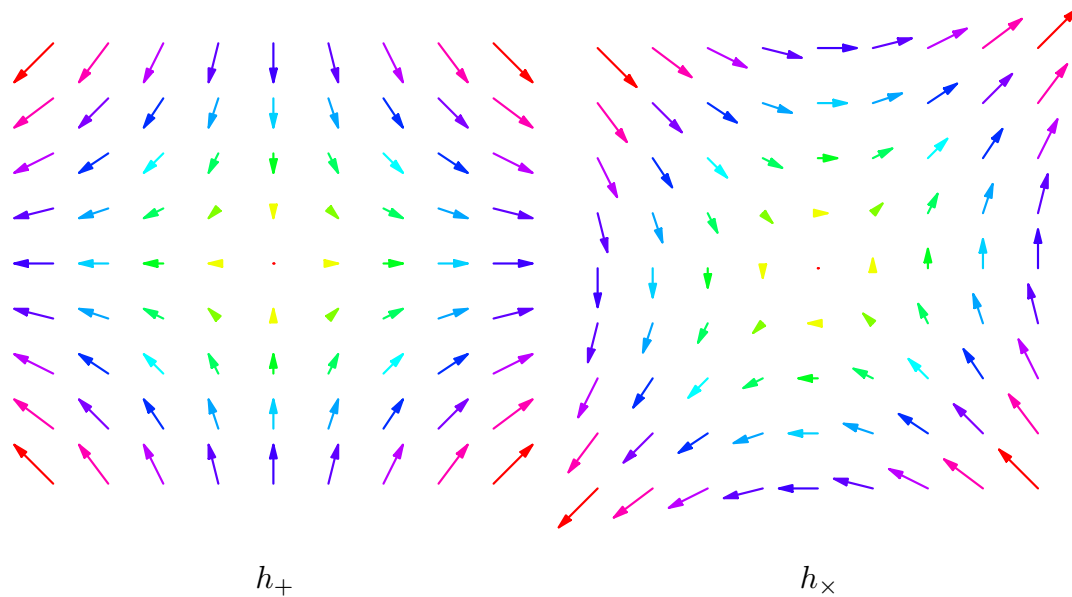


Figure 1.1: The quadrupole force field of a gravitational wave which is propagating perpendicularly to the figure plane. Its effect in the perpendicular plane is expressed as ‘stretching’ in one axis and ‘compressing’ in the orthogonal axis. Note that its effect increases with distance. The two (linear) polarizations, h_+ and h_\times , are shown (see 2.1.1 for their definition).

oscillation: it is dipole for an electro-magnetic wave and quadrupole for a gravitational wave. The ‘quadrupole’ nature of a gravitational wave is illustrated in Fig. 1.1.

The existence of gravitational waves was predicted by Einstein [7] shortly after the development of his theory of gravitation, the General Theory of Relativity (often abbreviated as General Relativity). In today’s physics, this theory is believed as the correct expression of gravitational phenomena (in a ‘classical’ sense), through the stringent experimental tests in the past decades. (A good historical review can be found in [8].) Among them, however, there is only one experiment which is related to the phenomenon of gravitational waves.

Hulse and Taylor, who received the Nobel Prize in 1993 for this result, confirmed the existence of a gravitational wave, though indirectly, through the observation of the binary pulsar PSR 1913+16 [9, 10]. This pulsar is losing its orbital energy by emitting a part of it as gravitational waves, and thus its period is decreasing with time. This decay rate is in very good agreement with what is expected from General Relativity.

The direct detection of a gravitational wave is, however, still left as an unattained goal for experiments. Its difficulty simply arises from the fact that a gravitational wave's interaction is so faint. Among the four fundamental interactions known today, the gravitational interaction is the weakest by many orders of magnitude. This means that a gravitational wave is not only difficult to detect, but also difficult to generate.

Just to get an impression of its faintness, let us consider the artificial generation of a gravitational wave. Assuming a rod with a mass M and a length ℓ rotating at an angular velocity ω around an axis perpendicular to the rod, the total power emitted as a gravitational wave (at an angular frequency 2ω) is given by [6]

$$P_{\text{GW}} = \frac{2 G_{\text{N}} M^2 \ell^4 \omega^6}{45 c^5} \approx 7.5 \cdot 10^{-26} \text{ W} \times \left[\frac{M}{1000 \text{ kg}} \right]^2 \left[\frac{\ell}{1 \text{ m}} \right]^4 \left[\frac{\omega/2\pi}{1 \text{ kHz}} \right]^6 \quad (1)$$

from the quadrupole formula of radiation. The quoted values correspond to a rod made of iron with a radius of ~ 0.2 m.

The above value should be compared with the total energy of the rotation, which is given by

$$\mathcal{E}_{\text{rot}} = \frac{1}{2} M \ell^2 \omega^2 \approx 1.3 \cdot 10^{10} \text{ J} \times \left[\frac{M}{1000 \text{ kg}} \right] \left[\frac{\ell}{1 \text{ m}} \right]^2 \left[\frac{\omega/2\pi}{1 \text{ kHz}} \right]^2. \quad (2)$$

Thus, the rod emits its rotational energy by a fraction of

$$\eta_{\text{period}} = \frac{4\pi G_{\text{N}} M \ell^2 \omega^3}{9 c^5} \approx 9.5 \cdot 10^{-39} \times \left[\frac{M}{1000 \text{ kg}} \right] \left[\frac{\ell}{1 \text{ m}} \right]^2 \left[\frac{\omega/2\pi}{1 \text{ kHz}} \right]^3 \quad (3)$$

per rotational period.

Considering the inverse process, the efficiency of detection must be of the same order as the efficiency of emission in eq. (3), though the actual number will depend on the detection scheme. This is the reason why the gravitational wave was thought to be of only theoretical existence rather than of real life, until Weber initiated experimental efforts to detect gravitational waves of astronomical origin (see 1.2) with his pioneering work [11].

He used a different type of detector from interferometers, called a *resonant type* detector. This ‘bar’ shaped antenna is made of aluminum with a high Q -value in order to improve the signal-to-noise ratio using its resonance. After his announcement of having detected gravitational waves in 1969 [12]—which at present is believed not to be a real signal, though—many groups around the world built similar detectors (see Ref. [13] and references therein), some with improved technologies. The main effort in the second generation resonant detectors was to cool the antenna to low temperature in order to reduce its thermally excited motion.

After decades of development, resonant detectors nowadays use the state-of-art level technologies and some achieved the best sensitivity to gravitational waves among any sorts of detectors. Their prospects for the future, however, seem less promising because of the ‘standard quantum limit’ which arises before reaching the interesting sensitivity level considering the expected signal amplitude (see eq. (5) in 1.2). Although this ‘limit’ may be overcome using a method called *back action evasion*, the technology is still to be developed. The above fact makes interferometers seem more attractive as a future device to detect gravitational waves. In addition, the narrow-band nature of a resonant detector seems less appealing for an astronomical observatory (see 1.2).

There are also other types of experimental efforts to detect gravitational waves, i.e. spaced-based ones such as *Doppler tracking* and *pulsar timing*. Both of them, as well as the interferometers in space, are aiming at gravitational waves in a much lower frequency range, from tenths of a hertz even down to microhertz. Thus, they are complementary rather than competitive to ground-based detectors such as interferometers, which are aiming at several tens of hertz to a few kilohertz

1.2: Expected signal

From eqs. (1) and (3), the artificial generation of a gravitational wave with a detectable strength seems to be impossible. Or, one may lose any hope of detecting any kind of gravitational waves. Before being discouraged, however, one should look at these equations a little more carefully. The gravitational wave power as well as the efficiency of emission depends on the mass, the length, and the period.

What we know as the most massive objects are stellar objects. A typical mass for such an object may be of the order of a solar mass,

$$M_{\odot} \approx 2.0 \cdot 10^{30} \text{ kg} \quad (4)$$

or even much heavier. Thus, we can expect much stronger gravitational effects among such objects. There is, however, another question: whether such a massive object will be accelerated enough so that the quadrupole moment varies rapidly. As can be seen in eq. (1), the gravitational wave power also depends on the time scale of the motion. If all the stellar objects are as quiescent as the sun, we cannot expect any strong gravitational waves from them.

Fortunately, there are several phenomena (‘events’), such as supernovae, expected in astronomical physics in which huge amounts of masses are accelerated within a relatively short time scale. According to numerical simulations, these events seem likely to produce signals with detectable strengths (see Fig. 1.2).

These events are also of great interest from the viewpoint of astrophysics. Furthermore, very little information from these events can be obtained through other observational media, that is, electro-magnetic waves in various frequency ranges. Only gravitational waves can bring the information from such ‘exotic’ events, due to their weakness of interaction. Electro-magnetic waves interact strongly with the gas surrounding the core and the information on the core will be lost while passing through the gas region.

This is why large scale interferometers are often called ‘observatories’—they are expected not only to be a ‘detector’ just to confirm the existence of gravitational waves, but also to be an instrument which provides various types of informations about the universe. For this purpose, interferometers are more adequate than resonant detectors because of their broad-band nature and the flexibility to alter their parameters such as the detector bandwidth.

It may be interesting to note that the astronomical events expected to produce signals are caused by gravitation. Since the gravitational force is proportional to the mass of the object, it can accelerate massive objects relatively easily (compared with other kinds of forces). Thus, the gravitational wave detection is by itself of interest to gravitational physics, and the signals observed are also valuable information on gravitational phenomena.

The drawbacks of astronomical sources of gravitational waves are their uncertainty, non-reproducibility, and their distance. The former two features require statistical treatment of the detection process, as is briefly discussed

in 3.1.1. What affects the signal strength is the distance between the source and the detector. In order to achieve a reasonable event rate (\sim several per year) of detectable signals, we must be able to observe as far out as the Virgo Cluster ($\sim 15 \text{ Mpc} \approx 4.6 \cdot 10^{23} \text{ m}$) where thousands of galaxies exist.

The gravitational wave amplitude h (see 2.1.1 for its definition) from such a distant source will be around

$$h \approx 1.4 \cdot 10^{-21} \times \left[\frac{\mathcal{E}_{\text{GW}}}{10^{-2} M_{\odot} c^2} \right]^{\frac{1}{2}} \left[\frac{\omega_{\text{g}}/2\pi}{1 \text{ kHz}} \right]^{-1} \left[\frac{t_{\text{d}}}{1 \text{ ms}} \right]^{-\frac{1}{2}} \left[\frac{x}{15 \text{ Mpc}} \right]^{-1}, \quad (5)$$

where \mathcal{E}_{GW} is the amount of energy emitted as the gravitational wave (measured by 10^{-2} of a solar mass energy $M_{\odot} c^2 \approx 1.8 \cdot 10^{47} \text{ J}$), $\omega_{\text{g}}/2\pi$ is its typical frequency, t_{d} is the duration time, and x is the distance from the source. The quoted values correspond to a fairly large burst of supernova occurring in the Virgo Cluster. Thus, this number is a ‘yardstick’ of the amplitude of a gravitational wave which we expect to detect.

Although the above number is discouragingly small, it does not mean this gravitational wave is ‘weak’. Actually, a huge amount of energy is carried by this gravitational wave. The power flux \mathcal{J}_{GW} on the earth of a gravitational wave is given by

$$\mathcal{J}_{\text{GW}} \approx 0.32 \text{ W/m}^2 \times \left[\frac{\omega_{\text{g}}/2\pi}{1 \text{ kHz}} \right]^2 \left[\frac{h}{10^{-21}} \right]^2. \quad (6)$$

This means that the smallest gravitational wave we may be able to observe ($h \approx 10^{-22}$) is as ‘bright’ as the full moon in terms of power flux.

The possible events which are expected to produce gravitational waves with detectable strengths and frequencies are extensively discussed in Refs. [8] and [14]. Here we briefly list up some of them, in order to quote the requirements on the detector to observe them.

A supernova, or more generally a gravitational collapse, will produce a gravitational wave of burst type, expected to have one or a few cycles within a few milliseconds, although the exact waveform is poorly understood due to the complexity of the process. Thus, a broad-band response of up to a few kilohertz would be appropriate to observe this kind of signals. The amplitude of the signal depends on the amount of energy emitted as the gravitational wave, which is also not well-understood. Numerical modelings suggest that 10^{-6} to

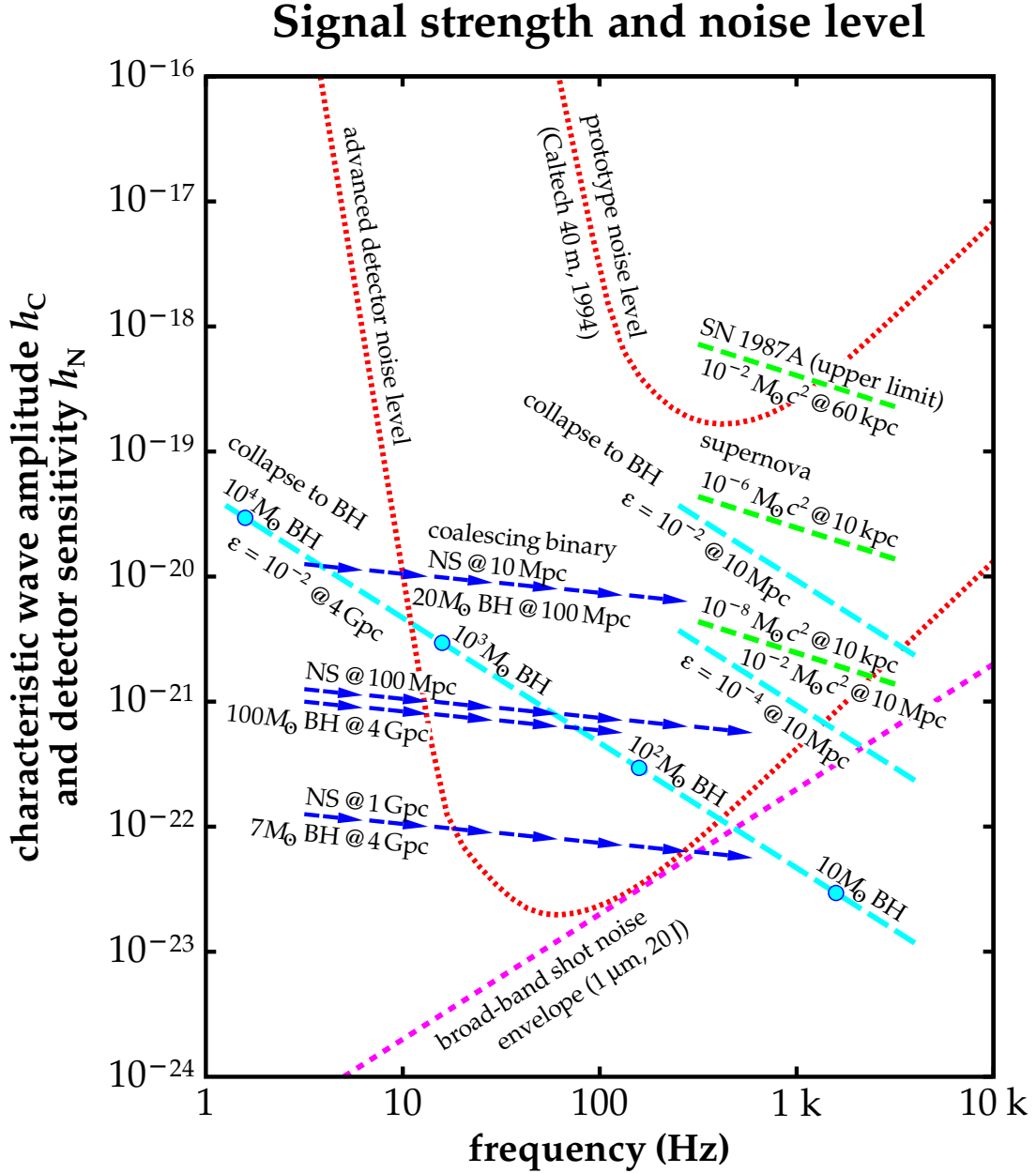


Figure 1.2: The expected signal strength and the frequency range (data are taken from Ref. [1] excepting the updated prototype noise level). For comparison, expected total noise level in an advanced detector is also shown. The former is represented by the characteristic wave amplitude $h_C \sim h_{\text{peak}} \times \sqrt{N_{\text{osc}}}$ where h_{peak} is the peak amplitude and N_{osc} is the number of oscillation. The noise level is represented by $h_N \sim \tilde{h} \sqrt{\Delta f_{\text{BW}}}$. For more precise discussion of these definitions, see Ref. [8].

10^{-4} of solar mass energy will be emitted as the gravitational wave if the system is axisymmetric, and probably more for the non-axisymmetric case.

Binary stars, consisting of ‘compact’ stars such as white dwarfs, neutron stars, and black holes, are another type of gravitational wave sources. In such a system, as PSR 1913+16 does, the two stars lose their orbital energy by emitting gravitational waves while they are rotating at a gradually decreasing period (up to a few milliseconds), and in the end the two stars will coalesce. In the last few minutes, the frequency of the gravitational wave due to the rotation reaches our observable range, and a burst type wave is expected at the moment of coalescence.

The advantage of this type of source is that the waveform in the last few minutes (‘chirp’) is well understood and a good signal-to-noise ratio can be expected by using appropriate templates. In this case a detector with a good low-frequency performance, i.e. both a good (spectral) sensitivity and a low cut-off frequency, is desirable since the ‘chirp’ changes its frequency slowly at lower frequencies and rapidly at higher frequencies. The burst at the very last moment may be better observed with a broad-band detector like the one for supernovae.

There are several known pulsars with periods from which we can expect a gravitational wave signal in the observable frequency range. The amplitude of the signal from such a system depends on its quadrupole moment, and is usually expected to be small. In this case, however, it is possible to improve the signal-to-noise ratio by integrating the signal for a long period. Furthermore, it is also possible to improve the (spectral) sensitivity of a detector by using a narrow-band configuration, as is discussed in 3.4.2. Thus, it may be possible to observe (or at least to set an upper limit for) the gravitational wave from a known pulsar.

In a way similar to the 3 K background of electro-magnetic radiation, it is expected that there is stochastic background radiation of gravitational waves. It is a reminiscence of those produced in the early history of the universe, such as the ones produced by cosmic strings. In this case the correlation of the outputs from two or more detectors must be examined in order to distinguish the signal from the noise of a detector, since the signal itself has random nature. Similar to the pulsar case, narrow-banding of the detectors, probably at a low frequency

where the signal is expected to be the strongest, may be useful for detecting this type of signal.

1.3: Possible noise

As is indicated by eq. (5), the expected signals from stellar objects are still quite small. There would be no difficulty to detect them if the output of a detector contained only signals and nothing else. In fact, however, there are many possible mechanisms which can cause outputs similar to those attributed to gravitational waves. All such outputs not produced by gravitational waves are specified as *noise*.

The ‘detection’ of a gravitational wave must be done in the presence of noise, i.e. what can be detected is a signal that is distinguishable from the noise. (Further discussion of this point can be found in 3.1.1.) Thus, the noise must be kept as small as possible, in order to increase the probability of detecting the gravitational wave signal. This can be accomplished either (or both) by reducing the origin of the noise or by making the detector insensitive to the noise source.

Even with such efforts, however, there will be remaining noise which limits the probability of signal detection. Some of the noise sources must be considered as practical limitations within the given environmental conditions, and then others must be suppressed so that they will be less significant than these limitations. Here we review the possible mechanisms which cause the noise in the output of an interferometer (see Fig. 1.3).

What we observe as an output from an interferometer is the changes in the photo-current of a photo-detector. As is discussed in Chapter 2, a gravitational wave causes a differential phase shift between the two arms of an interferometer, and this is converted to a change in the photo-current. Any mechanism which causes a similar effect will be a noise source. The possible noise sources in an interferometer can be classified into two categories: one is those which cause an ‘actual’ optical path length change in a similar way to a gravitational wave, and the other is those which cause a spurious signal through the detection process.

The former includes physical path length fluctuations due to the motion of optical components or of the beam position, excited either seismically or thermally. Even the light which measures the position causes motions of the optics

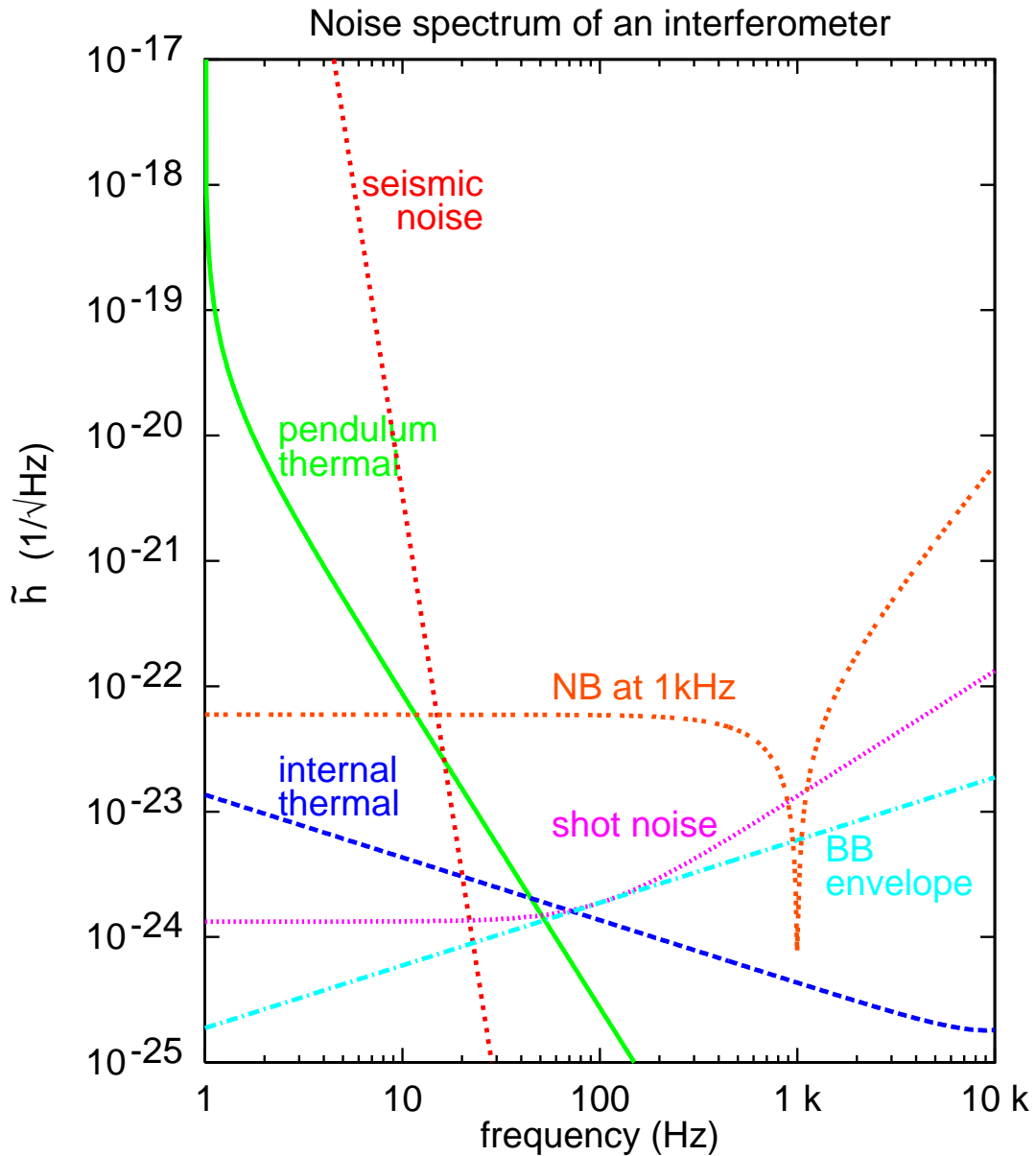


Figure 1.3: The expected level of limiting noise sources in an advanced detector. The seismically excited motion of the optical components, remaining after isolation system, is expected to be the limiting noise source in the lowest frequency region. In higher frequency region, thermally excited motion of either pendulum or internal motion of test mass will dominate the noise. Shot noise will become significant above ~ 100 Hz. For broad-band configurations, the best sensitivity achievable is determined by the amount of energy stored in the optical system (shown as the ‘envelope’). The peak sensitivity may be improved by using a narrow-band configuration, however only within a limited frequency range.

by the radiation pressure fluctuation, from which (and shot noise described below) the standard quantum limit arises (see Ref. [15]). In addition, the optical path length may change also by the refractive index fluctuations of the residual gas or of the optical substrates.

Even without any change in the optical path length, there will be spurious signals caused at various points in the detection process. The signals due to fluctuations in frequency or amplitude of the light source are typical examples of this sort of noise sources. Another example is the coupling of scattered light with the main beam. In any case, these affect the output through the imperfection of the interferometer.

The statistical error in the photon counting process, often called *shot noise*, is also of this type, but present even in an ideal interferometer. It is expected that this noise will be the limiting factor in the most interesting frequency range, above the lower cut-off frequency determined by other noise sources. The discussion in the following chapters assumes shot noise to be the *only* limiting noise source, unless otherwise noted. This is to concentrate on the fundamental limitation of each optical configuration.

In a ground-based experiment, the motions of test masses due to seismic excitations are inevitable. The typical ground motion, measured at a reasonably quiet site, has a spectrum of

$$\tilde{\delta x}_{\text{seis}} \approx 10^{-7} \text{ m}/\sqrt{\text{Hz}} \times \left[\frac{f}{1 \text{ Hz}} \right]^{-2}. \quad (7)$$

If this directly moves a test mass, it will cause a much stronger effect than that expected from a gravitational wave. Mechanical filters, made of pendulums and lead-and-rubber stacks, are in general employed in order to isolate the test masses from this disturbance.

The isolation system using these low-pass filters can be made, for given spectra of seismic motion, such that gravitational waves in the frequency range above several tens of hertz can be observed [16]. Because of its steep rise towards lower frequencies, this noise practically determines the lower cut-off frequency of the observable bandwidth.

Even if a test mass is perfectly isolated from external excitations, it will still be moving due to its own thermal energy. According to the equipartition theorem, each vibrational mode has a kinetic energy of $\frac{1}{2}k_{\text{B}}\mathcal{T}$ where k_{B} and \mathcal{T}

are the Boltzmann constant and the absolute temperature, respectively. The vibrational modes relevant to the test mass are the internal modes, the pendulum modes, and the ‘violin’ modes of the wire suspending the test mass [17].

In either case, the common remedy is to make the quality factor (Q -value) high so that most of the thermal energy will be concentrated within a narrow frequency range near the resonance. By making the resonance frequency outside the observation bandwidth, this effect will be less significant within the observation bandwidth. The thermally excited pendulum mode may, however, still become a lower limit of the observation bandwidth when the seismic excitation is successfully isolated.

The refractive index fluctuation of the residual gas also causes the change in the optical path length [18]. Thus, the whole optical path must be housed in a vacuum chamber, with the pressure of the order of $\sim 10^{-6}$ Pa. The refractive index fluctuation in optical substrates is expected to be of a negligible level, at least for ground-based experiments. It may, however, arise in space interferometers in which case much lower frequencies are of interest.

The fluctuation in the light source, either in amplitude or phase or spatial mode, appears in the interferometer output when there is an asymmetry in the interferometer [19]. Although the asymmetry is usually made to be small, the remainder often requires a very high stability of the light source. The coupling of the fluctuation to the interferometer output may also depend on the detection scheme and other factors (see Appendix F).

Although only a little attention is often paid to it in literature, ‘artificial’ noise sources also pose serious problems. These can cause either path length fluctuations or spurious signals, depending on the mechanism through which the noise becomes apparent in the output of a detector. For example, the control system required to keep the interferometer at the operating point moves a test mass by an actuator. The electronics of this control system and of the actuator inevitably have noise, which also move the test mass. Another example is the noise which will be added after the photo-detection, which cannot be separated from the photo-current. Since these noise sources are not considered as the limiting factors, all of them must be suppressed by careful engineering work.

Chapter 2

Principle of gravitational wave detection

2.0: Introduction

The very first thing to be considered in an optically coupled detector of gravitational waves is how the detector interacts with gravitational waves. As is discussed in this chapter, a gravitational wave will induce phase shift on the light travelling perpendicularly to the wave's incident direction. It is this phase shift that we need to detect in order to obtain the signal from an optically coupled detector.

The question is how we can detect the phase shift of the light, since the output from a photo-detector does not contain the phase information of the light. Thus, it is necessary to convert the phase shift to some other form which can be detected, i.e. the change in the photo-current which is proportional to the averaged light power. For this purpose, an application of the technique known as heterodyne detection is useful, although a slight modification is required.

Although there are several methods proposed in order to increase the amount of the phase shift (which are discussed in Chapters 3 and 4), the expected phase shift due to a gravitational wave is very faint. Thus, the conversion and the detection must be as efficient as possible to achieve good sensitivity. In this chapter, the principle of detecting the phase shift of light is discussed, and some of the schemes to realize it are presented.

2.1: Effect of gravitational waves

2.1.1: Gravitational waves in General Relativity

In metric theories including the General Theory of Relativity, the 'four-dimensional distance' (also called *proper distance* or *proper time*) between two points in space time with coordinate values¹⁾ x^μ and $x^\mu + dx^\mu$ in a certain

¹⁾ In this section, superscripts are used not to mean powers but as suffixes, unless they appear after parentheses. The indices follow Einstein's convention, i.e. these which appear as both super- and subscripts mean that they vary from 0 to 3 and the sum of these terms must be taken.

coordinate frame is given by

$$(ds)^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (\mu, \nu = 0, 1, 2, 3), \quad (1)$$

where $g_{\mu\nu}$ is the *metric tensor*. This represents all the information about the ‘curved’ space time.

In the General Theory of Relativity, the metric tensor is determined by the *energy-momentum tensor* $T_{\mu\nu}$ according to the Einstein equation

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G_N}{(c)^4} T_{\mu\nu} \quad (2)$$

(the cosmological term is assumed to be zero). Here the scalar curvature R and the Ricci tensor $R_{\mu\nu}$ are related to $g_{\mu\nu}$ through the Riemann tensor $R^\lambda{}_{\xi\mu\nu}$ and Christoffel’s symbol $\Gamma^\lambda{}_{\mu\nu}$ as follows:

$$\begin{cases} R &= g^{\mu\nu} R_{\mu\nu} \\ R_{\mu\nu} &= R^\xi{}_{\mu\nu\xi} \\ R^\lambda{}_{\xi\mu\nu} &= \partial_\mu \Gamma^\lambda{}_{\xi\nu} - \partial_\nu \Gamma^\lambda{}_{\xi\mu} + \Gamma^\kappa{}_{\xi\nu} \Gamma^\lambda{}_{\kappa\mu} - \Gamma^\kappa{}_{\xi\mu} \Gamma^\lambda{}_{\kappa\nu} \\ \Gamma^\lambda{}_{\mu\nu} &= \frac{1}{2} g^{\lambda\kappa} (\partial_\mu g_{\nu\kappa} + \partial_\nu g_{\mu\kappa} - \partial_\kappa g_{\mu\nu}), \end{cases} \quad (3)$$

where $\partial_\mu = \partial/\partial x^\mu$, the partial differentiation with respect to x^μ . Thus, eq. (2) is a set of complicated non-linear partial differential equations.

For the vacuum, i.e. the space in which neither energy nor momentum is present, the right-hand-side of eq. (2) is a zero matrix. Then the Minkowski metric, given as

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{pmatrix}, \quad (4)$$

can be one of the solutions for $g_{\mu\nu}$ in eq. (2). In the region where the energy (mass) density is small, this is a good first order approximation of the actual metric $g_{\mu\nu}$. Thus, in such cases, the solution of eq. (2) can be treated as a perturbation from the Minkowski metric, i.e.

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad (5)$$

where $|h_{\mu\nu}| \ll 1$ represents the amount of the perturbation. In the following discussion, only the first order terms of $h_{\mu\nu}$ are considered, and thus the indices are raised and lowered by $\eta_{\mu\nu}$ (or $\eta^{\mu\nu}$) instead of $g_{\mu\nu}$ ($g^{\mu\nu}$).

Under this condition, the Einstein equation eq. (2) can be approximated by the ‘linearized’ equations for $h_{\mu\nu}$,

$$\begin{aligned} \square h^{\mu\nu} + \partial^\mu \partial^\nu h - \partial^\nu \partial_\kappa h^{\kappa\mu} - \partial^\mu \partial_\kappa h^{\kappa\nu} \\ - \eta^{\mu\nu} (\square h - \partial_\kappa \partial_\lambda h^{\kappa\lambda}) = 2 \frac{8\pi G_N}{(c)^4} T^{\mu\nu}. \end{aligned} \quad (6)$$

Here

$$\begin{cases} \partial^\mu = \eta^{\mu\nu} \partial_\nu = \eta^{\mu\nu} \partial / \partial x^\nu \\ h = \eta^{\mu\nu} h_{\mu\nu} \\ \square = \partial^\mu \partial_\mu = \eta^{\mu\nu} \partial_\mu \partial_\nu \end{cases} \quad (7)$$

are used for simplicity.

As mentioned above, the right-hand-side of eq. (6) is a zero matrix for the vacuum, and then this equation will be a wave equation. Thus there are solutions of transverse waves which propagate at the speed of light. By choosing the transverse-traceless coordinate system, the wave propagating along the x -axis can be written as

$$h_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & h_{22} & h_{23} \\ 0 & 0 & h_{32} & h_{33} \end{pmatrix}. \quad (8)$$

Here non-zero components are related by

$$\begin{cases} h_{22} = -h_{33} = h_+(x-ct) \\ h_{23} = h_{32} = h_\times(x-ct), \end{cases} \quad (9)$$

where h_+ and h_\times represent two independent components usually called *polarizations*.

2.1.2: The effect on free masses

A mass with no external force other than gravity falls freely, i.e. moves along the geodesic of the manifold which represents the space time curved by the existence of gravity. In the transverse-traceless coordinate system, a freely falling mass has fixed coordinate values. This does not mean, however, that the *spatial distance* between the two freely falling masses does not change, since the spatial coordinate frame varies with time.

Chapter 2: Principle of gravitational wave detection

Let us consider the effect of the gravitational wave described above on an optical measurement of the spatial distance. The light travelling between the two masses has the null proper time, i.e.

$$(ds)^2 = g_{\mu\nu} dx^\mu dx^\nu = -(c dt)^2 + g_{ij} dx^i dx^j = 0 \quad (i, j = 1, 2, 3). \quad (10)$$

Concentrating on the h_+ polarization for simplicity, the spatial metrics are perturbed as

$$g_{ij} = \delta_{ij} + h_{ij} \quad \begin{cases} h_{22} = -h_{33} = h_+(t) \\ h_{ij} = 0 \quad (\text{otherwise}). \end{cases} \quad (11)$$

Then eq. (10) reduces to

$$(ds)^2 = -(c dt)^2 + (dx)^2 + (1+h_+)(dy)^2 + (1-h_+)(dz)^2 = 0. \quad (12)$$

For the light travelling along the z -axis,

$$\left(\frac{dz}{dt}\right)^2 = \frac{(c)^2}{1-h_+} \quad (13)$$

can be obtained by inserting $dx = dy = 0$ into eq. (12).

If the coordinate values of the z -axis between the two masses differ by ℓ , the spatial distance between the two masses measured as a round trip time of light will be

$$\begin{aligned} 2\ell &= \int_0^\ell dz - \int_\ell^0 dz = \int_{t-t_r}^t \frac{dz}{dt} dt \\ &= \int_{t-t_r}^t \frac{c}{\sqrt{1-h_+(t)}} dt = c \int_{t-t_r}^t \left\{ 1 + \frac{1}{2}h_+(t) + O(|h_+|^2) \right\} dt, \\ t_r &\simeq \frac{2\ell}{c} - \frac{1}{2} \int_{t-t_r}^t h_+(t) dt \\ &\simeq \frac{2\ell}{c} - \frac{1}{2} \int_{t-2\ell/c}^t h_+(t) dt \quad (\text{for } |h_+| \ll 1), \end{aligned} \quad (14)$$

i.e. the retarded time will be modulated by the gravitational wave. If we assume the speed of light is constant, then this effect is equivalent to the motion of one of the test masses; or if we assume the distance is constant, then this effect is equivalent to a modulation of the speed of light.

The light travelling along the y -axis experiences the same effect from the h_+ polarization but with an opposite sign, i.e. $dy/dt = c/\sqrt{1+h_+(t)}$. Thus, the

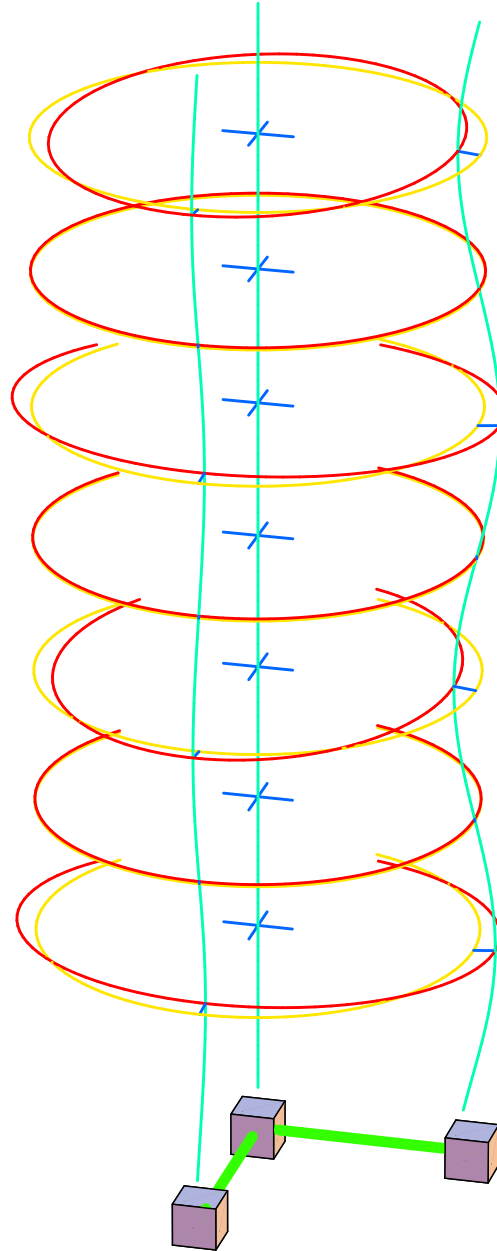


Figure 2.1: An illustration of the effect of a sinusoidal gravitational wave on an optical measurement. A gravitational wave changes the distances of the masses on the plane perpendicular to its propagation axis. The changes of the distances are opposite in orthogonal axes in this plane. The heavier lines represent the ellipses which are distorted from circles (shown by the lighter lines) by the gravitational wave.

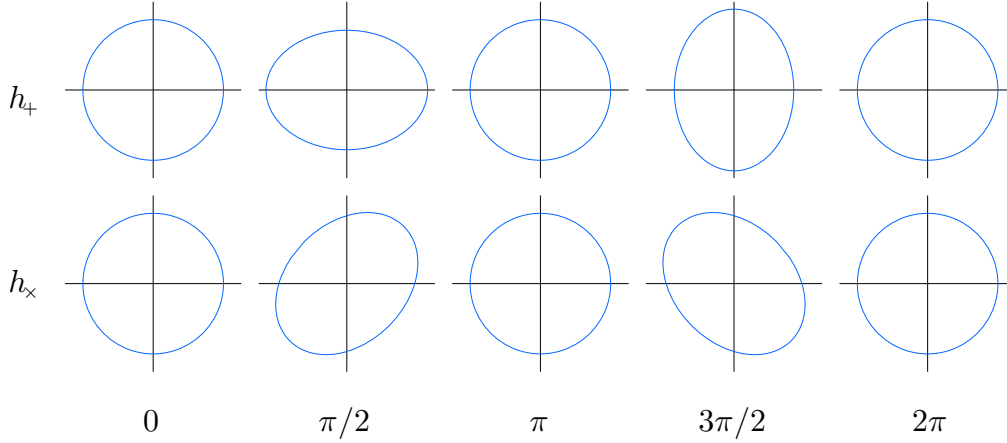


Figure 2.2: A schematic view of the deviation caused by a gravitational wave. Test particles originally placed on a circle change their relative positions elliptically, with major axes determined by the polarizations of the gravitational wave. The axes of h_+ and h_\times polarizations differ by $\frac{\pi}{4}$ radian (45 degrees).

relative positions of the test masses which are initially placed on a circle are distorted to an ellipse whose major axes are the y - and z -axes. This situation is illustrated in Fig. 2.1 and Fig. 2.2.

The other polarization component h_\times has a similar effect but with major axes which are $\frac{\pi}{4}$ radian (45 degrees) different from those of the h_+ polarization. The effects of h_+ and h_\times polarizations are schematically shown in Fig. 2.2. These two polarizations are independent of each other and any gravitational wave can be decomposed into these two. There is another possibility of choosing two independent polarizations as ‘circular’ ones (see [6]).

In general cases, the polarization of the gravitational wave will be a mixture of the two independent (either linear or circular) polarizations. Furthermore, the propagation axis of the gravitational wave is not necessarily perpendicular to the plane defined by the test masses. Thus, the amount of the effect due to the gravitational wave in general depends on the composition of polarizations and the relative angle between the propagation axis of a gravitational wave and the plane defined by the test masses. In the following discussion, however, such effects will be ignored and the optimum case is assumed for simplicity.

2.1.3: Optical measurement of gravitational wave effects

Let us consider the effect of a gravitational wave on the optical measurement in more detail. The phase change incurred during the return trip along the y - or z -axis is given by

$$\phi(t) = \omega_0 t_r \simeq \frac{2\omega_0 \ell}{c} \pm \frac{\omega_0}{2} \int_{t-2\ell/c}^t h_+(t) dt, \quad (15)$$

where ω_0 is the angular frequency of the light. Since the first term in the above expression is constant and does not depend on the gravitational wave, we will consider the second term, the *phase shift* $\delta\phi$ from the constant phase change, as the effect due to the gravitational wave.

Assuming a sinusoidal gravitational wave with an angular frequency ω_g and a peak amplitude h_0

$$h_+(t) = h_0 \cos(\omega_g t), \quad (16)$$

then eq. (15) can be rewritten as

$$\begin{aligned} \delta\phi(x) &\simeq \frac{\omega_0}{2} \int_{t-2\ell/c}^t h_0 \cos(\omega_g t) dt \quad (\text{for } |h_0| \ll 1) \\ &\simeq \frac{h_0}{2} \frac{\omega_0}{\omega_g} \left\{ \sin \omega_g t - \sin[\omega_g(t-2\ell/c)] \right\} \\ &\simeq h_0 \frac{\omega_0 \sin(\omega_g \ell/c)}{\omega_g} \cos[\omega_g(t-\ell/c)]. \end{aligned} \quad (17)$$

Thus, a gravitational wave modulates the phase of the light at its frequency. The modulation index is mainly determined by the light frequency and the distance between the two test masses. Since the expected amplitudes of gravitational waves are so faint (typically $\sim 10^{-21}$), this modulation index is extremely small ($\sim 10^{-9}$ radian) in spite of the huge ratio of optical and gravitational wave frequencies ($\omega_0/\omega_g \sim 10^{12}$). This justifies the use of an approximation of a weak phase modulation (see 2.2.1).

In more general cases, any gravitational wave can be considered as a mixture of many different Fourier components. Each Fourier component independently has the effect described above on the phase of the light, unless non-linear effects become significant. The expected gravitational waves are so faint, as mentioned above, that non-linearity can be ignored. Thus, the effect from a single Fourier component is mainly considered in the following discussion.

The modulation index due to a gravitational wave $h_0(\omega_0/\omega_g) \sin(\omega_g \ell/c) \simeq h_0 \omega_0 \ell/c$ is almost proportional to the (initial) distance ℓ between the two test

masses, as long as it is shorter than half of the wavelength of the gravitational wave, i.e. $\omega_g \ell / c \lesssim \pi/2$. Thus, a long distance is desirable, e.g. ~ 75 km for a gravitational wave with the frequency of 1 kHz, and even longer ones for those with lower frequencies. This is, however, an unrealistic value for ground-based experiments. Thus, several methods to achieve a similar effect within a manageable size have been proposed. They are thoroughly discussed in Chapters 3 and 4.

Although eq. (17) was derived for a single-bounce measurement, the situations are similar in more complicated configurations, although the apparent effects may not necessarily be pure phase modulation. The remainder of this chapter treats another problem of how to detect this phase shift of the light, including such non-standard modulations.

2.2: Modulation, sidebands, and their detection

2.2.1: Amplitude/phase modulation

Before discussing the detection of an optical phase, it is helpful to consider modulations in a general sense. Thus, in this section the general theory of modulation and its detection are discussed, and its application to optical problems is separately treated in the following section.

In general, any oscillations of an angular frequency ω_0 is characterized by its (real) amplitude A and the phase ϕ as

$$A \cos(\omega_0 t + \phi) = \Re \left\{ A e^{i(\omega_0 t + \phi)} \right\} = \Re \left\{ a e^{i\omega_0 t} \right\}. \quad (18)$$

Here $a = A e^{i\phi}$ is the *complex amplitude* which represents the real amplitude and the phase unified in a complex number. This corresponds to regarding the oscillation as the projection of a rotation to the real axis. We use this complex amplitude to express the oscillation in the left-hand-side of eq. (18) and take its real part ($\Re\{\dots\}$) whenever it is necessary.

It is also useful to visualize this complex amplitude as a vector in a complex plane; this is what is called the *phasor diagram*. In a phasor diagram, a summation (subtraction) of complex amplitudes can be reduced to a composition (decomposition) of corresponding vectors. This is especially useful in understanding the behavior of light, as we will see later.

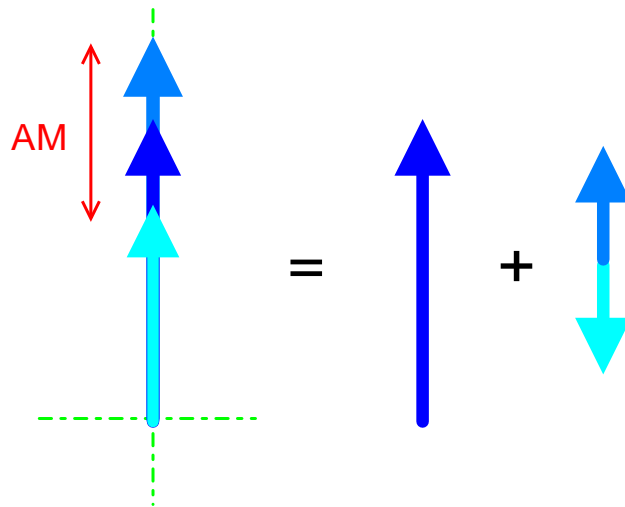


Figure 2.3: The vector which represents the amplitude modulation (AM) can be separated into two parts, the ‘static’ part and the small vector with the ‘swinging’ length. The former represents the carrier (on which our viewpoint is fixed) and the latter represents the sum of the upper and lower sidebands (see Fig. 2.4). In the amplitude modulation case, these two are parallel to each other.

An oscillation with an sinusoidal amplitude modulation of an index m and a modulation (angular) frequency ω_m is represented as

$$\begin{aligned}
 a_{\text{AM}} e^{i\omega_0 t} &= a_0 [1 + m \cos(\omega_m t)] e^{i\omega_0 t} \\
 &= a_0 \left[1 + \frac{m}{2} (e^{i\omega_m t} + e^{-i\omega_m t}) \right] e^{i\omega_0 t} \\
 &= a_0 \left[e^{i\omega_0 t} + \frac{m}{2} e^{i(\omega_0 + \omega_m)t} + \frac{m}{2} e^{i(\omega_0 - \omega_m)t} \right].
 \end{aligned} \tag{19}$$

It is seen that there are three components, each oscillating at a different frequency. In the last expression, the first term represents the *carrier*, the component with the original frequency without the modulation. The second and the third terms, called the *upper* and *lower sidebands* respectively, represent the components induced by the modulation. Each sideband has the frequency shifted up or down from that of the carrier by the modulation frequency. Both sidebands have the same amplitude which is determined by that of the carrier and the index of the modulation.

In a phasor diagram, we usually fix our ‘eyes’ (or the viewpoint) to the carrier frequency so that the vector that represents the carrier looks stationary (i.e. not rotating). Other components with different frequencies, such as the upper or lower sideband, rotate at the frequency which correspond to the differences from that of the carrier.

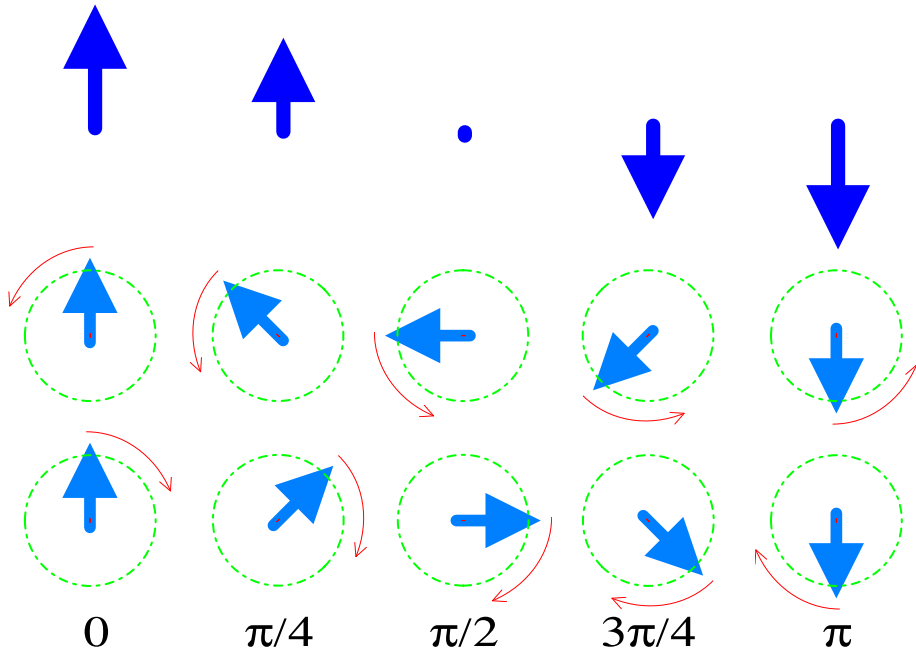


Figure 2.4: The swinging vector can be further decomposed into two oppositely rotating vectors. In the figure, the top line represents the summed vector whose length varies with time; the second and the third line show the oppositely rotating two vectors, which represent the upper and lower sidebands respectively; and the bottom line indicates the phase of rotation ($= \omega_m t$). This is simply a visualization of the famous equation $2 \cos \phi = e^{i\phi} + e^{-i\phi}$. Note that the upper and lower sidebands have the same length.

The oscillation with an amplitude modulation will be expressed by a vector with varying length at the modulation frequency, as is shown in Fig. 2.3. This can be decomposed into two parts, a static vector which represents the carrier and a vector with the ‘swinging length²⁾’ which is the sum of the two rotating sidebands.

The latter may require a little more explanation. A pair of vectors of the same length rotating at the same frequency but in opposite directions, like the upper and lower sidebands here, compose a vector with the swinging length when they are summed together (see Fig. 2.4). There are two chances in a rotating period for the two rotating vectors to coincide, e.g. $\omega_m t = 0$ and π (mod 2π). At these moments the summed vector has the maximum length of the swing at the two opposite directions. During the rest of the period, the

²⁾ Here ‘swinging length’ means that the length varies periodically from positive to negative through zero.

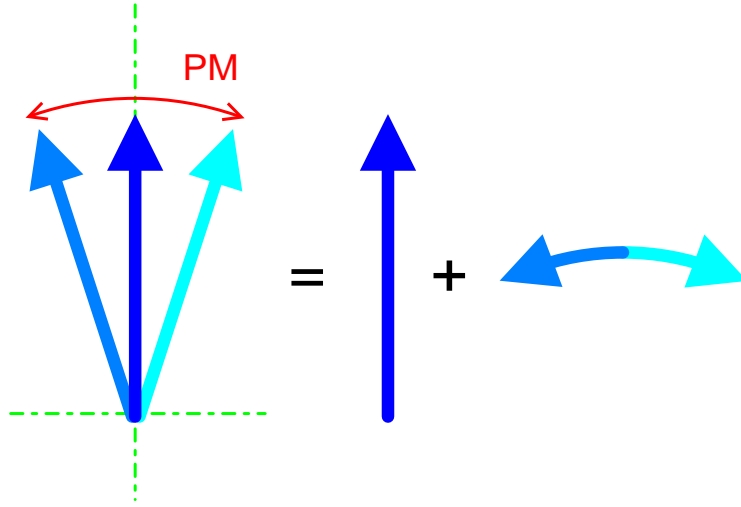


Figure 2.5: In a similar way to the case of an amplitude modulation, the vector which represents the phase modulation (PM) can be separated into two parts, the ‘static’ part and the small vector with the ‘swinging’ length. The difference from the amplitude modulation case is the relative angle between the swinging vector and the static vector; in the amplitude modulation case they are parallel to each other, whereas they are (almost) orthogonal in the phase modulation case. Exactly speaking, the small vector changes its relative angle in the phase modulation case. This can be ignored, however, if the maximum length is small enough compared with the static vector, which corresponds to a weak modulation.

summed vector will be on the same axis as the maxima, with a sinusoidally swinging length.

It should be noted that the sum of a pair of rotating vectors with the same amplitude but rotating in opposite directions has a definite axis. In this case the axis coincides with the real axis, but this actually depends on the relative phase between the two rotating vectors, as we will see shortly.

A complex amplitude with a phase modulation of index m and modulation (angular) frequency ω_m is given by

$$\begin{aligned}
 a_{\text{PM}} e^{i\omega_0 t} &= a_0 e^{im \cos(\omega_m t)} e^{i\omega_0 t} \\
 &= a_0 e^{i\omega_0 t} \left[J_0(m) + \sum_{l=1}^{\infty} i^l J_l(m) (e^{il\omega_m t} + e^{-il\omega_m t}) \right] \\
 &= a_0 e^{i\omega_0 t} \left[J_0(m) + iJ_1(m) (e^{i\omega_m t} + e^{-i\omega_m t}) + O(m^2) \right] \\
 &\simeq a_0 \left[e^{i\omega_0 t} + i \frac{m}{2} e^{i(\omega_0 + \omega_m)t} + i \frac{m}{2} e^{i(\omega_0 - \omega_m)t} \right] \quad (\text{for } |m| \ll 1).
 \end{aligned} \tag{20}$$

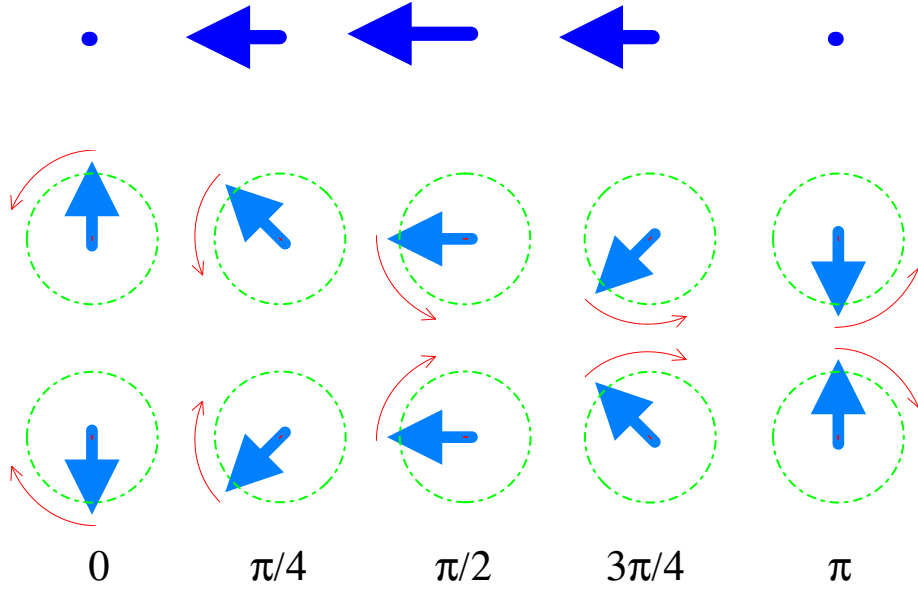


Figure 2.6: The same as Fig. 2.4, but with a different relative phase between the upper and lower sidebands. This corresponds to $2i \sin \phi = e^{i\phi} - e^{-i\phi}$, i.e. the phase between the two rotating vectors differ by $\angle(-1)=\pi$ from the previous case. Another difference is that the axis of the swinging vector is $\frac{\pi}{2}$ different from the previous case, as is indicated by the ‘ i ’ in the equation.

Here $J_l(m)$ are the Bessel functions of the first kind, defined by

$$\begin{aligned}
 J_l(\xi) &= \sum_{j=0}^{\infty} \frac{(-1)^j}{j!(j+l)!} \left(\frac{\xi}{2}\right)^{l+2j} & J_0(\xi) &= 1 - \left(\frac{\xi}{2}\right)^2 + \frac{1}{4}\left(\frac{\xi}{2}\right)^4 - \frac{1}{36}\left(\frac{\xi}{2}\right)^6 + \dots \\
 &\simeq \frac{1}{l!} \left(\frac{\xi}{2}\right)^l + o(\xi^l). & J_1(\xi) &= \left(\frac{\xi}{2}\right) - \frac{1}{2}\left(\frac{\xi}{2}\right)^3 + \frac{1}{12}\left(\frac{\xi}{2}\right)^5 - \dots \\
 & & & \vdots
 \end{aligned} \tag{21}$$

In the following discussion, we will assume a weak modulation ($|m| \ll 1$) unless otherwise noted. In this case, the only terms to be considered are the zero-th and the first order ones.

Under this approximation, one may notice that the last expression in eq. (20) looks quite similar to the one for the amplitude modulation eq. (19). In both cases there are components for the carrier and for the upper and lower sidebands. What is different is the ‘ i ’ in front of the sidebands’ terms, which represents the relative angle of $\angle i = \frac{\pi}{2}$ radian between the carrier and the sidebands’ axis.

This difference becomes clearer when one considers the phasor diagram. The oscillation with a weak phase modulation is expressed in a phasor diagram by the vector varying its angle at the modulation frequency, as is shown in Fig. 2.5. This can be decomposed into a static vector and a swinging component, as in the case of amplitude modulation. The relative angle between the static vector and the swinging one is, however, different from the previous case—they are orthogonal to each other. This is the effect of ‘ i ’ in eq. (20), and is illustrated in Fig. 2.6

The similarity between amplitude and weak phase modulations may be noticed; both are composed of a static vector and a swinging (small) vector, each represents the carrier and the sum of the two sidebands, respectively. In either case, the swinging vector can be further decomposed into the two oppositely rotating vectors which represent the upper and lower sidebands. The only difference between amplitude and phase modulations is the relative angle between the static vector and the swinging one.

2.2.2: Suppressed carrier and single sideband

There are some variations from amplitude and phase modulations that are worth mentioning here. From the above discussion, one may notice that the information is actually contained not in the ‘carrier’ but in the ‘sidebands’. Most of the energy of the oscillation is, however, present in the carrier component. In other words, the energy used to keep the information is only a small fraction of the total energy of the oscillation.

This leads one to the idea of using the sidebands without any carrier, to make the oscillation more efficient. This is what is called *suppressed carrier*. In this case only the sidebands induced by the modulation are used and the carrier frequency component is eliminated. By using this, the total energy required to keep the same amount of information can be reduced by a significant factor.

Another important feature is that the difference between the amplitude and the weak phase modulations will disappear when the carrier is eliminated. The difference is the relative angle between the carrier and the swinging components. If the carrier is eliminated, there is no way to distinguish the two cases—both are composed of only a component with the swinging length. Thus we do not distinguish ‘suppressed-carrier amplitude modulation’ and ‘suppressed-carrier (weak) phase modulation’, and both are referred to as *suppressed-carrier double sideband*, or simply suppressed carrier.

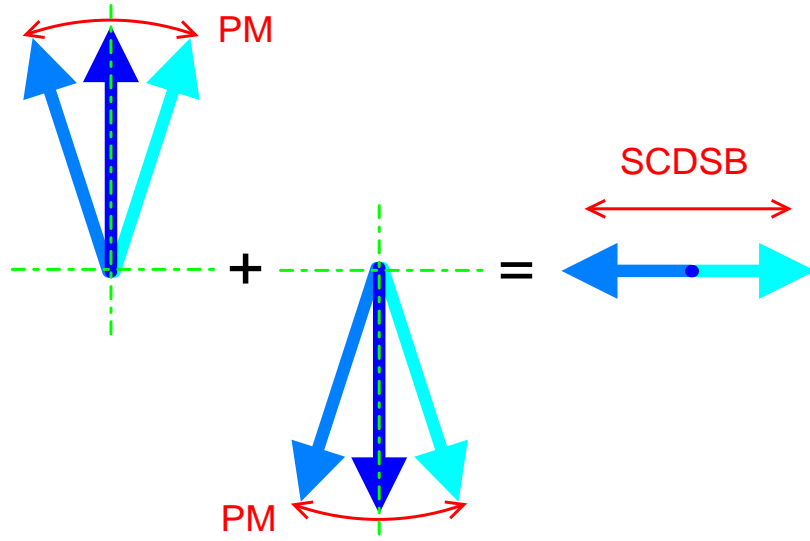


Figure 2.7: A phasor diagram of a balanced modulator which produces suppressed-carrier double sideband (SCDSB). By subtracting two oppositely modulated oscillations (here the case of phase modulation (PM) is shown), only the modulation-induced sidebands can be obtained.

The reason for the name ‘double sideband’ is that both the upper and lower sidebands remained. One may notice, however, from eqs. (19) and (20), that they are symmetric and contain in principle the same information. Thus, it may be possible to eliminate one of them, without losing any information. This is referred to as *suppressed-carrier single sideband* or simply single sideband.

The question left is how to ‘suppress’ the carrier. One of the ways to realize this is the use of a *balanced modulator*. This uses two modulators, each modulating the identical carrier but in opposite polarity, and then subtract the two modulated oscillations (see Fig. 2.7). As a result, the carrier components in each oscillation cancel out and the sidebands add up due to the opposite polarity of their modulations:

$$\begin{aligned}
 & a_0 \left\{ e^{i\omega_0 t} + \epsilon \frac{m}{2} [e^{i(\omega_0 + \omega_m)t} + e^{i(\omega_0 - \omega_m)t}] \right\} \\
 & - a_0 \left\{ e^{i\omega_0 t} - \epsilon \frac{m}{2} [e^{i(\omega_0 + \omega_m)t} + e^{i(\omega_0 - \omega_m)t}] \right\} \\
 & = \epsilon m a_0 [e^{i(\omega_0 + \omega_m)t} + e^{i(\omega_0 - \omega_m)t}] = 2\epsilon m \cos(\omega_m t) a_0 e^{i\omega_0 t} \quad (22)
 \end{aligned}$$

where $\begin{cases} \epsilon = 1 = e^0 & \text{for amplitude modulation} \\ \epsilon = i = e^{i\pi/2} & \text{for (weak) phase modulation.} \end{cases}$

Although ϵ represents the orientation of the swinging vector, it has little meaning unless there is another vector to measure the relative angle.

To produce a single sideband, one must use a filter since in this case the component to be eliminated has a finite bandwidth, which makes the cancellation much more difficult than in the single Fourier component case. In addition, the filter must be relatively sharp in order to select only one of the sidebands.

2.2.3: Heterodyne and homodyne techniques

In the field of electrical communications, *heterodyne technique* and its variant *homodyne technique* are commonly used to handle modulated signals. This can be used for either amplitude or phase modulation, or even suppressed carrier or single sideband. In the following discussion, however, we will assume an amplitude or phase modulation to concentrate on its principle.

Let's assume a modulated oscillation $A_c(t) \exp\{i[\omega_c t + \phi(t)]\}$, where $A_c(t) = A_0[1 + m(t)]$, $m(t)$ and $\phi(t)$ represent the amplitude and phase modulations respectively, and ω_c is the carrier frequency. This is multiplied by a *local oscillator* $A_{LO} \exp\{i[\omega_{LO} t + \psi]\}$ where ψ represents the relative phase between the modulated oscillation and the local oscillator:

$$\begin{aligned} & \Re\left\{A_c(t)e^{i[\omega_c t + \phi(t)]}\right\} \times \Re\left\{A_{LO}e^{i[\omega_{LO} t + \psi]}\right\} \\ &= A_c(t) \cos\{\omega_c t + \phi(t)\} \times A_{LO} \cos\{\omega_{LO} t + \psi\} \\ &= \frac{A_c(t)A_{LO}}{2} \left\{ \cos[(\omega_c - \omega_{LO})t + \phi(t) - \psi] + \cos[(\omega_c + \omega_{LO})t + \phi(t) + \psi] \right\}. \end{aligned} \quad (23)$$

Here the real part of each complex amplitude is taken because of the multiplication process. As a result, two modulated oscillations with the frequencies of $\omega_c \pm \omega_{LO}$ show up.

In eq. (23), both components have the same amplitude and phase information as the 'old' (original) oscillation before being multiplied. If these two components can be separated from each other, each component represents an oscillation with the same modulation as that before being multiplied, but with a different carrier frequency. The new carrier frequency will be either the sum or the difference between the frequencies of the old carrier and the local oscillator. This is what is called the heterodyne technique.

If the local oscillator has the same frequency as the old carrier ($\omega_{LO} = \omega_c$), the difference of these two frequencies will be zero hertz, i.e. the one of the new

carriers will be DC. In this special case called the homodyne technique, the first term in the above equation will be

$$\begin{aligned} \frac{A_c(t)A_{LO}}{2} \cos[\phi(t) - \psi] &= \frac{A_c(t)A_{LO}}{2} [\cos \psi \cos \phi(t) + \sin \psi \sin \phi(t)] \\ &\simeq \begin{cases} \pm A_c(t)A_{LO}/2 & \text{for } |\phi(t) - \psi| \simeq 0 \text{ or } \pi \\ \pm [A_c(t)A_{LO}/2] \sin \phi(t) & \text{for } \psi \simeq \pm \frac{\pi}{2}. \end{cases} \end{aligned} \quad (24)$$

Thus the signals in phase or amplitude modulation will show up at DC by choosing an appropriate relative phase ψ .

This difference in the appropriate relative phase can be understood if one recalls the phasor diagrams of amplitude and phase modulated oscillations. The relative angle between the carrier and the swinging vectors is parallel in the amplitude modulation case and orthogonal in the phase modulation case. The local oscillator to ‘detect’ the signal must be parallel to the swinging vector and thus its angle relative to the old carrier must differ by $\frac{\pi}{2}$ between the two cases.

There are several advantages in using the heterodyne (or homodyne) technique. One of them is that by choosing an appropriate frequency for the local oscillator, the new carrier can be an intermediate frequency which is convenient for filtering and, moreover, to avoid the noise in some frequency ranges. Some noise sources tend to have less significance at higher frequencies, so one can expect less disturbance from them than at lower frequencies.

Another point to be mentioned is that the amplitude of the output signal is proportional to the amplitude of the local oscillator. Thus, the use of heterodyne (or homodyne) technique not only converts the frequency range but also amplifies the signal amplitude. This is of advantage in trying to avoid the effects from extra noise sources that are independent of the amplitude of the local oscillator.

Note, however, that these advantages do not mean an improvement of the intrinsic signal-to-noise ratio of the original signal. The noise inherent in the original signal will be treated in exactly the same way as the signal. Thus, there will be no difference in the best achievable signal-to-noise ratio. Heterodyne (homodyne) techniques are useful only to avoid the reduction of the signal-to-noise ratio due to ‘extra’ noise sources.

2.3: Optical phase detection

2.3.1: Photo-detection

Although light is a sort of electro-magnetic radiation, there are big differences from radiation of lower frequency, e.g. radio-frequency. One of them is that the electro-magnetic field of light oscillates at an optical frequency (~ 500 THz). There is no detector which can reproduce an electrical signal proportional to the *field amplitude*, whereas for radio-frequency radiation the current in an antenna is proportional to the field amplitude.

What we detect optically is the photo-current in a photo-detector, which is proportional to the *power* averaged over a relatively short time period (~ 1 ns). The phase information of the field amplitude will be lost in this averaging process. Thus, in order to obtain the phase information of a field oscillating at an optical frequency (this is what is referred to as the ‘phase of light’), it must be converted to some other form that can be detected, i.e. either the phase or the amplitude of the photo-current.

To represent the electro-magnetic field of a light beam, it has become customary to use a complex amplitude, whose real part is proportional to the electric field amplitude. Although an actual beam has a finite cross-section, the complex amplitude of the beam is usually defined in a sectionally-integrated form so that it will give the total power carried by the beam when it is squared. Thus, the ‘instantaneous’ power carried by the beam is proportional to

$$\Re\{ae^{i\omega_0 t}\} \times \Re\{ae^{i\omega_0 t}\} = |a|^2 \cos^2(\omega_0 t + \phi) = |a|^2 \frac{1 + \cos(2\omega_0 t + 2\phi)}{2}, \quad (25)$$

which contains a DC term and a term oscillating at twice the optical frequency.

In practice, however, what a photo-detector detects is the *averaged* light power, as mentioned above. Then the oscillating term in eq.(25) disappears and the total (averaged) power of the oscillation is given by

$$P = |a|^2/Z_0 = \frac{a \cdot a^*}{Z_0}, \quad (26)$$

i.e. by squaring the absolute value of the complex amplitude. Here $1/Z_0$ is the proportional constant and $Z_0 = \sqrt{\mu_0/\epsilon_0} \approx 376.7 \Omega$ is the characteristic impedance of the vacuum. For a more precise discussion on the definition of a complex amplitude, see Appendix G.

Eq.(26) shows another advantage of using complex amplitudes: the averaged power detected by a photo-detector can be obtained only by multiplying

the complex amplitude by its complex conjugate. Since photo-detectors do not respond to optical frequencies, the main oscillating part $e^{i\omega_0 t}$ can be completely neglected. This makes the calculation much easier than that using the sinusoidal expression of oscillations.

This also means that in a phasor diagram what is of interest is the length of the vector, whose square is proportional to the photo-current. The orientation of the vector, which represents the phase of the light, will have no significance in the photo-current.

Let us consider the detection of the amplitude or phase modulated optical field³⁾. As mentioned above, only the complex amplitude needs to be considered.

$$\begin{aligned} PZ_0 &= |a_0 [1 + m(t)] e^{i\phi(t)}|^2 = |a_0|^2 [1 + 2m(t) + m^2(t)] \\ &\simeq |a_0|^2 [1 + 2m(t)] \quad (\text{for } |m(t)| \ll 1). \end{aligned} \quad (27)$$

Here $m(t)$ and $\phi(t)$ represent the amplitude and phase modulations, respectively. From the last expression it can be seen that the amplitude modulation signal can be detected directly but the phase modulation signal cannot.

This direct detection process, known as the *square-law detection*, converts the amplitude modulation signal on light to the photo-current at DC range. This can be understood as the multiplication in the photo-detection process working as homodyne with zero relative phase. More precisely, the carrier component of a in eq. (26) acts as the local oscillator to detect the sideband components of a^* , and vice versa. Since these two processes are symmetric, we usually need to consider only one of them.

In this direct detection process, the phase modulation signal cannot be detected since the local oscillator is orthogonal to the sum of the two sidebands. To understand this clearly, again the use of a phasor diagram is helpful. As mentioned above, what is detected by a photo-detector is the averaged power which is proportional to the square of the length of the vector. In an amplitude modulation case, the length of the vector varies by the modulation, whereas only the orientation of the vector varies and not its length in a phase modulation case. Thus, the latter does not cause any change in the photo-current.

In order to apply heterodyne or homodyne techniques to a wider variety of circumstances, it is necessary to multiply the oscillation by a local oscillator

³⁾ We will consider the modulations on the field amplitude, not on the power, since these are what are expected as the effect from gravitational waves. One must not confuse this with modulations on the light power, which are sometimes discussed in the literature (e.g. Ref. [16]).

with an arbitrary frequency and in an arbitrary relative phase. This can be realized by *adding* the local oscillator to the oscillation before its detection. The detected power in this case is given by

$$\begin{aligned}
 PZ_0 &= |a_1 + a_2|^2 = (a_1 + a_2) \times (a_1 + a_2)^* \\
 &= |a_1|^2 + |a_2|^2 + a_1 a_2^* + a_1^* a_2 = |a_1|^2 + |a_2|^2 + 2\Re\{a_1 a_2^*\} \\
 &= |a_1|^2 + |a_2|^2 + 2|a_1 a_2| \cos(\angle a_1 - \angle a_2).
 \end{aligned} \tag{28}$$

In the last expression, the first and the second term represent the averaged power of the oscillation and the local oscillator respectively, and thus do not contain the phase information.

On the other hand, the third *interfering term* is the linear product of the two amplitudes and it keeps the phase information. This can be detected by choosing an appropriate relative phase. Note, however, that what is multiplied by a_1 is not a_2 but its complex conjugate a_2^* . In addition, $a_1 a_2^*$ is proportional to the complex amplitude of the detected photo-current.

We have so far considered only the spatially-integrated complex amplitude. In practice, however, the electro-magnetic field of the light is distributed within a finite cross-section. Thus, it is necessary to superimpose the two field amplitudes properly, i.e. both beams must have the same spatial distribution such as the radius, the divergence, the propagating direction, and the spatial (TEM) mode.

If there are differences in such spatial features between the two optical fields, only the components coinciding with each other interfere and the rest does not. As a result, the interfering term will be smaller than that given in eq. (28):

$$\begin{aligned}
 PZ_0 &= |a_1|^2 + |a_2|^2 + 2\mathcal{C}_0 \Re\{a_1 a_2^*\} \\
 &= |a_1|^2 + |a_2|^2 + 2\mathcal{C}_0 |a_1 a_2| \cos(\angle a_1 - \angle a_2),
 \end{aligned} \tag{29}$$

where $\mathcal{C}_0 \leq 1$ represents the finite contrast⁴⁾ of the interference. Note that this should not be confused with the *apparent visibility* \mathcal{C} given by

$$\begin{aligned}
 \mathcal{C} &= \frac{P_{\max} - P_{\min}}{P_{\max} + P_{\min}} = \frac{2\mathcal{C}_0 |a_1 a_2|}{|a_1|^2 + |a_2|^2} = \mathcal{C}_0 \mathcal{C}_{\max} \\
 \text{here } \mathcal{C}_{\max} &= \frac{2|a_1 a_2|}{|a_1|^2 + |a_2|^2} \quad : \text{ the ideal visibility.}
 \end{aligned} \tag{30}$$

⁴⁾ The words ‘contrast’ and ‘visibility’ are often used for the same meaning. In this paper, the latter is used for ‘the apparent visibility’ and the former for ‘the reduction from the ideal visibility \mathcal{C}_{\max} which is determined by the amplitude difference between the two interfering beams’.

If the two (real) amplitudes are not equal, the apparent visibility is limited by the ideal visibility \mathcal{C}_{\max} .

2.3.2: Local oscillator light

There is a question of how to superimpose, or ‘add’, the local oscillator. This can be done using a beamsplitter with a finite reflectivity ρ and transmittance τ ($\tau, \rho \leq 1$). From the conservation of power, reflectivity and transmittance of a beamsplitter must satisfy the relation

$$\rho^2 + \tau^2 = 1 - \mathcal{A}, \quad (31)$$

where \mathcal{A} is the power loss at the beam-splitting surface. In the following discussion the loss is assumed to be small ($\mathcal{A} \ll 1$).

A beamsplitter has two inputs and two outputs, and both of the two outputs a_{\oplus} and a_{\ominus} contain a fraction of each input (see Fig. 2.8).

$$\begin{cases} a_{\oplus} = \rho a_1 + i\tau a_2 = \rho \cdot a_1 + \tau \cdot ia_2, \\ a_{\ominus} = i\tau a_1 + \rho a_2 = i[\tau \cdot a_1 - \rho \cdot ia_2], \end{cases} \quad (32)$$

where a_1 and a_2 are the field amplitudes at the two inputs of the beamsplitter, and $i = e^{i\pi/2}$ in front of τ stands for the phase difference between the transmitted and reflected light⁵⁾. As can be seen in eq. (32), we can obtain the sum and the difference of the two amplitudes a_1 and ia_2 , with an additional weighting of τ or ρ . The two outputs of a beamsplitter are different in the relative phase between the two components by $\angle(-1) = \pi$, due to the phase shift at the beam-splitting surface.

The last question in applying heterodyne or homodyne techniques to an optical frequency is the source of the local oscillator. Since the bandwidth of a photo-detector is no more than ~ 1 GHz, the local oscillator must have a frequency which differs from that of the carrier by less than 1 ppm ($\Delta\omega/\omega_0 \ll 10^{-6}$). An absolute stability within this tolerance is difficult to realize for any source of light.

In the case of homodyne, this problem can be solved by using, for the local oscillator, light picked off from the same source as the carrier. In this case both

⁵⁾ Actually, both transmitted and reflected light have phase shifts relative to the input light at the beam-splitting surface. In our case, however, only the difference between the two is relevant, and thus the phase shift relative to the input light is ignored. Then, as long as the power loss is small enough, the relative phase between the two is always $\frac{\pi}{2}$ radian.

with an arbitrary frequency and in an arbitrary relative phase. This can be realized by *adding* the local oscillator to the oscillation before its detection. The detected power in this case is given by

$$\begin{aligned}
 PZ_0 &= |a_1 + a_2|^2 = (a_1 + a_2) \times (a_1 + a_2)^* \\
 &= |a_1|^2 + |a_2|^2 + a_1 a_2^* + a_1^* a_2 = |a_1|^2 + |a_2|^2 + 2\Re\{a_1 a_2^*\} \\
 &= |a_1|^2 + |a_2|^2 + 2|a_1 a_2| \cos(\angle a_1 - \angle a_2).
 \end{aligned} \tag{28}$$

In the last expression, the first and the second term represent the averaged power of the oscillation and the local oscillator respectively, and thus do not contain the phase information.

On the other hand, the third *interfering term* is the linear product of the two amplitudes and it keeps the phase information. This can be detected by choosing an appropriate relative phase. Note, however, that what is multiplied by a_1 is not a_2 but its complex conjugate a_2^* . In addition, $a_1 a_2^*$ is proportional to the complex amplitude of the detected photo-current.

We have so far considered only the spatially-integrated complex amplitude. In practice, however, the electro-magnetic field of the light is distributed within a finite cross-section. Thus, it is necessary to superimpose the two field amplitudes properly, i.e. both beams must have the same spatial distribution such as the radius, the divergence, the propagating direction, and the spatial (TEM) mode.

If there are differences in such spatial features between the two optical fields, only the components coinciding with each other interfere and the rest does not. As a result, the interfering term will be smaller than that given in eq. (28):

$$\begin{aligned}
 PZ_0 &= |a_1|^2 + |a_2|^2 + 2\mathcal{C}_0 \Re\{a_1 a_2^*\} \\
 &= |a_1|^2 + |a_2|^2 + 2\mathcal{C}_0 |a_1 a_2| \cos(\angle a_1 - \angle a_2),
 \end{aligned} \tag{29}$$

where $\mathcal{C}_0 \leq 1$ represents the finite contrast⁴⁾ of the interference. Note that this should not be confused with the *apparent visibility* \mathcal{C} given by

$$\begin{aligned}
 \mathcal{C} &= \frac{P_{\max} - P_{\min}}{P_{\max} + P_{\min}} = \frac{2\mathcal{C}_0 |a_1 a_2|}{|a_1|^2 + |a_2|^2} = \mathcal{C}_0 \mathcal{C}_{\max} \\
 \text{here } \mathcal{C}_{\max} &= \frac{2|a_1 a_2|}{|a_1|^2 + |a_2|^2} \quad : \text{ the ideal visibility.}
 \end{aligned} \tag{30}$$

⁴⁾ The words ‘contrast’ and ‘visibility’ are often used for the same meaning. In this paper, the latter is used for ‘the apparent visibility’ and the former for ‘the reduction from the ideal visibility \mathcal{C}_{\max} which is determined by the amplitude difference between the two interfering beams’.

If the two (real) amplitudes are not equal, the apparent visibility is limited by the ideal visibility \mathcal{C}_{\max} .

2.3.2: Local oscillator light

There is a question of how to superimpose, or ‘add’, the local oscillator. This can be done using a beamsplitter with a finite reflectivity ρ and transmittance τ ($\tau, \rho \leq 1$). From the conservation of power, reflectivity and transmittance of a beamsplitter must satisfy the relation

$$\rho^2 + \tau^2 = 1 - \mathcal{A}, \quad (31)$$

where \mathcal{A} is the power loss at the beam-splitting surface. In the following discussion the loss is assumed to be small ($\mathcal{A} \ll 1$).

A beamsplitter has two inputs and two outputs, and both of the two outputs a_{\oplus} and a_{\ominus} contain a fraction of each input (see Fig. 2.8).

$$\begin{cases} a_{\oplus} = \rho a_1 + i\tau a_2 = \rho \cdot a_1 + \tau \cdot ia_2, \\ a_{\ominus} = i\tau a_1 + \rho a_2 = i[\tau \cdot a_1 - \rho \cdot ia_2], \end{cases} \quad (32)$$

where a_1 and a_2 are the field amplitudes at the two inputs of the beamsplitter, and $i = e^{i\pi/2}$ in front of τ stands for the phase difference between the transmitted and reflected light⁵⁾. As can be seen in eq. (32), we can obtain the sum and the difference of the two amplitudes a_1 and ia_2 , with an additional weighting of τ or ρ . The two outputs of a beamsplitter are different in the relative phase between the two components by $\angle(-1) = \pi$, due to the phase shift at the beam-splitting surface.

The last question in applying heterodyne or homodyne techniques to an optical frequency is the source of the local oscillator. Since the bandwidth of a photo-detector is no more than ~ 1 GHz, the local oscillator must have a frequency which differs from that of the carrier by less than 1 ppm ($\Delta\omega/\omega_0 \ll 10^{-6}$). An absolute stability within this tolerance is difficult to realize for any source of light.

In the case of homodyne, this problem can be solved by using, for the local oscillator, light picked off from the same source as the carrier. In this case both

⁵⁾ Actually, both transmitted and reflected light have phase shifts relative to the input light at the beam-splitting surface. In our case, however, only the difference between the two is relevant, and thus the phase shift relative to the input light is ignored. Then, as long as the power loss is small enough, the relative phase between the two is always $\frac{\pi}{2}$ radian.

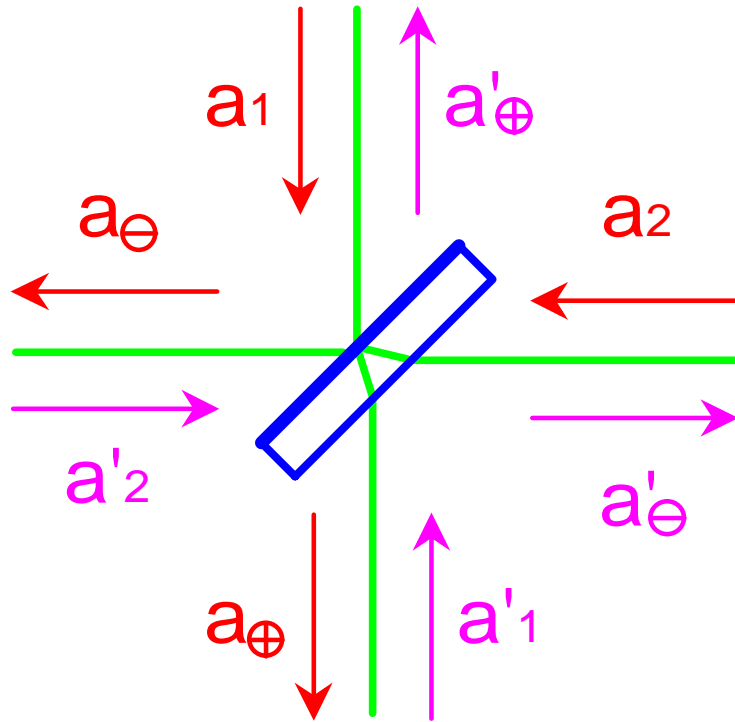


Figure 2.8: A beamsplitter couples the electro-magnetic field in the two different axes. Considering all the directions, there are four inputs and four outputs in total. They can, however, be separated into two groups, which are indicated by a and a' in the figure. Each group satisfies the relations in eq. (32), and does not couple with the other group. Due to the phase shift upon the reflection, the two outputs of a beamsplitter corresponds to the sum and the difference of the two input fields, respectively, with the weightings determined by the reflectivity and transmittance of the beamsplitter.

have the same frequency and the problem associated with coherency is reduced. Furthermore, in realistic situations, there may be problems associated with the fluctuations of the carrier, either in the amplitude or the phase. Using the same source for the local oscillator as the carrier, the common mode rejection effect will reduce these problems.

Considering the above advantages, it may be good if the local oscillator for heterodyne can be obtained from the same light source as the carrier with the frequency altered by a small fraction. This can be realized by using either an acousto-optic or an electro-optic modulator. The former can directly shift the optical frequency and the latter induces an amplitude or a phase modulation whose sidebands can then be used as the local oscillators. In the latter case, both the upper and lower sidebands act as the local oscillators simultaneously. This results in a difference from the standard heterodyne technique which will be further discussed in 2.3.3.

2.3.3: Optical heterodyne

Now let us consider the phase detection of the modulated light. We assume an ideal interference, i.e. $\mathcal{C}_0=1$, for simplicity. By superimposing a local oscillator on the phase modulated light before the photo-detection, the phase information of the light is converted to the interfering term of the photo-current and can be detected:

$$\begin{aligned}
 PZ_0 &= |a_S e^{i\omega_0 t} + a_{LO} e^{i\omega_{LO} t}|^2 \\
 &= |a_S|^2 + |a_{LO}|^2 + 2\Re\left\{a_S a_{LO}^* e^{i(\omega_0 - \omega_{LO})t}\right\} \\
 &= |a_S|^2 + |a_{LO}|^2 + 2|a_S a_{LO}| \cos[(\omega_0 - \omega_{LO})t + \angle a_S - \angle a_{LO}].
 \end{aligned} \tag{33}$$

Here a_S and a_{LO} represent the amplitudes measured *after* the beamsplitter, e.g. $a_S = i\tau a_1$.

If the local oscillator has the same frequency as the carrier $\omega_0 = \omega_{LO}$, this becomes homodyne and the relative phase $\phi \equiv \angle a_S - \angle a_{LO}$ must be chosen appropriately, $\phi \simeq \pm \frac{\pi}{2}$ (see previous section 2.2.3). Then the variation δP in detected power will be proportional to the variation in phase $\delta\phi = \phi \mp \frac{\pi}{2}$

$$\begin{aligned}
 \delta P Z_0 &\simeq \pm 2|a_S a_{LO}| \sin(\delta\phi) && \text{(for } \phi \simeq \pm \frac{\pi}{2}\text{)} \\
 &\simeq \pm 2|a_S a_{LO}| \delta\phi && \text{(for } |\delta\phi| \ll 1\text{)}.
 \end{aligned} \tag{34}$$

On the other hand, if heterodyne is used, i.e. $\omega_0 \neq \omega_{LO}$, the relative phase has no importance since it varies with time. In this case the signal appears as a phase modulation on the detected photo-current which is oscillating at the difference frequency $|\omega_0 - \omega_{LO}|$. The original signal ($\delta\phi$) can be recovered by using an appropriate electric heterodyne technique.

It is worth describing how these situations are visualized in a phasor diagram. As is discussed in 2.2.1, the sum of the two sidebands due to the phase modulation are orthogonal to the carrier. Thus, adding the local oscillator orthogonally to the carrier (represented by the relative phase $\phi \simeq \pm \frac{\pi}{2}$) means it is parallel to the sum of the two sidebands. This effectively results in an ‘amplitude modulation’ on the local oscillator which can be detected.

In the homodyne method, the vector for the local oscillator looks static since it rotates at the same rate as the carrier, to which our ‘eyes’ (viewpoint) are fixed. Thus, there is a definite relative angle between the old carrier and the local oscillator and it must be adjusted to be orthogonal in order to detect the phase modulation signal.

If the heterodyne technique is used, the vector for the local oscillator rotates at the frequency difference between the local oscillator and the old carrier, i.e. their relative angle varies from 0 to 2π radian with time. Thus, the local oscillator becomes parallel to the sum of the two modulation sidebands twice in a period. (This is why the initial orientation is less important.) It is possible to synchronously detect this periodic signal by heterodyning electrically.

As is mentioned in 2.3.2, the modulation-induced sidebands are often used as the local oscillators in the optical phase detection. In either amplitude or phase modulation, both the upper and lower sidebands are produced and they are usually not separated. In this case both act as local oscillators simultaneously, which makes a significant difference from the standard heterodyne technique. Let us consider this case here. To avoid the confusion with signal sidebands and the sidebands used as the local oscillators, we will refer to the latter as upper and lower local oscillators.

If the upper and lower local oscillators are produced either by an amplitude or a phase modulation, both have the same amplitude and the same frequency spacing (in opposite directions) from the carrier frequency. This frequency spacing is the modulation frequency used to produce the local oscillators. In the phasor diagram, the vectors which represent the two local oscillators rotate in opposite directions at this modulation frequency. As is discussed in 2.2.1, adding these two rotating vectors together results in a vector with the ‘swinging’ length. Since this swinging vector varies only its length but not its orientation, it has a definite axis in the phasor diagram.

When this swinging vector is used as the local oscillator, its behavior is actually closer to the homodyne case rather than to the heterodyne case—the swinging vector must be set parallel to the signal sidebands that are to be detected. Since this swinging local oscillator does not change its orientation, its relative angle to the carrier component of the signal must be fixed, as in homodyne case.

For example, if one intends to detect a phase modulation signal, the local oscillator must be set parallel to the sideband due to the phase modulation, i.e. the local oscillator must be orthogonal to the carrier component of the signal. On the other hand, if the local oscillators are produced by a phase modulation, its un-modulated component (this is what may be called the ‘carrier’ of the local oscillator, but we will not use it for this purpose to avoid the possible confusion with the carrier component of the signal) is orthogonal to the swinging local

oscillator (modulation-induced sidebands). Thus, the carrier component of the signal and the un-modulated component of the local oscillator must be parallel to each other.

Let us calculate the signal in this case, assuming a simple case where only a signal light and two local oscillators are present. (In practice, there will be more components in addition to these, but they will not affect the dominant signal. See 2.4.6 for more detail.) When the phase modulated light $a_s e^{i\omega_0 t}$ is added to two local oscillators $a_{LO} e^{i(\omega_0 \pm \omega_a)t}$, the averaged power detected is given by

$$\begin{aligned}
 PZ_0 &= \left| a_s e^{i\omega_0 t} + \frac{1}{2} a_{LO} [e^{i(\omega_0 + \omega_a)t} + e^{i(\omega_0 - \omega_a)t}] \right|^2 \\
 &= \left| a_s e^{i\omega_0 t} + a_{LO} e^{i\omega_0 t} \cos \omega_a t \right|^2 \\
 &= |a_s|^2 + |a_{LO}|^2 [\cos(\omega_a t)]^2 + 2\Re\{a_s a_{LO}^* \cos(\omega_a t)\} \\
 &= |a_s|^2 + \frac{1}{2} |a_{LO}|^2 [1 + \cos(2\omega_a t)] + 2|a_s a_{LO}| \cos(\omega_a t) \cos(\angle a_s - \angle a_{LO}).
 \end{aligned} \tag{35}$$

The second line is similar to the homodyne case but with an additional modulation at ω_a . (This factor represents the ‘swing’ of the local oscillator.) This results in the second term in the bottom line (which represents the power of the local oscillator) modulated at $2\omega_a$, and the last (interfering) term at ω_a .

The signal is contained in the interfering term that is modulated at ω_a . Since the other terms are oscillating at other frequencies ($2\omega_a$ and DC), it is possible to separate the main interfering term from all the others by using appropriate filters:

$$P_{\omega_a} Z_0 = 2|a_s a_{LO}| \cos(\omega_a t) \cos(\angle a_s - \angle a_{LO}). \tag{36}$$

(This can be thought of as the suppressed-carrier double sideband, with the carrier frequency of ω_a , *not* of the field amplitude but of the detected power or equivalently the photo-current.)

To obtain the phase modulation signal, it is necessary to use an electric homodyne technique, i.e. to multiply this by another local oscillator with an angular frequency ω_a . This results in

$$\begin{aligned}
 I_{\text{out}} &\propto A \cos(\omega_a + \psi) \times 2|a_s a_{LO}| \cos \omega_a t \cos(\angle a_s - \angle a_{LO}) \\
 &\propto A |a_s a_{LO}| \cos(\angle a_s - \angle a_{LO}) [\cos \psi + \cos(2\omega_a + \psi)].
 \end{aligned} \tag{37}$$

Again by using an appropriate filter, it is possible to eliminate $2\omega_a$ term in order to reproduce the signal. Since this corresponds to the homodyne technique, the relative phase ψ must be chosen appropriately, i.e. $\psi \simeq 0$ or π .

2.3.4: General double sidebands

We have discussed the detection of well-defined types of modulation like amplitude or phase modulation. In general, it is possible for the upper and lower sidebands to have different amplitude and an arbitrary phase relative to each other. Here, we extend our discussion to such ‘general cases’ for completeness.

As we have seen in 2.2.1, a single Fourier component of a signal is contained in the upper and lower sidebands whose frequencies are shifted up or down from that of the carrier by the signal frequency. In amplitude or phase modulation case, these two sidebands have the same amplitude and their sum has the axis parallel or orthogonal to the carrier.

In general cases, however, two sidebands can have different amplitudes and phases, a_+ and a_- expressed in complex amplitudes. In the phasor diagram the vectors that represent these two sidebands rotate in opposite direction at the signal frequency ω_g , as $a_+e^{i\omega_g t}$ and $a_-e^{-i\omega_g t}$ (we can ignore the carrier frequency in the phasor diagram). This is illustrated in Fig. 2.9.

As in the amplitude or phase modulation case, there are two chances in a rotating period for the two vectors to coincide their orientations, although in this case their lengths may be different.

$$\begin{aligned} \angle(a_+e^{i\omega_g t}) - \angle(a_-e^{-i\omega_g t}) &= \angle a_+ - \angle a_- + 2\omega_g t = 0 \pmod{2\pi}, \\ \text{or equivalently} \quad \omega_g t &= \frac{1}{2}(\angle a_- - \angle a_+) \pmod{\pi}. \end{aligned} \quad (38)$$

Thus, the two chances are actually the opposite directions of an axis whose orientation is given by

$$\psi_{\text{ax}} = \frac{1}{2}(\angle a_- + \angle a_+). \quad (39)$$

The sum of the two vectors will have the maximum length at these moments.

During the rest of the period, however, the summed vector will have different orientation from this axis, differently from the amplitude or phase modulation case. When each vector is orthogonal to the main axis, the two rotating vectors are opposite to each other and their sum is the shortest. In general, the two have different lengths and hence the summed length is not necessarily zero. The locus of the summed vector is actually an ellipse whose major axis is given by eq. (39). The amplitude and phase modulation are special cases of this.

The question is how much of the signal we can detect when the two signal sidebands have different lengths. As is discussed in 2.3.3, a local oscillator detects the sidebands parallel to it. In the heterodyne case, it is not so critical

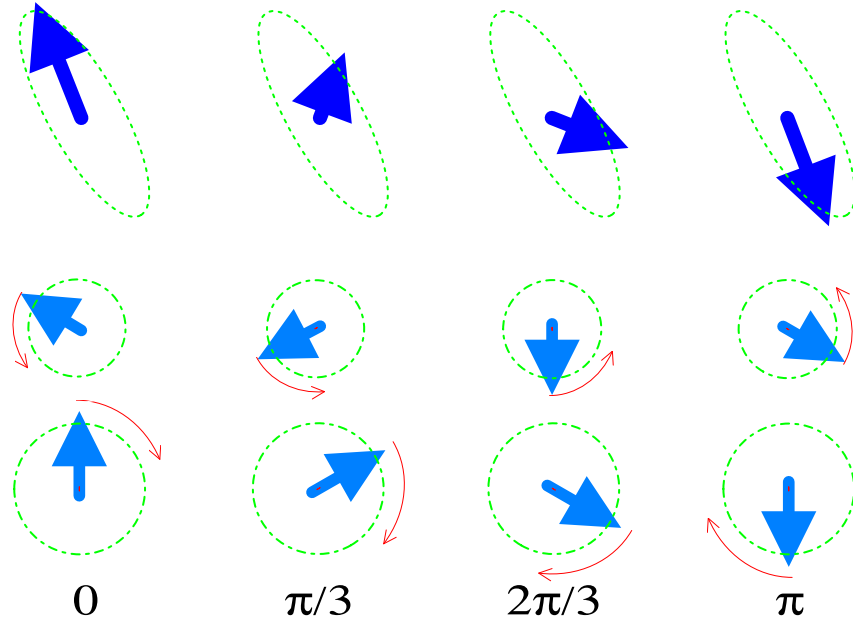


Figure 2.9: In general cases, the upper and lower sidebands have different amplitudes with arbitrary relative phase. Then the locus of the summed vector is an ellipse rather than a straight line (the straight line in the pure amplitude or phase modulation case can be thought of as a special case of an ellipse). The major axis of the ellipse is given by eq. (39) and is not necessarily horizontal or vertical. To maximize the amount of the detected signal, the local oscillator must have the same orientation as this major axis.

since the local oscillator varies the relative angle to the sidebands anyway. Thus, it is expected that there will be chances periodically that the local oscillator becomes parallel to the major axis of the ellipse which represents the locus of the sidebands⁶⁾.

In homodyne and ‘swinging’ local oscillator (see 2.3.3) cases, however, the relative angle between the local oscillator and the signal sidebands is static. Then the projection of the sidebands’ locus to the local oscillator axis is detected. Thus, to maximize the amount of the detected signal, it is necessary to optimize the orientation of the local oscillator. (Since it has no effect on the

⁶⁾ Actually this may require a little more discussion. Both the local oscillator and the signal sidebands rotate in different manner, and the local oscillator may not be parallel to the major axis of the ellipse when the sum of the sidebands is present on this axis. In practice, however, the local oscillator is chosen so that its rotation period is much shorter than that of the signal. Then the signal sidebands look almost static compared with the local oscillator and it is expected that the local oscillator becomes parallel to the major axis while the signal sidebands is close to it.

optimization, we will ignore the swinging of the local oscillator length in the following discussion.)

The best orientation of the local oscillator would be the major axis of the ellipse, which is given by eq. (39). The maximum amplitude of the signal when the local oscillator is optimized is expected to be $|a_+| + |a_-|$, the sum of the lengths of the two vectors, from an intuitive consideration.

This can be justified as follows. The detected power after adding a local oscillator $a_{\text{LO}}e^{i\omega_0 t} = A_{\text{LO}}e^{i(\omega_0 t + \psi)}$ (which has the same frequency as the carrier) is given by

$$\begin{aligned}
 PZ_0 &= \left| a_+ e^{i(\omega_0 + \omega_g)t} + a_- e^{i(\omega_0 - \omega_g)t} + A_{\text{LO}} e^{i(\omega_0 t + \psi)} \right|^2 \\
 &= |a_+|^2 + |a_-|^2 + A_{\text{LO}}^2 + 2\Re\{a_+ a_-^* e^{i2\omega_g t}\} \\
 &\quad + 2A_{\text{LO}}\Re\{a_+ e^{i(\omega_g t - \psi)}\} + 2A_{\text{LO}}\Re\{a_- e^{-i(\omega_g t + \psi)}\} \\
 &= |a_+|^2 + |a_-|^2 + A_{\text{LO}}^2 + 2|a_+ a_-| \cos(2\omega_g t + \angle a_+ - \angle a_-) \\
 &\quad + 2A_{\text{LO}} \left[|a_+| \cos(\omega_g t + \angle a_+ - \psi) + |a_-| \cos(\omega_g t - \angle a_- + \psi) \right].
 \end{aligned} \tag{40}$$

The interfering term of interest can be rewritten as

$$\begin{aligned}
 P_{\text{int}}Z_0 &= 2A_{\text{LO}}\Re\{a_+ e^{i(\omega_g t - \psi)}\} + 2A_{\text{LO}}\Re\{a_- e^{-i(\omega_g t + \psi)}\} \\
 &= 2A_{\text{LO}}\Re\{a_+ e^{i(\omega_g t - \psi)}\} + 2A_{\text{LO}}\Re\{a_-^* e^{i(\omega_g t + \psi)}\} \\
 &= 2A_{\text{LO}}\Re\{(a_+ e^{-i\psi} + a_-^* e^{i\psi}) e^{i\omega_g t}\} \\
 &= 2A_{\text{LO}} |a_+ e^{-i\psi} + a_-^* e^{i\psi}| \cos\left[\omega_g t + \angle(a_+ e^{-i\psi} + a_-^* e^{i\psi})\right].
 \end{aligned} \tag{41}$$

Thus, $(a_+ e^{-i\psi} + a_-^* e^{i\psi})$ can be thought as being the complex amplitude of the detected power or equivalently the photo-current.

When the orientation ψ of the local oscillator is chosen to be $\psi = \psi_{\text{ax}}$ given in eq. (39), then the signal amplitude will be proportional to

$$\begin{aligned}
 a_+ e^{-i\psi} + a_-^* e^{i\psi} &= |a_+| e^{i(\phi_+ - \psi)} + |a_-| e^{-i(\phi_- - \psi)} \\
 &= |a_+| e^{i(\phi_+ - \phi_-)/2} + |a_-| e^{-i(-\phi_+ + \phi_-)/2} \\
 &= (|a_+| + |a_-|) e^{i(\phi_+ - \phi_-)/2},
 \end{aligned} \tag{42}$$

where $\phi_j = \angle a_j$ ($j = +$ or $-$). Thus, the maximum amplitude of the signal is proportional to the sum of the amplitudes of the two sidebands.

The above discussion concludes the maximum signal for a single Fourier component, but it is not the whole story. What is to be detected is actually a mixture of different Fourier components. The optimum angle for the local oscillator may be different for each component, i.e. $\psi_{\text{ax}} = \psi_{\text{ax}}(\omega)$, depending on the relative angle between $a_+(\omega_g)$ and $a_-(\omega_g)$ (see eq. (39)). The frequency response of the signal is, however, given by

$$G_{\text{sig}}(\omega) \propto [a_+(\omega)e^{-i\psi} + a_-^*(\omega)e^{i\psi}] , \quad (43)$$

where the relative angle ψ between the local oscillator and the signal is unique for all the Fourier components. Thus, it must be chosen to achieve the best compromise of all the Fourier components.

In our application of an interferometric gravitational-wave detector, this is determined by the frequency response of the optical configuration used. In narrow-band configurations, the upper and lower sidebands experience different resonance conditions in the detector, resulting in different complex amplitudes of the two sidebands. Further discussion on this subject can be found in Appendix F, with some examples to show this effect. In general, however, this effect is small and we assume ‘phase modulation signal detection’ in the following discussion.

2.4: A Michelson interferometer as a phase detector

2.4.1: Differential method—homodyne

Now let us consider the actual configuration for an optical phase detector. We assume an ideal interferometer, i.e. an interferometer with zero loss and a perfect contrast, for simplicity. Let us start from the simplest one, which is a Michelson interferometer operated at the mid-fringe.

In a Michelson interferometer, the injected light is divided by a beamsplitter (usually equally) into the two arms, reflected back from each arm, and then recombined at the same beamsplitter, as is shown in Fig. 2.10. (If the beams are recombined at another beamsplitter, it will be a Mach-Zehnder interferometer whose behavior is very similar to that described here.) There are two outputs from the beamsplitter, each containing a fraction of the light from each arm, as described in 2.3.2. The light from the two arms interfere with each other and this interfering pattern represents the phase information of the two arms.

2.4: A Michelson interferometer as a phase detector

Assuming that the same beamsplitter is used for both splitting and recombining, the two outputs of the interferometer can be written as

$$\begin{aligned} P_{\oplus}/P_0 &= (\tau\rho)^2 + (\rho\tau)^2 + 2(\rho\tau)^2 \cos \phi \\ P_{\ominus}/P_0 &= (\rho\rho)^2 + (\tau\tau)^2 - 2(\rho\tau)^2 \cos \phi , \end{aligned} \quad (44)$$

where ϕ represents the phase difference between the two arms. These reduce to

$$\begin{aligned} P_{\oplus}/P_0 &= \frac{1}{2}(1 + \cos \phi) \\ P_{\ominus}/P_0 &= \frac{1}{2}(1 - \cos \phi) , \end{aligned} \quad (45)$$

when a beamsplitter with $|\tau|^2 = |\rho|^2 = \frac{1}{2}$ is used as usual.

The change in the detected power δP in one of the outputs due to the phase change is given by

$$\frac{\delta P}{P_0} = \pm \frac{\sin \phi_0}{2} \delta \phi , \quad (46)$$

where $\delta \phi = \phi - \phi_0$ is the deviation from the initial operating point ϕ_0 . It is best to choose $\phi_0 \simeq \pm \frac{\pi}{2}$ to maximize the output signal⁷⁾. This corresponds to the mid-fringe⁸⁾, i.e. the center of the maximum and the minimum of the interference, as can be seen in eqs. (44) and (45). Since the signal appears at both outputs symmetrically, taking the difference of the two outputs gives the maximum signal, and thus this is often called the *differential method*.

Although it is not an indispensable requirement, the two arms are usually set to have equal optical path lengths (in a macroscopic sense). The advantage of this is that a considerable common mode rejection effect of the fluctuations in light source can be expected, and also that the problems associated with the spatial mode matching can be avoided. Note, however, that the phase change common in the two arms cannot be detected due to the common mode rejection effect. For gravitational wave detectors, this can be avoided by setting the two arms orthogonal so that the light in each arm experiences a different effect due to a gravitational wave.

⁷⁾ Actually what is to be optimized is not the output signal but the signal-to-noise ratio of the detected signal. In this case, however, the result is the same since the amount of the shot noise is determined by the total light power detected, which is a constant in this case. It can also be shown that choosing the beamsplitter as described in the text gives the best signal-to-noise ratio.

⁸⁾ It is by accident that the power at the two outputs are equal when this operating point is chosen. If the beamsplitter is different from the assumption here, the powers at the two outputs may be different from each other. In this case, some offset and weighting may be required before taking the difference of the two outputs.

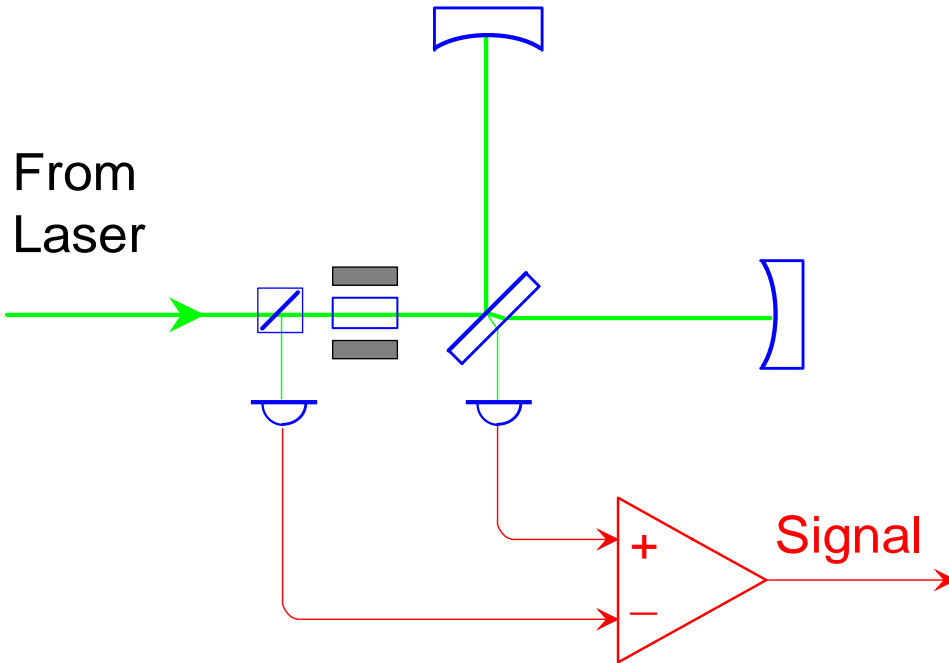


Figure 2.10: A schematic diagram of the differential method. The relative phase information of the Michelson interferometer is taken as the difference of the powers in the two outputs. The Michelson interferometer must be operated at the mid fringe, i.e. the center of the maximum and the minimum of the interference, to maximize the amount of the signal.

Let us illustrate this scheme in a phasor diagram. The light is phase modulated differentially in the two arms, which are represented by the two vectors changing their orientations. As described in 2.2.1, each vector can be decomposed into the static part and the swinging part orthogonal to the carrier. On the other hand, the initial phase difference of $\pm\frac{\pi}{2}$ between the beams from the two arms means that the two vectors are orthogonal to each other. Thus, the static part from one arm is parallel to the swinging part from the other arm and works as the local oscillator to detect it.

In this case, the local oscillator (the static part in one of the arms) has the same frequency as the carrier (that in the other arm), this can be thought of as a homodyne technique. The roles of the two arms are, however, symmetric and the distinction between the carrier and the local oscillator has less significance. The interferometer responds only to the phase difference between the two arms, as is shown in eqs. (44) and (45).

2.4: A Michelson interferometer as a phase detector

Assuming the same condition as above, let us evaluate this in the ‘sidebands’ viewpoint. The amplitude in each arm is written as

$$a_{\text{arm}} = \frac{1}{\sqrt{2}} a_0 e^{i\omega_0 t} \left[1 \pm i \frac{\phi_g}{2} (e^{i\omega_g t} + e^{-i\omega_g t}) \right], \quad (47)$$

where ϕ_g and ω_g are index and frequency of the phase modulation respectively, and the sign (\pm) in front of the sideband terms is $+$ for one arm and $-$ for the other, representing the differential feature.

The amplitudes from the two arms are superimposed with relative phase of $\pm \frac{\pi}{2}$ at the beamsplitter. Then the detected power will be

$$\begin{aligned} P_{\oplus} &= \frac{P_0}{2^2} \left| \left[1 + i \frac{\phi_g}{2} (e^{i\omega_g t} + e^{-i\omega_g t}) \right] + i \left[1 - i \frac{\phi_g}{2} (e^{i\omega_g t} + e^{-i\omega_g t}) \right] \right|^2 \\ &= \frac{P_0}{2} \left[1 + \frac{\phi_g}{2} (e^{i\omega_g t} + e^{-i\omega_g t}) \right]^2 = \frac{P_0}{2} \left[1 + 2\phi_g \cos(\omega_g t) + O(\phi_g^2) \right] \\ P_{\ominus} &= \frac{P_0}{2} \left[1 - \frac{\phi_g}{2} (e^{i\omega_g t} + e^{-i\omega_g t}) \right]^2 = \frac{P_0}{2} \left[1 - 2\phi_g \cos(\omega_g t) + O(\phi_g^2) \right]. \end{aligned} \quad (48)$$

This gives the same result as inserting $\phi = -\frac{\pi}{2} + 2\phi_g \cos(\omega_g t)$ into eq. (45). Note that the effective modulation index is twice that in one arm due to the differential (‘push-pull’) feature. (The higher order terms of ϕ_g will disappear if one considers higher harmonics of the phase modulation.)

Since this is the simplest configuration, many experiments including the first prototype of the interferometric gravitational-wave detectors built by Forward [21] have employed this scheme. In this configuration, however, all the injected light power must be detected by the photo-detectors, and it is a very tough condition in a real interferometric gravitational-wave detector because of the high power required to achieve a good sensitivity.

To alter the detection scheme to the standard heterodyne technique, the frequency of the light in one of the arms must be shifted to a slightly different frequency. This would be realized by using acousto-optic modulators. This does not, however, help the problem described above and the losses in the modulators make it less attractive than the homodyne technique.

2.4.2: Dark fringe operation—balanced modulator

To relieve the problem of high light power at the photo-detectors, the operation of a Michelson interferometer at the minimum of the interference, the so-called *dark fringe*, was proposed [22]. In this case only one of the two outputs of an interferometer, e.g. P_{\oplus} in eq. (44), is used. Then the initial phase must be chosen as $\phi_0 \simeq \pi$ to make the detecting output to be a dark fringe.

As can be seen in eq. (44), the output power varies quadratically with the relative phase around this operating point. This means that the output power does not represent the phase directly. Thus, an additional means to detect the phase information is necessary. Before discussing that, however, it is worth illustrating this from the view of a phasor diagram.

The vectors that represent the light from the two arms are the same as in the previous case. What is different is the relative angle between the two vectors—in this case they are opposite to each other, as is indicated by the initial phase difference of π . When these two vectors are summed together, their static parts cancel each other and the swinging parts add up if (and only if) they are differentially modulated in the two arms.

As one might notice, this actually works as a balanced modulator, shown in Fig. 2.7. The output from this balanced modulator is, as described in 2.2.2, suppressed-carrier double sideband. In this case, each arm is considered as a modulator with opposite polarity⁹⁾, and the beamsplitter takes the difference between the two. Incidentally, a gravitational wave with appropriate orientation and polarization modulates the light in two orthogonal arms oppositely, as described in 2.1.3.

In order to suppress the carrier component, the light from the two arms must have the same amount of the carrier component. From this requirement, the beamsplitter is usually¹⁰⁾ chosen to be $|\tau|^2 = |\rho|^2 = \frac{1}{2}$. When the carrier component is successfully suppressed, the light power leaking out of the interferometer can be significantly reduced.

⁹⁾ If only one of the arms is modulated, it can be considered as a composition of a common mode modulation and a differential modulation. Only the latter is detected in this type of interferometer.

¹⁰⁾ Actually it is possible to suppress the carrier component with different parameters. It will, however, give a worse signal-to-noise ratio than that obtained with the parameters described in the text.

Under this condition, the output amplitude of this balanced modulator can be written as

$$\begin{aligned} a_{\text{bm}} &= \frac{1}{2} a_0 e^{i\omega_0 t} \left\{ \left[1 + i \frac{\phi_g}{2} (e^{i\omega_g t} + e^{-i\omega_g t}) \right] - \left[1 - i \frac{\phi_g}{2} (e^{i\omega_g t} + e^{-i\omega_g t}) \right] \right\} \\ &= i \frac{\phi_g}{2} (e^{i\omega_g t} + e^{-i\omega_g t}) a_0 e^{i\omega_0 t} = i \phi_g \cos(\omega_g t) a_0 e^{i\omega_0 t}. \end{aligned} \quad (49)$$

As can be found in the above equation, there is no carrier component and only upper and lower sidebands are present.

The most important feature of a balanced modulator is the separation of the carrier and the signal, i.e. only the differential sidebands leak out of the interferometer from one of the outputs and the carrier which contains most of the injected power exits from the other. This will allow us to re-use the rejected carrier in order to increase the effective light power incident on the interferometer by a scheme called *power recycling*. This will be thoroughly discussed in Chapter 3, together with other variations from this scheme.

Power recycling, or one of the other variations, is currently considered as being indispensable in the future interferometric gravitational-wave detectors with good sensitivities. From this reasoning, we will neglect the differential method in the following discussion and concentrate on the dark-fringe operation.

The optical suppressed-carrier double sideband cannot, however, be detected directly. Thus it is necessary to use some kind of local oscillator in order to detect them. As is discussed in 2.3.3, the local oscillator must be parallel to the sidebands to be detected (unless standard heterodyne is used). There are several schemes proposed to introduce the local oscillator to such a Michelson interferometer. Some of them are discussed in the following part.

2.4.3: External modulation

The most straightforward way of introducing the local oscillator to detect the suppressed-carrier double sideband is to make an independent interferometer for this purpose. The local oscillator is taken separately from the same light source as the main interferometer, and it is made to interfere with the output of the main interferometer. Thus, the optical layout looks more like a Mach-Zehnder interferometer which has a Michelson interferometer in one of the arms,

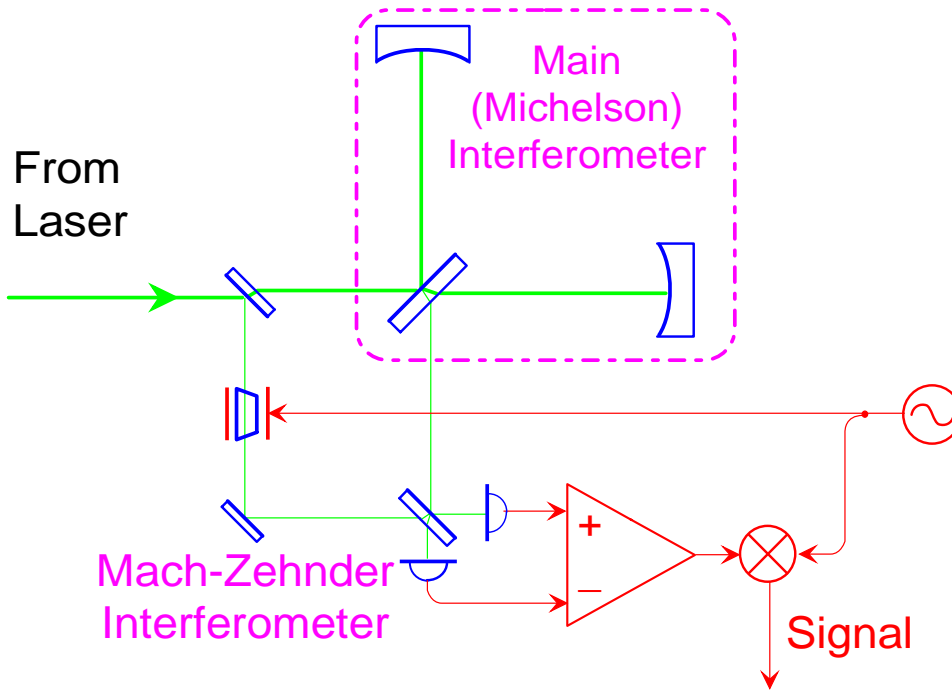


Figure 2.11: A schematic diagram of the external modulation scheme (the case of the local oscillators induced by a phase modulation is shown). The whole system looks like a Mach-Zehnder interferometer with a Michelson interferometer in one of its arms. The main (Michelson) interferometer is operated at a dark fringe and its output is made to interfere with the local oscillator beam, and the interfering outputs give the signal. If the local oscillators are shifted in their frequency, the signal is obtained by demodulating the photo-current.

as is shown in Fig. 2.11. Since the local oscillator is taken out from the outside of the main interferometer, this scheme is called the *external modulation*¹¹⁾ [23, 24].

In this scheme, the frequency of the local oscillator and the phase relative to the main interferometer can be chosen freely. If the original frequency from the light source is used, it will be homodyne; if the frequency is shifted (by, for example, an acousto-optic modulator), then it will be standard heterodyne; if the modulation-induced sidebands (produced by, for example, an electro-optic modulator) are used, then the local oscillator will have a swinging length. In practice, however, it is difficult to produce pure double sideband and usually the

¹¹⁾ This naming may cause confusion for some people. As described in 2.1.3, a gravitational wave will impose a phase modulation on the light, not depending on its detection scheme. The ‘modulation’ used here is, as discussed in the text, merely to produce the local oscillator(s) to detect the sidebands produced by the phase modulation due to a gravitational wave. Thus, this may be better to be named ‘external local oscillator’. Furthermore, in the ‘external modulation’ case a ‘modulation’ is actually not an absolute requirement. In this paper, however, the author followed the conventional notation in this field.

2.4: A Michelson interferometer as a phase detector

carrier frequency component or higher order sideband components are present together with the actual local oscillator. This effect is discussed in Appendix H.

Let us calculate the signal for this case. The output of the main interferometer is assumed to be the suppressed carrier, which is described in eq. (49). This is superimposed with an externally taken local oscillator a_{LO} by an auxiliary beamsplitter with the reflectivity ρ_z and the transmittance τ_z :

$$\begin{aligned} a_{\oplus} &= \rho_z [i\phi_g \cos(\omega_g t) a_0 e^{i\omega_0 t}] + i\tau_z a_{\text{LO}} e^{i\omega_0 t} \\ a_{\ominus} &= i\tau_z [i\phi_g \cos(\omega_g t) a_0 e^{i\omega_0 t}] - \rho_z a_{\text{LO}} e^{i\omega_0 t} . \end{aligned} \quad (50)$$

There are two outputs from this Mach-Zehnder interferometer and both of them can be used for signal detection.

Assuming the homodyne case, i.e. the local oscillator has the original frequency from the light source, the detected powers at the two outputs are

$$\begin{aligned} P_{\oplus} Z_0 &= (\rho_z |a_{\text{LO}}|)^2 + [\tau_z \phi_g \cos(\omega_g t)]^2 + 2\rho_z \tau_z \phi_g \cos(\omega_g t) |a_0 a_{\text{LO}}| \cos \psi \\ P_{\ominus} Z_0 &= (\tau_z |a_{\text{LO}}|)^2 + [\rho_z \phi_g \cos(\omega_g t)]^2 - 2\rho_z \tau_z \phi_g \cos(\omega_g t) |a_0 a_{\text{LO}}| \cos \psi , \end{aligned} \quad (51)$$

where $\psi = \angle a_0 - \angle a_{\text{LO}}$ is the relative angle between the suppressed carrier and the local oscillator. The first and the second term are the power of the suppressed carrier and the local oscillator respectively, and are not much of our interest. The third, the interfering term provides the signal proportional to the original modulation $\phi_g \cos(\omega_g t)$.

The magnitude of the output signal is determined by the local oscillator amplitude $|a_{\text{LO}}|$ and the relative phase ψ . As is discussed in 2.3.3, the relative phase must be chosen as 0 or π , i.e. ‘parallel’ to each other, to maximize the output signal. Thus, this requires additional control systems as well as additional optics in order to maintain the optimum signal detection.

When the single local oscillator whose frequency is shifted by ω_a is used, $a_{\text{LO}} e^{i\omega_0 t}$ in eq. (50) must be replaced by $a_{\text{LO}} e^{i(\omega_0 + \omega_a)t}$. Then the interfering term in eq. (51) becomes

$$\delta P Z_0 = 2\rho_z \tau_z \phi_g \cos(\omega_g t) |a_0 a_{\text{LO}}| \cos(\omega_a t + \psi) . \quad (52)$$

In this case, the initial phase ψ has little meaning since the relative phase varies with time anyway. The signal appears as an amplitude modulation on the oscillation at ω_a .

When the modulation-induced double sidebands are used as the local oscillators, $a_{\text{LO}}e^{i\omega_0 t}$ in eq. (50) will have an additional oscillating factor $\cos(\omega_a t)$. Then the interfering term in eq. (51) becomes

$$\delta P Z_0 = 2\rho_z \tau_z \phi_g \cos(\omega_g t) |a_0 a_{\text{LO}}| \cos(\omega_a t) \cos(\psi) . \quad (53)$$

Thus, again the output signal appears as an amplitude modulation on this additional oscillation. In this case, however, the relative phase factor $\cos\psi$ remains static and thus it must be optimized.

In either case, the original signal $\phi_g \cos(\omega_g t)$ can be recovered by (electrically) demodulating the photo-current, i.e. by multiplying it by the new ‘carrier’ frequency ω_a , in a similar way to that described in eq. (37).

2.4.4: Internal modulation

In spite of its intuitiveness, the external modulation is in practice a relatively complicated scheme. In order to superimpose the local oscillator onto the signal sidebands, an additional interferometer is required. It may be easier to realize if the same (or a similar) thing can be done within a single interferometer. This can be done, although with some restrictions, by a scheme called the *internal modulation*, which was actually proposed [22] earlier than the external modulation and commonly used in the current prototypes [25, 26].

The principle of this scheme is, in short, to use the main Michelson interferometer as the phase detector as well as the balanced modulator. This requires a local oscillator which is already superimposed onto the signal sidebands when they leak out of the interferometer. Then the local oscillator must have a different frequency from that of the carrier to be distinguishable from it. Thus this will employ either standard heterodyne technique or a ‘swinging local oscillator’.

Since the Michelson interferometer works as a balanced modulator for the phase modulation, for example due to a gravitational wave, it is also possible to use the same interferometer in order to produce the local oscillator by imposing an artificial modulation. In order to make the best use of the balanced modulator, the modulation should be differential in the two arms of the Michelson interferometer. Since what we want to detect is the sidebands induced by a phase modulation, the modulation to produce the local oscillator must also be a phase modulation to make the local oscillator ‘parallel’ to the signal sidebands.

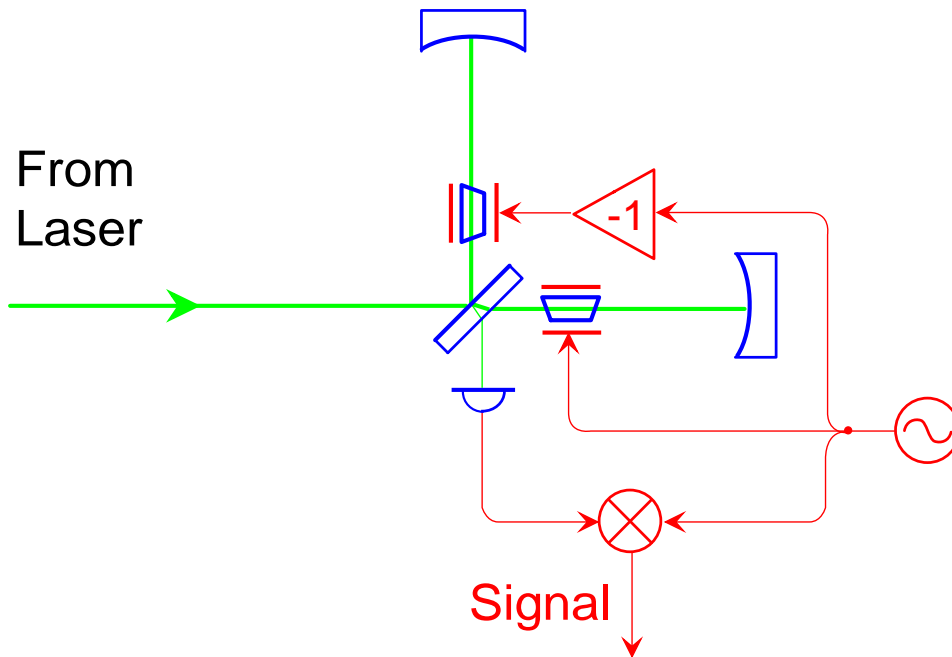


Figure 2.12: A schematic diagram of the internal modulation scheme. The beams in the two arms are differentially phase modulated and the photocurrent is demodulated in order to recover the signal. The Michelson interferometer must be operated at a dark fringe.

In the end, the optical system looks like Fig. 2.12: a Michelson interferometer is artificially phase modulated at a frequency well above the expected signal frequencies, differentially in the two arms. One of the outputs from the interferometer is kept near a dark fringe, and it contains the signal sidebands as well as the artificial sidebands which work as the local oscillators. When these are detected by a photo-detector, the interfering term yields the modulation frequency component proportional to the relative phase of the interferometer.

In the above description, the use of standard heterodyne technique is excluded, since it requires a single local oscillator with a frequency different from the carrier and it is difficult to realize. Similarly, only the phase modulation components of the signal can be detected, even if the optimum detection phase (see the discussion in 2.3.4) is different from it. These limitations do not, however, affect our application so much, since what we usually measure is the phase modulation component anyway.

Since the local oscillators are produced by a balanced modulator, there is no carrier components remaining in the detecting output. There will, however, be higher harmonics left and will affect the signal as well as noise, as discussed in Appendix H. In addition, the fact that the dominating component among the

detected power is the swinging local oscillator means that it is not stationary. This adds a further complication to the shot noise consideration. This topic is discussed in Ref. [27] in detail.

2.4.5: Pre-modulation

While the internal modulation is a convenient and useful scheme, there will be a few practical problems associated with the Pockels cell used as the phase modulators when it is used in full scale interferometric gravitational-wave detectors. In such an interferometer with a few kilometers of arm length, the beam radius will be as large as several centimeters (see eq. (G14) in Appendix G). Such a crystal with a big enough aperture is very difficult to manufacture, considering the demands on the optical quality. Furthermore, the high light power inside an advanced interferometer may cause damage to the crystal by a small absorption inside it.

One of the ways to avoid this problem is the use of the external modulation described in 2.4.3, however at the price of a little more complication. There is an alternative way, the *pre-modulation* proposed by Schnupp [28, 29], to introduce the local oscillator onto the signal sidebands leaking out of a Michelson interferometer without any internal modulator.

In this scheme, the local oscillators are produced by phase-modulating the beam from the light source *before* it enters the interferometer, and both the carrier and the local oscillator components are injected to the interferometer (see Fig. 2.13). If the interferometer is perfectly symmetric, the local oscillators will not appear at the detecting output due to the common mode rejection effect. The novelty of this scheme is to put a small asymmetry on the lengths of the two arms in order to make the local oscillator leak out of the interferometer. This effectively results in the same effect as the internal modulation.

In a simple Michelson interferometer with a small asymmetry $\Delta\ell$ in the arm lengths, the amplitude of the local oscillator, expressed in the ‘effective’ modulation index, by the pre-modulation is determined by the amount of asymmetry and the modulation frequency, as

$$m_{\text{eff}} = m_a \sin(\omega_a \Delta\ell / c) . \quad (54)$$

By choosing the modulation frequency to produce the local oscillators much higher than the expected signal frequency, i.e. $\omega_a \gg \omega_g$ (which is a condition usually satisfied), it is possible to set the asymmetry to have negligible effect on

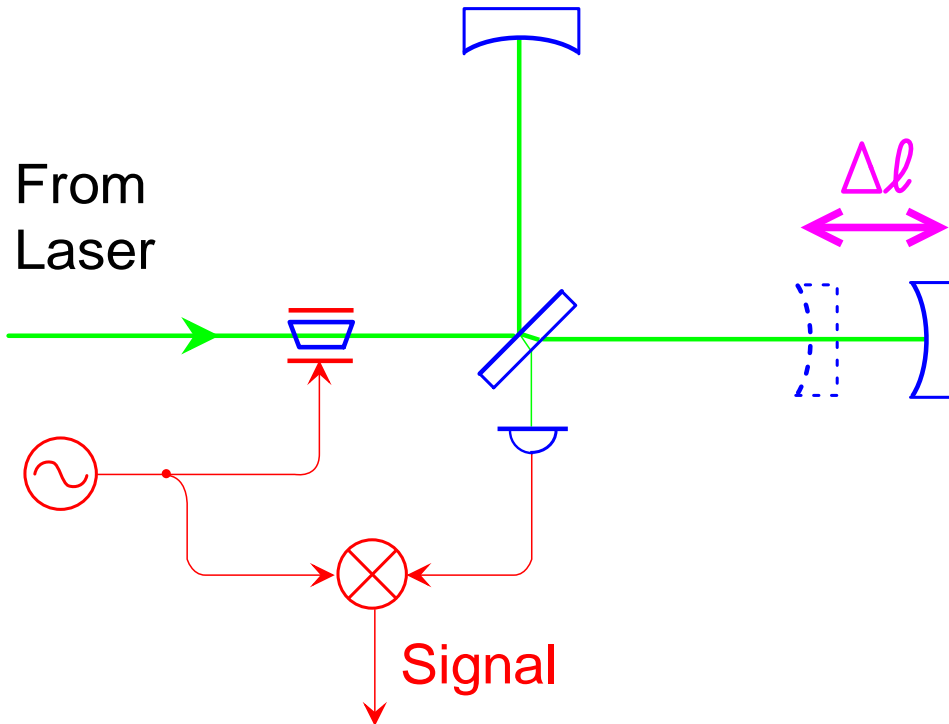


Figure 2.13: A schematic diagram of the pre-modulation scheme. In this case the beam is phase modulated before entering the Michelson interferometer, whose arm lengths have an intentional difference $\Delta\ell$. Due to this difference, a part of modulation-induced sidebands leaks out of the interferometer operated at a dark fringe, together with the sidebands due to differential modulation inside the interferometer.

the signal sidebands. (Note that replacing ω_a in eq. (54) by ω_g gives the fraction of the signal sidebands that will not be detected.)

In practice, however, the main interferometer will have either Fabry-Perot cavities or delay lines in the arms and also have power recycling and so on incorporated, as described in Chapter 3. Thus, the effect in more realistic cases must be carefully considered, taking the actual configuration into account (for example, Ref. [30] treats a specific possible scheme in detail). Here we restrict ourselves just to mention the principle of its operation.

Since this scheme has a similar effect as the internal modulation, only the phase modulation signal can be detected in this simple configuration described above. In this case, however, the local oscillator is produced outside the interferometer and there are less restriction in producing it than in the internal modulation case. Thus, there are some possibilities of using different types of local oscillators, e.g. a single local oscillator, by using somewhat more complicated optical configuration.

Chapter 3

Various configurations

3.0: Introduction

In Chapter 2, we have discussed the signal that is expected to be produced by gravitational waves and also the principle of its detection. The sensitivity of a detector is, however, determined also by the noise in the detector output. Among various noise sources, the statistical error in the photon counting process is expected to be one of the limiting factors. As a result, a simple Michelson interferometer requires an exceptionally high-power light source as well as impractically long arm lengths in order to achieve good sensitivity for the detection of gravitational waves.

Several methods have been proposed to improve the sensitivity of such an interferometer by using somewhat more complicated optical configurations. The aim of this chapter is to compare these configurations and to highlight the similarities and the differences. For that purpose, one versatile theorem which can explain most of their characteristics is presented. Then this is applied here to each configuration to evaluate its sensitivity, except for the case of *resonant sideband extraction* which is treated separately in Chapter 4.

It should be noted that the results here are derived without any consideration of non-linear effects, such as thermal distortions, which may become important for large scale interferometric gravitational-wave detectors. Thermal effects are discussed in Chapter 5, and the relevant sensitivity of each configuration is estimated there. The discussion here is, however, helpful in developing a general understanding of the various configurations.

3.1: The sensitivity theorem

3.1.1: Sensitivity

As briefly reviewed in 1.3, there will be noise contributions from various sources in an output from a detector, and a gravitational wave signal must be compared with them. Almost all the noise sources are of statistical, or random,

nature. Thus, the ‘detection’ of a gravitational wave signal must also be treated statistically. In such a case, the noise is represented by the ‘typical’, or averaged, deviation of the output from the static value. (If the static value is different from that in the ideal case, it may be treated separately as an *offset*.) The ratio between the signal amplitude and this ‘typical’ deviation is called *signal-to-noise ratio*, by which the criteria of ‘detection’ are discussed.

The signal-to-noise ratio does not only depend on the amount of signal and noise, but also on the filters used to extract the signal from the observed data. The filters can be applied either in the frequency domain or in the time domain. The latter is more often called the *template* of the waveform. According to the *matched filter theory*, the template having the same form as the signal gives the best signal-to-noise ratio. It should be noted that the output of an interferometer has frequency dependence determined by the configuration employed, and this dependence must be taken into account in the data analyses.

The output from one detector must be examined in coincidence with those from others, in order to avoid the error of the first kind (false alarm). More detectors used in coincidence reduce the probability of the error of the second kind (missing events) for a certain tolerance of a possible false alarm. Then a lower signal-to-noise ratio is sufficient in one detector, and this results in a better probability of detection with a given amount of noise in a detector. Another possibility to improve the detectability is taking the correlation between the data from multiple detectors. In our application of gravitational wave detection, either scheme will be necessary since the signal-to-noise ratio is not expected to be very high.

All of the factors described above are involved in the ‘detection’ of a gravitational wave [8, 14]. For the ‘detectability’, i.e. the probability of detecting a signal, all these factors are important and should be taken into account.

The aim of this paper is, however, to compare the possible performance of various optical configurations for individual interferometric gravitational-wave detectors. For this purpose, only the amount of noise in a detector must be considered, i.e. the ‘sensitivity of a detector’ must be represented by its noise level. Thus, the amplitude of a signal that gives unity signal-to-noise ratio, i.e. the amplitude of a noise-equivalent signal, would be the appropriate way to express the sensitivity of a detector.

Since shot noise (which is assumed to be the only noise in the following discussion) is broad-band in nature, its root-mean-square, or *rms*, amplitude

depends on the bandwidth of observation. To avoid introducing the discussion of such an observation process, it is convenient (or necessary) to express the noise by its spectral density. Thus, the sensitivity of a detector, which should be equivalent to its noise, must also be described in a spectral form.

As a consequence, the word ‘sensitivity’, symbolized as $\tilde{h}(\omega_g)$, implies ‘*the spectrum of signal amplitude equivalent to that of the noise in the detector*’ throughout this paper. It should be emphasized again that this is a characteristic of a single detector, and the actual detection of a signal depends on other parameters as well.

3.1.2: Normalized response

Because the response of a detector to a gravitational wave depends on the frequency of its Fourier components, the sensitivity is in general a function of signal frequency. Since the spectrum of the shot noise is ‘white’, or frequency-independent, the frequency dependence of the sensitivity can be obtained by considering the amplitude of the signal at the output. Thus, the frequency dependence of the sensitivity can also be represented by its *transfer function*, the ratio of the input and the output signals. In our case, the input signal is the amplitude of the gravitational wave and the output signal is, virtually, the photo-current in the photo-detector.

The transfer function is, however, not so convenient in comparing the various configurations for interferometric gravitational-wave detectors, because it depends not only on the configuration employed but also on the power of the illuminating light. Fortunately, the latter does not affect the frequency dependence of the detector but just changes the overall sensitivity. Thus, it is possible to ‘normalize’ the response in order to remove the dependence on the input power. This *normalized response* is the most useful measure in comparing the characteristics of various configurations.

In this paper, both the sensitivity and the normalized response are used in discussing the characteristics of each configuration. The former is convenient in making a rough estimate of the detector performance for a chosen configuration, and the latter is useful in comparing the detailed properties of different configurations.

These two quantities are, of course, related to each other. The phase fluctuation $\delta\tilde{\phi}$ equivalent to shot noise is given by [32]

$$\begin{aligned}\delta\tilde{\phi} &= \sqrt{\frac{2\hbar\omega_0}{P_0}} \\ &\approx 2\pi \cdot 10^{-10} \text{ rad}/\sqrt{\text{Hz}} \times \left[\frac{P_0}{1 \text{ W}} \right]^{-\frac{1}{2}} \left[\frac{\lambda}{1 \mu\text{m}} \right]^{-\frac{1}{2}},\end{aligned}\tag{1}$$

where ω_0 and P_0 are the angular frequency and the power of the incident light from the source, respectively. The sensitivity $\tilde{h}(\omega_g)$ is obtained by dividing eq. (1) by the frequency response $\mathbf{G}(\omega_g)$, thus

$$\tilde{h}(\omega_g) = \frac{1}{2|\mathbf{G}(\omega_g)|} \sqrt{\frac{2\hbar\omega_0}{P_0}}.\tag{2}$$

There is an additional factor of 1/2 which is explained in Appendix A.

The normalized response would, however, be useful only when the optical system is linear, i.e. when it does not change with input power. The *loss* in the optics is one of the factors to determine the frequency response, and it may change with the light power when thermal effects become significant. In this chapter, such non-linear effects are ignored to concentrate on understanding of each configuration. Some of the non-linear problems are discussed in Chapter 5.

3.1.3: Simplified theorem

Calculating the frequency response of a configuration to estimate the sensitivity with given parameters was carried out in the previous literature (see, e.g. [33, 34]), and found not to be an easy task. The general interest is, however, not the detailed frequency dependence of the sensitivity but in most of the cases a rough estimation of the achievable sensitivity at its peak. It is found by the author that this can be obtained by a simple formula, with only a few parameters to be considered.

If the sensitivity of an interferometric gravitational-wave detector is limited by shot noise, the peak sensitivity \tilde{h}_0 to gravitational waves expressed as a linear spectral density can be given by

$$\tilde{h}_0 \gtrsim \sqrt{\frac{2\hbar\lambda}{\pi c} \frac{\Delta f_{\text{BW}}}{\mathcal{E}}},\tag{3}$$

or,

$$\tilde{h}_0 \gtrsim 3.3 \cdot 10^{-24} / \sqrt{\text{Hz}} \times \left[\frac{\lambda}{1 \mu\text{m}} \right]^{\frac{1}{2}} \left[\frac{\mathcal{E}}{20 \text{ J}} \right]^{-\frac{1}{2}} \left[\frac{\Delta f_{\text{BW}}}{1 \text{ kHz}} \right]^{\frac{1}{2}}.\tag{4}$$

Here \mathcal{E} is the light energy stored in the optical system, λ the wavelength of the light, and Δf_{BW} the *detector bandwidth*. The last quantity means the approximate frequency range in which the spectral sensitivity is close to that at the peak, and should not be confused with the *observation bandwidth*, which is a result of filtering of the data to extract the signal.

It should be noticed that the relation in eq. (3) is not an exact but an approximate one, as is indicated by the inequality symbol. It could be exact if the interferometer were ideal and if the bandwidth were appropriately defined. In reality, however, the sensitivity is easily degraded from that given by this equation due to various losses in the optical system. In addition, if the system utilizes only one of the sidebands produced by the phase modulation due to the gravitational wave, then the sensitivity may be worse by a factor of $\sim\sqrt{2}$ due to the inefficiency in the signal detection process. With the exception of this case, any reasonable configuration has in general low enough loss to achieve a sensitivity close to the value given by eq. (3).

For a given detector, the wavelength is determined by the laser employed as the light source, which means it cannot be changed so easily to improve the sensitivity. On the other hand, the bandwidth is determined by the requirements of the signal being sought. (How broad a detector bandwidth should be for a certain type of signal is another question which is not discussed here.) It should anyway be fixed in our comparison of detectors using different configurations. As a consequence, the *energy* is the only parameter which can be changed to improve the sensitivity.

If the energy determines the sensitivity of an interferometric gravitational-wave detector, how much of it can be stored in the optical system? For a standard Fabry-Perot cavity, the maximum energy that can be stored with given parameters of length ℓ of the cavity, illuminating power P_0 from the light source, and maximum power reflectivity¹⁾ (ρ_{max})² obtainable for the mirrors, is uniquely determined as

$$\begin{aligned} \mathcal{E}_{\text{max}} &= P_0 \frac{\ell}{c} \frac{(\rho_{\text{max}})^2}{1 - (\rho_{\text{max}})^2} \\ &\approx 20 \text{ J} \times \left[\frac{P_0}{50 \text{ W}} \right] \left[\frac{\ell}{3 \text{ km}} \right] \left[\frac{1 - (\rho_{\text{max}})^2}{25 \text{ ppm}} \right]^{-1}. \end{aligned} \quad (5)$$

¹⁾ In this and next chapters, the loss per reflection on a surface of a mirror is assumed to be constant, i.e. not depending on the reflectivity of the mirror, and thus the maximum reflectivity should be obtained by unity minus (constant) loss. In practice, this may not be exactly true (see Appendix B). If the dependence is small, however, the following discussion will not be affected though some of the exact relations should be taken as approximations.

This is because the light experiences, at least, a fractional loss of $1 - (\rho_{\max})^2$ upon a reflection in each passage through the cavity.

With efficient use of the illuminating light power (by a method called *power recycling* which is described in 3.3) and if one can ignore the loss at the beam-splitter (which is a reasonable approximation in many cases), it can be shown that equation eq. (5) is also valid for interferometric gravitational-wave detectors. In this case, ℓ should be taken as the arm length of the interferometer. The formulae eqs. (3) and (5) can be combined to yield the optimum sensitivity achievable with given parameters, as can be found in Ref. [35]. Eq. (3) is valid, however, even for non-optimized cases, as we will see in the rest of this chapter

A result of this theorem is that any configuration storing the same amount of energy will have the same sensitivity when optimized for the same bandwidth. This is an extension of the results from previous analyses [33, 34]:

“There are no significant differences in each configuration, as long as they are optimized for the same bandwidth.”

Although this has been derived only for configurations which had been proposed by that time, the theorem eq. (3) is stated as a general relationship which must *always* be satisfied. The physical background of this theorem will be discussed in 3.5.

Before doing that, it is interesting to present a review of the configurations proposed so far, along with the application of the theorem to each configuration considered.

3.2: Delay lines and Fabry-Perot cavities

3.2.1: Multi-reflection schemes (optical delay-lines)

As we have seen in Chapter 2, a simple Michelson interferometer must have an arm length of the order of $\ell \sim c/\omega_g$ when it is optimized for a frequency $\omega_g/2\pi$. This means that an arm length of ~ 50 km is required for an interferometer optimized for 1 kHz gravitational wave, and even longer arm lengths for lower frequencies. This is an unrealistic value for ground-based experiments.

To realize a longer optical path length within a manageable size, folding the path by multiple reflections (see Fig. 3.1), especially the use of optical delay lines [36, 37], has been proposed [22] and successfully tested [25, 26]. The normalized response of an interferometer with $N - 1$ reflections (where N represents

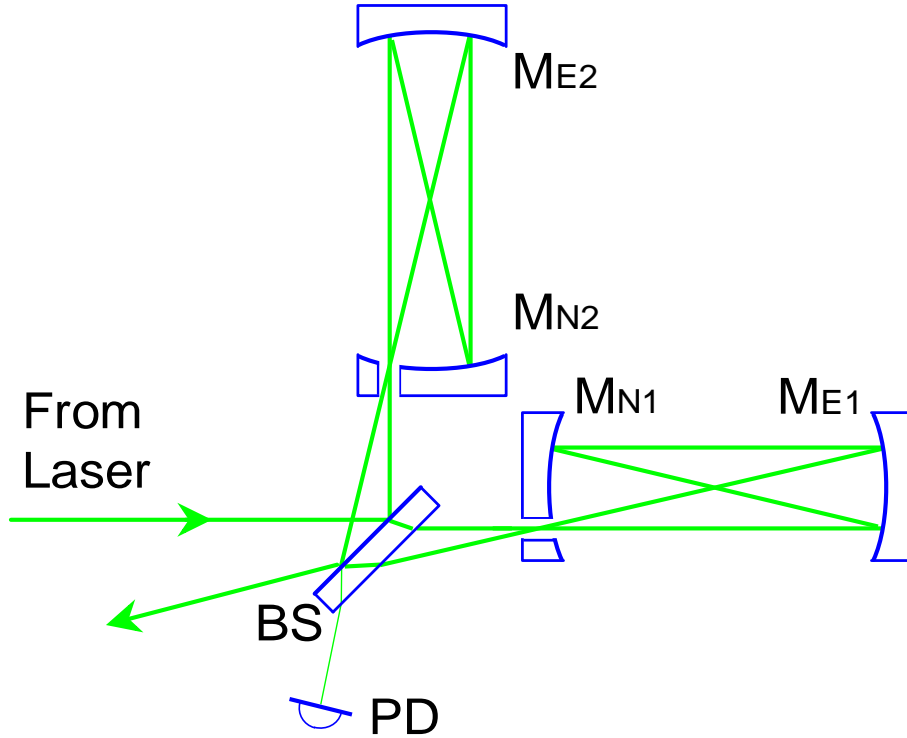


Figure 3.1: Optical layout of an interferometer with a Herriott delay line in each arm ($N=4$ is shown). The light from the source is divided by a beam splitter into two arms. Each arm contains two (concave) mirrors facing each other, which compose an *optical delay line*. The light is injected through a hole on the near-end mirror, and is reflected back and forth inside the delay line. After $N-1$ reflections (where N stands for the number of transits in the delay line), the light exits the delay line through the same hole again and interferes with the light from the other arm.

the number of transits or beams, e.g. $N=2$ for a single-bounce interferometer) in the arms of length ℓ was found to be almost the same as a single-bounce interferometer with arm length $N\ell/2$ and a mirror reflectivity of ρ^{N-1} instead of ρ .

The storage time of a photon in one arm with mirrors of reflectivities ρ_n and ρ_e (stand for near and far end mirrors respectively) is given by

$$\begin{aligned}
 t_s &= \frac{\ell}{c} (1 + \rho_e^2) \sum_{j=1}^{N/2} (\rho_n \rho_e)^{2(j-1)} \\
 &= \frac{\ell}{c} \frac{1 - (\rho^2)^N}{1 - \rho^2} \quad (\text{for } \rho_n = \rho_e = \rho) \\
 &\simeq N\ell/c \quad (\text{for } \rho^2 \cong 1).
 \end{aligned} \tag{6}$$

The normalized response of this type of interferometer has a sharp cut-off at $\omega_c/2\pi = c/2N\ell$, thus the bandwidth can be estimated as

$$\Delta f_{\text{DL}} = c/2N\ell \simeq 1/2t_s. \quad (7)$$

Thus, the bandwidth is approximately given by a half of the reciprocal of the storage time if $\rho^2 \cong 1$.

It is interesting to consider this from our new viewpoint. Assuming sufficiently low losses of the mirrors, the $N-1$ reflections cause an increase in storage time by a factor of $\sim N/2$. The energy stored is in this case given by the product of the illuminating power and the storage time, thus increased by the same factor $\sim N/2$. This leads to an improvement by $\sim\sqrt{N/2}$ in terms of sensitivity. On the other hand, the bandwidth becomes narrower by a factor of $N/2$, which yields another factor of $\sim\sqrt{N/2}$ in sensitivity. These two effects result in the expected improvement by $\sim N/2$ in total.

To estimate the sensitivity of this type of interferometer, the only thing to be done is to evaluate the energy in the interferometer. If we choose parameters that realize a 1 kHz bandwidth, then $N\ell$ has to be 150 km and we get

$$\begin{aligned} \mathcal{E}_{\text{DL}} &= P_0 \times t_s \simeq P_0 \times \frac{N\ell}{c} \\ &\approx 25 \text{ mJ} \times \left[\frac{P_0}{50 \text{ W}} \right] \left[\frac{\Delta f_{\text{DL}}}{1 \text{ kHz}} \right]^{-1}. \end{aligned} \quad (8)$$

Inserting these values into eq. (4), we can estimate the sensitivity as

$$\tilde{h}_{\text{DL}} \approx 0.9 \cdot 10^{-22}/\sqrt{\text{Hz}} \times \left[\frac{\lambda}{1 \mu\text{m}} \right]^{\frac{1}{2}} \left[\frac{P_0}{50 \text{ W}} \right]^{-\frac{1}{2}} \left[\frac{\Delta f_{\text{DL}}}{1 \text{ kHz}} \right]. \quad (9)$$

This corresponds to the classical expression for the sensitivity of this type of interferometer [25],

$$\tilde{h}_{\text{DL}} \simeq \frac{1}{N\ell} \sqrt{\frac{\hbar c}{\pi} \cdot \frac{\lambda}{P_0}}. \quad (10)$$

Note that the sensitivity only depends on the total path length $N\ell$, not on the arm length ℓ .

Though this sensitivity may be sufficient to detect some rare types of gravitational waves, it is not satisfactory for an astronomical observatory. It may be possible, in principle, to improve the sensitivity by increasing the optical path length, though this restricts the observation to lower frequencies. A practical limitation, however, arises from the demands on mirror sizes, whose areas

are required to be almost proportional to the total path length [32]. Even the value quoted above requires mirrors with ~ 0.90 m diameter, which seems to be an upper limit (technically and financially) for manufacturing high quality mirrors.

3.2.2: Fabry-Perot cavities

After initial experiments pursued multi-reflection schemes intensively, the use of Fabry-Perot cavities was proposed [38, 39] and some experiments have employed this configuration [40, 41]. In this configuration a Michelson interferometer contains two Fabry-Perot cavities²⁾, one in each arm, as is shown in Fig. 3.2. The primary benefit of this was that the mirror size required can be made considerably smaller than that for multi-reflection schemes, by superimposing all the reflection spots. In addition, the problem of spurious signals caused by the scattered light in delay line type interferometer [42] can be avoided by using Fabry-Perot cavities in the arms. This, however, requires each cavity to be controlled in order to keep it resonant with the incoming light.

For investigating the behavior of a Fabry-Perot cavity, it is useful to introduce several parameters similar to those for the delay lines (hereafter ‘delay line’ will be used to represent all of the multi-reflection schemes, because the performance does not depend on how the multiple reflections are realized). Among them are the storage time, the bandwidth, and the energy stored inside.

The definition of the storage time in a cavity is not so obvious as in delay lines, because each photon may experience a different storage time. In this paper, the word ‘storage time’ will imply ‘the average time one photon stays in the cavity without any excitation from outside’. This means that it is defined by the power, not the amplitude, and also that it is not defined for the illuminating light which stays longer than this storage time, as we will see shortly.

By this definition, the storage time in a cavity can be calculated using amplitude reflectivities of the coupling mirror ρ_c and of the reflecting mirror ρ_r , as

$$t_s = \frac{\ell}{c} \frac{1 + \rho_r^2}{1 - \rho_c^2 \rho_r^2} \quad (11)$$

$$\simeq \frac{1}{\Delta f_{\text{FSR}}} \times \frac{1}{1 - \rho_c^2 \rho_r^2} \quad (\text{for } \rho_r^2 \simeq 1),$$

²⁾ To avoid any possible confusions, the word ‘(Michelson) interferometer’ is *exclusively* used for the two arm arrangement, either with delay lines or Fabry-Perot cavities in the arms. A Fabry-Perot arrangement, such as the one inside a single arm here, will be referred to as a ‘(Fabry-Perot) cavity’.

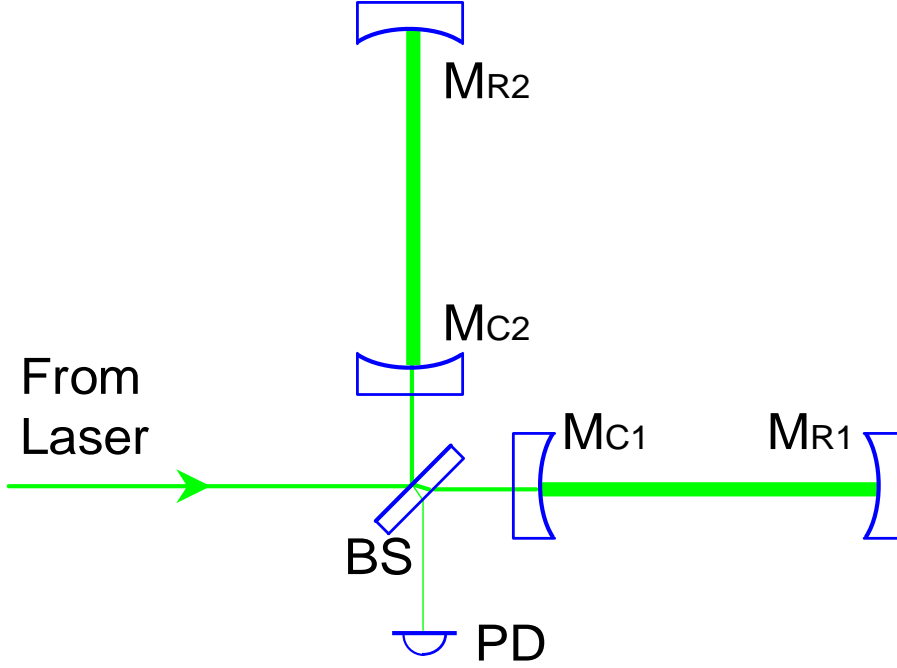


Figure 3.2: Optical layout of an interferometer with a Fabry-Perot cavity in each arm. The difference from the Fig. 3.1 is that the optical delay line in each arm is replaced by a Fabry-Perot cavity. The reflected light from the cavity in each arm interferes with that from the other arm at the beam-splitter. A detailed analysis of an isolated Fabry-Perot cavity can be found in 4.2.2.

where $\Delta f_{\text{FSR}} = c/2\ell$ is the free-spectral-range of the cavity. This can be obtained by taking the limit $N \rightarrow \infty$ of eq. (6), the storage time for delay lines, with different reflectivities for the two mirrors (ρ_n and ρ_e are replaced by ρ_c and ρ_r , respectively).

The amount of energy stored in the cavity is given by (see eqs. (4.6) and (4.8) in 4.2.2)

$$\mathcal{E}_{\text{FP}} = \mathcal{G}_{\text{FP}} P_0 \times \frac{\ell (1 + \rho_r^2)}{c} = \frac{P_0 \ell}{c} \times \frac{\tau_c^2 (1 + \rho_r^2)}{(1 - \rho_c \rho_r)^2}, \quad (12)$$

where \mathcal{G}_{FP} represents the equivalent power enhancement inside the cavity. (This reduces to eq. (5) by choosing $\rho_c = (1 - \mathcal{A}_c) \rho_r$, where \mathcal{A}_c is the loss of the coupling mirror and $\rho_r^2 = 1 - \mathcal{A}_c$ is assumed.) Note that the interchange of two mirrors results in a different amount of energy stored.

The *linewidth* of a cavity is in general measured by the full width at half maximum Δf_{FWHM} , i.e. the spacing of the two frequencies where the power transmittance becomes half of that at the peak of resonance. For gravitational wave

detectors the cavities are usually resonant with the illuminating light and the effective detector bandwidth will be only half of this linewidth, $\Delta f_{\text{FP}} = \Delta f_{\text{FWHM}}/2$. This corresponds to 3 dB bandwidth in terms of the amplitude.

Another parameter often used to characterize a cavity is the *finesse* \mathcal{F} , the ratio of the free-spectral-range to the linewidth, i.e. $\mathcal{F} = \Delta f_{\text{FSR}}/\Delta f_{\text{FWHM}}$. In most of the interesting cases this is related to other parameters by

$$\left. \begin{aligned} \mathcal{F} &\simeq \frac{\pi\sqrt{\rho_c\rho_r}}{1-\rho_c\rho_r} \\ t_s &\simeq \frac{\mathcal{F}}{2\pi\Delta f_{\text{FSR}}} \\ \mathcal{E}_{\text{FP}} &\simeq \frac{2P_0\mathcal{F}}{\pi\Delta f_{\text{FSR}}} \end{aligned} \right\} \quad \text{for } 0 \ll \rho_c \leq \rho_r \cong 1. \quad (13)$$

Using the above approximations, another useful relation can be obtained:

$$\Delta f_{\text{FP}} = \Delta f_{\text{FSR}}/2\mathcal{F} \simeq 1/4\pi t_s. \quad (14)$$

Comparing this with the corresponding relation eq. (7) for delay lines, the bandwidth for Fabry-Perot cavities is less by a factor of $\sim 2\pi$ compared with that for delay lines with the same storage time.

As long as the above approximations are valid, the sensitivity of this type of interferometer is characterized by the finesse. The energy stored is proportional to the finesse, and the bandwidth is inversely proportional to the finesse. Combining these two relations results in the sensitivity being proportional to the finesse.

It is interesting to note that the energy stored in a cavity is not simply related to the storage time, i.e.

$$\mathcal{E}_{\text{FP}} = P_0 t_s \times \frac{\tau_c^2(1 + \rho_c\rho_r)}{1 - \rho_c\rho_r} \neq P_0 t_s. \quad (15)$$

This is because the *amplitude* just after the coupling mirror is the result of interference between two amplitudes, one from transmission of the injected and the other from reflection of the internal amplitude, which is in turn the result of the interference. In other words, the injected light experiences a higher reflectivity at the coupling mirror due to the interference between the freshly coming light and that stored inside the cavity. Due to this effect, the injected light will stay inside the cavity longer than the ‘storage time’ defined above.

The most interesting case occurs when the coupling mirror has a relatively high transmittance and the reflecting mirror has a high reflectivity. In this case the amplitude inside the cavity is almost twice (or four times in terms of power)

as much as the value one might expect from the storage time. This is exactly the case for an interferometric gravitational-wave detector using Fabry-Perot cavities, in which their finesse are kept low to retain a certain bandwidth. We will see the effect of this later.

Again to estimate the sensitivity, let us evaluate the energy stored in the interferometer. Using the definition given above, the energy is given by

$$\begin{aligned}\mathcal{E}_{\text{FP}} &\simeq P_0 \times \frac{2\mathcal{F}}{\pi\Delta f_{\text{FSR}}} \simeq P_0 \times \frac{1}{\pi} \frac{1}{\Delta f_{\text{FP}}} \\ &\approx 16 \text{ mJ} \times \left[\frac{P_0}{50 \text{ W}} \right] \left[\frac{\Delta f_{\text{FP}}}{1 \text{ kHz}} \right]^{-1}.\end{aligned}\tag{16}$$

Again, the sensitivity does not depend on the arm length ℓ explicitly but on the bandwidth which is determined by the combination of arm length and finesse, although the arm length ℓ should be reasonably long in order to satisfy the conditions in eq. (13),

Comparing the results of eq. (8) for the delay lines and eq. (16) for the Fabry-Perot cavities, it can be seen that a little less energy can be stored in Fabry-Perot cavities, when both configurations are adjusted to have the same bandwidth. Thus, a sensitivity worse by a factor of ~ 1.25 is expected. However, one will find that the sensitivity is actually even worse, by a factor of ~ 1.57 instead of 1.25. The reason for this will become clear when the two schemes are compared in more detail. For that purpose, introducing the frequency responses of the two schemes is helpful.

3.2.3: Frequency responses

The transfer function from gravitational wave signal to phase shift of the light returned from arms can be obtained by considering the process of the conversion (see Chapter 2). This phase shift of the light during the retarded time t_r is given by

$$\delta\phi(t) = \int_{t-t_r}^t \frac{1}{2} \omega_0 h(t) dt.\tag{17}$$

Taking the Laplace transformations of both sides yields

$$\begin{aligned}\mathcal{L}\{\delta\phi(t)\} &= \frac{\omega_0}{2} \mathcal{L}\left\{ \int_{t-t_r}^t h(t) dt \right\} \\ &= \frac{\omega_0}{2} \left[\mathcal{L}\left\{ \int_0^t h(t) dt \right\} - \mathcal{L}\left\{ \int_0^{t-t_r} h(t) dt \right\} \right] \\ &= \frac{\omega_0}{2} \frac{1 - e^{-st_r}}{s} \mathcal{L}\{h(t)\}.\end{aligned}\tag{18}$$

The ratio of Laplace transformations of phase shift $\delta\phi(t)$ and gravitational wave amplitude $h(t)$ gives the transfer function, thus

$$\begin{aligned} X_{h \rightarrow \phi}(\omega) &= \frac{\mathcal{L}\{\delta\phi(t)\}}{\mathcal{L}\{h(t)\}} = \frac{\omega_0}{2} \frac{1 - e^{-i\omega t_r}}{i\omega} \\ &= \frac{\omega_0}{2} t_r e^{-i\omega t_r/2} \frac{\sin(\omega t_r/2)}{\omega t_r/2}, \end{aligned} \quad (19)$$

when expressed in angular frequency ω .

As we have discussed in Chapter 2, phase modulation produces upper and lower sidebands, each having the same amplitude relative to the carrier of $J_1(\phi) \simeq \phi/2$ for $\phi \ll 1$ (see eq. (2.20)). It is these sidebands that we detect. Since the amplitudes of the sidebands must be proportional to that of the carrier, it is convenient to normalize the transfer function using the relative amplitude a/a_0 of the sidebands.

For a delay line with arm length ℓ and N beams in the arms, the retarded time is given by $t_r = N\ell/c$. The resultant sideband amplitude will be reduced by a factor of ρ^{N-1} due to the finite reflectivities of the mirrors. Combining these results, the normalized response of a delay line type interferometer is given by:

$$\begin{aligned} G_{\text{DL}}(\omega) &= \rho^{N-1} \times \frac{1}{2} X_{h \rightarrow \phi}(\omega) = \frac{\rho^{N-1}}{2} \frac{\omega_0}{2} \frac{1 - e^{-i\omega N\ell/c}}{i\omega} \\ &= \frac{\omega_0}{2} \frac{N\ell}{2c} \rho^{N-1} e^{-i\omega N\ell/2c} \frac{\sin(\omega N\ell/2c)}{\omega N\ell/2c}. \end{aligned} \quad (20)$$

For the case of Fabry-Perot cavities it is convenient to consider the reflected light as a sum of many contributions which experienced different storage times in the cavity:

$$G_{\text{FP}}(\omega) = -\tau_c^2 \sum_{n=1}^{\infty} \left[\rho_c^{n-1} \rho_r^n \times \frac{1}{2} X_{2n}(\omega) \right], \quad (21)$$

where $X_{2n}(\omega)$ is the transfer function for light that experienced $2n$ transits in the cavity and can be obtained by inserting $t_r = 2n \times \ell/c = nt_a$ into eq. (19). Using this, eq. (21) can be expanded as

$$\begin{aligned} G_{\text{FP}}(\omega) &= -\tau_c^2 \sum_{n=1}^{\infty} \frac{\rho_c^{n-1} \rho_r^n}{2} \frac{\omega_0}{2} \frac{1 - e^{-in\omega t_a}}{i\omega} \\ &= -\frac{\omega_0}{i4\omega} \frac{\tau_c^2}{\rho_c} \sum_{n=1}^{\infty} (\rho_c \rho_r)^n (1 - e^{-in\omega t_a}) \\ &= -\frac{\omega_0}{i4\omega} \frac{\tau_c^2}{\rho_c} \left(\frac{\rho_c \rho_r}{1 - \rho_c \rho_r} - \frac{\rho_c \rho_r e^{-i\omega t_a}}{1 - \rho_c \rho_r e^{-i\omega t_a}} \right) \\ &= \frac{-\tau_c^2}{(1 - \rho_c \rho_r)(1 - \rho_c \rho_r e^{-i\omega t_a})} \times \frac{\rho_r \omega_0}{4} \frac{1 - e^{-i\omega t_a}}{i\omega}, \end{aligned} \quad (22)$$

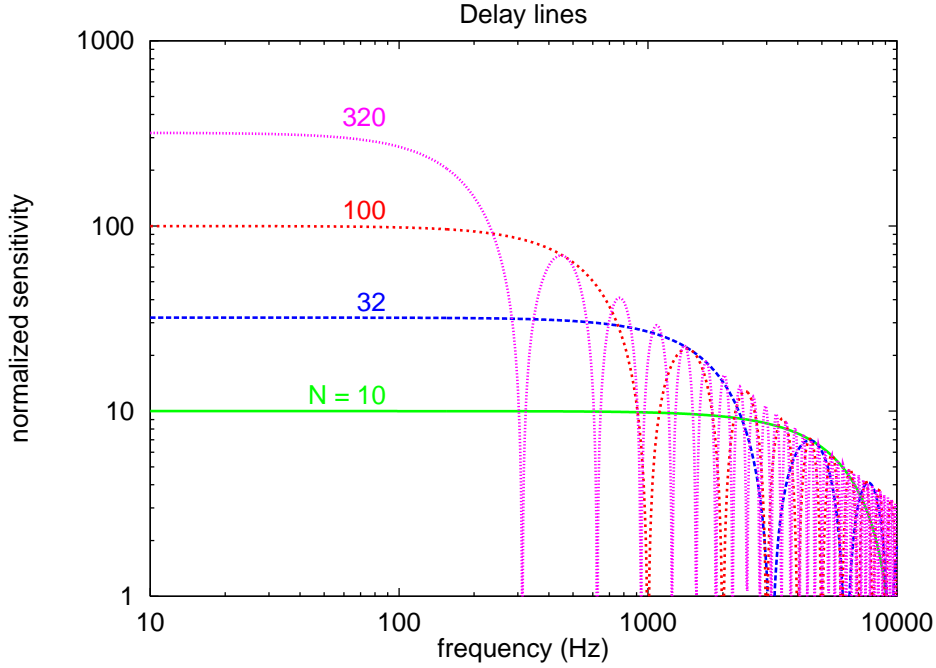


Figure 3.3: The normalized response of interferometers with delay lines in the arms. Each plot corresponds to a different number of transits in the arm. For each plot the sensitivity is almost flat from DC to a certain cut-off frequency which is determined by the storage time in the arms. This gives the approximate bandwidth. In the higher frequency range, notches (frequencies where sensitivity is zero) appear periodically. The height of the peaks between the notches decrease with the frequency ($\propto 1/\omega$). As for the unit of the vertical axis and the assumption in plotting these curves, see Appendix B.

where $t_a=2\ell/c$ is the transit time of one return trip in the arm. Note that this can be separated into two factors, the second of which has the same form as that for a single bounce Michelson interferometer. The first factor can be thought of as the effect of the Fabry-Perot cavity.

The normalized responses of interferometers with either delay lines or Fabry-Perot cavities are shown in Fig. 3.3 and Fig. 3.4, respectively. Each showing several plots which are different in storage time. The frequency response of a delay line has a sharp cut-off after the flat region in the low frequency range and many notches at higher frequencies. In contrast, that for a Fabry-Perot cavity has a gentle curve around the ‘shoulder’ of the peak without any notches inside the relevant frequency range. The sensitivity at a particular frequency does not depend so much on the storage time if the bandwidth (determined by the storage time) is broader than $\Delta f_{\text{FP}} \lesssim \omega_g/4\pi$.

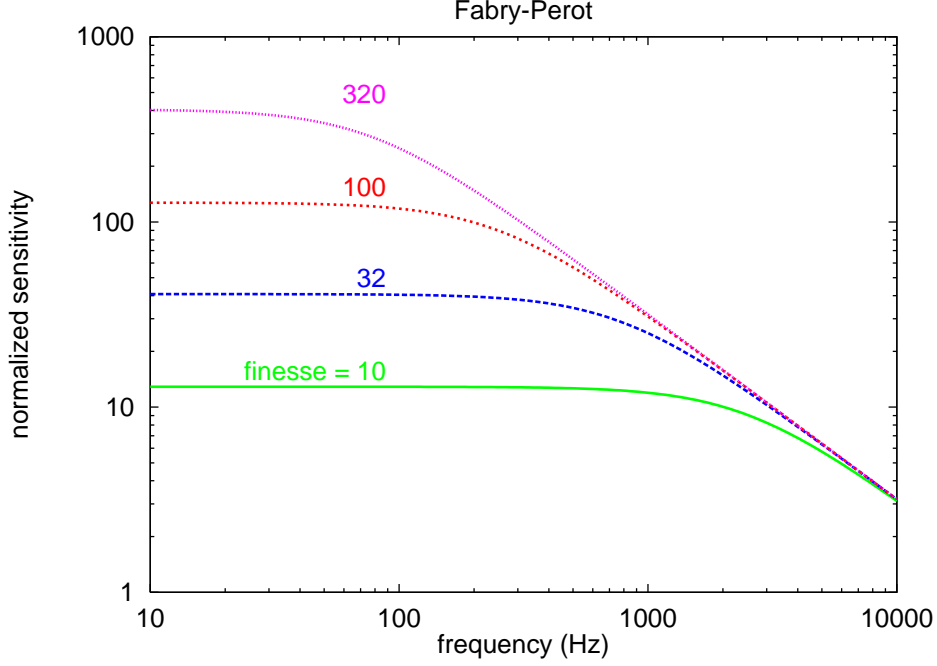


Figure 3.4: The normalized response of interferometers with Fabry-Perot cavities in the arms. Each plot corresponds to different finesse of the cavity in the arm. Similarly to the delay line case Fig. 3.3, the sensitivity is flat from DC to a certain cut-off frequency which is determined by the storage time in the arms. The difference is that there is no significant notch (within the frequency range shown) and the curve around the cut-off frequency is more gentle than in the delay line case. This makes the definition of the approximate bandwidth more difficult; it is common to define it at the frequency where the height is $1/\sqrt{2}$ or -3 dB of that at the peak (in this case, DC). The slope in the higher frequency range is, again, proportional to $1/\omega$.

It is known that a simple Fabry-Perot cavity with reasonably high finesse ($\mathcal{F} \gg 1$) behaves like a one-pole low-pass filter. This can be derived from eq. (22) assuming $(\omega/2\pi)/\Delta f_{\text{FSR}} \ll 1$,

$$\begin{aligned}
 G_{\text{FP}}(\omega) &\simeq -\frac{\tau_c^2}{(1-\rho_c\rho_r)[1-\rho_c\rho_r(1-i\omega t_a)]} \times \frac{\rho_r\omega_0}{4} \frac{1-(1-i\omega t_a)}{i\omega} \\
 &= -\frac{\tau_c^2/(1-\rho_c\rho_r)}{1-\rho_c\rho_r+i\rho_c\rho_r\omega t_a} \times \frac{\rho_r\omega_0 t_a}{4} \\
 &= -\frac{\tau_c^2 \rho_r \omega_0 t_a}{4(1-\rho_c\rho_r)^2} \times \left[1 + i\frac{\rho_c\rho_r}{1-\rho_c\rho_r}\omega t_a\right]^{-1}.
 \end{aligned} \tag{23}$$

This corresponds to the transfer function of a low-pass filter, with a pole at $\Delta f_{\text{FSR}}/(1-\rho_c\rho_r) \simeq \Delta f_{\text{FSR}}/2\mathcal{F} = \Delta f_{\text{FP}}$ and a DC gain of $-\frac{\tau_c^2 \rho_r}{4(1-\rho_c\rho_r)^2}\omega_0 t_a$.

The reason for the difference in the two configurations can be understood when the process of storing photons is considered. In case of a delay line, all the photons experience the same storage time which is optimum for a gravitational wave signal with frequency $\omega_g/2\pi = 1/2t_s$. On the other hand, for a Fabry-Perot cavity some of the photons may exit after just a few transits between the two mirrors and some may experience a longer storage time than the average. This makes the storage less efficient for the gravitational wave signal with a particular frequency, because neither shorter nor longer storage times are appropriate for that frequency.

3.2.4: Comparison of delay-lines and Fabry-Perot cavities

Initially it was thought that an interferometer with Fabry-Perot cavities in the arms has a sensitivity similar to that with delay lines, if both have the same storage time. In fact, there are a couple of interesting differences. When one attempts to compare the two configurations, the following problem arises: *Under which condition is it appropriate to make a comparison?* Some of the parameters should be adjusted in order to make a meaningful comparison. The relations among these parameters are, however, different in the two schemes, i.e. if one of the parameters were made to be the same then the others might not be the same as exemplified in eqs. (6)–(8) and (11)–(15).

Let us see this by taking the storage time as an example. If both configurations have the same storage times, the energy stored in a Fabry-Perot system is nearly four times as high as that in a delay line system, due to the interference at the coupling mirror. Furthermore, the bandwidths in the two configurations differ by a factor of 2π . As a result, a Fabry-Perot system has a much narrower bandwidth and higher peak sensitivity than a delay line system. Obviously this is not a good basis on which to make a comparison.

One possibility for making a better comparison is, as has been done above, to make both schemes have the same bandwidth. We have already seen the result of this case. Another interesting possibility would be to compare two configurations assuming that both store the same amount of energy in the system. Let us consider the frequency responses for this case, assuming, for simplicity, no loss.

Fig. 3.5 shows the frequency responses for the delay line and Fabry-Perot cases, both storing the same amount of energy. Under this condition, the ‘bandwidth’ of a delay line configuration is broader by a factor of ~ 1.57 compared

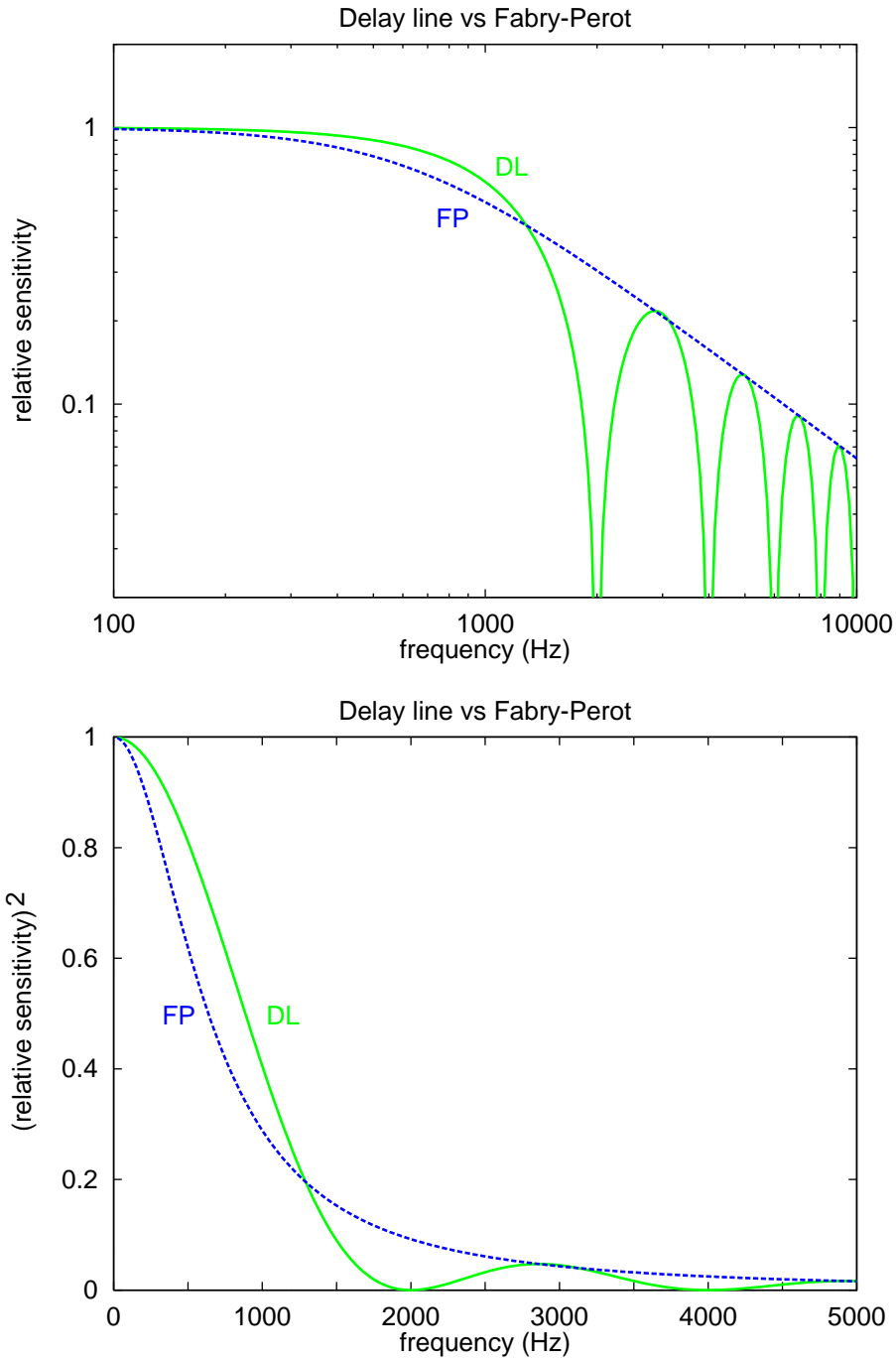


Figure 3.5: The relative sensitivity of interferometers with delay lines and Fabry-Perot cavities in the arms, assuming the same amount of energy being stored in both cases. The upper picture shows it in logarithmic scale and the lower one is the *squared* response in linear scale. The peak sensitivity at DC is the same in both cases (the vertical scale is normalized with respect to this peak sensitivity). The one for the delay line has a somewhat (~ 1.57) broader bandwidth and thus the sensitivity around the cut-off frequency is better. On the other hand, the one using Fabry-Perot cavities has a better sensitivity in the higher frequency range due to the notches of delay line response. One may notice in the lower picture that the area surrounded by the curve and the axes are the same in both cases.

with the Fabry-Perot one, while the peak sensitivity at DC is the same in both cases. On the other hand, the sensitivity outside the ‘bandwidth’ is better in the Fabry-Perot case because of the notches in the delay line response.

As one might expect from these plots, the area enclosed by the curve and the axes is the same in both cases. More precisely speaking, the frequency integration of the squared response is the same if both have the same amount of energy stored. This is the most important result of this comparison and we will discuss this point further in 3.5.1.

Because the area and the peak height are the same in both cases, it would be useful to find the ‘width’ of a box-shaped function with the same height and area. It is by accident that the ‘approximate bandwidth’ of the delay line response is exactly the same as this box width, whereas that of the Fabry-Perot response is, as mentioned above, narrower by about ~ 1.57 . This explains the inconsistency we found in 3.2.2; the inappropriate definition of the bandwidth of the Fabry-Perot response caused the overestimation of the sensitivity by a factor of $\sim \sqrt{1.57} \approx 1.25$.

It may be worth pointing out that this box width corresponds to what is called ‘*noise bandwidth*’ [43] in the field of filter designing. It is known in this field that the noise bandwidth for a one-pole low-pass filter is given by

$$\Delta f_{\text{NB}} = \frac{\pi}{2} \Delta f_{3\text{dB}} . \quad (24)$$

As is mentioned in 3.2.3, the frequency response of a Fabry-Perot cavity is similar to that of a one-pole low-pass filter. Thus, the relation between noise bandwidth and the 3 dB bandwidth shown above is also a good approximate relation for a Fabry-Perot cavity, explaining the factor of 1.57 in the above discussion. Using the above relation and eq. (16), the sensitivity of interferometers with Fabry-Perot cavities in the arms can be written as

$$\tilde{h}_{\text{FP}} \approx 1.5 \cdot 10^{-22} / \sqrt{\text{Hz}} \times \left[\frac{\lambda}{1 \mu\text{m}} \right]^{\frac{1}{2}} \left[\frac{P_0}{50 \text{ W}} \right]^{-\frac{1}{2}} \left[\frac{\Delta f_{\text{FP}}}{1 \text{ kHz}} \right]. \quad (25)$$

Comparing eqs. (9) and (25), the sensitivity with a given bandwidth seems to be better in an interferometer with delay lines. Note, however, that because of the notches at higher frequencies, delay lines must have the storage time chosen according to the highest frequency $\omega_g/2\pi$ of interesting gravitational wave spectra, given by $\Delta f_{\text{DL}} \simeq \omega_g/2\pi$. On the other hand, Fabry-Perot cavities can increase the storage time (and the amount of energy stored) without sacrificing

the sensitivity at high frequencies. In this respect Fabry-Perot cavities are advantageous over delay lines, although this conclusion is not valid when power recycling is adopted.

3.3: Power recycling

3.3.1: Principle of power recycling

The sensitivities expected from interferometers using delay line eq. (9) and Fabry-Perot cavities eq. (25) are not so promising for gravitational wave observatories. The realistic possibilities left to improve the sensitivity are to increase the light power and to make the bandwidth narrower. The latter is, however, difficult for delay lines as is described in 3.2.1. The former seems also difficult when one recalls the requirements on the stabilities for the light source and the present technology of lasers.

The possibility of increasing the *effective* light power incident on the interferometer with given light power from the source has been proposed independently by Drever [38, 39] and by Schilling as one of the solutions for this problem. The key concept of this is to utilize the wasted power reflected from a low loss interferometer by using a little more complicated configuration. This method is called *power recycling*.

As is discussed in Chapter 2, interferometric gravitational-wave detectors are in general operated at a *dark fringe*, or interference minimum, where in principle the sensitivity can be the best. Under this condition, all the carrier goes to the non-detecting output (carrier port) and all the differential signal, that a gravitational wave is expected to produce, goes to the detecting output (signal port). It should be emphasized that the light leaking out of the interferometer through the carrier port is not used at all.

If the losses in the interferometer are low enough, most of the injected light power leaks out of the interferometer through the carrier port. By superimposing this output onto the ‘fresh’ light coming from the source, the effective light power incident on the interferometer can be increased.

If the interferometer is on-axis, i.e. the beam returning from the interferometer has the same axis as that of the incoming beam, power recycling can be realized by putting a mirror (*power recycling mirror*) between the beamsplitter and the light source, as is shown in Fig. 3.6. This corresponds to forming a

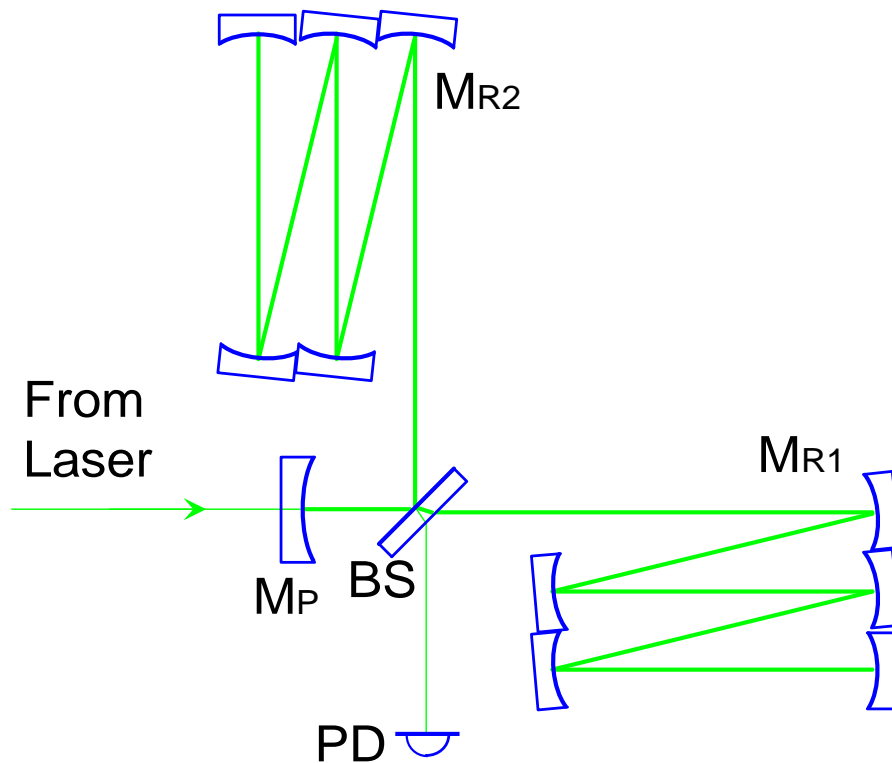


Figure 3.6a: Optical layout of a power recycled interferometer with multi-reflections in the arms. The interferometer is operated at a *dark fringe*, i.e. the detecting output is kept at a minimum of interference. The additional mirror M_P (*power recycling mirror*) placed between the light source and the beamsplitter effectively composes a cavity (*power recycling cavity*) with the interferometer.

cavity (*power recycling cavity*) with the interferometer as one of the cavity mirrors. This cavity should, of course, be kept resonant with the incoming beam to achieve the maximum improvement in the sensitivity.

If the interferometer is off-axis, it is necessary to use a couple of mirrors to bring the outgoing beam back to the axis of the incoming one. This forms a ring-type cavity, resulting in the same effect on the performance of the interferometer with, however, a little more complexity.

The effect of power recycling is equivalent to the use of a more powerful light source to illuminate the interferometer, i.e. it does not affect the shape of the transfer function of the interferometer. This is because the power recycling cavity contains only the carrier light but all differential signals escape at the beamsplitter of the interferometer. Here is the most innovative point (from the present viewpoint) of power recycling: *It is possible to set different storage times for the carrier and for the signal sidebands.*

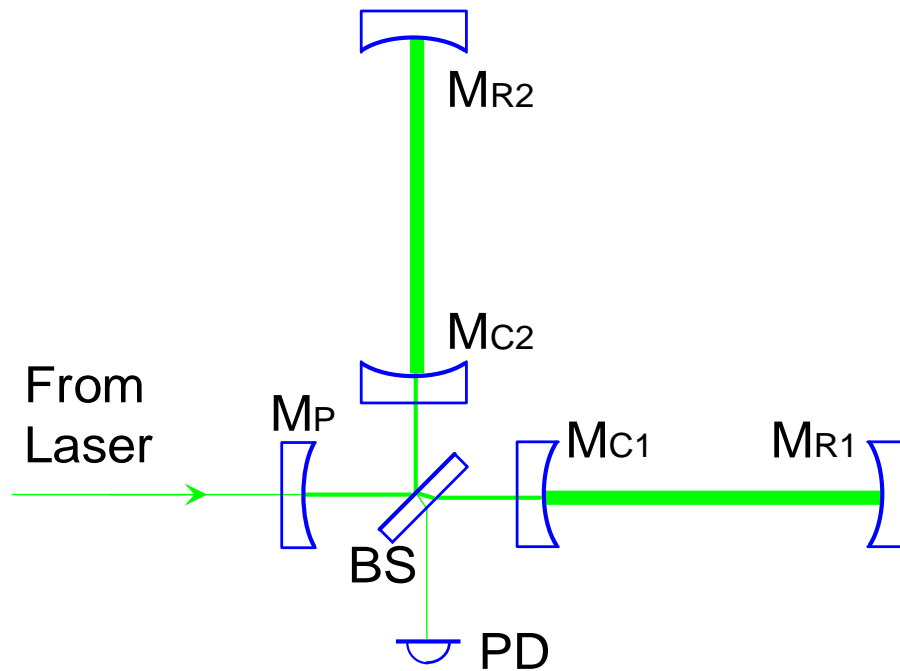


Figure 3.6b: Optical layout of a power recycled interferometer with Fabry-Perot cavities in the arms. As in the multi-reflection case, the power recycling mirror M_P is added in order to compose the power recycling cavity with the interferometer. In this case, the power recycling cavity is a *three-mirror coupled cavity* which is analyzed in 4.2.4.

Formerly, increasing the storage time in the arms always involved narrowing the detector bandwidth, without any improvement in sensitivity at frequencies outside the bandwidth. By using power recycling, it is possible to increase the storage time for the carrier without narrowing the detector bandwidth. On the other hand, the improvement in sensitivity due to power recycling is proportional only to the square root of the increased storage time for the carrier. This is different from the case of an increased storage time in the arms, i.e. in delay lines or Fabry-Perot cavities.

So from now on we must separate the two storage times, one for the carrier and the other for the signal. The former determines the energy inside the interferometer (for a given illuminating power), and the latter determines the bandwidth. As can be seen from eq. (3), these parameters determine the sensitivity in principle.

3.3.2: Power recycling with delay lines

Since the shot noise limited sensitivity is inversely proportional to the square root of the power, the normalized response of a power recycled interferometer can be obtained by multiplying the normalized response without power recycling with the square root of *equivalent power gain*, which is the ratio of injected power to that inside the power recycling cavity³⁾. The reflectivity of the power recycling mirror and the losses in the interferometer determine the equivalent power gain and the associated improvement in sensitivity by power recycling.

To realize the maximum equivalent power gain, the power transmittance of the recycling mirror must be chosen as (see eqs. (4.6) and (4.8) in 4.2.2)

$$\begin{aligned} |\tau_{\text{PR}}|^2 &= 1 - \mathcal{A}_{\text{PR}} - |\rho_{\text{PR}}|^2 = (1 - \mathcal{A}_{\text{PR}}) [1 - (1 - \mathcal{A}_{\text{PR}})(1 - \mathcal{A}_{\text{int}})] \\ &\simeq \mathcal{A}_{\text{int}} + \mathcal{A}_{\text{PR}}, \end{aligned} \quad (26)$$

where \mathcal{A}_{PR} is the loss of the power recycling mirror, and \mathcal{A}_{int} is the whole loss inside the interferometer. This choice provides an equivalent power gain \mathcal{G}_{PR} of

$$\mathcal{G}_{\text{PR}} = \frac{1}{\mathcal{A}_{\text{int}} + \mathcal{A}_{\text{PR}}/(1 - \mathcal{A}_{\text{PR}})} \simeq \frac{1}{\mathcal{A}_{\text{int}} + \mathcal{A}_{\text{PR}}}, \quad (27)$$

which will be limited by the loss in the interferometer.

The causes of the loss in the interferometer are the finite reflectivities of the mirrors, imperfect interference, and scattering and absorption in the traversed optical substrates. Thus the loss in the interferometer is given by

$$1 - \mathcal{A}_{\text{int}} = (\rho_{\text{arm}})^2 \times (1 - \mathcal{A}_{\text{BS}}) \times \frac{1 + \mathcal{C}}{2}, \quad (28)$$

where $(\rho_{\text{arm}})^2$ is the power reflectivity of the arm, \mathcal{A}_{BS} the loss at beamsplitter, and \mathcal{C} the apparent visibility of the interferometer. In the ideal case the last two factors can be unity. If each factor is close to unity (which we will assume), an approximation

$$\mathcal{A}_{\text{int}} \simeq [1 - (\rho_{\text{arm}})^2] + \mathcal{A}_{\text{BS}} + \frac{1 - \mathcal{C}}{2} \quad (29)$$

can be used.

Assuming a fixed arm length, increasing the storage time in the arms will require more reflections in the arms. Due to the finite reflectivity of the mirrors,

³⁾ The term ‘power inside a cavity’ may not be strictly defined, since the field inside the cavity is a sort of ‘standing wave’ and there is no energy being transported. Nevertheless, the number of photons hitting on the surface of mirrors per unit time can be considered as the definition of the power inside the cavity.

this results in an increase of the loss in the arms, which in turn results in a reduction of the equivalent power gain. In the end, the light energy stored in the system scarcely depends on the storage time in the arms, as long as the losses in the arms dominate over other losses. Under this condition, the resultant energy stored in the system will be close to the value obtained by eq. (5).

Let us first consider the delay line case. The equivalent power gain can be obtained as

$$\mathcal{G}_{\text{DL}}^{\text{PR}} = \frac{1}{\frac{1}{1-\mathcal{A}_{\text{PR}}} - (1-\mathcal{A}_{\text{int}})} = \frac{\rho^2}{1 - (\rho^2)^N (1-\mathcal{A}_{\text{BS}}) \frac{1+\mathcal{C}}{2}}, \quad (30)$$

where ρ is the reflectivity of the delay line mirrors which must be chosen as high as possible, $1-\rho^2 = \mathcal{A}_{\text{min}}$, and the loss of the power recycling mirror is also assumed to be the minimum, $\mathcal{A}_{\text{PR}} = \mathcal{A}_{\text{min}} = 1-\rho^2$. The amount of energy in a power recycled interferometer with delay lines is given by

$$\mathcal{E}_{\text{DL}}^{\text{PR}} = P_{\text{BS}} \times t_{\text{s}} = \mathcal{G}_{\text{DL}}^{\text{PR}} P_0 \times \frac{\ell}{c} \frac{1 - \rho^{2N}}{1 - \rho^2}. \quad (31)$$

The normalized response of this interferometer can be obtained by multiplying that of delay line by the square root of equivalent power gain:

$$G_{\text{DL}}^{\text{PR}}(\omega) \Big|_{h \rightarrow a/a_0} = \sqrt{\mathcal{G}_{\text{DL}}^{\text{PR}}} \times G_{\text{DL}}(\omega) \Big|_{h \rightarrow a/a_0}. \quad (32)$$

The loss in delay line is determined by the reflectivity of the mirrors as $\mathcal{A}_{\text{DL}} = 1-(\rho^2)^{N-1}$. Assuming the ideal condition (i.e. the minimum loss for each mirror but no loss at beamsplitter and perfect visibility), the maximum energy stored is

$$\mathcal{E}_{\text{DL}}^{\text{max}} = \frac{P_0 \ell}{c} \frac{1 - \rho^{2N}}{1 - \rho^2} \frac{\rho^2}{1 - \rho^{2N}} = \frac{P_0 \ell}{c} \frac{\rho^2}{1 - \rho^2}, \quad (33)$$

which reduces to eq. (5). As we expected, this does not depend on N , the number of transits in the delay line. In practice, however, the finite loss at the beamsplitter and the imperfect contrast increase \mathcal{A}_{int} and reduce the amount of

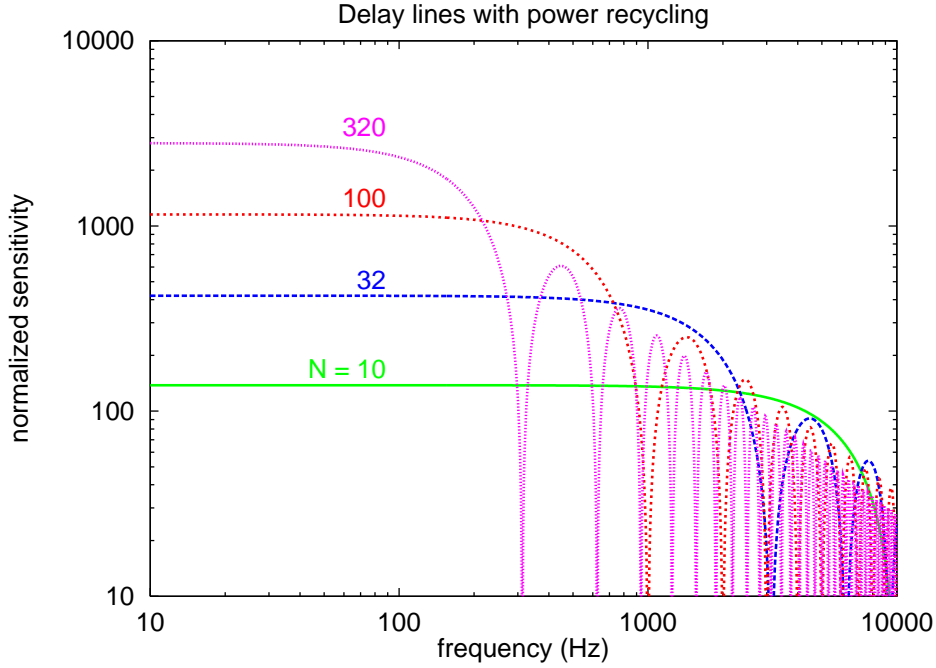


Figure 3.7: The normalized response of power recycled interferometers with delay lines in the arms. Each plot corresponds to different number of transits in the arm. The shape of the curve does not differ from that without power recycling. The peak sensitivity is, however, improved by a significant factor which depends on the number of reflections in the arms. The unit of the vertical axis and the assumption in plotting these curves are the same as those for non-power recycled interferometers, and described in Appendix B.

energy to less than this value. This will be discussed in 3.3.3 with the maximum energy in Fabry-Perot case.

Fig. 3.7 shows the normalized responses of power recycled interferometers with delay lines, with different numbers of reflections in the arms. It is obvious that there is an optimum signal storage time for the signal with a particular frequency range. Assuming reasonably low loss mirrors and a long enough arm length, the best sensitivity for a signal with frequency $\omega_g/2\pi$ will be obtained by choosing the bandwidth as

$$\Delta f_{\text{DL}} \approx 1.35 \cdot \omega_g/2\pi, \quad (34)$$

and the optimized sensitivity at that frequency will be

$$\tilde{h}_{\text{DL}}^{\text{opt}}(\omega_g) \approx 4.9 \cdot 10^{-24}/\sqrt{\text{Hz}} \times \left[\frac{\lambda}{1 \mu\text{m}} \right]^{\frac{1}{2}} \left[\frac{\mathcal{E}}{20 \text{ J}} \right]^{-\frac{1}{2}} \left[\frac{\omega_g/2\pi}{1 \text{ kHz}} \right]^{\frac{1}{2}}. \quad (35)$$

In practice, however, this is difficult to realize because of the limitation on the mirror size. Thus, for lower frequencies we must consider the use of Fabry-Perot cavities.

3.3.3: Power recycling with Fabry-Perot cavities

In a similar way to the delay line case, the energy stored in a power recycled interferometer with Fabry-Perot cavities can be obtained as

$$\mathcal{E}_{\text{FP}}^{\text{PR}} = \mathcal{G}_{\text{FP}}^{\text{PR}} P_0 \times \frac{\ell}{c} \frac{\tau_c^2 (1 + \rho_r^2)}{(1 - \rho_c \rho_r)^2}, \quad (36)$$

and its normalized response as

$$G_{\text{FP}}^{\text{PR}}(\omega) = \sqrt{\mathcal{G}_{\text{FP}}^{\text{PR}}} \times G_{\text{FP}}(\omega), \quad (37)$$

where $\mathcal{G}_{\text{FP}}^{\text{PR}}$ is the equivalent power gain which is determined by the losses in the interferometer as in eq. (27). When Fabry-Perot cavities are used in the arms, the total loss inside the interferometer will be

$$\begin{aligned} \mathcal{A}_{\text{int}} &= 1 - (\rho_{\text{FP}})^2 (1 - \mathcal{A}_{\text{BS}}) \frac{1 + \mathcal{C}}{2} \\ &= 1 - \left[\frac{\rho_c - (1 - \mathcal{A}_c) \rho_r}{1 - \rho_c \rho_r} \right]^2 (1 - \mathcal{A}_{\text{BS}}) \frac{1 + \mathcal{C}}{2} \\ &\simeq 1 - \left[\frac{\rho_c - (1 - \mathcal{A}_c) \rho_r}{1 - \rho_c \rho_r} \right]^2 + \mathcal{A}_{\text{BS}} + \frac{1 - \mathcal{C}}{2}, \end{aligned} \quad (38)$$

where \mathcal{A}_c is the loss of the coupling mirror, and ρ_c and ρ_r are the reflectivities of it and the reflecting mirror, respectively. Using this, the equivalent power gain can be written as

$$\mathcal{G}_{\text{FP}}^{\text{PR}} = \frac{1 - \mathcal{A}_{\text{PR}}}{1 - \left[\frac{\rho_c - (1 - \mathcal{A}_c) \rho_r}{1 - \rho_c \rho_r} \right]^2 (1 - \mathcal{A}_{\text{PR}}) (1 - \mathcal{A}_{\text{BS}}) \frac{1 + \mathcal{C}}{2}}. \quad (39)$$

Let us assume an optimum case in which the Fabry-Perot reflecting mirror has the highest reflectivity possible, $\rho_r = \rho$, the minimum loss for the Fabry-Perot coupling mirror and the power recycling mirror $\mathcal{A}_c = \mathcal{A}_{\text{PR}} = 1 - \rho^2$, no loss at the

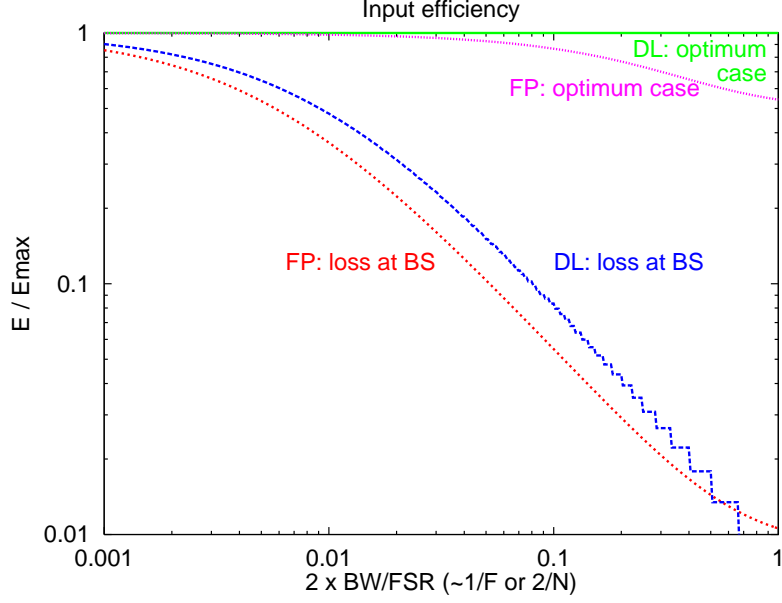


Figure 3.8: The efficiency of storing energy in each configuration. The vertical axis is the *input efficiency*, the amount of stored energy normalized by the maximum value given in eq. (5). The (visible) top curve corresponds to the power recycled interferometer with Fabry-Perot cavities assuming the optimum case, i.e. the minimum loss for each mirror and no loss at the beamsplitter and inside the coupling mirror substrate. In this case the amount of energy stored can be close to the maximum value. As shown in eq. (33), the input efficiency of a power recycled interferometer with delay lines is unity in the optimum case. In practice, however, the amount of stored energy is always limited by the loss at the beamsplitter and inside the coupling mirrors (for Fabry-Perot case), as is illustrated by the lower curves. The horizontal axis is twice the bandwidth normalized by the free-spectral-range, and its reciprocal corresponds to the finesse or half the number of transits in the arm. The assumptions are the same as those to plot the normalized responses, and are fully described in Appendix B.

beamsplitter $\mathcal{A}_{\text{BS}}=0$, and the perfect visibility $\mathcal{C}=1$. Then the amount of energy stored will be given by

$$\begin{aligned}
 \mathcal{E}_{\text{FP}}^{\text{max}} &= \frac{P_0 \ell}{c} \times \frac{\rho^2}{1 - \rho^2 \left(\frac{\rho_c - \rho^3}{1 - \rho_c \rho} \right)^2} \times \frac{\tau_c^2 (1 + \rho^2)}{(1 - \rho_c \rho)^2} \\
 &= \frac{P_0 \ell}{c} \times \frac{\rho^2 \tau_c^2 (1 + \rho^2)}{(1 - \rho_c \rho)^2 - \rho^2 (\rho_c - \rho^3)^2} \\
 &= \frac{P_0 \ell}{c} \frac{\rho^2}{1 - \rho^2} \times \frac{\rho^2 - \rho_c^2}{1 - 2\rho_c \rho + \rho^4},
 \end{aligned} \tag{40}$$

where $\tau_c^2 = 1 - \mathcal{A}_c - \rho_c^2 = \rho^2 - \rho_c^2$ was used. In this result, there is an additional factor to the value in eq. (5), which will be referred to as the *input efficiency*.

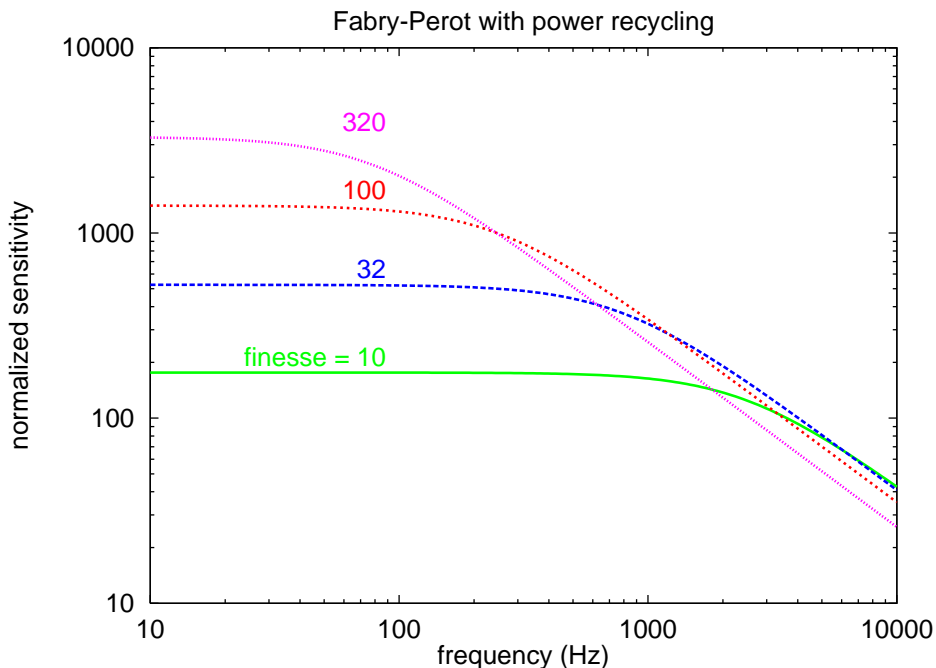


Figure 3.9: The normalized response of power recycled interferometers with Fabry-Perot cavities in the arms. Each plot corresponds to different finesse in the arm. Again, the shape of the curve does not differ from that without power recycling. The peak sensitivity is improved by power recycling and it can be seen that there is an optimum finesse for a particular frequency range.

For a general purpose, it is appropriate to define the input efficiency as the ratio of the amount of energy stored relative to the maximum value given in eq. (5). Fig. 3.8 shows the dependence of the input efficiencies on the bandwidth, assuming the optimum case and some more realistic cases. Even in the optimum case, this factor is always less than unity for the interferometers with Fabry-Perot cavities; in other words, the amount of energy stored is always less than the maximum value given in eq. (5). The dependence is, however, relatively small and almost negligible in a realistic condition $\mathcal{F} \gtrsim 20$ or $\rho_c^2 \gtrsim 0.75$.

In practice, the input efficiency will be more strongly affected by the loss at the beamsplitter, as can be seen in Fig. 3.8. This is true for interferometers either with delay lines or Fabry-Perot cavities. When Fabry-Perot cavities are present in the arms, however, it will be more significant since there are losses in the substrate of Fabry-Perot coupling mirrors which have the same effect as the loss at the beamsplitter. The value given in eq. (5) is the *maximum* number and it may be a good approximation if

$$\mathcal{A}_{\text{BS}} + \frac{1-\mathcal{C}}{2} \ll 1 - (\rho_{\text{arm}})^2, \quad (41)$$

i.e. the loss in the arms dominates the total loss of the interferometer.

Fig. 3.9 shows the normalized responses of power recycled interferometers with Fabry-Perot cavities, each plot with a different reflectivity for the coupling mirror. As in the delay line case, there is an optimum signal storage time for the signal with a particular frequency. For the signal with frequency $\omega_g/2\pi$, it is best to choose the bandwidth of Fabry-Perot cavity

$$\Delta f_{\text{FP}} \approx \omega_g/2\pi, \quad (42)$$

and the optimized sensitivity at that frequency will be

$$\tilde{h}_{\text{FP}}^{\text{opt}}(\omega_g) \approx 5.9 \cdot 10^{-24}/\sqrt{\text{Hz}} \times \left[\frac{\lambda}{1 \mu\text{m}} \right]^{\frac{1}{2}} \left[\frac{\mathcal{E}}{20 \text{ J}} \right]^{-\frac{1}{2}} \left[\frac{\omega_g/2\pi}{1 \text{ kHz}} \right]^{\frac{1}{2}}, \quad (43)$$

as long as the condition eq. (41) is satisfied. This is a little worse than the value obtained for delay lines. However, the sensitivities at other frequencies are better in the Fabry-Perot case due to the gentle curve around the shoulder of the peak.

3.4: Other schemes

3.4.1: Signal recycling: Broad-band operation

Signal recycling proposed by Meers [44] is another way to improve the sensitivity of a Michelson interferometer. The basic idea is to increase the signal storage time at the expense of a bandwidth narrower than that determined by the storage time in the arms. The fact that the signal storage time can be increased more than the storage time in the arms means that it is not limited by the loss in the arms. The carrier storage time could also be increased by using the power recycling scheme. Thus, the best sensitivity might be obtained by employing both signal recycling and power recycling. When these two schemes are combined together, it is often called *dual recycling* because both carrier and signal sidebands are recycled.

The optical layout of this scheme is shown in Fig. 3.10. If the interferometer is on-axis, signal recycling is realized by putting another mirror, the *signal recycling mirror*, at the detecting output (signal port) of the interferometer before the photo-detector. This is quite similar to the case of power recycling, with the only difference being the position of the additional mirror. In this case a cavity (*signal recycling cavity*) is formed in which the signal sidebands will

be stored instead of the carrier. If the interferometer is off-axis, it should be possible to form a ring-type cavity for the signal sidebands by using a couple of mirrors, although the help from an extra beam for the alignment would be necessary. The effect on the sensitivity does not depend on how the recycling is realized.

Signal recycling can in principle be incorporated into interferometers using either delay lines or Fabry-Perot cavities in the arms. In the former case the cavity storing the signal sidebands will be a ‘split Fabry-Perot’, a cavity with one mirror being an interferometer locked to a dark fringe. As long as the difference in the lengths of the two arms is negligible, this behaves as a usual Fabry-Perot cavity.

With Fabry-Perot cavities present in the arms, the situation becomes more complicated. The cavity storing the signal sidebands is again a split cavity which is composed of one mirror and a pair of Fabry-Perot cavities in the two arms of a Michelson interferometer. Even if we assume the two arms including the cavities are identical and thus the ‘split’ part is negligible, the rest will behave as a ‘*three-mirror coupled cavity*’—a cavity composed of three mirrors aligned to an optical axis.

Fortunately this can be approximated by a simple Fabry-Perot cavity, resulting in a performance similar to that of the delay line case. This is true only when the Fabry-Perot cavities in the arms have a bandwidth much broader than that of the coupled cavity. This point will be discussed in more detail in 4.3.2 describing resonant sideband extraction, where this condition is not satisfied.

As a consequence, the behavior of a dual-recycled interferometer (either with delay lines or Fabry-Perot cavities in the arms) is similar to that of an interferometer with Fabry-Perot cavities. The difference from that of Fabry-Perot case is that the signal recycling mirror does not affect the carrier and thus there is no enhancement of the power inside the cavity. On the other hand, the amount of energy stored does not change from that without the signal recycling mirror, which is given in eq. (31) or eq. (36). Thus, the sensitivity can be estimated when the bandwidth is determined.

The normalized response of a dual-recycled interferometer can be obtained by ‘generalizing’ the normalized response of an interferometer with Fabry-Perot cavities eq. (22) as follows:

$$G_{\text{DR}}(\omega) = \sqrt{\mathcal{G}_{\text{arm}}} \times \frac{-\tau_s}{1 - \rho_s \rho_a e^{-i\omega t_r}} \times \frac{\rho_a \omega_0}{4} \frac{1 - e^{-i\omega t_r}}{i\omega}, \quad (44)$$

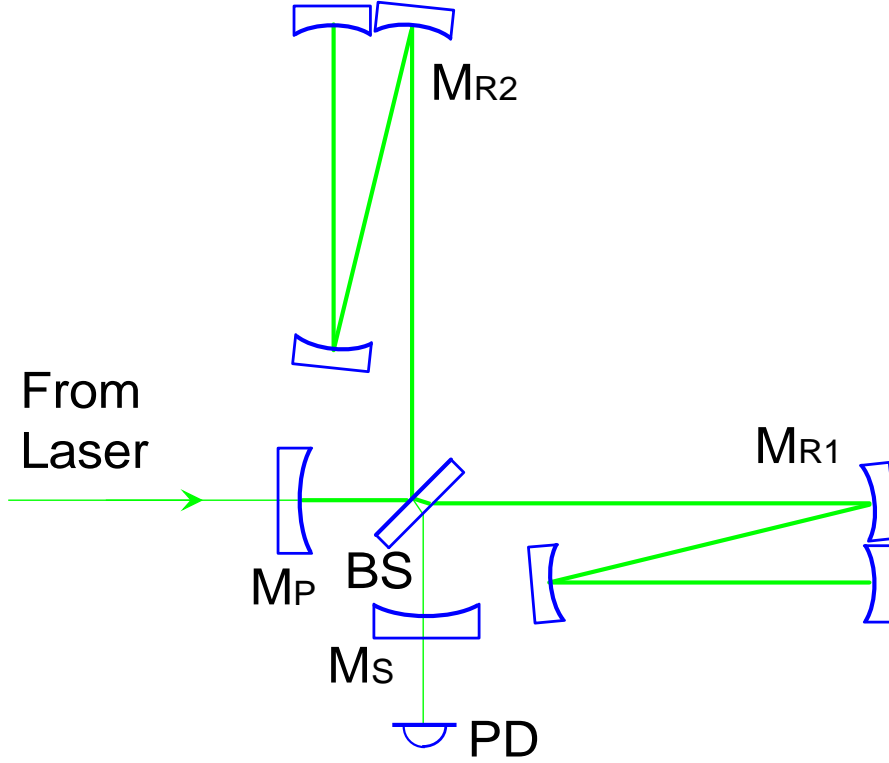


Figure 3.10a: Optical layout of a dual-recycled interferometer with multi-reflections in the arms. There is an additional mirror M_S to that of power-recycled interferometer, between the beamsplitter and the photo-detector.

where \mathcal{G}_{arm} is the ratio of the power injected to that in the arm⁴⁾. (Note that for a Fabry-Perot cavity $\mathcal{G}_{\text{arm}} = \left(\frac{\tau_c}{1-\rho_c\rho_r}\right)^2$ and eq. (44) reduces to eq. (22).) Here the transit time t_a is replaced by the retarded time $t_r = N\ell/c$ to allow the multi-reflection in the arm. The transmittance and reflectivities of cavity mirrors τ_c , ρ_c and ρ_r are also replaced by τ_s , ρ_s and ρ_a , respectively. The meanings of these parameters will be explained later.

This result can be separated into three factors. The first one represents the power enhancement in the arm, either by power recycling or by Fabry-Perot cavity in the arm. The last one represents the effect of the optical path length, which is determined by the arm length and the number of reflections. The second one left can be considered as the ‘generic’ effect of the cavity in which the signal sidebands are stored, after eliminating the (additional) effect of power enhancement inside the cavity. (As for this factor, we discuss it more in 4.3.1.)

The interferometers without power recycling or signal recycling can be thought of as a dual-recycled interferometer in which the power recycling mirror

⁴⁾ Here we consider the total power in the two arms. See also Appendix A.

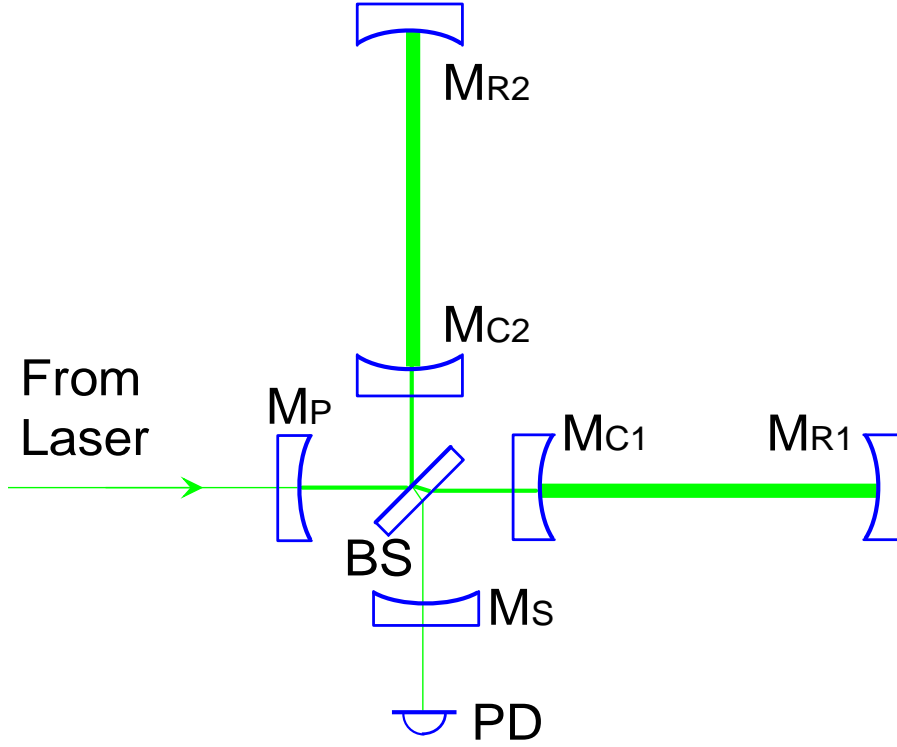


Figure 3.10b: Optical layout of a dual-recycled interferometer with Fabry-Perot cavities in the arms. As in the multi-reflection case, there is an additional mirror M_S .

or signal recycling mirror has a transmittance of unity. Thus, the generalized normalized response eq. (44) represents in a unified form all the normalized responses described so far, as stated by Meers [34].

The ρ_a , ρ_s and τ_s in eq. (44) are the reflectivities and the transmittance of the mirror to compose the cavity in which the signal sidebands are stored. When signal recycling is incorporated into an interferometer with delay lines, the signal recycling cavity is composed of the arm with multi-reflection and the signal recycling mirror. Note, however, that there is a beamsplitter inside that cavity and we must take its losses into account, i.e. $\rho_s \neq \rho_{SR}$. Then the parameters will be

$$\begin{aligned} t_r &= Nl/c, & \rho_s &= \rho_{SR}(1-\mathcal{A}_{BS})\sqrt{\frac{1+C}{2}}, \\ \rho_a &= \rho^{N-1}, & \tau_s &= \tau_{SR}\sqrt{(1-\mathcal{A}_{BS})\frac{1+C}{2}}, \end{aligned} \quad (45)$$

and $\mathcal{G}_{arm} = \mathcal{G}_{DL}^{FR}$ given in eq. (30).

By inserting these parameters into eq. (44), the normalized response of the dual-recycled interferometer with delay lines can be obtained. The detector

bandwidth of this system is also determined by using these parameters. The free-spectral-range and the finesse of this signal recycling cavity are $\Delta f_{\text{FSR}} = c/N\ell$ and $\mathcal{F} \simeq \frac{\pi\sqrt{\rho_s\rho_a}}{1-\rho_s\rho_a}$, respectively, and then the bandwidth will be $\Delta f_{\text{BW}} = \Delta f_{\text{FSR}}/2\mathcal{F}$. It should be noted, however, that there could be cases where the reflectivity of the signal recycling mirror is so low that the approximate relations in eq. (13) are not valid.

When Fabry-Perot cavities are present in the arms, it is appropriate first to consider one pair of mirrors, and then add the third one. Contrary to what one might expect, it is easier to consider the combination of the signal recycling mirror and the Fabry-Perot coupling mirror first, since this cavity is anti-resonance in signal recycling and its response is almost frequency independent.

Thus, ρ_s and τ_s in eq. (44) will be the reflectivity and the transmittance of this ‘compound mirror’. Again, considering the losses at the beamsplitter, they will be (see eq. (4.8) in 4.2.2)

$$\begin{aligned}\rho_s &= \frac{\rho_c + \rho_{\text{SR}}(1-\mathcal{A}_c)(1-\mathcal{A}_{\text{BS}})\sqrt{\frac{1+\mathcal{C}}{2}}}{1 + \rho_c\rho_{\text{SR}}(1-\mathcal{A}_{\text{BS}})\sqrt{\frac{1+\mathcal{C}}{2}}} \\ \tau_s &= \frac{\tau_c\tau_{\text{SR}}\sqrt{(1-\mathcal{A}_{\text{BS}})\frac{1+\mathcal{C}}{2}}}{1 + \rho_c\rho_{\text{SR}}(1-\mathcal{A}_{\text{BS}})\sqrt{\frac{1+\mathcal{C}}{2}}}.\end{aligned}\tag{46}$$

The reflectivity of the other mirror of the signal recycling cavity and the retarded time are given as

$$\rho_a = \rho_r, \quad t_r = 2\ell/c.\tag{47}$$

The \mathcal{G}_{arm} is the equivalent power gain in the arms, thus it is the product of that given in eq. (39) and that of the Fabry-Perot cavity in the arm,

$$\begin{aligned}\mathcal{G}_{\text{arm}} &= \mathcal{G}_{\text{FP}}^{\text{PR}} \times \mathcal{G}_{\text{FP}} = \mathcal{G}_{\text{FP}}^{\text{PR}} \times \left(\frac{\tau_c}{1 - \rho_c\rho_r}\right)^2 \\ &= \frac{(\tau_c)^2(1 - \mathcal{A}_{\text{PR}})}{(1 - \rho_c\rho_r)^2 - [\rho_c - (1-\mathcal{A}_c)\rho_r]^2(1-\mathcal{A}_{\text{PR}})(1-\mathcal{A}_{\text{BS}})\frac{1+\mathcal{C}}{2}} \\ &\lesssim \frac{1}{2} \frac{(\rho_{\text{max}})^2}{1 - (\rho_{\text{max}})^2}.\end{aligned}\tag{48}$$

This factor in fact shows the same dependence as the input efficiency described in Fig. 3.8 found in 3.3.3.

The normalized responses of a dual-recycled interferometer have the same form as that of the power-recycled interferometer with Fabry-Perot cavities, thus the shape of them with different bandwidths, set by different reflectivities of the signal recycling mirrors, are the same as those we have discussed in 3.4.3. As can be seen in Fig. 3.9, the peak sensitivity is improved by increasing the signal storage time, with a corresponding decrease in bandwidth. Note that increasing the storage time results in decreasing the sensitivity at frequencies outside the bandwidth.

As is described in 3.1.3, the sensitivity achievable is determined by the amount of energy stored, when the bandwidths set are the same. The amount of energy that can be stored does not differ so much, as can be seen in Fig. 3.8, when it is optimized by using power recycling. Thus, the difference between the configurations must be taken from more practical considerations.

One of the advantages of using dual recycling is that the requirement on the mirror size can be reduced when incorporated into interferometers with delay lines. This is especially important when the interferometer is intended for gravitational waves at low frequencies. Another advantage of dual recycling in operation is that the detector bandwidth can be altered by a minimum effort, i.e. only by changing the reflectivity of the signal recycling mirror.

The other advantages of signal recycling are the mode cleaning effect of the signal recycling cavity and the use of narrow banding. The former will be discussed in 5.2.5 in connection with thermal distortion problems. Here we consider the latter, the narrow-band response of signal recycling.

3.4.2: Signal recycling: Narrow-band operation

As we have seen in 3.3, the sensitivity at a certain frequency is uniquely determined by the energy stored in the system, either for Fabry-Perot or delay line type configurations. If one can make the sensitivity of a detector with a given stored energy better only at a certain frequency even with a narrower bandwidth, it may be useful for particular purposes, such as the detection of the signal from a known pulsar. Such configurations, where the bandwidth is narrow compared with the frequency of the peak sensitivity, i.e. $\Delta f_{\text{BW}} \ll \omega_g / 2\pi$, will be referred to as *narrow-band* configurations.

As Meers stated in his paper [44], signal recycling can be used also in a narrow-band mode, i.e. the sensitivity peak can be moved away from zero hertz. This can clearly be seen if one considers the effect of the signal recycling mirror

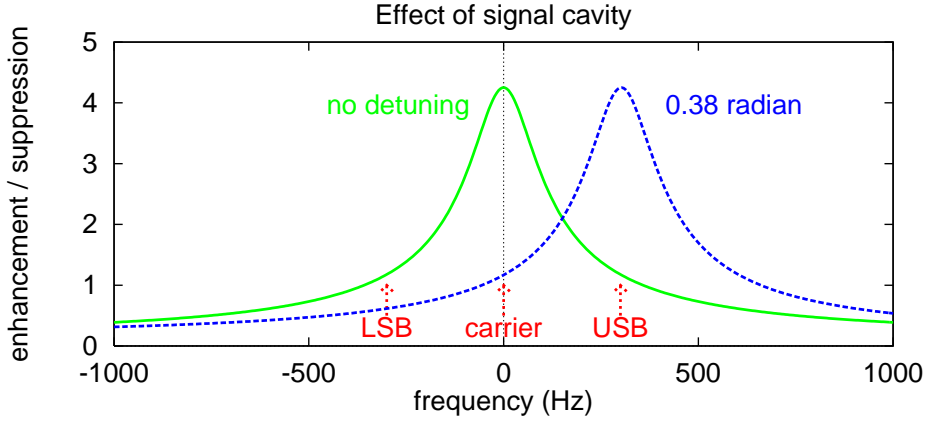


Figure 3.11: The schematic of the effect of signal cavity. It *enhances* the sidebands that are resonant, and *suppresses* the off-resonant ones. Detuning the signal cavity (by shifting the position of the signal recycling mirror by a fraction of a wavelength) moves the peak frequency away from DC to another frequency. In this case, however, only one of the sidebands is resonant and thus the peak sensitivity will be worse by a factor of ~ 2 . On the other hand, the full linewidth of the signal cavity is effective for the detector bandwidth.

in the frequency domain, i.e. the sidebands. As is discussed in 2.1.3 and 2.2.1, a gravitational wave produces upper and lower sidebands, each of which has the frequency separation from the carrier by the gravitational wave frequency. These sidebands will be stored in the signal recycling cavity composed of the interferometer and the signal recycling mirror at the signal port. This cavity ‘enhances’ the sidebands that are resonant, and ‘suppresses’ the off-resonant ones. (See eq. (44) and Fig. 3.9.)

In the broad-band operation of signal recycling, the peak of the resonance and thus of the enhancement is located at the carrier frequency and spread symmetrically to both upper and lower frequencies. This peak can be moved to another frequency by moving the position of the signal recycling mirror by a fraction of a wavelength. This corresponds to tuning the resonant frequency of the signal recycling cavity away from the carrier to another frequency (see Fig. 3.11).

The normalized response of narrow-band operation of signal recycling can be expressed by adding the detuning factor⁵⁾ $\exp(i\psi)$ to eq. (44), that of broad-

⁵⁾ In the derivation of eq. (22), ρ_r can be a complex number. In signal recycling, this does not affect the equivalent power gain \mathcal{G} in eq. (44) since the signal recycling cavity contains only sidebands. It is also expected that there will be another phase factor to τ_s , i.e. $\tau_s \exp i\psi_\tau$. This will, however, be ignored since it has no effect on the response unless it has frequency dependence.

band case:

$$\begin{aligned}
G_{\psi}(\omega) &= \sqrt{\mathcal{G}} \times \frac{-\tau_s}{1 - (\rho_s e^{i\psi}) \rho_a e^{-i\omega t_r}} \times \frac{\rho_a \omega_0}{4} \frac{1 - e^{-i\omega t_r}}{i\omega} \\
&= \sqrt{\mathcal{G}} \times \frac{-\tau_s}{1 - \rho_s \rho_a e^{-i(\omega t_r - \psi)}} \times \frac{\rho_a \omega_0}{4} \frac{1 - e^{-i\omega t_r}}{i\omega}.
\end{aligned} \tag{49}$$

With the detuning of ψ (measured in radian), the peak frequency of resonance will be

$$\omega_p/2\pi = \psi/2\pi t_r = (\psi/2\pi)(c/N\ell) = (\psi/2\pi)\Delta f_{\text{FSR}}. \tag{50}$$

Although the linewidth of the signal recycling cavity does not change when it is detuned, the detector bandwidth may be twice as much as that without detuning, i.e. roughly equal to the linewidth. This is because both sides of the resonance peak are effective for the bandwidth, whereas only one half is effective in the Fabry-Perot response or in the broad-band case, as discussed in 3.2.2.

On the other hand, detuning reduces the sensitivity by a factor of ~ 2 because only one of the sidebands is resonant with the signal recycling cavity, if the cavity linewidth is narrow compared with the peak frequency. Then the effect of the other sideband can be ignored and the modulation can be considered as single sideband. As a consequence, the sensitivity in this case is $\sim \sqrt{2}$ worse than that expected from eq. (3). This can be understood as we are still detecting “where the other sideband would be”, even though there is only noise instead of signal (see the footnote in 3.5.1).

When the two sidebands have different amplitudes, the sensitivity after detection will depend on the component we detect, as is discussed in 2.3.4. There will be an optimum component for the detection, depending on each configuration. In this paper, however, it is assumed that we will detect the phase modulation component. The differences with other schemes are usually small, and further discussion on this topic can be found in Appendix F.

When the phase modulation component is detected, the normalized responses of the upper and lower sidebands should be summed as

$$\begin{aligned}
\mathbf{G}_{\text{PM}}(\omega_g) &= \mathbf{G}_{h \rightarrow a/a_0}(\omega_g) + \mathbf{G}_{h \rightarrow a/a_0}^*(-\omega_g) \\
&= 2 \times \mathbf{G}_{\text{BB}}(\omega_g) \quad \text{for the broad-band case.}
\end{aligned} \tag{51}$$

Using the above definition, the normalized response of narrow-band signal recycling can be obtained as is shown in Fig. 3.12. Similar to the broad-band case,

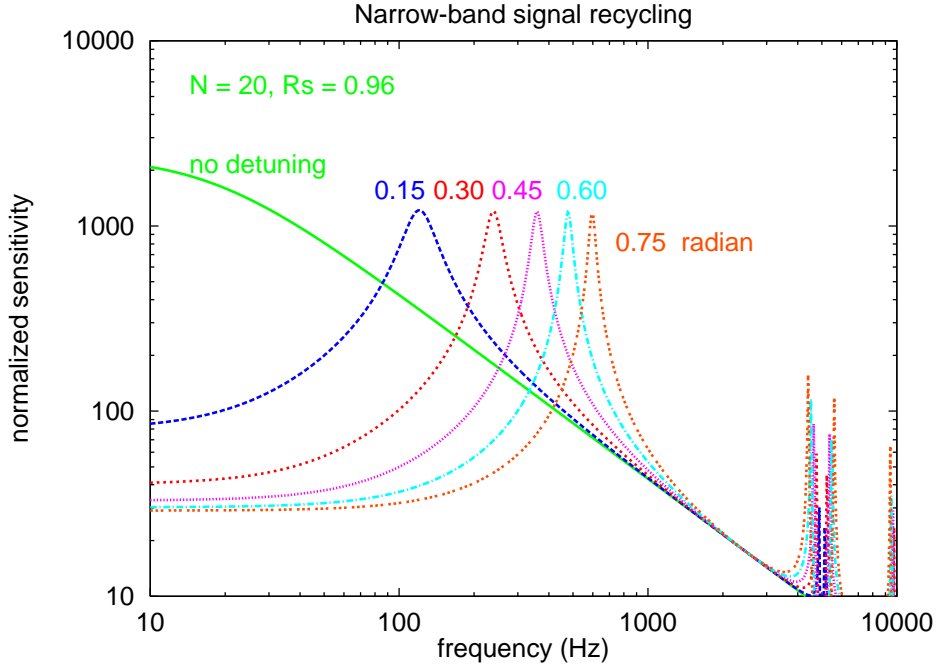


Figure 3.12: The normalized response of dual-recycled interferometers with delay lines in the arms operated in *narrow-band* mode. Each plot corresponds to a different amount of detuning for the signal extraction mirror position (key indicates the amount of detuning measured in radian). For each curve, the peak sensitivity and the bandwidth are almost constant, excepting the ‘no detuning’ case. Only the peak frequency moves according to the detuning. In the ‘no detuning’ case, the peak sensitivity is better by a factor of two than in the other cases, but the bandwidth is narrower by a factor of two. Note, however, that the peak sensitivity is obtained only in the low-frequency region where other noise sources may dominate the sensitivity.

narrowing the bandwidth increases the sensitivity at the peak at the expense of poorer sensitivity at other frequencies. What is different from the broad-band case is the location of the peak sensitivity. Note that the peak sensitivity can be better than that determined by eqs. (35) and (42), as it was expected.

The detuning ψ represents the amount of phase shift due to the displacement Δx of the signal recycling mirror. If delay lines are present in the arms, this is given by

$$\psi = 2k\Delta x = 2(2\pi/\lambda)\Delta x, \quad (52)$$

where 2 is due to the path length change by the displacement. When Fabry-Perot cavities are used in the arm, however, the dependence of the detuning on the displacement of the signal recycling mirror is much more complicated.

The ‘compound mirror’ (composed of the signal recycling mirror and the Fabry-Perot coupling mirror) actually has a *complex reflectivity* ρ_s which represents the amplitude reflectivity and the phase shift upon the reflection in a complex number. Thus, ρ_s and ψ are to be defined as

$$\rho_s = |\rho_s(\Delta x, \omega)|, \quad \psi = \angle \rho_s(\Delta x, \omega). \quad (53)$$

Similarly, the transmittance of the compound mirror τ_s also has a complex transmittance τ_s . Furthermore, these complex reflectivity and transmittance values have frequency dependence (see 4.2.2).

These effects are, however, relatively small as long as the resonance peak frequency is within the linewidth of the Fabry-Perot cavity in the arm, i.e. $\omega_p/2\pi \lesssim \Delta f_{\text{FP}}$ (see Fig. 4.12 in 4.3.2 and the discussion there). Thus, we will assume a constant real reflectivity multiplied by the tuning factor, as in the delay line case. Then the response will be the same as that for delay line case. The full expression of a three-mirror coupled cavity will be treated in 4.3.2.

3.4.3: Detuned recycling

Detuned recycling, proposed by Vinet et al. [33], is another way to realize a narrow-band response. This is useful only for interferometers with Fabry-Perot cavities in the arms. The optical layout in this configuration is exactly the same as for a power recycled interferometer with Fabry-Perot cavities in the arms (see Fig. 3.6b); the only difference being in the tuning of the Fabry-Perot cavities. While in standard configurations the cavities are resonant with the carrier as well as the sidebands, in detuned recycling they are resonant with only one of the sidebands.

Under this condition, most of the carrier light from the beamsplitter side is reflected at the coupling mirror of the cavity, i.e. the light power inside the cavity is much less than that hitting the coupling mirror (roughly by a factor of $\sim 1/\mathcal{F}$). When power recycling is incorporated, however, this can yield higher power recycling gain. In the end, the amount of energy inside the cavity can be equal to the maximum value given in eq. (5), if the losses at the beamsplitter and the Fabry-Perot coupling mirror can be neglected.

The cavity storing the signal sidebands is the simple Fabry-Perot cavity, i.e. not a coupled or split cavity, with a detuning ψ (if the cavities in the two arms are set identically). This means that the frequency response of the

detuned recycled interferometer should have the same form as the one for the dual recycled interferometer operated in narrow-band mode.

The sensitivity will, however, be worse in this configuration because the condition mentioned above is impossible to satisfy when the coupling mirror has a finite loss. The typical energy stored will be

$$\mathcal{E}_{\text{detuned}} \lesssim \mathcal{E}_{\text{max}} \times \frac{\mathcal{A}_{\text{min}}}{\mathcal{F}(\mathcal{A}_{\text{c}} + \mathcal{A}_{\text{BS}} + \frac{1-\mathcal{C}}{2})}, \quad (54)$$

where \mathcal{E}_{max} is the maximum energy given in eq. (5) and \mathcal{A}_{min} is the minimum loss per reflection. From this result (which is still an overestimation), it is expected that the amount of stored energy will be much less than the maximum case.

Furthermore, the amount of energy stored will in practice be limited by the loss at the beamsplitter, resulting in poorer sensitivity than that of a dual-recycled interferometer. This will especially be important when one considers thermal effects, because in this configuration the power inside the substrate of the beamsplitter must be enormously high.

Because of this practical difficulty, the author will pay no particular attention to this configuration in the following discussion, although this does not mean the worth of this idea is denied.

3.4.4: Synchronous recycling

The first idea for realizing a narrow-band configuration called *synchronous recycling* was proposed by Drever [38, 39]. This is quite different from all other configurations, as can be seen in Fig. 3.13. While the others are based on the Michelson-type layout, one would better consider this as a variant of a Fabry-Perot cavity with a ‘kink’ at its center. Each side of the kink corresponds to an arm of a Michelson interferometer, and it must contain either a delay line or a Fabry-Perot cavity to set an appropriate storage time in the arm.

The total (‘effective’, if Fabry-Perot cavities are present in the arms) cavity length is set to be the wavelength of the gravitational wave of the target frequency. The bandwidth is determined by the storage time in the whole cavity. Since the storage time in *one arm* is matched to a half of the gravitational wave period, a longer storage time improves the peak sensitivity by experiencing the effects from the gravitational wave constructively during successive travelling in the two arms. This, however, sacrifices the sensitivity outside the bandwidth

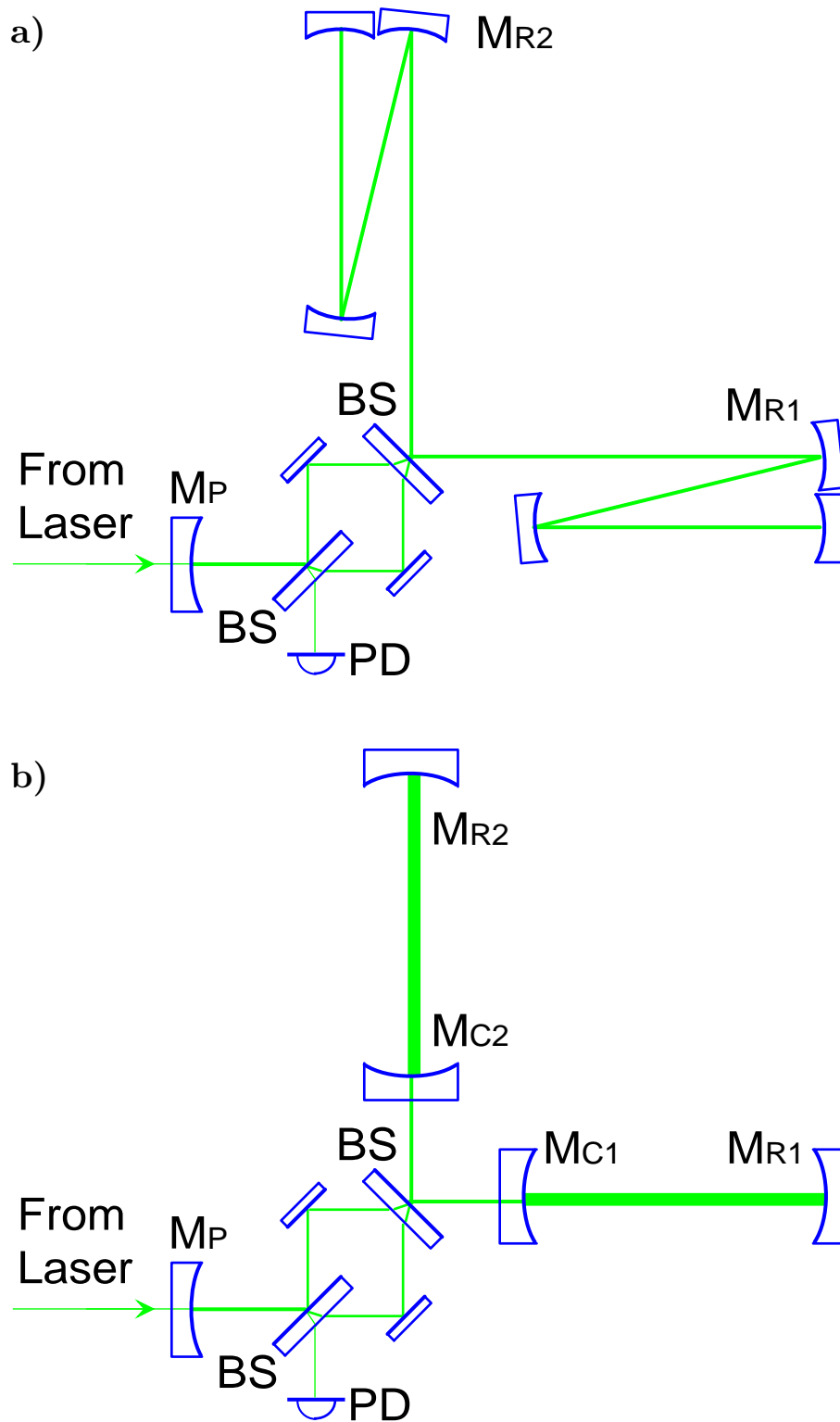


Figure 3.13: Optical layout of interferometers using synchronous recycling, with a) multi-reflections or b) Fabry-Perot cavities in the arms. Differently from other configurations, the two arms are directly combined by a beam-splitter, and light is injected to the combined arms through this beam-splitter.

due to the partial cancellation of the gravitational wave effects, as in all other narrow-band configurations.

As a consequence from its exotic optical layout, the carrier and the sidebands are stored in the same cavity, which is the biggest difference from other configurations. This does not necessarily mean that both have the same storage time, which is true when delay lines are used in the arms. When Fabry-Perot cavities are present in the arms, however, the resonance feature of the whole coupled cavity can be altered by adjusting the distance between the two Fabry-Perot cavities in the arms. This on the other hand means that the choice of parameters is not as obvious as in the case of using delay lines in the arms⁶⁾, though we do not consider it in detail here.

In spite of the differences described above, the sensitivity of the detector using this configuration expected from the previous analyses [46, 33] is similar to those of others. This means that this configuration also follows the sensitivity theorem described in 3.1.3. Furthermore, the amount of energy that can be stored, which determines the sensitivity, does not improve by using this configuration, since it is limited by the dissipation in the arms. Thus, we do not analyze this in detail here, assuming (without any proof) the theorem is valid also for this configuration.

3.5: More about the theorem

3.5.1: Complete theorem

We now return to the theorem presented in eq. (3) of section 3.1.3. This is actually a simplified form, and we discuss it here in more detail, summarizing the pieces of information that appeared through the review. Note, however, that the simplified form is valid in most cases and is useful for many purposes. The main object of this section is to help understand the physical background of the theorem.

As mentioned in 3.2.4, the frequency integration of the squared transfer function is proportional to the energy stored in the optical system in good approximation. This was actually indicated in eq. (3): If one can approximate the squared sensitivity of a configuration by a ‘box’ shape, i.e. a constant sensitivity within a certain band and no sensitivity outside this band, its height can be

⁶⁾ It is proposed by Schilling [41] to use the cavities with higher finesse than that determined by the signal storage time, similar to resonant sideband extraction described in Chapter 4, in order to avoid thermal distortion problems.

obtained by dividing the area by the width. What is stated in eq. (3) is that the area is proportional to the amount of light energy stored in the optical system.

In more detail, the frequency integration of the power of the *sidebands* produced by a gravitational wave is *exactly* proportional to the amount of light energy stored. Some parts of the sidebands will, however, be lost before they are detected, due to the imperfect reflectivity of the mirrors, the inevitable loss in the optical substrates, the finite contrast of the interference, and so on. As a result, the amount of the sidebands *detected* will be less than that produced by the gravitational wave, degrading the sensitivity of the detector. This is what will be referred to as *output efficiency*, the fraction (in power) of the amount of sidebands detected to those produced. Note that this quantity can have frequency dependence.

Furthermore, there is also degradation of the sensitivity by the efficiency in the detection process, which will be referred to as *detection efficiency* (this must also be defined in power). This is because the amount of the signal as well as that of noise depend on the detection scheme. The best sensitivity will be achievable by maximizing the detected signal and by minimizing the detected noise.

In the double sideband case, it is possible to make the most efficient detection scheme, i.e. the only noise detected is that which has the same frequency as the signal. The single sideband case is much worse since the signal present in only one of the sidebands, while noise at both sidebands is detected⁷⁾. As a result, the sensitivity will be worse roughly by a factor of $\sim\sqrt{2}$ in this case. In general the achievable detection efficiency is somewhere between a half and unity.

Further degradation of the sensitivity in the detection process can occur due to the use of inappropriate schemes. These should, however, be treated separately from the discussion here which is concentrating on the *optimum* performance of a configuration.

Combining the above considerations, eq. (3) can be rewritten in an ‘integral’ form

$$\int_0^\infty \frac{df}{\eta_o \eta_d [\tilde{h}(2\pi f)]^2} = \frac{\pi c}{2\hbar\lambda} \eta_i \mathcal{E}_{\max} , \quad (55)$$

⁷⁾ This disadvantage is not a principal limitation, i.e. this can be overcome if one can detect a single sideband efficiently.

or in a ‘simplified’ form for the peak sensitivity \tilde{h}_0

$$\tilde{h}_0 = \frac{1}{\sqrt{\eta_i \eta_o \eta_d}} \times \sqrt{\frac{2\hbar\lambda}{\pi c} \frac{\Delta f_{\text{NB}}}{\mathcal{E}_{\text{max}}}}. \quad (56)$$

Here Δf_{BW} is replaced by Δf_{NB} to make the definition of the bandwidth clearer, and η_i , η_o and η_d are the efficiencies of input, output and detection, respectively (note that η_o and η_d can have frequency dependence). The efficiencies η are always less than unity. In many reasonable configurations, the input efficiency η_i is order of unity, the output efficiency η_o is close to unity, and the detection efficiency η_d is close to unity for double sideband cases and around half for single sideband cases. Here $\eta_o \eta_d \eta_i \mathcal{E}_{\text{max}}$ can be considered as ‘effective stored energy’ which determines the sensitivity of the configuration.

The above is the complete form of the theorem. The remainder of this chapter is devoted to the investigation of the physical background of the theorem, although a complete proof may require a satisfactory theory of quantum gravity and its interaction with QED. Thus, the following discussion should be considered not as an exact, but rather a qualitative one. Some additional information which did not fit into the main text can also be found in Appendix E.

3.5.2: Shot Noise

For the understanding of the sensitivity theorem, it is good to start by investigating the cause of the shot noise, because this is one of the most important factors that determines the sensitivity.

The origin of this noise can be understood as the *zero-point energy* of the vacuum state which has energy equal to half that of a photon in that state, i.e. $\hbar\omega_0/2$. This corresponds to the amplitude noise called *vacuum fluctuations* $\tilde{a}_v = \sqrt{\hbar\omega_0/Z_0}$. To be detectable, the amplitude of the signal sideband should be larger than these vacuum fluctuations; or, in terms of photons, at least one photon should be produced by the signal to be detected. These vacuum fluctuations have a constant amplitude with random phase throughout the entire frequency range.

What we can detect is the amplitude modulation of the light; therefore upper and lower sidebands must be taken into account. Due to the randomness of the relative phase between the two, their amplitudes should be summed quadratically to produce the single frequency noise component in the detected signal. An even better way is to consider the *noise power* which should be summed linearly.

Although the aim of this chapter is to discuss how to improve the signal-to-noise ratio by increasing the signal sidebands, there still is the other possibility of reducing the noise. This does not necessarily mean that one must reduce the zero-point energy. Generally, we use only one component of conjugate observables in signal detection. Thus it should be possible to reduce the noise in the measuring component at the expense of increasing noise in the conjugate observable. This method is called *squeezing*, and some theoretical as well as experimental research has been carried out in this field [15, 47] with some successful results. However, this will not be considered throughout this paper and the ‘standard’ shot noise is assumed to be the limiting noise above a certain frequency.

Because the shot noise has constant amplitude and random phase for all frequencies, the sensitivity of a detector has the same frequency dependence as the sideband amplitude. The sensitivity limit, or noise-equivalent signal, should be given by the reciprocal of the transfer function from the gravitational wave signal to the sideband amplitude. In practice, however, there is the problem of how to detect the sidebands. This topic is thoroughly discussed in Chapter 2.

3.5.3: Energy, or number of photons

Though it appears quite unusual to express the sensitivity of an interferometer by the energy, it is not so surprising if one recalls that this corresponds to the number of photons in the system. If one considers the detection process of a gravitational wave as an interaction between *gravitons* and photons, it is quite natural that their numbers determine the sensitivity⁸⁾.

It is interesting to remember that a light energy of 20 J corresponds to 10^{20} photons of $1\ \mu\text{m}$ wavelength, which sounds like a huge number. On the other hand, even the smallest gravitational wave detectable contains an enormous number of gravitons, as can be seen in eq. (57). This is due to the huge power the wave carries and the small energy per graviton due to the low frequency of the wave (assuming the energy of one graviton is given by $\hbar\omega_g$).

However, the sensitivity of an interferometric gravitational-wave detector may still be limited by the statistical error in the photon counting process, i.e. only a few photons are expected to interact with a gravitational wave. This

⁸⁾ Actually it is not necessary to involve gravitons. It can be any sort of perturbation to the electro-magnetic field. If it occurs at a mirror, however, the number of interaction per unit time must be also taken into account. Then the light ‘power’ rather than the light ‘energy’ will be a better representation of the sensitivity.

indicates the extremely low probability of the interaction between gravitons and photons. One can understand this by recalling the smallness of the characteristic strength of gravitational interaction $\alpha_G = G_N M_p^2 / \hbar c \simeq 5.9 \cdot 10^{-39}$, which corresponds to the fine structure constant for electro-magnetic interaction $\alpha_{EM} = e^2 / 4\pi\epsilon\hbar c \simeq 1/137$.

Assuming a gravitational wave pulse with a single peak of duration time t_g , the number of gravitons per unit area \mathcal{N}_{GW} can be estimated as [48]

$$\begin{aligned} \mathcal{N}_{GW} &\simeq \frac{t_g}{\hbar\omega_g} \frac{c^3}{16\pi G_N} \langle \dot{h}_+^2 + \dot{h}_\times^2 \rangle \simeq \frac{1}{\hbar\omega_g} \frac{c^3}{16\pi G_N} \frac{4h^2}{t_g} \\ &\approx 3.1 \cdot 10^{24} / \text{m}^2 \times \left[\frac{h}{10^{-22}} \right]^2. \end{aligned} \quad (57)$$

The values quoted correspond to the amplitude of a gravitational wave scarcely detectable with $N_{ph} \simeq 10^{20}$ photons of $\lambda = 1 \mu\text{m}$ wavelength, and the probability of interaction

$$\begin{aligned} &\mathcal{N}_{GW} \times (ct_g)^2 \times N_{ph} \times \alpha_G \times \left(\frac{2\pi\hbar c/\lambda}{M_p c^2} \right)^2 \\ &\simeq 3.1 \cdot 10^{24} / \text{m}^2 \times (3 \cdot 10^5 \text{m})^2 \times 1 \cdot 10^{20} \times 5.9 \cdot 10^{-39} \times \left(\frac{2.0 \cdot 10^{-19} \text{J}}{1.5 \cdot 10^{-10} \text{J}} \right)^2 \\ &\sim 1, \end{aligned} \quad (58)$$

results in an order of magnitude estimation consistent with the previous hypothesis: at least one photon must interact with the gravitational wave in order to be detectable. In the above equation, $(ct_g)^2$ represents the ‘effective area’ of the detector, and the last factor is the (squared) ratio of the typical energy of a proton $M_p c^2$ to that of a photon $2\pi\hbar c/\lambda$, which with α_G represents the ‘characteristic strength of gravitational interaction between two photons’.

3.5.4: Cavity effect (boundary condition)

Actually it is more difficult to understand why the detector bandwidth affects the spectral sensitivity. This should not be confused with the fact that the root-mean-square, or *rms*, sensitivity will be improved by narrowing the observation bandwidth, in which the contribution of quadratically increasing stochastic noise becomes less significant than the linearly increasing signal by integrating over a long time period. This is quite a common feature for *periodic* signals such as those from pulsars.

What is being discussed here is totally different. Narrowing the *detector bandwidth* improves the spectral sensitivity which affects not only periodic signals but also broad-band type signals such as stochastic background gravitational waves. The numbers of gravitons and photons involved are the same, but the sensitivity is different—why is this?

The only reasonable explanation is that the probability of interaction between them depends on the detector bandwidth which is set by, for example, the cavities in the arms. For the understanding of this phenomenon it is helpful to remember some of the aspects of quantum mechanics.

According to *Fermi's golden law* [49], the probability of the transition between two states due to a small perturbation in the potential can be given by

$$d\wp_{fi} = \frac{2\pi}{\hbar} |\langle \psi_f | \widehat{U}_p | \psi_i \rangle|^2 \varrho(E_f) dE_f . \quad (59)$$

Here $|\psi_i\rangle$ and $|\psi_f\rangle$ represent the initial and final states respectively, \widehat{U}_p the perturbation, $\varrho(E_f)dE_f$ the number of quantum states at the energy of final state E_f , and $d\wp_{fi}$ the (differential) probability of transition from i -state to f -state.

In the above equation, the probability of the transition depends not only on the perturbation and the initial and final states but also on the *density* of the modes (or states) around the final state. It should also be noted that the initial and final states are represented by the state vectors when no perturbation exists, i.e. the perturbation does not change the *eigenmodes* of the potential but causes the transition between two eigenmodes.

For the photon in a cavity the states are determined by its energy or frequency and by the spatial distribution, often called TEM (transverse electromagnetic) mode. Since we are only interested in the difference in change in energy, we assume the light to be in the fundamental mode. In this case the modes are characterized uniquely by the frequency.

It is known from cavity QED that the density of the modes near resonance is approximately proportional to the quality factor, or Q -value, of the cavity. If this Q -value is unity, this corresponds to the vacuum state, and the density of the modes is that for the vacuum state. On the other hand, the total number of modes cannot change with the existence of a cavity or by the Q -value of the cavity. Thus, the density of the modes off-resonance must be less than that of the vacuum state.

This can explain what we have seen—increasing the finesse means a higher Q -value, and thus provides an enhancement of the transition to the modes near resonance, while the transition to the other modes are suppressed, or prohibited. In our case, the cause of the perturbation is the gravitational wave, and it generates the upper and lower sidebands from the carrier light.

Thus, the gravitational wave induced sidebands within the cavity bandwidth are enhanced according to the Q -value, resulting in an increase of the detectable signal at that frequency. The signals outside the bandwidth are, however, suppressed for the same reason. The total area of the frequency response curve should not depend on the finesse or the Q -value, as the total number of the modes does not depend on these. In short, the cavity has the effect of ‘concentrating’ the interaction due to a perturbation into a limited bandwidth.

3.5.5: The arm length effect

Naturally the next question arises: How can the effects of multiple reflections and the importance of the optical path length be explained? These also affect the bandwidth, but in a manner different from the cavities, as we have seen earlier. It seems that these effects can be attributed to different origins. This will become clear when one considers a cavity a little more carefully.

Because a cavity shows a periodic response in frequency, also the sensitivity should repeat its response for frequencies apart by one free-spectral-range, if the cavity effect described above is the only one to determine the sensitivity. This is not the case; in fact, in a simple cavity only one resonance peak is significant and the others are almost negligible.

As we have seen in 3.2.3 the frequency response of cavities can be separated into two parts, one of which has the same form as that for simple Michelson interferometers. The cavity effect can explain the other part which is dominant for most of the realistic cases.

The effect of multiple reflections can be thought of as being equivalent to just a long optical path length, as we have discussed in 3.2.1. The remaining problem is the effect of the optical path length after splitting and interfering of the two beams, which seems to have a sort of boundary effect.

In this case the total number of the modes is proportional to the optical path length. This is because one mode occupies volume of $2\pi\hbar$ in the ‘phase

space', i.e. $\Delta x \Delta p = 2\pi\hbar$. Here Δ does not mean the error in the measurement but the width of the mode.

It is natural to assume that the width of the mode in physical length space is of the order of the optical path length $N\ell$, i.e. $\Delta x \simeq N\ell$. Then the width in momentum space or in frequency space will be

$$\begin{aligned}\Delta p &= 2\pi\hbar/\Delta x \simeq 2\pi\hbar/N\ell \\ \Delta f &\simeq \frac{1}{2\pi}\Delta\omega \simeq \frac{1}{2\pi}\frac{c}{\hbar}\Delta p \simeq c/N\ell,\end{aligned}\tag{60}$$

thus the width of one mode is inversely proportional to the optical path length. This results in the density of the mode, which is given by the number of the modes within unit frequency spacing, being proportional to the optical path length. Note that the Δf is the full width of positive and negative frequencies and the half of it corresponds to the detector bandwidth.

The notches at higher frequency can be thought of as the 'nodes' where the density of the modes is null. In addition the modes between the notches can be considered as the 'higher order modes'. In such a higher order mode the amplitude will be smaller than the fundamental mode even if they have the same amounts of power, because of the nodes—in general, modes with more nodes have more power assuming constant amplitude.

Chapter 4

Resonant sideband extraction

4.0: Introduction

According to the theorem described in Chapter 3, the sensitivity of an optical system is determined by two parameters: the amount of light energy stored in the system and the bandwidth set by the optical configuration—it does not matter how the bandwidth is set. When power recycling is incorporated in an interferometer, the amount of energy that can be stored in the system is likely to be limited by dissipation in the arms.

Resonant sideband extraction, proposed by the author, is another method of setting the desired bandwidth. While some non-ideal aspects of real optics, such as mirror sizes or thermal distortions, limit the usefulness of particular configurations, this new configuration is advantageous when thermal distortion becomes significant. In this chapter, however, a linear model is treated in order to concentrate on the principle of operation. Non-linear problems including thermal effects are discussed in Chapter 5.

As is discussed in Chapter 3, the frequency response of an optical configuration is determined by the cavity in which the signal sidebands are stored. In the case of resonant sideband extraction, this will be a three-mirror coupled cavity which cannot be approximated by a Fabry-Perot cavity. Thus, it is indispensable for the understanding of resonant sideband extraction to know how a three-mirror cavity behaves. For this purpose, a useful method to analyze cavities is presented.

4.1: The principle of resonant sideband extraction

4.1.1: Basic concept

We have seen in 3.4.1 that the key point of signal recycling is to make the detector bandwidth *narrower* than that determined by the storage time in the arms. This improves the peak sensitivity at the expense of narrower bandwidth. Naturally the next question arises: *Is it possible to make the bandwidth broader*

than that determined by the optics in the arms? The answer was found to be yes, but only if Fabry-Perot cavities are used to set the storage time in the arms. This is the basic concept of *resonant sideband extraction*.

As is discussed in 4.2.2, the reflectivity of a cavity can be lower than that of either of the mirrors used to compose the cavity. By applying this fact, the reflectivity of the Fabry-Perot coupling mirror can be made lower by placing an additional mirror appropriately. This results in a shorter storage time for the Fabry-Perot cavity and so makes its bandwidth broader.

The expense for the broader bandwidth is a deterioration of the peak sensitivity from that would be obtained without the additional mirror. However, it is not sensible to discuss which configuration is better with or without the additional mirror. Instead, comparisons should be made among the configurations that are designed to give similar bandwidths.

According to the theorem described in 3.1.3, the sensitivity of an optical system is determined by the amount of light energy stored in the system and the bandwidth set by the optical configuration. The amount of energy stored does not differ so much among any configurations when power recycling is incorporated appropriately, since it is limited by the dissipation of the optics in the arms. Thus, this new configuration brings no improvement in terms of sensitivity, in the ideal case.

In practice, however, there are advantages in increasing the storage time in the arms. For instance, this decreases the amount of light power which must be transmitted through the optical substrates of the beamsplitter and Fabry-Perot coupling mirrors in order to obtain the same amount of energy stored in the arms. This will then reduce the thermally induced distortion inside these substrates.

In other words, if thermal distortion becomes the limiting factor, then the light power which can be transmitted through the optical substrates may be restricted to a maximum value. Then the amount of energy that can be stored under this condition can be higher in resonant sideband extraction than in other configurations, and thus better sensitivity is expected with resonant sideband extraction. This subject is discussed in more detail in Chapter 5.

4.1.2: Optical configuration

The optical layout for resonant sideband extraction is shown in Fig. 4.1. A Michelson type interferometer having a Fabry-Perot cavity in each arm is supplemented by an additional mirror (\mathbf{M}_S in Fig. 4.1) put between the interferometer and the photo-detector which will be referred to as *signal extraction mirror*. There may be another mirror (\mathbf{M}_P) between the beamsplitter and the light source for power recycling, but it is not intrinsic to the operation of resonant sideband extraction.

The beam from the light source is injected to a beamsplitter (possibly through the power recycling mirror) and divided into the two arms. The cavity in each arm is set to be resonant with the incident light. The reflected beams from the two cavities interfere with each other at the beamsplitter. There are two interfering beams, one goes out in the direction towards the light source and the other in the direction towards the photo-detector. The latter is partially reflected back to the interferometer by the signal extraction mirror. The former may also partially be reflected back to the interferometer when power recycling is incorporated.

The relative phase between the two arms of the Michelson interferometer is adjusted so that the beam which goes out in the direction towards the light source will be the maximum of interference. Then the other must be the minimum, the so-called dark fringe. Under this condition, only the sidebands which are differentially modulated in the arms appear at the dark fringe, and the unmodulated carrier¹⁾ goes back in the direction towards the light source. Thus the former will be referred to as the *signal port* and the latter as the *carrier port*.

Although only one of them are used in usual applications, there are actually two inputs into an interferometer, resembling the two outputs described above. In resonant sideband extraction (and in signal recycling), the sidebands reflected at the signal port re-enter the interferometer through this second (usually unused) input. The additional mirror at the signal port is called signal extraction mirror (or signal recycling mirror).

Just as most of the light from the source leaks out of the interferometer from the carrier port, the sidebands entering through the second input exit

¹⁾ Actually any common mode components go back in the direction towards the light source. We will ignore this, however, since we are only interested in the differential signal (sidebands) and the carrier.

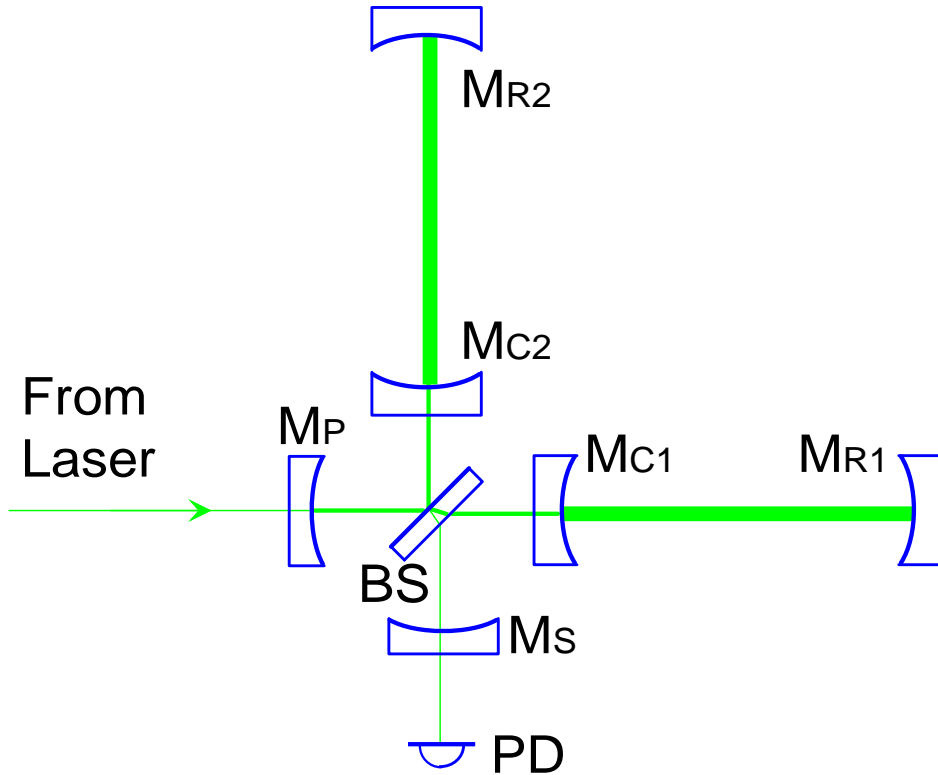


Figure 4.1: The optical layout of resonant sideband extraction. A Michelson interferometer having a Fabry-Perot cavity in each arm is supplemented by an additional mirror, the *signal extraction mirror* (M_S), between the beamsplitter (BS) and the photo-detector (PD). Another additional mirror for power recycling, the *power recycling mirror* (M_P), is also shown.

through the signal port (unless they are differentially modulated—such second order effects are so weak and will be ignored). Thus, the differential sidebands are stored in a cavity composed of the interferometer and the mirror at the signal port. This cavity (*signal cavity* for convenience) determines the frequency response of the configuration (see the discussion in 3.4.1 and 4.3.1).

Assuming the two arms are identical and the interferometer stays exactly at a dark fringe, one can disregard the beamsplitter and the carrier port. Then the signal cavity can be thought of as being composed of the arm and the mirror at the signal port. For resonant sideband extraction and signal recycling with Fabry-Perot cavities in the arms, this will be a three-mirror coupled cavity (see Fig. 4.2).

The coupled cavity is composed of a Fabry-Perot reflecting mirror, a Fabry-Perot coupling mirror, and a signal extraction mirror (or a signal recycling mirror). We will refer to the combination of the first and the second mirrors as the arm cavity, and that of the second and the third mirrors as the *signal*

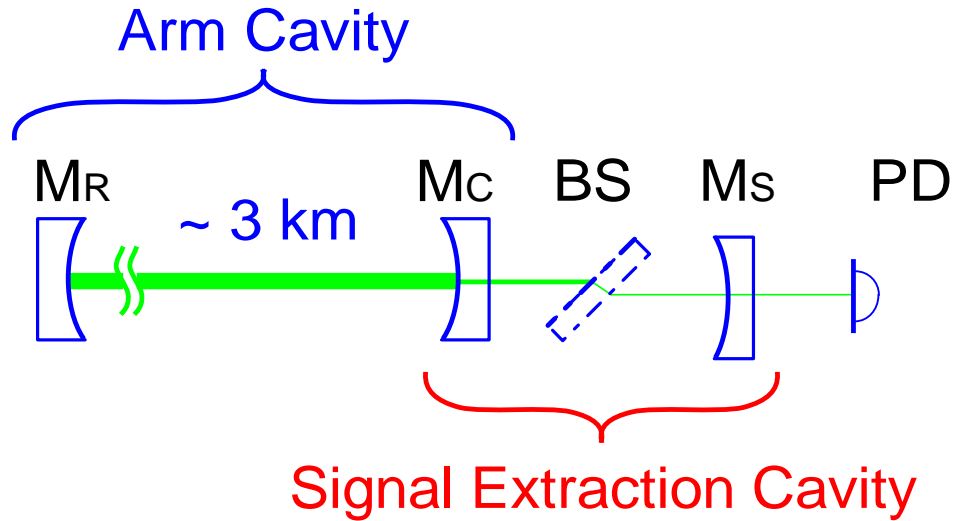


Figure 4.2: The sidebands induced by a differential modulation will be stored in the cavity composed of the arm and the mirror at the signal port. In the case of resonant sideband extraction, this *signal cavity* is a three-mirror coupled cavity, assuming the two arms are identical and the Michelson interferometer stays at a dark fringe. The mirrors which compose the three-mirror coupled cavity are the reflecting mirror (M_R) and the coupling mirror (M_C) of the cavity in the arm (*arm cavity*), and the signal extraction mirror (M_S). The combination of the latter two mirrors is referred to as the *signal extraction cavity*.

extraction cavity (in signal extraction case). Note that the order of optics is based on the direction from the arm cavity towards the photo-detector, not from the light source.

4.1.3: Brief discussion

According to the sensitivity theorem, the sensitivity of an optical system is determined by the amount of light energy stored and the noise bandwidth set by the optical configuration. Now, the question is how to estimate the bandwidth of a three-mirror coupled cavity. When Fabry-Perot cavities are used in the arms, the bandwidth is determined by its length and the reflectivities of the two mirrors. Thus, we start with approximating a three-mirror coupled cavity by a Fabry-Perot cavity.

In general, the detector bandwidth of a configuration is inversely proportional to the signal storage time (although the proportionality constant may differ among configurations). From practical considerations, the arm cavity is

thought to be longer than the signal extraction cavity. Thus we can assume that the storage time is dominated by the arm cavity length rather than the signal extraction cavity length.

The two ‘mirrors’ at both sides of this tentative cavity will be the reflecting mirror of the arm cavity and the ‘compound mirror’ which is actually the signal extraction cavity. The reflectivity of the reflecting mirror must be as high as possible in order to maximize the amount of stored energy. Thus, the factor that, in practice, determines the storage time of the coupled cavity is the reflectivity of the compound mirror.

The reflectivity of the compound mirror is in turn determined by the reflectivities of the two mirrors and the spacing between them, i.e. the tuning of the compound mirror. As is described in 4.2.2, the reflectivity of a cavity can be either higher or lower than that of one of the component mirrors, depending on the tuning condition of the cavity. For the present purpose, this means that there is an arbitrariness for the choice of the reflectivity of the Fabry-Perot coupling mirror to realize the required storage time.

By tuning the compound mirror to be anti-resonant, its reflectivity will be higher, resulting in a longer storage time and thus a narrower bandwidth than those of the arm cavity. This corresponds to signal recycling with Fabry-Perot cavities in the arms. In this case the linewidth of the arm cavity must be broader than the desired detector bandwidth, and the signal recycling mirror must be chosen to achieve an appropriate bandwidth when the compound mirror is anti-resonant.

On the other hand, tuning the compound mirror to be resonant results in a lower reflectivity and thus a shorter storage time. Then the bandwidth is broader than that of the arm cavity, as required in resonant sideband extraction. The linewidth of each arm cavity must be chosen to be narrower than the desired detector bandwidth and the signal extraction mirror must be chosen appropriately. More detailed discussion on the choice of the parameters can be found in 4.4.

It should be noted, however, that the tuning of a cavity depends on the light frequency. Not only the power reflectivity but also the phase shift upon the reflection from a cavity depends on the tuning. Thus the compound mirror actually has a frequency-dependent *complex reflectivity* whose absolute value and argument represent the amplitude reflectivity and the phase shift upon the reflection, respectively.

This complex reflectivity may have poor effects in signal recycling with Fabry-Perot cavities since the tuning is anti-resonant and the compound mirror has little frequency dependence (see 4.2.3 and Fig. 4.8). In resonant sideband extraction, however, this has a strong effect because the tuning of the signal extraction cavity is resonant and its complex reflectivity strongly depends on the light frequency (unless its length is very short). Thus, we must treat the full expression of a three-mirror coupled cavity for a more precise discussion.

The characteristics of three-mirror coupled cavities are, however, poorly understood so far. It may worth spending some time to analyze three-mirror coupled cavities in a general sense, since it is an essential part of resonant sideband extraction.

4.2: Three-mirror coupled cavities

4.2.1: Optics matrices

In general, a cavity is composed of multiple (partially transmitting) mirrors aligned to make the optical path which forms a ‘loop’ (or loops), i.e. a part of the light may trace the same path more than once. For simplicity, only ‘on-axis²⁾’ cases are considered in the following discussion, although it is possible to construct a ring-type cavity in which mirrors are not aligned to a single axis. In both cases the behavior is the same in principle.

To concentrate on the frequency-dependent response of the cavity, only a single spatial mode will be considered here. (Further discussion on the higher order modes is treated in Chapter 5.) Thus, it is assumed that the mirrors have appropriate curvatures for a stable mode to exist and that the incident light is matched to the mode. With these assumptions, the resonance feature (tuning) is determined by the cavity length and the frequency of the light.

If all the optics are aligned to a single axis, the field at any point of the axis can be considered as the sum of two propagating waves: one has the same direction of propagation as the incident light from the light source and the other has the opposite:

$$\mathbf{a} = a_{\rightarrow} e^{i(\omega t - kx)} + a_{\leftarrow} e^{i(\omega t + kx)} = \begin{pmatrix} e^{i(\omega t - kx)} & e^{i(\omega t + kx)} \end{pmatrix} \begin{pmatrix} a_{\rightarrow} \\ a_{\leftarrow} \end{pmatrix}. \quad (1)$$

²⁾ Here ‘on-axis’ is used to mean the optics are aligned so that the reflected beam has the same path as the incoming beam.

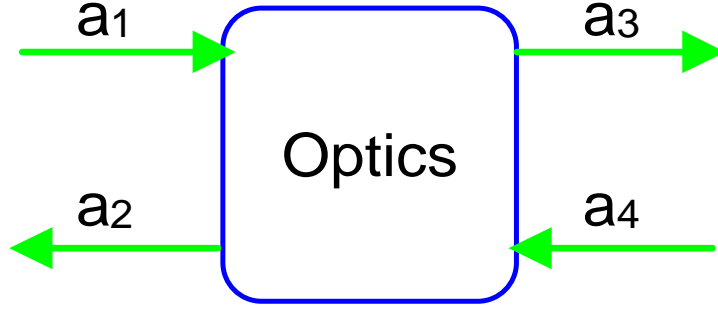


Figure 4.3: There are two waves on an optical axis propagating oppositely to each other, which can be represented by a two-dimensional complex vector. An optical component couples the fields at either side of it, thus it can be represented by a 2×2 matrix.

Then the amplitudes of the two waves can be treated as a two-dimensional complex vector.

Any optical component (or combination of them) on the axis couples the two vectors at either side of it, as is shown in Fig. 4.3. Thus, it is represented by a 2×2 matrix. For example, a finite distance imposes a phase delay on the light passing through it and some absorption (if there is) of the light through the medium. In this case there is no coupling between the two oppositely propagating waves. Thus, the matrix to represent a finite length ℓ will be

$$\mathbf{S} = \begin{pmatrix} e^{-(\gamma+ik)\ell} & 0 \\ 0 & e^{(\gamma+ik)\ell} \end{pmatrix}, \quad (2)$$

where γ and $k = 2\pi/\lambda = \omega/v$ are the absorption and the wave number per unit length in the medium, respectively.

A mirror couples the fields at either side of it by its (amplitude) reflectivity ρ and transmittance τ . The outgoing wave from one side of a mirror is the sum of partial reflection of the wave from the same side and partial transmission of the wave from the other side. This gives the matrix for a mirror \mathbf{M} as

$$\mathbf{M} = \frac{-1}{i\tau} \begin{pmatrix} 1-\mathcal{A} & -\rho \\ \rho & -1 \end{pmatrix} \quad (3)$$

where $\mathcal{A} = 1-\rho^2-\tau^2$ is the power loss per reflection due to absorption and scattering.

Any combination of mirrors aligned to a single axis can in principle be represented using matrices of the two types described above. Usually light is injected from one side and nothing from the other. From the two equations, the amplitudes of outgoing waves from both sides relative to the incident wave can be obtained. Let us take a Fabry-Perot cavity as an example.

4.2.2: Fabry-Perot cavities

A Fabry-Perot cavity is composed of two mirrors aligned to an axis and separated by a finite distance, as is shown in Fig. 4.4. When light is injected along the optical axis of the cavity through the *coupling mirror* (**Mirror1** in Fig. 4.4), a part of the light enters the cavity and will be reflected back and forth inside the cavity. The many contributions from different numbers of reflections compose the light field inside the cavity.

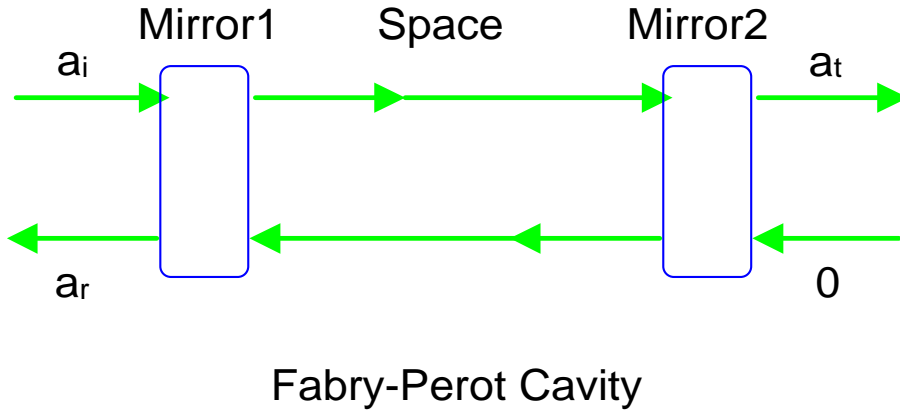


Figure 4.4: A Fabry-Perot cavity is composed of two mirrors separated by a finite distance, which can be represented by the product of three matrices. In usual applications, light is injected from one side (\mathbf{a}_i) and nothing from the other ($\mathbf{0}$). This makes it possible to solve the reflected and transmitted amplitudes ($\mathbf{a}_r, \mathbf{a}_t$) relative to the incident amplitude.

The reflection at the coupling mirror is, however, actually an interference between the two amplitudes: one is the reflection of the internal amplitude of the cavity and the other is the transmission of the incident light. Furthermore, the internal amplitude is in turn a result of the interference at the coupling mirror. This ‘loop’ makes the analysis of a cavity complicated.

The advantage of using matrices described above is that it is not necessary to worry about the actual form of the internal amplitude but to consider the relation between the amplitudes. By this method, a Fabry-Perot cavity can be represented as a combination of two mirrors and the space between them:

$$\begin{aligned}
 \mathbf{M}_2 \mathbf{S} \mathbf{M}_1 &= \frac{-1}{\tau_1 \tau_2} \begin{pmatrix} 1 - \mathcal{A}_2 & -\rho_2 \\ \rho_2 & -1 \end{pmatrix} \begin{pmatrix} e^{-ik\ell} & 0 \\ 0 & e^{ik\ell} \end{pmatrix} \begin{pmatrix} 1 - \mathcal{A}_1 & -\rho_1 \\ \rho_1 & -1 \end{pmatrix} \\
 &= \frac{-1}{\tau_1 \tau_2} \begin{pmatrix} -\rho_1 \rho_2 e^{ik\ell} + (1 - \mathcal{A}_1)(1 - \mathcal{A}_2) e^{-ik\ell} & \rho_2 e^{ik\ell} - \rho_1 (1 - \mathcal{A}_2) e^{-ik\ell} \\ -\rho_1 e^{ik\ell} + \rho_2 (1 - \mathcal{A}_1) e^{-ik\ell} & e^{ik\ell} - \rho_1 \rho_2 e^{-ik\ell} \end{pmatrix}, \quad (4)
 \end{aligned}$$

assuming the medium between the two mirrors to be the vacuum, which is lossless with speed of light c .

Considering the fact that the light is injected from one side of a cavity and nothing from the other side, the fields at either side of the cavity are related by

$$\begin{pmatrix} a_t \\ 0 \end{pmatrix} = M_2 S M_1 \begin{pmatrix} a_i \\ a_r \end{pmatrix}, \quad (5)$$

where a_i , a_t , and a_r are the complex amplitudes of the incident wave, the transmitted wave, and the reflected wave, respectively.

Solving these equations yields the amplitudes of the transmitted and the reflected waves relative to the incident wave, i.e. the (complex) reflectivity and the transmittance of the cavity. The results are

$$\begin{aligned} \rho_{\text{FP}}(\phi) &= \frac{a_r(\phi)}{a_i} = \frac{\rho_1 - (1 - \mathcal{A}_1)\rho_2 e^{-i2\phi}}{1 - \rho_1 \rho_2 e^{-i2\phi}} \\ \tau_{\text{FP}}(\phi) &= \frac{a_t(\phi)}{i a_i} = \frac{i \tau_1 \tau_2 e^{-i\phi}}{1 - \rho_1 \rho_2 e^{-i2\phi}}, \end{aligned} \quad (6)$$

where $\phi = k\ell = \omega\ell/c \pmod{2\pi}$ is the *tuning parameter* of the cavity. This tuning parameter represents the fractional part of the cavity length measured in wavelengths.

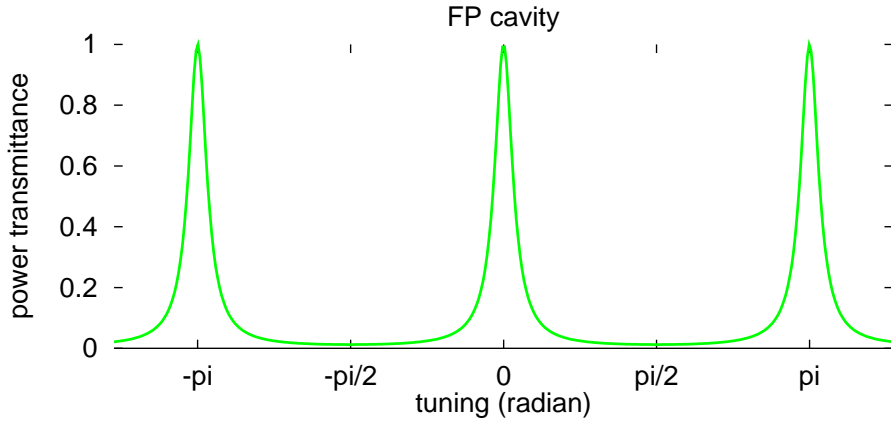


Figure 4.5: The power transmittance of a Fabry-Perot cavity as a function of its tuning. Both mirrors are assumed to have power reflectivities of ≈ 0.8 and no loss, thus the cavity has a finesse of ≈ 28 and the coupling is matched.

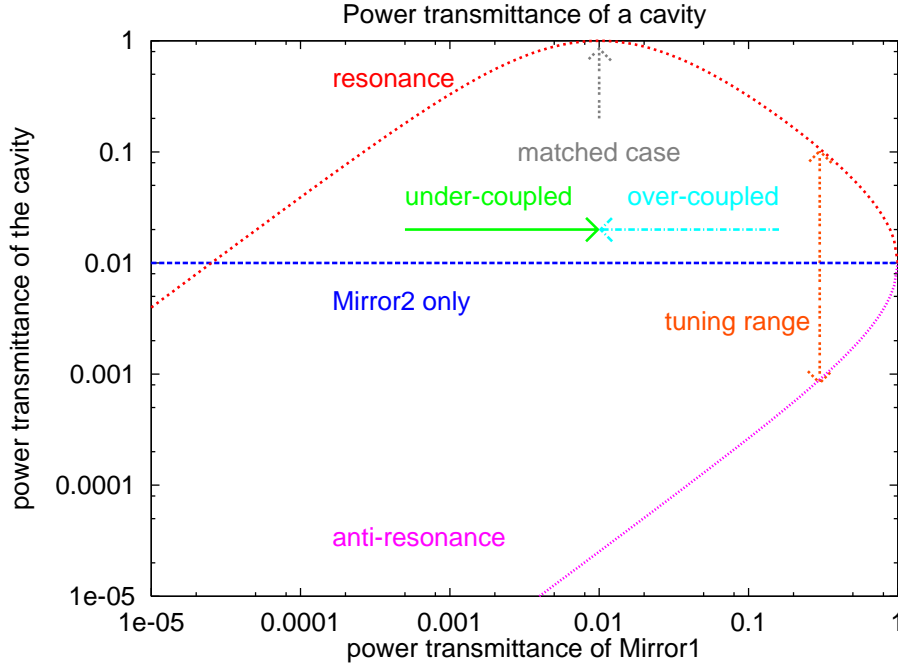


Figure 4.6: The power transmittance of a Fabry-Perot cavity as a function of the power transmittance of the coupling mirror, while the other mirror is fixed (zero loss is assumed). The cases when the cavity is resonant or anti-resonant are shown. The power transmittance of the cavity in any tuning condition must be somewhere between these two cases. The power transmittance of the cavity can be either more or less than that of the fixed mirror, depending on the tuning. The coupling is *matched* when the power transmittance at a resonance is the maximum (of unity if no loss). Otherwise the cavity is classified either as *over-coupled* (the coupling mirror is more transmissive than in matched case) or as *under-coupled* (less transmissive).

The response of a Fabry-Perot cavity is periodic with ϕ by a modulus 2π , as can be seen from eq. (6). As a consequence, it is also periodic with the light frequency ω by a modulus determined by the cavity length

$$\Delta\omega = 2\pi/(\ell/c) = 2\pi c/\ell. \quad (7)$$

This corresponds to twice the free-spectral-range Δf_{FSR} of the cavity. Thus, the cavity length changes the tuning parameter in a microscopic sense ($\ll \lambda$) and its frequency dependence in a macroscopic sense ($\gg \lambda$).

Fig. 4.5 shows the dependence of the power transmittance of a cavity on the tuning parameter ϕ . The light will be resonant, i.e. the power transmittance is the maximum, when $\phi = 0$ or π , and on anti-resonant, the minimum, when

$\phi = \pm \frac{\pi}{2}$. The amplitude reflectivities and transmittances of the cavity when resonant and anti-resonant are given by

$$\begin{aligned} \rho_{\text{FP}}(0) = \rho_{\text{FP}}(\pi) &= \frac{\rho_1 - (1 - \mathcal{A}_1)\rho_2}{1 - \rho_1\rho_2} & \tau_{\text{FP}}(0) = -\tau_{\text{FP}}(\pi) &= \frac{i\tau_1\tau_2}{1 - \rho_1\rho_2} \\ \rho_{\text{FP}}\left(\frac{\pi}{2}\right) = \rho_{\text{FP}}\left(-\frac{\pi}{2}\right) &= \frac{\rho_1 + (1 - \mathcal{A}_1)\rho_2}{1 + \rho_1\rho_2} & \tau_{\text{FP}}\left(\frac{\pi}{2}\right) = -\tau_{\text{FP}}\left(-\frac{\pi}{2}\right) &= \frac{\tau_1\tau_2}{1 + \rho_1\rho_2}. \end{aligned} \quad (8)$$

The power reflectivity of the cavity in an arbitrary tuning condition must be somewhere between these two cases.

Fig. 4.6 shows the dependence of the power transmittance of a cavity (when it is resonant or anti-resonant) on the power transmittance of the coupling mirror, assuming a fixed reflectivity for the reflecting mirror. It is seen that the power transmittance can be either increased or decreased from that of the fixed mirror by putting another mirror to compose a cavity and by adjusting the tuning appropriately. This feature is used by resonant sideband extraction and signal recycling with Fabry-Perot cavities in the arms.

4.2.3: Under- and over-coupling

It is interesting to note, in Fig. 4.6, that the power transmittance of a cavity on resonance has a maximum value (which is unity if there is no loss) given by a choice of the reflectivity of the one of the mirrors. Either increasing or decreasing the transmittance (or reflectivity) of the coupling mirror (**Mirror1**) results in less power transmittance. In other words, there are two possible values for the coupling mirror reflectivity to realize a certain (non-unity) power transmittance of the cavity on resonance (as long as this transmittance is higher than that of the reflecting mirror **Mirror2**).

What is described above is actually shown in eq. (8). The reflectivity of a cavity on resonance is zero when $\rho_1 = (1 - \mathcal{A}_1)\rho_2$, the so-called *matched case*. Either increasing or decreasing the reflectivity of one of the mirrors results in more reflection of the incident power, but whose amplitude is opposite in sign. To classify the two cases we call the cavity *under-coupled* when $\rho_1 > (1 - \mathcal{A}_1)\rho_2$ and *over-coupled* when $\rho_1 < (1 - \mathcal{A}_1)\rho_2$. The difference between the two cases will be clear when one considers the phasor diagram of the reflected light from a cavity.

For this purpose, it is useful to rewrite the complex reflectivity in eq. (6) as follows:

$$\rho_{\text{FP}}(\phi) = \frac{\rho_1 - (1 - \mathcal{A}_1)\rho_2 e^{-i2\phi}}{1 - \rho_1\rho_2 e^{-i2\phi}} = \rho_1 + \rho_2 \frac{(i\tau_1)^2 e^{-i2\phi}}{1 - \rho_1\rho_2 e^{-i2\phi}}, \quad (9)$$

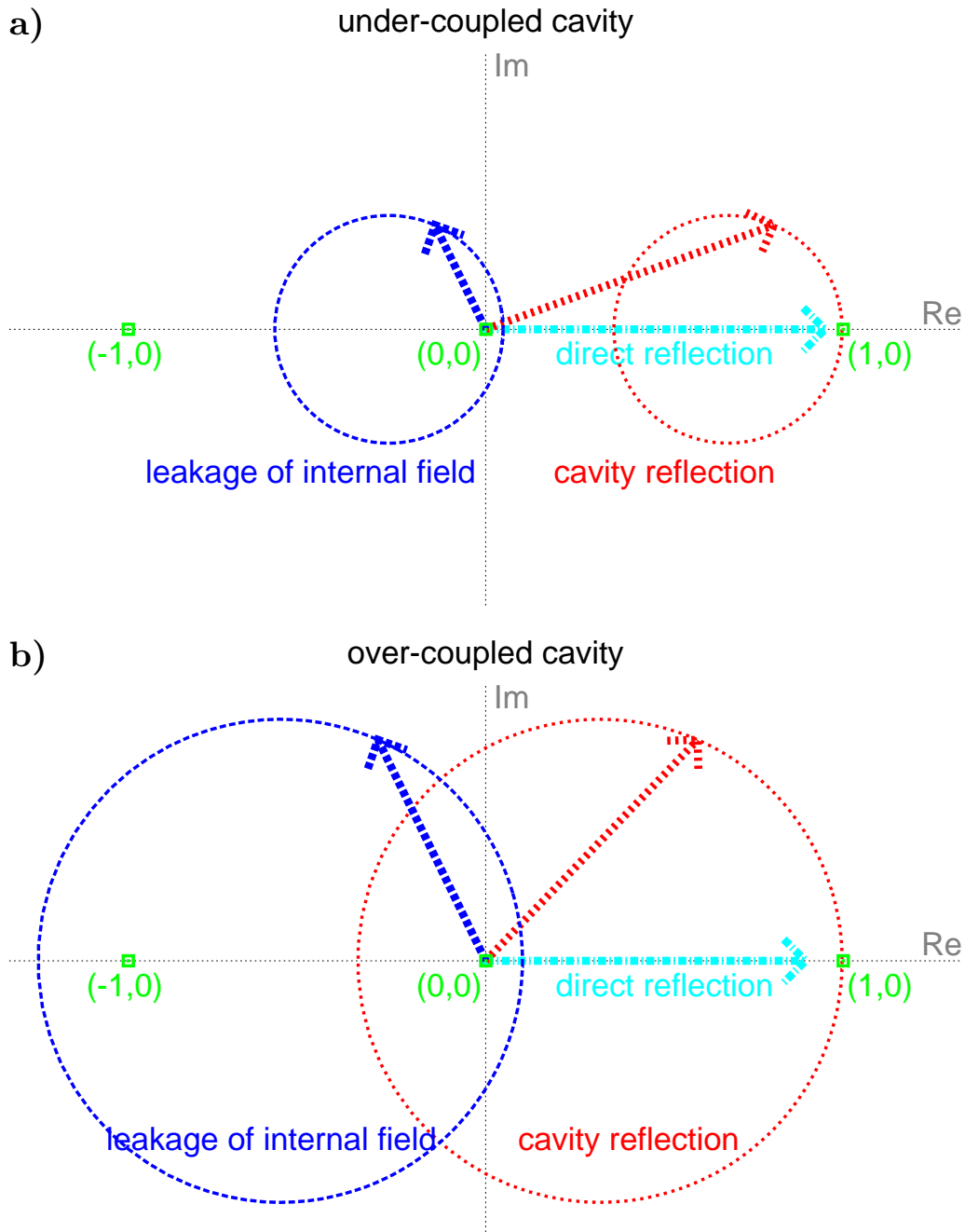


Figure 4.7: The phasor diagram of the light reflected by a cavity, **a)** an under-coupled case and **b)** an over-coupled case. In the under-coupled case, the locus of reflected light along the change of the tuning condition always stays in the right half of the complex plane. On the other hand, its locus in the over-coupled case encloses the origin. This makes a difference in the phase of the ‘identical’ vectors. In either case the locus is a circle, though its orientation (angle measured from the positive half of the real axis) does not directly represent the tuning condition.

where $1 - \mathcal{A}_1 = (\rho_1)^2 + (\tau_1)^2$ was used. In the last expression, the first term represents the field directly reflected at the front surface of the coupling mirror and does not depend on the tuning of the cavity. On the other hand, the second term represents the field leaking out of the cavity, whose amplitude as well as phase varies with tuning. (Note the similarity between the second term and the transmittance of the cavity found in eq. (6).) The reflected light from a cavity can be thought as a sum of these two fields.

In a phasor diagram, the first (constant) term in eq. (9) is expressed as a fixed length vector (we take this direction as the positive part of the real axis) and the second (tuning-dependent) term as a rotating vector whose length varies with the tuning of the cavity, as is shown in Fig. 4.7. When at resonance, the rotating vector has the opposite direction to the fixed one so that their sum (which represents the light reflected from the cavity) is minimized. On the other hand, the vectors have the same direction when at anti-resonance and their sum is maximized.

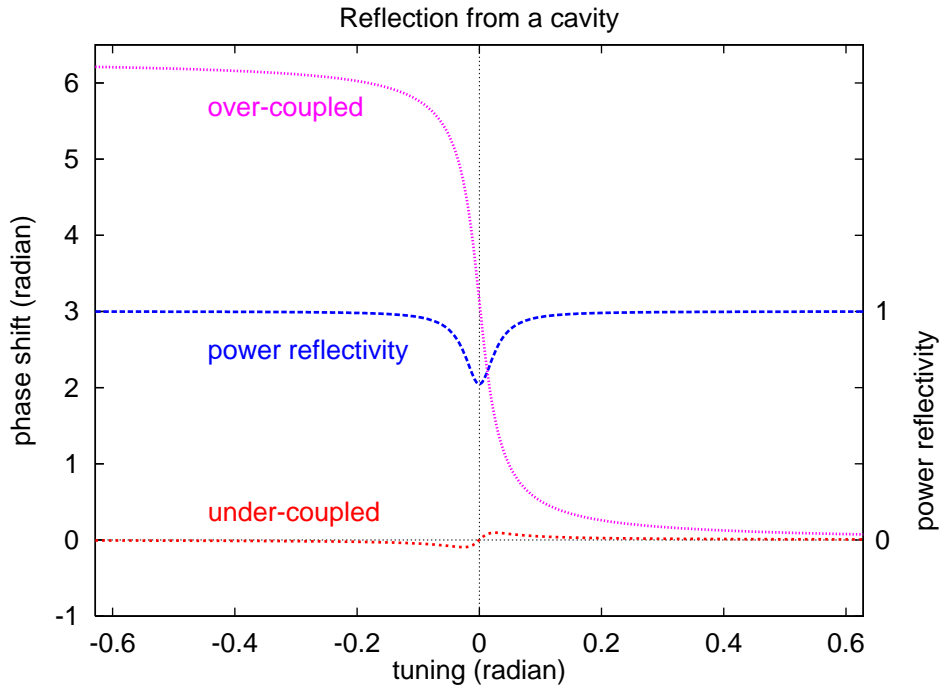


Figure 4.8: The phase shifts of reflected light from under- and over-coupled cavities. The power reflectivity of these cavities, the same in both cases, is also shown. It can be seen that the phase shift from an over-coupled cavity changes by 2π around each resonance, whereas that from an under-coupled cavity just ‘fluctuates’ near zero around the resonance. It should also be noticed that the gradient of phase shift from an under-coupled cavity can be positive near a resonance.

In the case of an under-coupled cavity, the rotating vector is always shorter than the fixed one. This makes their sum always remain in the right half of the complex plane. Thus, the minimum of the summed vector has the same direction as the fixed one. For an over-coupled cavity, however, the rotating vector is longer than the fixed one when on resonance. Thus, the locus of the sum encloses the origin of the complex plane through the change of the tuning parameter ϕ from 0 to π .

The above distinction makes a big difference when one considers the phase shift of the reflected light around a resonance. In Fig. 4.7, the phase shift upon the reflection is represented by the angle of the summed vector measured from the positive half of the real axis. This is shown in Fig. 4.8 as a function of the tuning of the cavity. The phase shift increases by an additional 2π at each resonance³⁾ for an over-coupled cavity, whereas it just ‘fluctuates’ around zero for an under-coupled cavity. The phase shifts at resonance differ by π between the two cases, representing their opposite directions.

Fig. 4.8 also clarifies the fact that the power reflectivity and the phase shift are almost constant for tunings far away from resonance, in both under- and over-coupled cavities. This allows a cavity to be approximated by a mirror for a relatively wide range of tuning as long as there is no resonance in this range.

4.2.4: Three-mirror coupled cavities

A three-mirror cavity can be analyzed in a similar way to a Fabry-Perot cavity, i.e. by solving (see Fig. 4.9)

$$\begin{pmatrix} a_t \\ 0 \end{pmatrix} = M_3 S_2 M_2 S_1 M_1 \begin{pmatrix} a_i \\ a_r \end{pmatrix}, \quad (10)$$

with (assuming the vacuum between the mirrors)

$$M_3 S_2 M_2 S_1 M_1 = \frac{-i}{\tau_1 \tau_2 \tau_3} \times \begin{pmatrix} 1-\mathcal{A}_3 & -\rho_3 \\ \rho_3 & -1 \end{pmatrix} \begin{pmatrix} e^{-ik\ell_2} & 0 \\ 0 & e^{ik\ell_2} \end{pmatrix} \begin{pmatrix} 1-\mathcal{A}_2 & -\rho_2 \\ \rho_2 & -1 \end{pmatrix} \begin{pmatrix} e^{-ik\ell_1} & 0 \\ 0 & e^{ik\ell_1} \end{pmatrix} \begin{pmatrix} 1-\mathcal{A}_1 & -\rho_1 \\ \rho_1 & -1 \end{pmatrix}. \quad (11)$$

³⁾ It may be surprising that there is a difference in the phases of vectors pointing in the same direction. This difference arises from the fact that the negative half of the real axis is the branching line of the logarithmic function and crossing it makes a difference of 2π in its imaginary part.

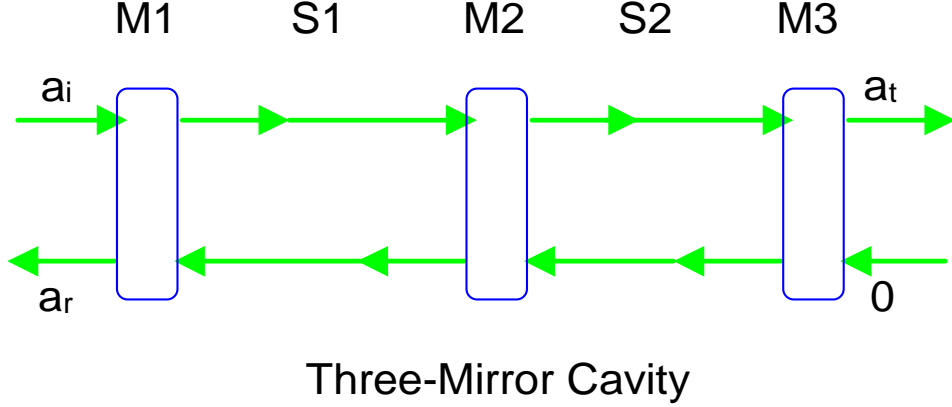


Figure 4.9: A three-mirror coupled cavity is composed of three mirrors, each separated by a finite distance. This can be represented by the product of five matrices. In a similar way to the Fabry-Perot case, the reflected and transmitted amplitudes (\mathbf{a}_r , \mathbf{a}_t) relative to the incident one (\mathbf{a}_i) can be obtained.

Assuming no loss for the mirrors ($\mathcal{A}_1=\mathcal{A}_2=\mathcal{A}_3=0$) for simplicity, the results are

$$\begin{aligned} \rho_{3M}(\phi_1, \phi_2) &= \frac{a_r(\phi_1, \phi_2)}{a_i} = -\frac{\rho_3 e^{-i\phi_2} (e^{-i\phi_1} - \rho_1 \rho_2 e^{i\phi_1}) - e^{i\phi_2} (\rho_2 e^{-i\phi_1} - \rho_1 e^{i\phi_1})}{\rho_3 e^{-i\phi_2} (\rho_2 e^{i\phi_1} - \rho_1 e^{-i\phi_1}) - e^{i\phi_2} (e^{i\phi_1} - \rho_1 \rho_2 e^{-i\phi_1})} \\ \tau_{3M}(\phi_1, \phi_2) &= \frac{a_t(\phi_1, \phi_2)}{i a_i} = \frac{\tau_1 \tau_2 \tau_3}{\rho_3 e^{-i\phi_2} (\rho_2 e^{i\phi_1} - \rho_1 e^{-i\phi_1}) - e^{i\phi_2} (e^{i\phi_1} - \rho_1 \rho_2 e^{-i\phi_1})}, \end{aligned} \quad (12)$$

where $\phi_1 = \omega l_1 / c \pmod{2\pi}$ and $\phi_2 = \omega l_2 / c \pmod{2\pi}$ are the tuning parameters. Due to these two independent tuning parameters, the behavior of a three-mirror coupled cavity is more complicated than a standard Fabry-Perot cavity. Let us calculate it taking the power transmittance as an example.

The power transmittance $|\tau_{3M}(\phi_1, \phi_2)|^2$ of a three-mirror coupled cavity can be plotted taking the two parameters ϕ_1 and ϕ_2 as the x - and y -axes, respectively. An example is shown in Fig. 4.10. Note that this pattern is determined only by the reflectivities of the mirrors. Since the function is periodic to both parameters, this pattern must fill the whole x - y plane like ‘tiles’, as is illustrated in Fig. 4.11. Any condition of this coupled cavity must be represented by a point somewhere in the x - y plane, and thus a point in a ‘tile’.

When the light frequency changes by $\Delta\omega$, the change in each tuning parameter is proportional to it; $\Delta\phi_j = \Delta\omega l_j / c$ ($j=1, 2$). Then the two tuning parameters are also proportional to each other, i.e. $\Delta\phi_1 / \Delta\phi_2 = l_1 / l_2$. This relation can be represented by an oblique line (bias) in the x - y plane filled by the tiles, as is shown in Fig. 4.11. This oblique line has the slope of l_1 / l_2 (which

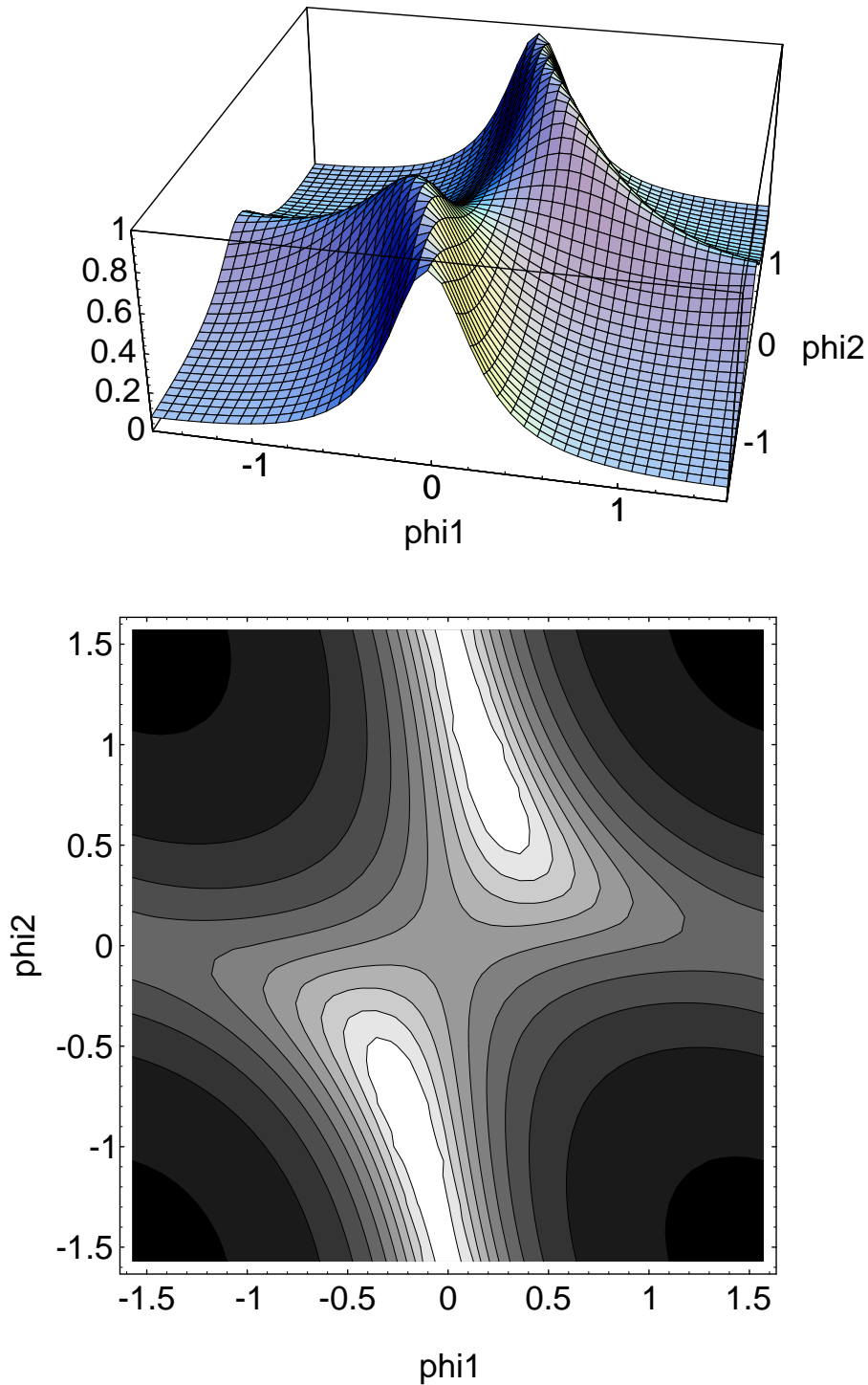


Figure 4.10: A sample of a ‘tile pattern’ which represents the power transmittance of a three-mirror coupled cavity. The two axes correspond to the two tuning parameters, respectively. The upper picture shows the three-dimensional view and the lower shows its contour map. This pattern is determined only by the power reflectivities of the mirrors; they are 0.5, 0.3, and 0.2, and zero loss is assumed.

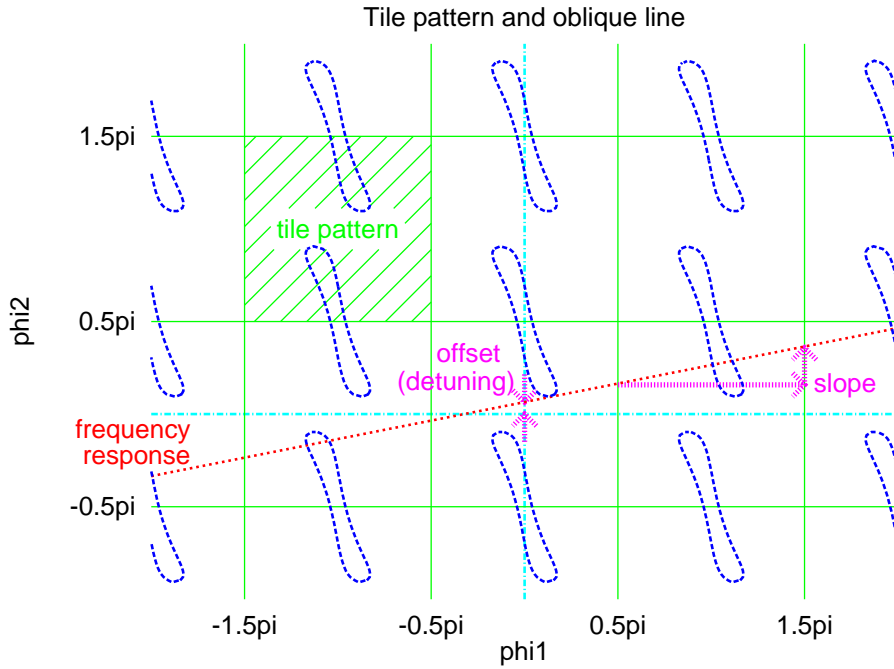


Figure 4.11: The response of a three-mirror coupled cavity is periodic to two tuning parameters, thus whole x - y plane is divided into small sections by a ‘mesh’. Each section contains the ‘tile pattern’ shown in Fig. 4.10, though only a simplified one is presented here. The light frequency changes the two tuning parameters proportionally to each other, thus the relation between them is expressed by an oblique line in this x - y plane. The offset and the slope of this oblique line are determined by the initial tuning condition and the ratio of the distances between the mirrors, respectively. The frequency response of the three-mirror coupled cavity is obtained as the ‘cross-section’ of this x - y plane along this oblique line. It will not be periodic to the light frequency unless the slope is a rational number.

is always positive) and is shifted by the offsets determined by the initial tuning conditions.

Thus, the light frequency dependence of the power transmittance of the three-mirror coupled cavity will be the ‘cross section’ of the x - y plane along the oblique line. If the slope ℓ_1/ℓ_2 is a rational number, then the frequency dependence will be periodic since it will pass through exactly the same point of a tile after crossing several tiles. Otherwise, however, it will never trace the same point in a tile.

Although the changes in the distances between the mirrors do not change the pattern of a tile, they alter the frequency dependence by changing the offsets and the slope of the oblique line. A microscopic change in one of the

4.3: Frequency response of resonant sideband extraction

mirror distances varies the corresponding offset, thus ‘shifts’ the position of the oblique line. Note that the motion of the center mirror makes changes in both distances simultaneously. On the other hand, a macroscopic change in one of the distances varies the frequency dependence of the corresponding tuning parameter and thus changes the slope of the oblique line.

It should be noted that $\phi_1=0$ or $\phi_2=0$, which are the resonance condition if the first or the third mirror is not present, are not necessarily the resonance condition of the three-mirror coupled cavity. This can be true in a few cases, such as $\phi_1=0$ or $\phi_2=\pi/2$, but not in general. Thus, excepting such cases, one must treat the full expression of a three-mirror coupled cavity shown in eq. (12).

4.3: Frequency response of resonant sideband extraction

4.3.1: The effect of the signal cavity

The frequency response of an interferometer with resonant sideband extraction can be obtained by inserting the reflectivity and the transmittance of the ‘compound mirror’ into the generalized response given in eq. (3.44).

$$G_{\text{SX}}(\omega) = \sqrt{\mathcal{G}_{\text{arm}}} \times \frac{i\tau_{\text{CM}}}{1 - \rho_{\text{CM}}\rho_{\text{r}}e^{-i\omega t_{\text{a}}}} \times \frac{\rho_{\text{r}}\omega_0}{4} \frac{1 - e^{-i\omega t_{\text{a}}}}{i\omega}. \quad (13)$$

As is discussed in 3.4.1, the frequency response is mainly determined by the second factor if the arm length is short compared with the wavelength of the gravitational wave. Let us consider this term in more detail.

The reflectivity and the transmittance of the compound mirror are those of a Fabry-Perot cavity, but the losses inside the cavity must be taken into account. Thus eq. (6) must be modified as

$$\begin{aligned} \rho_{\text{CM}}(\phi_{\text{s}}) &= \frac{\rho_{\text{c}} - \rho_{\text{SX}}(1 - \mathcal{A}_{\text{c}})(1 - \mathcal{A}_{\text{BS}})\sqrt{\frac{1+\mathcal{C}}{2}}e^{-i2\phi_{\text{s}}}}{1 - \rho_{\text{c}}\rho_{\text{SX}}(1 - \mathcal{A}_{\text{BS}})\sqrt{\frac{1+\mathcal{C}}{2}}e^{-i2\phi_{\text{s}}}} \\ \tau_{\text{CM}}(\phi_{\text{s}}) &= \frac{i\tau_{\text{c}}\tau_{\text{SX}}\sqrt{(1 - \mathcal{A}_{\text{BS}})\frac{1+\mathcal{C}}{2}}e^{-i\phi_{\text{s}}}}{1 - \rho_{\text{c}}\rho_{\text{SX}}(1 - \mathcal{A}_{\text{BS}})\sqrt{\frac{1+\mathcal{C}}{2}}e^{-i2\phi_{\text{s}}}} \end{aligned} \quad (14)$$

using accordingly changed notations. Here $\phi_{\text{s}} = \omega\ell_{\text{s}}/c + \psi$ is the tuning of the signal extraction cavity with $\psi = \omega_0\ell_{\text{s}}/c$ the initial detuning at the carrier frequency, \mathcal{C} the apparent visibility, \mathcal{A}_{c} the loss at the surface of Fabry-Perot coupling mirror, and \mathcal{A}_{BS} the losses at the beamsplitter and in the substrates of the Fabry-Perot coupling mirrors.

By inserting the reflectivity and transmittance of the compound mirror into eq. (13), the second factor can be expanded as follows:

$$\begin{aligned}
 G_{\text{SX}} &= \frac{i\tau_{\text{CM}}}{1 - \rho_{\text{CM}}\rho_{\text{r}}e^{-i2\phi_{\text{a}}}} = \frac{i\frac{\tau_{\text{s}}\tau_{\text{c}}e^{-i\phi_{\text{s}}}}{1 - \rho_{\text{s}}\rho_{\text{c}}e^{-i2\phi_{\text{s}}}}}{1 - \frac{\rho_{\text{c}} - \rho_{\text{s}}(1 - \mathcal{A}_{\text{c}})e^{-i2\phi_{\text{s}}}}{1 - \rho_{\text{s}}\rho_{\text{c}}e^{-i2\phi_{\text{s}}}}\rho_{\text{r}}e^{-i2\phi_{\text{a}}}} \quad (15) \\
 &= \frac{-\tau_{\text{s}}\tau_{\text{c}}e^{-i\phi_{\text{s}}}}{1 - \rho_{\text{s}}\rho_{\text{c}}e^{-i2\phi_{\text{s}}} - \rho_{\text{c}}\rho_{\text{r}}e^{-i2\phi_{\text{a}}} + (1 - \mathcal{A}_{\text{c}})\rho_{\text{s}}\rho_{\text{r}}e^{-i2(\phi_{\text{a}} + \phi_{\text{s}})}} ,
 \end{aligned}$$

where $\phi_{\text{a}} = \omega\ell_{\text{a}}/c$ is the tuning of the arm cavity, and the ‘effective’ reflectivity ρ_{s} and transmittance τ_{s} of the signal extraction mirror, defined as

$$\begin{cases} \rho_{\text{s}} = \rho_{\text{SX}}(1 - \mathcal{A}_{\text{BS}})\sqrt{\frac{1 + \mathcal{C}}{2}} \\ \tau_{\text{s}} = \tau_{\text{SX}}\sqrt{(1 - \mathcal{A}_{\text{BS}})\frac{1 + \mathcal{C}}{2}} \end{cases} , \quad (16)$$

are used for simplicity.

One may observe the similarity between this expression and the transmittance of a three-mirror coupled cavity found in eq. (12). Assuming no loss, the only difference is the additional factor $\tau_{\text{r}}e^{-i\phi_{\text{a}}}$. This difference can be understood as resulting from two causes: the sidebands produced inside the cavity do not traverse the reflecting mirror of the arm cavity, and the reference for phase measurements are different in the two cases. With the exception of these points, the resonance features in the two cases are identical.

As discussed in 2.1.3, the sidebands due to a gravitational wave (or any kind of modulation) are *produced* inside the arm cavity and their amplitudes are proportional to that of the carrier. The power of the sideband is a result of the multiple interference inside the signal cavity, and thus depends on the resonance condition of this cavity. It should be emphasized that the magnitude of the sideband produced is determined by the amplitude, not by the power. This is similar to the light injected into a cavity through one of the mirrors; in this case the amplitude inside the cavity is determined by that of the injected light and the transmittance of the coupling mirror.

The above similarity implies that a cavity does not distinguish the amplitude produced inside the cavity from that injected through one of the mirrors. The field leaking out of the cavity must then be the same in the two cases. The amplitude of the light reflected from a cavity is, however, different since it is a result of the interference of the amplitude directly reflected from the coupling

mirror with that leaking out of the cavity. For the sidebands produced inside the cavity, on the contrary, there will be no interference.

As a consequence of this similarity, all the discussion on the transmittance of a three-mirror coupled cavity found in 4.2.4 are also applicable for the signal cavity of resonant sideband extraction. Thus, we qualitatively discuss the characteristics of the signal cavity taking Fig. 4.10 as an example. In the figure, ϕ_1 and ϕ_2 correspond to ϕ_a and ϕ_s , respectively. Note, however, that the one shown in Fig. 4.10 is not necessarily a good example of parameters for gravitational wave detectors. More realistic choices of parameters are shown in 4.4.

4.3.2: Frequency response

As is discussed in 4.2.4, the power transmittance of a three-mirror coupled cavity can be represented by the pattern in a single ‘tile’ (like that in Fig. 4.10), whose shape is determined by the reflectivities of the mirrors. The frequency response of the cavity is obtained as the cross-section of a tile along an oblique line. The slope of this oblique line is determined by the ratio of distances between the mirrors. Since the arm cavity is short compared with expected wavelengths of gravitational waves (a practical limitation), we only need to consider the response within a single tile.

The horizontal and vertical widths of a tile correspond to the free-spectral-range of the arm cavity and the signal extraction cavity, respectively. It should be noted that these do not depend on the ratio of the two cavity lengths. Since the length of the arm cavity (and thus its free-spectral-range) are one of the (practically) fixed parameters in a gravitational wave detector, it is convenient to measure the linewidth of the resonance by projecting it to the horizontal axis.

In resonant sideband extraction, there is an additional condition that the light from the source is resonant with the (isolated) arm cavity. This means the carrier frequency satisfies

$$\phi_a(\omega_0) = 0 \tag{17}$$

(as is included in eq. (14)). This corresponds to the horizontal center in Fig. 4.10 ($\phi_1=0$ in the figure). The sideband produced inside the cavity must have a finite offset ϕ_a from the horizontal center. The amount of this offset is determined by the modulation frequency ω_m (e.g. the gravitational wave frequency) as

$$\phi_a = \omega_m \ell_a / c, \tag{18}$$

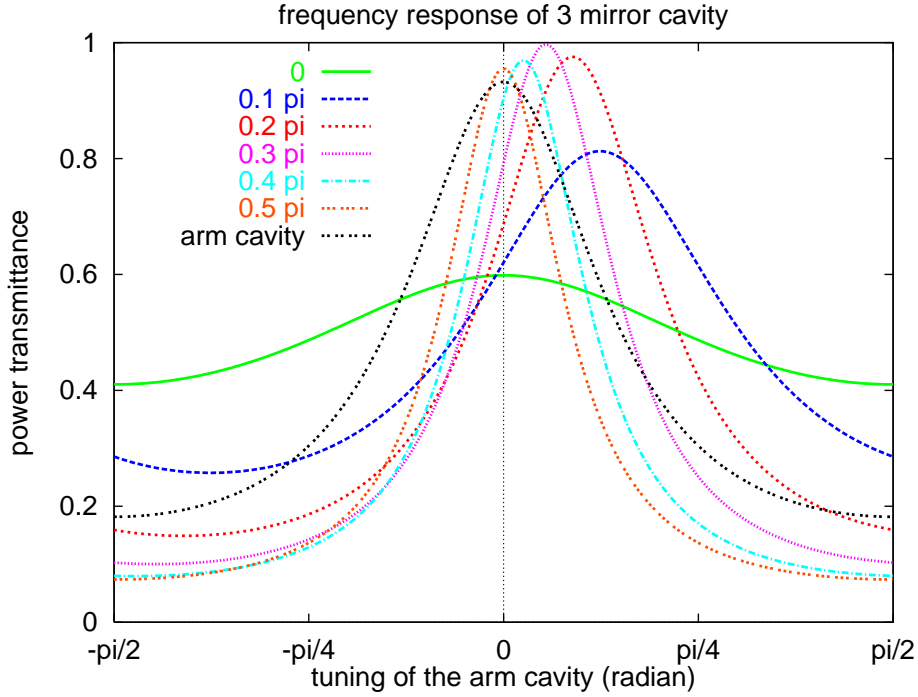


Figure 4.12: The power transmittance of a three-mirror coupled cavity with various amounts of detuning of the compound mirror. Each curve corresponds to a horizontal cross-section of Fig. 4.10 at a different vertical position. For comparison, the response with no mirror at the signal port is also shown. The horizontal axis is the tuning of the arm cavity, which is given by eq. (18) in the text. It can be seen that the linewidth as well as the peak frequency varies with the amount of detuning. With no detuning, a response with a resonance (centered at the carrier frequency) broader than that of the arm cavity is obtained.

where ℓ_a is the length of the arm cavity.

The parameters left undetermined are the slope of the oblique line and its vertical position. The former is assumed to be nearly horizontal, from the practical limitation that the compound mirror is short compared with the arm cavity. Thus, we initially consider the responses obtained by taking cross-sections along horizontal lines. This is to concentrate on the effect of the tuning of the compound mirror, which is represented by the vertical offset of the oblique line.

Fig. 4.12 shows the responses of a three-mirror coupled cavity, each with a different detuning of the compound mirror. It is assumed to be relatively short compared with the arm cavity. The horizontal axis is the tuning of the arm cavity ϕ_a given in eq. (18), and the carrier frequency corresponds the center, i.e. $\phi_a(\omega_0)=0$. For comparison, the response of the arm cavity without any mirror at the signal port is also shown.

4.3: Frequency response of resonant sideband extraction

When $\phi_s(\omega_0)=\frac{\pi}{2}$, i.e. the compound mirror is anti-resonant with the carrier, the coupled cavity is resonant with the carrier with a sharper response than that obtained without the additional mirror. This corresponds to signal recycling, as expected. When $\phi_s(\omega_0)=0$, i.e. the compound mirror is resonant with the carrier, the coupled cavity is also resonant with the carrier but with a much broader linewidth than that obtained without the additional mirror. This is what is expected in resonant sideband extraction.

It is worth mentioning the responses obtained with other tuning conditions ϕ_s of the compound mirror for the carrier. In general, the resonance peak frequency moves away from that of the carrier and the linewidth also varies with the amount of detuning. The response is not symmetric at either side of $\phi_a=0$, i.e. the upper and lower sidebands experience different resonance in the cavity.

These are natural consequences of the situation described in 4.2.3. The power reflectivity as well as the phase shift upon the reflection varies with the tuning condition. The former changes the linewidth of the coupled cavity and the latter changes the peak frequency. Furthermore, the tuning condition depends on the light frequency. These complexities make the intuitive understanding of a coupled cavity comparatively difficult.

It is also expected from Fig. 4.8 in 4.2.3 that the dependence on the tuning of the compound mirror is critical when it is near resonance and less critical near anti-resonance. This is because both the power reflectivity and the phase shift upon the reflection have strong dependence on the tuning only around a resonance, and otherwise they are almost constant. This makes it possible to approximate the response of signal recycling with Fabry-Perot cavities in the arms by that of a Fabry-Perot cavity, when the resonance peak frequency is within the linewidth of the arm cavity.

4.3.3: Coupling of the signal extraction cavity

The characteristics of a three-mirror coupled cavity described above actually depend on the reflectivities of the mirrors used. When mirrors with different reflectivities are chosen, its behavior will change in a qualitative way. Another example of the response, with different reflectivities of the mirrors from the previous one, is shown in Fig. 4.13 and Fig. 4.14, each corresponding to Fig. 4.10 and Fig. 4.12, respectively. One may notice the difference between Fig. 4.12 and Fig. 4.14, especially when there is no detuning.

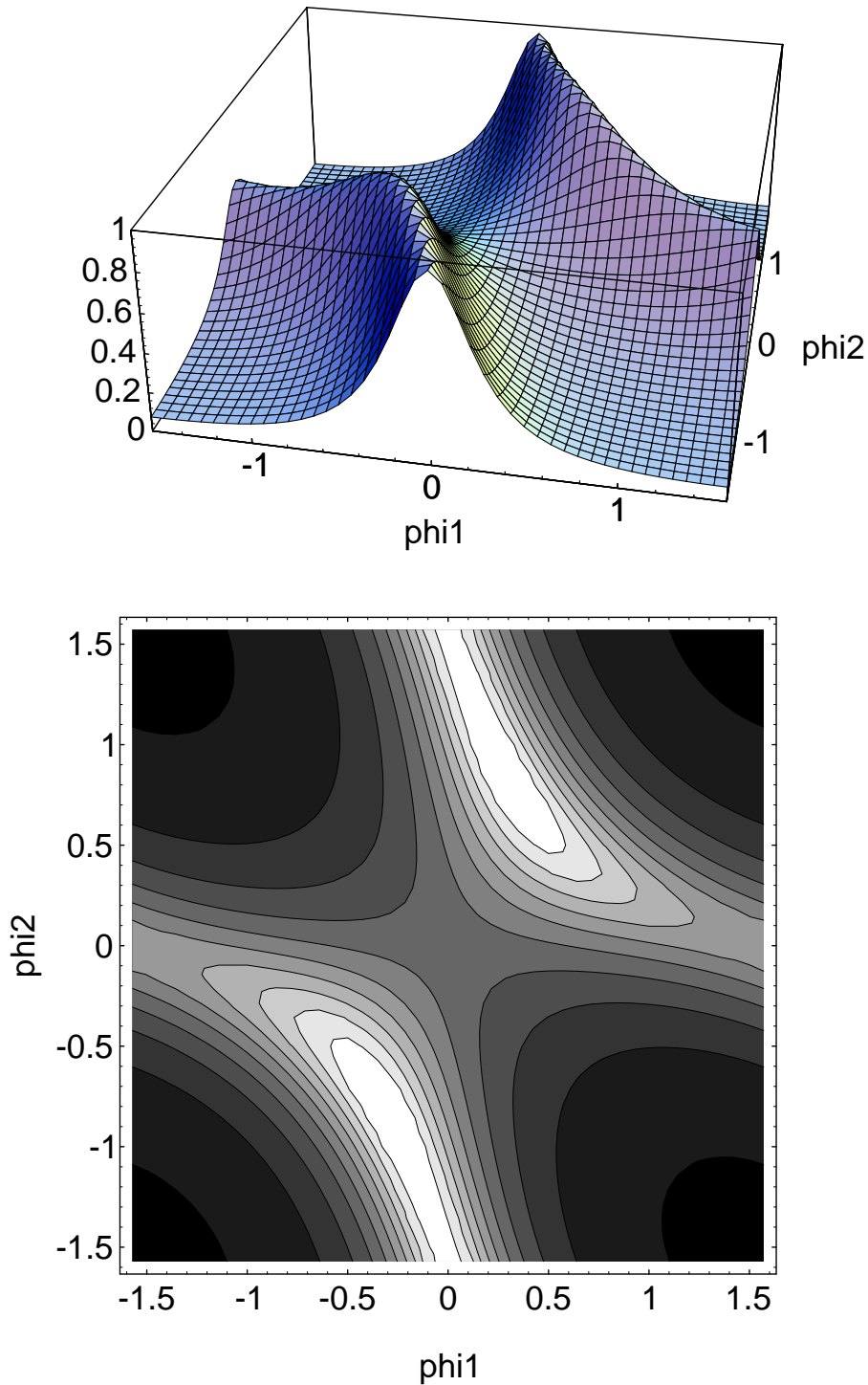


Figure 4.13: The same as Fig. 4.10, with different reflectivities of the mirrors. The power reflectivities of mirrors are 0.5, 0.2, and 0.3, and no loss is assumed. The latter two mirrors are interchanged from those in Fig. 4.10, to show the effect of the coupling of the compound mirror. The significant difference around $\phi_2 \simeq 0$ can be noticed.

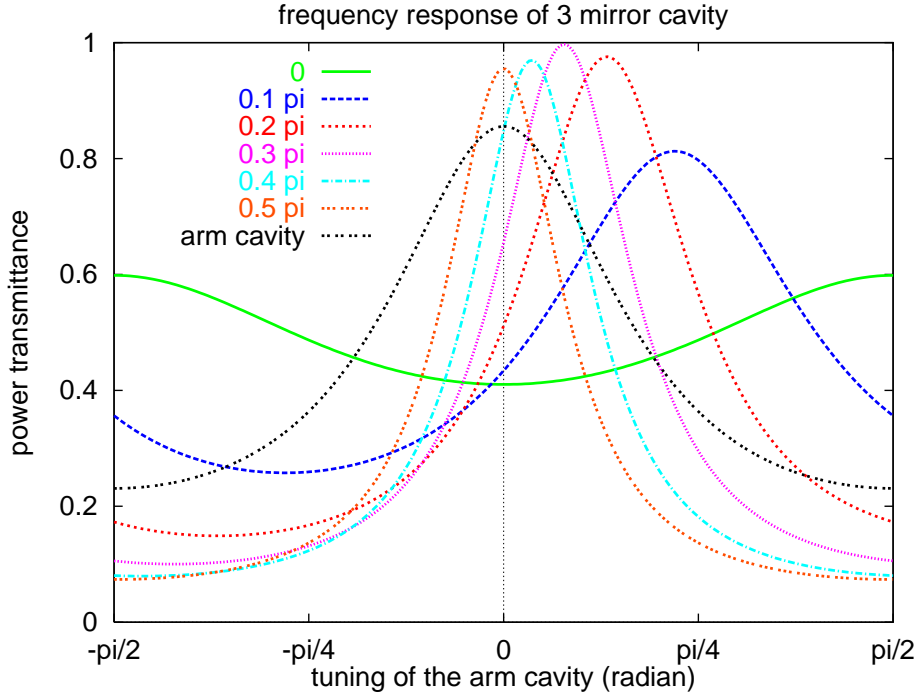


Figure 4.14: The same as Fig. 4.12, with different mirror reflectivities. Each curve corresponds to a horizontal cross-sections of Fig. 4.13 at a different vertical position. With an over-coupled signal extraction cavity, the carrier frequency is anti-resonant when there is no detuning.

The qualitative difference between the two cases is the coupling of the compound mirror. As is described in 4.2.3, there are two possibilities for coupling of a cavity, under- and over-coupling. The compound mirror, looking from the arm cavity side, must be one of these (excepting the matched case). Fig. 4.10 and Fig. 4.12 show the response of the three-mirror coupled cavity with an under-coupled compound mirror, whereas Fig. 4.13 and Fig. 4.14 show that with an over-coupled one.

The biggest difference appears when the compound mirror is resonant with the carrier. If the signal extraction cavity is under-coupled, the three-mirror coupled cavity will still be resonant with the carrier, with a relatively broad linewidth. If the compound mirror is over-coupled, however, it will be anti-resonant with the carrier. This is so because the reflected beam from an over-coupled cavity has opposite phase to that from a mirror (see Fig. 4.7 in 4.2.3 and discussion there).

Placing the third mirror (the signal extraction mirror) adds extra phase shift on the light reflected from the second mirror (the coupling mirror of the

arm cavity). Then the light which was resonant with the arm cavity (in this case, the carrier) will be anti-resonant with this three-mirror coupled cavity. Instead, the light which was anti-resonant with the isolated arm cavity will become resonant.

As a conclusion from the above discussion, it is appropriate to choose an under-coupled cavity as the signal extraction cavity for our application of resonant sideband extraction. In other words, the effective reflectivity of the signal extraction mirror must be chosen to be lower than that of the Fabry-Perot coupling mirror. Then the coupled cavity will have a broader resonance than that of the arm cavity, as desired.

4.3.4: Choosing the length of the signal extraction cavity

The last parameter to be considered is the macroscopic length of the compound mirror, i.e. the signal extraction cavity. As mentioned earlier, the ratio of the lengths of the signal extraction cavity to the arm cavity determines the slope of the oblique line along which the cross-section is taken. The oblique line must have a positive slope as described in 4.2.4, and is assumed to be nearly horizontal.

In signal recycling case where $\phi_s = \pm \frac{\pi}{2}$, it is expected that this slope does not affect the response so much by inspecting Fig. 4.10 and Fig. 4.13. Actually it can only make the linewidth narrower without improving the peak sensitivity. Thus, within the above constraints, elongating the compound mirror does not make any sense in signal recycling with Fabry-Perot cavities.

In the signal extraction case, however, this elongation has an interesting effect on the response. Assuming no detuning of the signal extraction cavity, both the arm cavity and the signal extraction cavity must be resonant with the carrier, thus the oblique line passes through the point $\phi_1 = \phi_2 = 0$, the origin. When the length of the signal extraction cavity is elongated (to keep the tuning condition constant, it must be done by steps of multiple of a wavelength), the slope becomes steeper. When the slope becomes steep enough, the line may cross the ‘skirts’ of the two peaks nearby (at upper-left and lower-right of the origin). Since the pattern of a ‘tile’ is symmetric around the origin, the frequency response is symmetric in upper and lower sidebands.

Fig. 4.15 shows the power transmittance of a three-mirror coupled cavity with various signal extraction cavity lengths while other conditions are fixed.

4.3: Frequency response of resonant sideband extraction

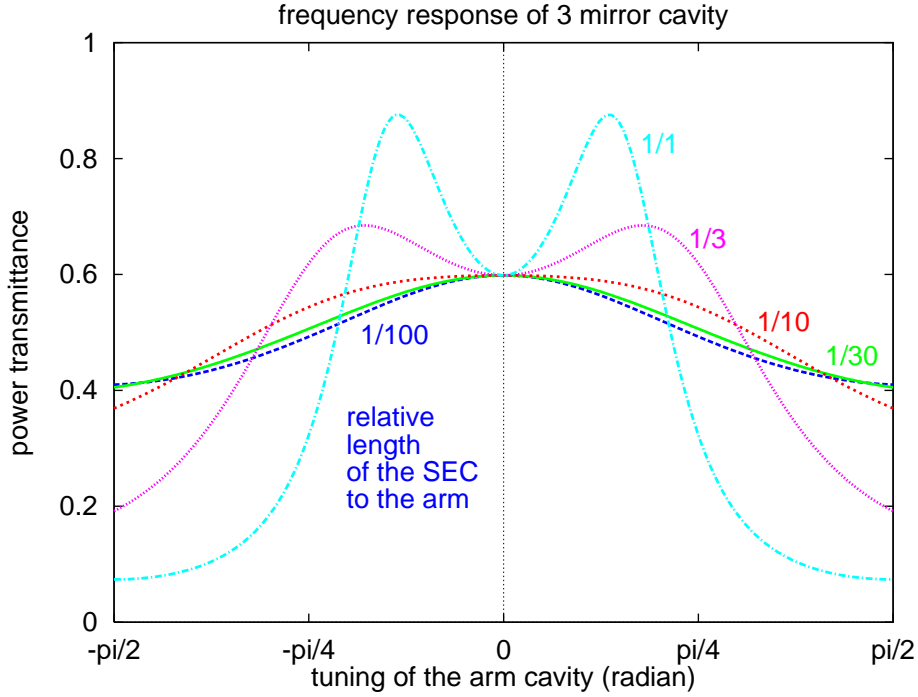


Figure 4.15: The power transmittance of a three-mirror coupled cavity with various signal extraction cavity lengths, assuming there is no detuning. Each curve corresponds to the cross-section of Fig. 4.10 along an oblique line which passes through the origin. The difference among the curves is the slope of the oblique line, or equivalently the ratio of the signal extraction cavity length to the arm cavity length, which is indicated by the key. By elongating the signal extraction cavity, a response with a resonance peak at each side of the carrier frequency can be obtained.

When the signal extraction cavity is elongated, the curve at the top of the resonance becomes flat, and the linewidth measured around its top gets broader. At the same time, however, the gradient at each side of the resonance gets steeper, and thus the height at off-resonance becomes lower. If the signal extraction cavity is further elongated, a resonance peak appears at the edge of each side of the flat region. Even further elongation of the signal extraction cavity makes the each resonance peak higher and narrower, and also makes two peak frequencies closer to the center ($\phi_s=0$).

It is interesting to note that these responses are similar to those of two-pole low pass filters, just as Fabry-Perot cavities behave like one-pole low pass filters. The general transfer function of a two-pole low pass filter is given as

$$G(\omega) = \frac{G_0}{1 - \left(\frac{\omega}{\omega_p}\right)^2 + i\frac{1}{Q} \left(\frac{\omega}{\omega_p}\right)}. \quad (19)$$

This transfer function is characterized by the Q , the quality factor. When the Q is high enough ($Q \gtrsim 1$), the height and the linewidth are approximately given by QG_0 and ω_p/Q , respectively.

The response of resonant sideband extraction in eq. (15) approximates, by keeping up to second order terms in ω , as

$$\begin{aligned}
 G_{\text{SX}}(\omega) &= \frac{-\tau_s \tau_c e^{-i\omega t_x/2}}{1 - \rho_s \rho_c e^{-i\omega t_x} - \rho_c \rho_r e^{-i\omega t_a} + (1 - \mathcal{A}_c) \rho_s \rho_r e^{-i\omega(t_a + t_x)}} \\
 &\simeq \frac{-\tau_s \tau_c e^{-i\omega t_x/2}}{1 - \rho_s \rho_c [1 - i\omega t_x - \frac{1}{2}(\omega t_x)^2] - \rho_c \rho_r [1 - i\omega t_a - \frac{1}{2}(\omega t_a)^2] \\
 &\quad + A \rho_s \rho_r [1 - i\omega(t_a + t_x) - \frac{1}{2}\omega^2(t_a + t_x)^2]} \\
 &\simeq \frac{-\tau_s \tau_c e^{-i\omega t_x/2}}{[1 - \rho_c(\rho_s + \rho_r) + A \rho_s \rho_r] + i\omega[\rho_c(\rho_s t_x + \rho_r t_a) - A \rho_s \rho_r(t_a + t_x)] \\
 &\quad - \frac{1}{2}\omega^2[A \rho_s \rho_r(t_a + t_x)^2 - \rho_c(\rho_s t_x^2 + \rho_r t_a^2)]} \\
 &\simeq \frac{G_0}{1 - \left(\frac{\omega}{\omega_p}\right)^2 + i\frac{1}{Q}\left(\frac{\omega}{\omega_p}\right)} e^{-i\omega t_x/2}
 \end{aligned} \tag{20}$$

where

$$\begin{aligned}
 G_0 &= -\tau_s \tau_c / B & A &= 1 - \mathcal{A}_c \\
 Q &= \frac{B/(\omega_p t_a)}{\rho_s(\rho_c - A\rho_r)\chi + \rho_r(\rho_c - A\rho_s)} & B &= 1 - \rho_s \rho_c - \rho_c \rho_r + A \rho_s \rho_r \\
 & & \chi &= t_x/t_a = \ell_s/\ell_a \\
 \omega_p &= \frac{\sqrt{2B}/t_a}{\sqrt{-\rho_s(\rho_c - A\rho_r)\chi^2 + 2A\rho_s \rho_r \chi - \rho_r(\rho_c - A\rho_s)}}.
 \end{aligned} \tag{21}$$

Thus, resonant sideband extraction behaves like a two-pole low pass filter. Note that the DC gain G_0 does not depend on χ .

For broad-band purposes where $Q \sim 1$, this approximation is valid. When the Q becomes high, however, it is no longer valid since the effects from higher order terms become significant. In general, they have an influence similar to reducing the Q value. Thus, replacing Q by

$$Q' \simeq \frac{B/(\omega_p t_a)}{\rho_s(\rho_c - A\rho_r)\chi + \rho_r(\rho_c - A\rho_s) - (\omega_p t_a)^2 \{\rho_s \rho_c \chi^3 + \rho_c \rho_r - A \rho_s \rho_r (1 + \chi)^3\} / 6} \tag{22}$$

provides a better approximation⁴⁾.

⁴⁾ Eq. (15) can be better approximated by keeping the terms up to third order of ω . Third order functions are, however, hard to solve and less convenient. Here Q' was obtained by using ω_p instead of ω in the third order terms. According to the results of simulations, this approximation provides a frequency response closer to the full expression than that using Q .

It can be seen that elongating the signal extraction cavity changes the resonance peak frequency and the quality factor simultaneously. The above approximation is convenient for estimating the bandwidth and the peak sensitivity of resonant sideband extraction, which are given by ω_p/Q' and G_0Q' , respectively.

The behavior of resonant sideband extraction can also be explained qualitatively by investigating the phase shift inside the three-mirror coupled cavity. This was adopted by the author in the former paper [53], and fully described again in Appendix C.

4.4: Application as gravitational wave detectors

4.4.1: Parameters in resonant sideband extraction

In the previous section, we discussed the characteristics of resonant sideband extraction qualitatively. Now we are ready to choose appropriate parameters to use it in a realistic interferometric gravitational-wave detector. We will assume that the detector will have an arm length of the order of a few kilometers and that its bandwidth must be set to somewhere between a few tens of hertz and a few kilohertz depending on the type of gravitational waves being sought.

The parameters which must be chosen to determine the frequency response of resonant sideband extraction are the arm length, the signal extraction cavity length, and the reflectivities of six mirrors which are the cavity coupling and reflecting mirrors, the power recycling mirror, and the signal extraction mirror. The length of the signal extraction cavity must be considered both in macroscopic and microscopic senses. The latter, which will be referred to as the tuning of the signal extraction cavity, is easy to change after the construction of the detector, and thus has lesser importance in the choice of parameters.

There is also the possibility of using folded Fabry-Perot cavities in the arms in order to achieve a longer storage time, and this may be useful in some cases (see Chapter 5). We will not treat this here, however, to concentrate on the basic procedure of choosing parameters. There is no difficulty in extending the following discussion to more complicated applications.

By assuming the two arms are identical, the number of mirrors whose reflectivities are to be chosen is reduced to four. Among these, the reflectivity of the reflecting mirror must be as high as possible to maximize the amount of energy stored in the interferometer. Furthermore, by the same reasoning the

power reflectivity of the power recycling mirror must uniquely be chosen to be (almost) the same as that of the interferometer (see 3.3.3).

The arm length may be determined by other reasons such as the location or the cost. Then the parameters left undetermined are the reflectivities of the Fabry-Perot coupling mirrors and the signal extraction mirror, and the length and the tuning of the signal extraction cavity. In the end, what is to be chosen is the signal extraction cavity appropriate for an interferometer with a certain arm length.

The criterion for choosing parameters must be to make the probability of detecting gravitational waves as high as possible. This, however, depends on the signal spectrum and related matters, as discussed in 3.1.2, and is beyond the scope of this paper. Thus we take examples to illustrate several possibilities and to demonstrate the procedure for choosing parameters, for use in different circumstances.

The examples taken here are a broad-band response, a narrow-band response, and a low frequency response. Each one is expected to be appropriate for a particular purpose; the first one is suitable for burst type waves in relatively short time scales, the second one is useful for the detection of periodic signals with known frequencies such as those from pulsars, and the last one is adequate for the signals from coalescing binaries.

4.4.2: Broad-band detector

First let us consider a relatively broad-band case. Once the arm length is determined, the free-spectral-range of the arm can be obtained as

$$\Delta f_{\text{FSR}} = c/2\ell_a \approx 50 \text{ kHz} \times \left[\frac{\ell_a}{3 \text{ km}} \right]^{-1}. \quad (23)$$

The bandwidth of the signal cavity must be twice the desired bandwidth, according to the discussion in 3.2.2. If standard Fabry-Perot cavities were used in the arms to achieve the desired bandwidth, this would require a finesse \mathcal{F} of

$$\mathcal{F} \simeq \frac{\Delta f_{\text{FSR}}}{2\Delta f_{\text{BW}}} \approx 250 \times \left[\frac{\ell_a}{3 \text{ km}} \right]^{-1} \left[\frac{\Delta f_{\text{BW}}}{100 \text{ Hz}} \right]^{-1}, \quad (24)$$

and this must be realized by the reflectivity of the coupling mirror ρ_{FP} :

$$\mathcal{F} \simeq \frac{\pi\sqrt{\rho_{\text{FP}}}}{1 - \rho_{\text{FP}}} \quad \text{or} \quad |\rho_{\text{FP}}|^2 \simeq 1 - \left(\frac{2\pi}{\mathcal{F}}\right) + \frac{1}{2}\left(\frac{2\pi}{\mathcal{F}}\right)^2 + O\left(\left(\frac{2\pi}{\mathcal{F}}\right)^3\right). \quad (25)$$

If we choose 1 kHz bandwidth for our example, $\mathcal{F} \approx 25$ and $|\rho_{\text{FP}}|^2 \approx 0.78$ are obtained.

In the case of resonant sideband extraction, the signal extraction cavity must have this order of power reflectivity when it is resonant, which is determined only by the reflectivities of the mirrors. There is still some freedom left in choosing the reflectivity of one of the mirrors, and that of the other can be determined from it.

Fabry-Perot coupling mirrors with low reflectivities may reduce the amount of energy stored, as can be seen in Fig. 3.5. This is because the amount of energy stored will not be limited by the dissipation in the arms but by the losses at the beamsplitter and the coupling mirror substrates. A more serious problem will arise with low finesse arm cavities when thermally induced distortions become significant. This problem is discussed further in Chapter 5.

To avoid these effects, the arm cavity must have a high enough finesse, typically $\gg 100$, while the exact number depends on the properties of the optics used. Here we will take $\mathcal{F} \approx 1,000$ as an example. Then the power reflectivity of the Fabry-Perot coupling mirror will be $|\rho_c|^2 \approx 0.994$ from eq. (25). This, with the loss due to imperfect contrast described below, determines the reflectivity of the interferometer and thus the reflectivity of the power recycling mirror as $|\rho_{\text{PR}}|^2 \approx 0.96$, which results in a power recycling gain of $\mathcal{G}_{\text{PR}} \approx 25$ (see 3.3.3).

The reflectivity of the signal extraction mirror must be chosen so that the power reflectivity of the signal extraction cavity will be the desired value, though the relation may not necessarily be exact.

$$|\rho_{\text{FP}}|^2 \simeq \left| \frac{\rho_c - (1 - \mathcal{A}_c)\rho_s}{1 - \rho_s\rho_c} \right|^2 \quad \text{or} \quad \rho_s \simeq \frac{\rho_c - \rho_{\text{FP}}}{(1 - \mathcal{A}_c) - \rho_c\rho_{\text{FP}}} \quad (26)$$

Here ρ_s is the effective reflectivity of the signal extraction mirror defined in eq. (16). There may be two solutions for this equation, but we must choose the one which satisfies $\rho_c > (1 - \mathcal{A}_c)\rho_s$ for a broad-band response according to the discussion in 4.3.3.

The losses inside the signal extraction cavity must be included in our assumption in order to obtain the reflectivity of the signal extraction mirror. Among them, the loss due to the imperfect contrast $(1 - \mathcal{C})/2$ is considered as being the major contribution. In our example, we will assume $\mathcal{C} \approx 0.99$ or $(1 - \mathcal{C})/2 \approx 5 \cdot 10^{-3}$ as a typical number. Although this looks like an unusually good number, it will be justified by the mode cleaning effect explained in

Chapter 5. With this assumption, the power reflectivity of the signal extraction mirror can be determined as $|\rho_{sx}|^2 \approx 0.91$ in our example. This results in the sensitivity at DC of $\tilde{h} \approx 4.6 \cdot 10^{-24}/\sqrt{\text{Hz}}$ assuming 50 W illuminating light power of 1.06 μm wavelength.

The length and the tuning of the signal extraction cavity are left undetermined. Here we will assume that there is no detuning, since it only narrows the bandwidth and is not very useful for broad-band purposes.

According to the discussion in 4.3.4, the macroscopic length of the signal extraction cavity has the effect of broadening (to some extent) the bandwidth measured at -3 dB points. This effect is, however, relatively small—typically by no more than a factor of two. Thus, it is important to choose a proper bandwidth without taking this effect into account, i.e. by the reflectivities of the mirrors.

If the signal of interest is well within the detector bandwidth, it may be appropriate to choose the flattest response so that the sensitivity inside the bandwidth will be optimized. This, however, sacrifices the sensitivity outside the bandwidth. It may be better to choose a response with less steep cut-off if the information contained in the higher frequency components of the signal is important as well⁵⁾. A response of either type can be obtained in resonant sideband extraction by adjusting the length of the signal extraction cavity.

If the signal extraction cavity is short enough, the response is identical to that for a Fabry-Perot cavity with the same 3 dB bandwidth. To obtain the flattest response, the Q in eq. (21) must be $\sim 1/\sqrt{2} < 1$. By solving eq. (21) for χ , the relative length of the signal extraction cavity to the arm cavity can be obtained.

$$\chi_Q = \frac{\ell_s}{\ell_a} \simeq \frac{\rho_c \rho_r - C \rho_r (\rho_c - A \rho_s) - \sqrt{(\rho_c \rho_r)^2 - C \rho_c \rho_r (\rho_c - A \rho_s) (\rho_s + \rho_r)}}{C \rho_s (\rho_c - A \rho_r)} \quad (27)$$

where

$$\begin{aligned} A &= 1 - \mathcal{A}_c \\ B &= 1 - \rho_s \rho_c - \rho_c \rho_r + A \rho_s \rho_r \\ C &= C(Q) = 1 + 2Q^2 \rho_s (\rho_c - A \rho_r) / B \end{aligned} \quad (28)$$

⁵⁾ One may say that then the detector bandwidth should be chosen to be broader. However, there is a trade-off between the bandwidth and the peak sensitivity, as can be seen eq. (3.3). Thus, there may be cases where the detector bandwidth is set narrower than that of the signal spectrum to achieve higher peak sensitivity with some interest on the signal outside the bandwidth.

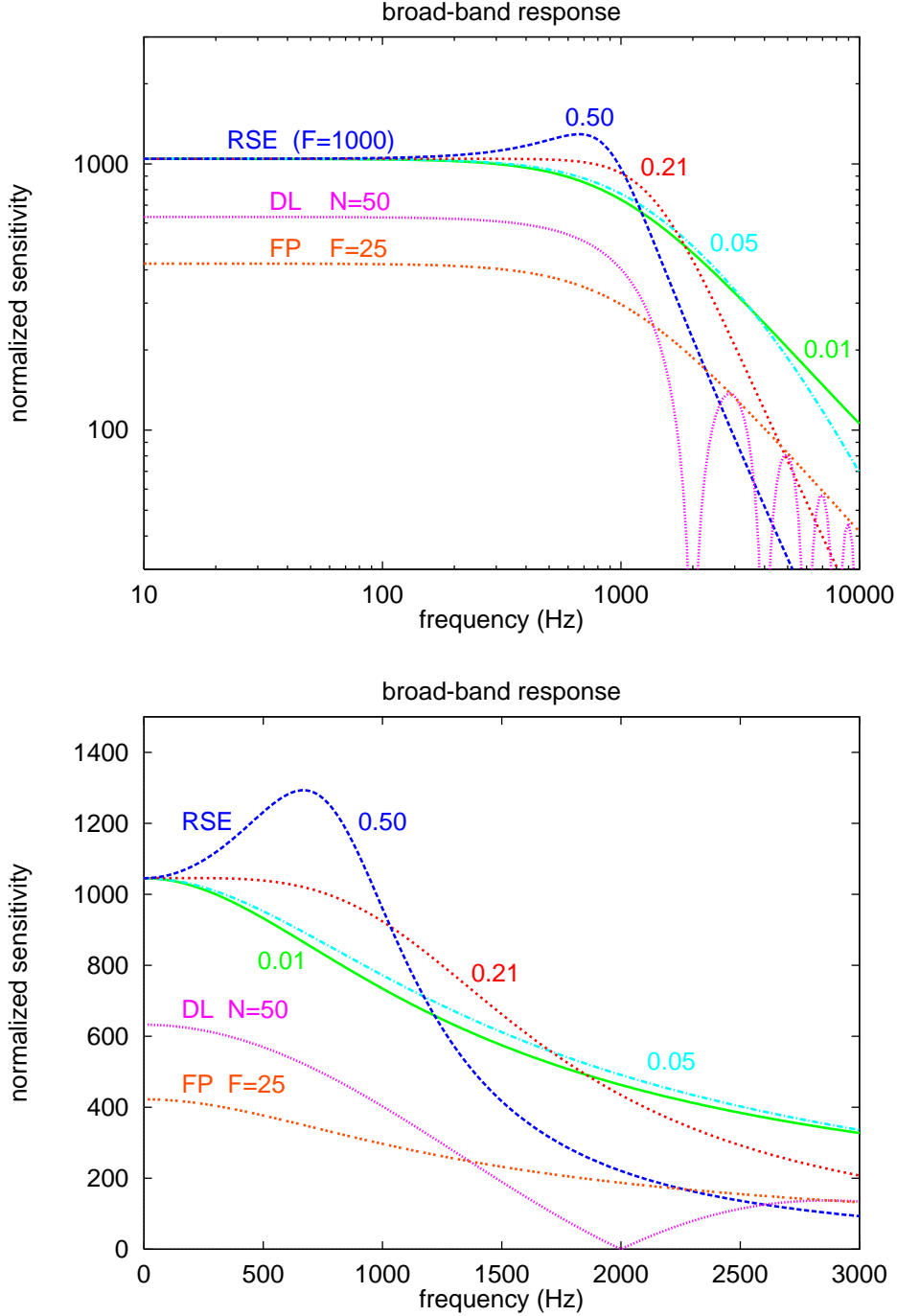


Figure 4.16: The sensitivities of interferometers optimized for 1 kHz bandwidth, in logarithmic and in linear scale. For resonant sideband extraction, the finesse of the arm cavity is $\approx 1,000$, the power reflectivity of the signal extraction mirror is ≈ 0.91 , and several curves with different signal extraction cavity lengths are shown (keys indicate the ratio of the signal extraction cavity length to the arm length). For comparison, the response of an interferometer with delay lines of $N \approx 50$ and that with Fabry-Perot cavities of $\mathcal{F} \approx 25$ are also shown. The definition of the vertical unit and the assumptions for the parameters can be found in Appendix B.

were used for simplicity. Although this relation is a little complicated, it is not difficult to calculate the required length for the signal extraction cavity to achieve the desired Q value.

In our example, the flattest response is obtained by choosing $\chi_Q \approx 0.21$, or $\ell_s \approx 620$ m. Although this looks unrealistically long, it should be possible to use one of the arms also for the signal extraction cavity by making a kink with a mirror, in principle. One may be afraid of the gravitational wave effect in such a long signal extraction cavity. There would be no problem, however, since there would be few photons to be affected by gravitational waves.

Fig. 4.16 shows the response of our example, with various lengths for the signal extraction cavity to illustrate their effects. The delay line type and Fabry-Perot type responses with similar bandwidths are also shown as references. This comparison is, however, assuming constant losses and is not valid when thermal distortion becomes significant.

4.4.3: Narrow-banding

As we have discussed in the previous section, there are two ways to make narrow-band responses in resonant sideband extraction. One is elongating the signal extraction cavity and the other is the detuning of the signal extraction cavity. Furthermore, it is also possible to combine the two methods. Which is the best way to realize the desired response?

For this problem, recalling the ‘tile pattern’ described in 4.2.4 is very helpful. The frequency response of resonant sideband extraction corresponds to the cross section of a tile along an oblique line. Elongating the signal extraction cavity makes the slope of the oblique line steeper and finally it crosses the peaks nearby, which results in a response with a resonance peak at each side of the carrier. On the other hand, detuning shifts the oblique line and it will cross a peak with a certain detuning, resulting in a response with a resonance peak at one side. Note that in either case the oblique line crosses the *same* peak.

As described in 4.2.4, the pattern of a tile is determined only by the reflectivities of the mirrors used. As a consequence, the height of the resonance peak for one of the sidebands does not depend on how the response is realized. In the case of elongation, both upper and lower sidebands will be at the resonance frequencies whereas in detuning only one of them is at the resonance frequency and the other is off-resonance. This makes a difference in sensitivity between the two cases by a factor of up to two.

Since the peak is the same in all the cases, its width will not depend so much on how the response is realized. From the shape in Fig. 4.10, it is expected that a less steep oblique line will yields a broader width of resonance. It is also expected from the figure that the combination of elongation and detuning has a weak effect unless the elongation has a considerable effect by itself.

In practice, however, narrowing the bandwidth by elongation often requires an unrealistic value for the length of the signal extraction cavity, especially when the bandwidth is narrow compared with the peak frequency. Thus, we will consider the detuning as being of primary interest in the following discussion.

Assuming a certain length for the signal extraction cavity, there must be an optimum amount of tuning for the target frequency. For simplicity, we will assume the signal extraction cavity to be short compared with the arm cavity ($\ell_s \ll \ell_a$). This means that the oblique line, along which the cross section of a tile is taken, is almost horizontal.

In a tile pattern, the one of the sidebands (we consider only one at this point) at the target frequency is represented by a vertical line whose horizontal position is determined by the arm length. We need to find the maximum point along this line, and the vertical position of the point represents the amount of tuning required.

It should be noted that this does not necessarily mean that this is the maximum point of the frequency response. In other words, the sensitivity at a certain frequency may not be the best when it is the top of the response. This arises from the fact that the peak frequency and the sensitivity at that frequency is related in a complicated way, as can be seen in eqs. (20) and (21), and cannot be chosen independently.

The optimum tuning ψ_0 for a target frequency ω_z can be obtained by investigating eq. (15), as

$$\psi_0 = \arctan \frac{[(1-\mathcal{A}_c) - (\rho_c)^2] \sin \omega_z t_a}{(\rho_c/\rho_r) [1 + (1-\mathcal{A}_c)(\rho_r)^2] - [(1-\mathcal{A}_c) + (\rho_c)^2] \cos \omega_z t_a}. \quad (29)$$

It is worth noting that this does not depend on the reflectivity of the signal extraction mirror. If the signal extraction cavity is infinitesimally short, the detuning of the signal extraction cavity at the carrier frequency must be this ψ_0 . If the signal extraction cavity has a finite length, the detuning at the carrier frequency is obtained by subtracting the phase delay in the signal extraction cavity, as

$$\psi_\chi = \psi_0 - \chi \omega_z t_a. \quad (30)$$

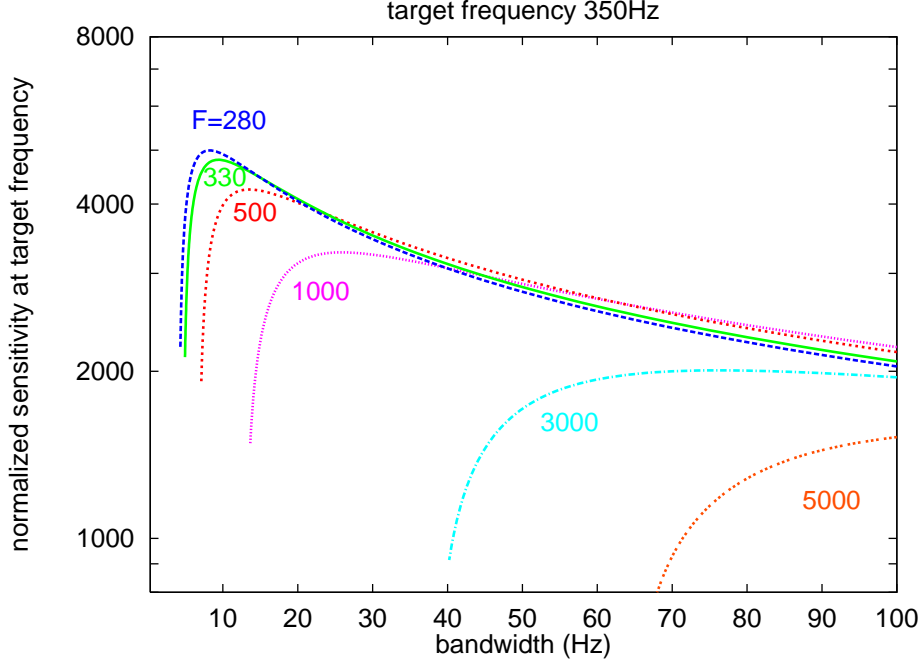


Figure 4.17: The relation between the normalized sensitivity and the bandwidth in resonant sideband extraction, detuned to optimize for a 350 Hz signal. Each curve represents a different finesse of the arm cavity. In general, narrowing the bandwidth improves the sensitivity until the minimum bandwidth determined by the finesse of the arm cavity. To achieve a very narrow bandwidth and a correspondingly improved sensitivity, it is necessary to reduce the finesse of the arm cavity.

We will, however, continue assuming an infinitesimally short one.

When the amount of detuning is set as described above, the peak frequency of the response and the bandwidth measured at the -3 dB point at both sides of the peak frequency, i.e. not centered at the target frequency, are approximately given by

$$\begin{aligned}\omega_p &\simeq \omega_z + \Im\{E(\omega_z, \psi_0)\}/t_a \\ \Delta f_{3\text{dB}} &\simeq |\Re\{E(\omega_z, \psi_0)\}|/(\pi t_a)\end{aligned}\quad (31)$$

where

$$E(\omega, \psi) = 1 - \frac{1 - \rho_s \rho_c e^{-i\psi}}{\rho_t e^{-i\omega t_a} [\rho_c - (1 - \mathcal{A}_c) \rho_s e^{-i\psi}]} . \quad (32)$$

These approximations are valid when $Q = \omega_p/(2\pi\Delta f_{3\text{dB}}) \gtrsim 1$, which will be satisfied in most narrow-band applications. The derivation of eqs. (29)–(33) can be found in Appendix D.

In general, a narrow-band response optimized for a target frequency tends to have a peak at a lower frequency than the target frequency. Thus, the target

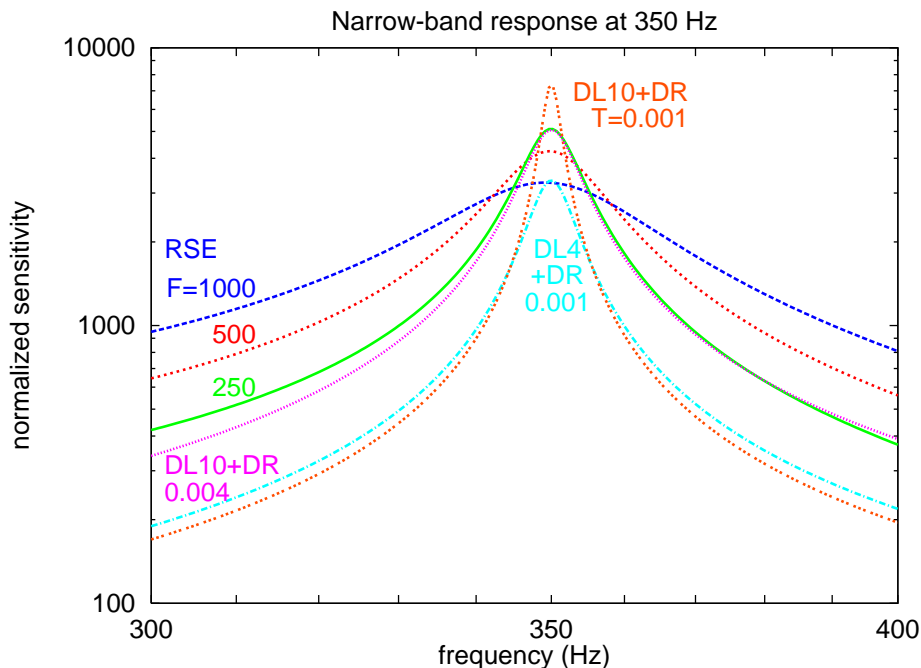


Figure 4.18: The sensitivities of the interferometers optimized at 350 Hz. For resonant sideband extraction, the three cases with different finesse (250, 500, 1,000) of the arm cavity are shown. The responses of dual recycled interferometers with 4 and 10 transits in the arms are also shown for comparison (keys indicate the power transmittance of the signal recycling mirror). In general, dual recycling can achieve better sensitivity as long as the storage time in the arm is reasonably long. Both configurations have similar response curves when both are optimized for the same bandwidth.

frequency used in the above calculation must be chosen as the upper limit of the bandwidth of interest. Note that the value of ρ_s is not the reflectivity of the signal extraction mirror but is related to it by eq. (16), and its value is limited by the losses inside the signal extraction cavity.

The most common requirement for a narrow-band purpose, such as the detection of a gravitational wave from a known pulsar, is to realize a certain bandwidth at a certain frequency. This must be achievable by choosing appropriate parameters for the reflectivities of mirrors and the amount of detuning, i.e. by solving

$$\Delta f_{3\text{dB}} = \Delta f_{3\text{dB}}(\omega_z, \rho_s, \rho_c, \psi(\omega_z, \rho_c)). \quad (33)$$

The above relations are, however, a little too complicated to be solved analytically.

It may be better to use plots of these functions taking the reflectivity of the signal extraction mirror as a parameter. An example, optimized for a 350 Hz

signal assuming an arm length of 3 km, is shown in Fig. 4.17 with several choices for the finesse of the arm cavity. From the dependence of bandwidth and peak sensitivity on the reflectivity of the signal extraction mirror, the relation between the two can be plotted. Note that this result may include signal recycling with Fabry-Perot cavities in the arms since there is no assumption which restricts the argument to resonant sideband extraction.

For example, if we assume the same arm cavity ($\mathcal{F} \approx 1,000$), a bandwidth of ≈ 40 Hz is obtained by choosing the reflectivity of the signal extraction mirror as $|\rho_{\text{sx}}|^2 \approx 0.989$ and the detuning as $\psi_0 \approx 0.14$ radian. With these choices the sensitivity at the target frequency (350 Hz) is $\tilde{h} \approx 1.6 \cdot 10^{-24}/\sqrt{\text{Hz}}$, assuming the same light source as the broad-band case (50 W with $1.06 \mu\text{m}$ wavelength). Narrowing the bandwidth further does not improve the sensitivity so much. Reducing the finesse of the arm cavity down to ~ 280 allows realization of a narrower bandwidth of ~ 10 Hz with a sensitivity improvement of ~ 1.5 .

As a general tendency, it is necessary to have a lower finesse for the arm cavity to achieve a narrow-band response, especially when the target frequency is relatively high. Once the finesse of the arm cavity enters the reasonable range, however, the sensitivity achievable for a given bandwidth (or vice versa) does not depend so much on the finesse of the arm cavity.

This is what is expected from the theorem described in 3.1.3. The stored energy in each case does not differ so much and the only differences are the effects from losses. Since both limits of low finesse or high finesse for the arm cavities increase the effects from losses at or around the beamsplitter, it may be reasonable to choose a moderate finesse (1,000~3,000) for the arm cavity.

4.4.4: Low frequency detector

When one considers a low frequency detector using resonant sideband extraction, there are two possibilities, broad-band type or narrow-band type response. The former has a bandwidth starting from DC up to a certain cut-off frequency, whereas the latter has a peak at a certain frequency with a bandwidth spread around that frequency. This choice arises since at low frequency range the bandwidth of a broad-band type response is relatively narrow and becomes comparable to that of the narrow-band type responses.

In this comparison, one must recall that the signal at very low frequencies cannot be observed anyway due to the noise caused by seismically or thermally excited motions of the mirrors (see 1.3). Then the broad-band type responses

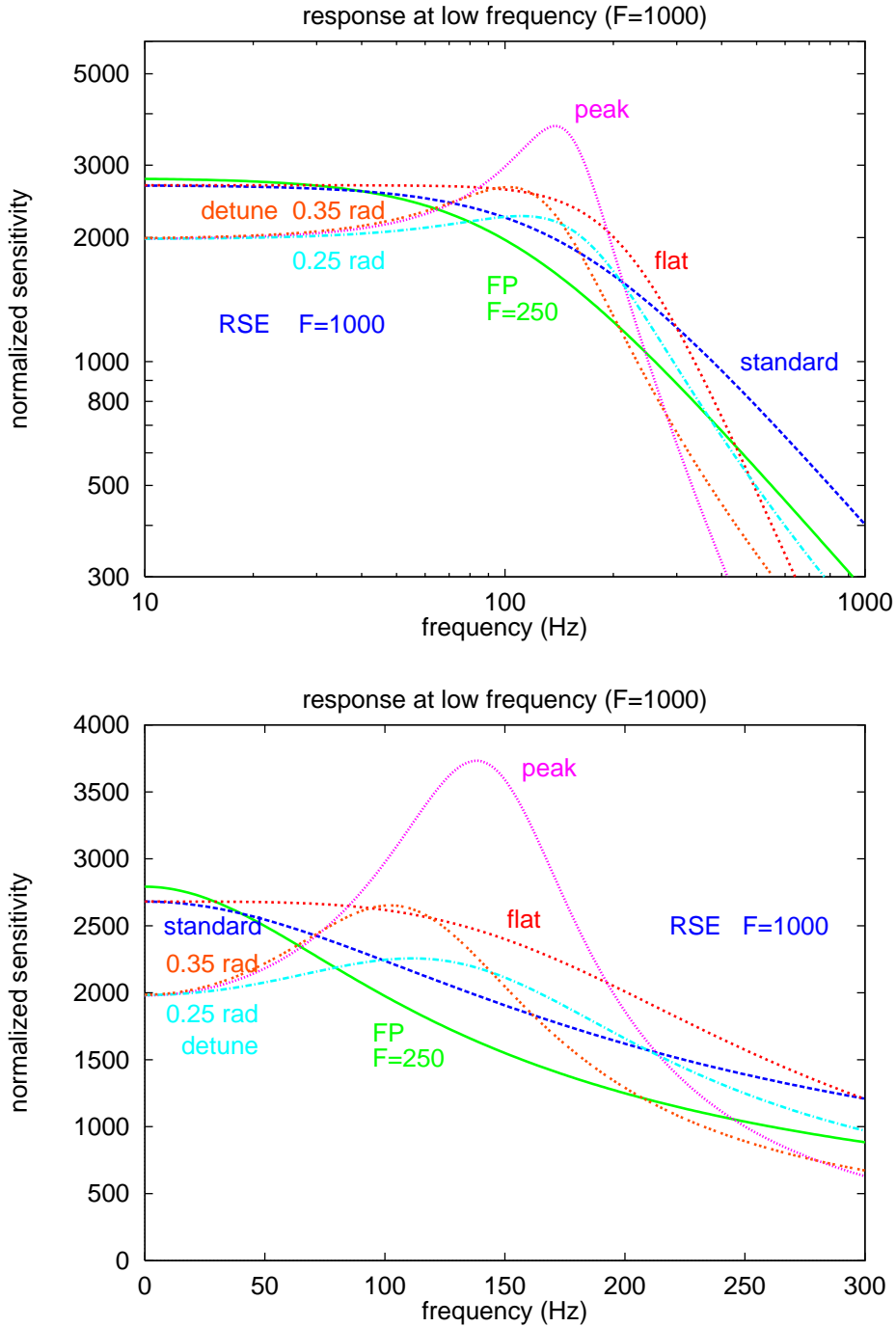


Figure 4.19: The responses of interferometers optimized at low frequencies, around 100 Hz–200 Hz, in logarithmic and in linear scale. For resonant sideband extraction, the same arm cavity as the broad-band case (Fig. 4.16) is assumed. The power reflectivity of the signal extraction mirror is ≈ 0.719 for ‘standard’ and ‘flat’ responses and ≈ 0.837 for detuning and ‘peaking’ by elongation. The amount of detuning is ≈ 0.25 and ≈ 0.35 radian, each optimizes ~ 200 Hz and 140 Hz. The response of an interferometer with Fabry-Perot cavities of $\mathcal{F} \approx 250$ is also shown for comparison.

effectively have the same response as that of narrow-band type ones, i.e. they ‘waste’ the sensitivity at the very low frequency. The lower cut-off frequency is determined by the mechanical design of the test masses and the suspension systems, and here we take the cut-off frequency to be ~ 100 Hz.

If the sensitivity is only determined by the amount of energy stored and the bandwidth set, it may be advantageous to have narrowed response just above the lower cut-off frequency so that nothing is lost. In general, however, the narrowed responses are not so efficient as the broad-band ones since only one of the sidebands is utilized in many cases. This is especially true when detuning is used to narrow the bandwidth.

As a typical example, let us consider responses with an effective bandwidth of ~ 100 Hz. To avoid the lower cut-off frequency, the broad-band type response must have ~ 200 Hz bandwidth. This requires an unrealistic value of the mirror size for delay lines, so we assume Fabry-Perot based configuration and resonant sideband extraction. On the other hand, the narrowed response must have ~ 100 Hz bandwidth centered at ~ 150 Hz, i.e. $Q \sim 1.5$.

The procedure to choose the parameters is the same as in the previous cases. The results are shown in Fig. 4.19. Both broad-band type and narrow-band type responses of resonant sideband extraction, with a Fabry-Perot response for comparison, are shown. Again, the same arm cavity as the broad-band case is assumed in resonant sideband extraction. The reflectivity of the signal extraction mirror is ≈ 0.719 for the ‘standard’ (neither detuning nor elongation) and ‘flat’ responses, and ≈ 0.837 for the detuning and ‘peaking’ by elongation. The amount of detuning is ≈ 0.25 radian when optimized at ~ 200 Hz and ≈ 0.35 radian for ~ 140 Hz.

For a broad-band type response, resonant sideband extraction shows somewhat better sensitivity than the simple Fabry-Perot, though the difference is smaller than that in the broad-band case. This is due to better efficiency of resonant sideband extraction in storing the light energy. ‘Flattening’ the response by elongating the signal extraction cavity may considerably improve the sensitivity in the frequency range of interest (100–200 Hz), though this requires an unrealistically long signal extraction cavity length ($\chi \gg 1$).

From the same reasoning, the narrow-band response by elongation is not so promising. Narrowing the bandwidth by detuning improves the sensitivity a little within the frequency range of interest, when appropriate parameters

are chosen. This is, however, less efficient than ‘peaking’ or ‘flattening’ by elongation since it utilizes only one of the sidebands.

It may be more useful if a flat or peaked response is obtained with a more reasonable signal extraction cavity length. In general, increasing the reflectivity of the Fabry-Perot coupling mirror has such an effect. In order to retain the detector bandwidth, however, the reflectivity of the signal extraction mirror must also be increased. This means that light stays in the signal extraction cavity longer, which increases the effect of losses inside it (such as that at or around the beamsplitter) and reduces the amount of sidebands that can be detected.

When arm cavities with a finesse of $\sim 15,000$ are used, a flat response can be obtained by elongating the signal extraction cavity length to a half of the arm cavity. Similarly a peaked response is obtained with the signal extraction cavity of the same length as the arm cavity. In these cases, however, the peak sensitivity will be worse than those shown in Fig. 4.19 by a factor of ~ 0.92 .

Chapter 5

Thermal distortion

5.0: Introduction

Up to here, we have been considering a linear model, i.e. the parameters of optical components are assumed to be constant and do not depend on other parameters. In practice, however, there are non-linearities due to the dependence among the parameters. Among them, the thermal effects are revealed as a severe problem which can limit the sensitivity [54–56]. In this chapter, the sensitivities with such effects are estimated.

A complete analysis using the heat conduction equation may, however, take too much effort to compare several configurations under various conditions, and may not be so convenient. Here we adopt the simplified approximation introduced by Winkler et al. [54]. Thus, all the discussion here should be taken as rough estimations. These still give an insight of the problems associated with thermally induced distortions.

Since all the discussions are merely rough estimates, the detailed transfer function of each configuration is not so meaningful. Thus we use the simplified theorem described in 3.1.3 to estimate the sensitivities of interferometers, ignoring the effect of output efficiency. This may add another factor (of up to ~ 1.05) onto the inaccuracy of the estimate. This inaccuracy should, however, not change the main arguments and actually has an advantage to make the discussion lucid.

5.1: Non-linear problems

5.1.1: Non-linearity

In previous chapters, the parameters of optical components are assumed to be constant; this treatment is referred to as a linear model. This simple model may, however, no longer be valid if there is dependence among the parameters. Such effects are classified as *non-linear* problems. Due to their complexities,

non-linear problems have previously been less considered. The use of the sensitivity theorem described in 3.1.3 provides a clearer view of these problems.

According to the theorem, the sensitivity is determined by the wavelength λ of the light, the noise bandwidth Δf_{NB} of the optical configuration, and the amount of light energy \mathcal{E} stored in the optical system, as

$$\tilde{h}_0 \gtrsim \sqrt{\frac{2\hbar\lambda}{\pi c} \frac{\Delta f_{\text{NB}}}{\mathcal{E}}} \quad (1)$$

(reprint of eq. (3.3)). Among the parameters, the noise bandwidth must be chosen to be appropriate for the desired observation purpose, and the wavelength is determined by the type of the laser used as the light source. Then the sensitivity is primarily determined by the amount of the light energy stored in the optical system.

The amount of light energy that can be stored in an interferometer is limited to a maximum value \mathcal{E}_{max} by the dissipation in the arms,

$$\mathcal{E}_{\text{max}} \simeq \frac{P_0 \ell}{c\mathcal{A}} \quad (2)$$

(reprint of eq. (3.5)). This is determined by the injected light power P_0 , the arm length ℓ of the interferometer, and the minimum loss \mathcal{A} per reflection from a mirror.

In the linear model, the loss is assumed to be a constant that does not depend on other parameters. Then the amount of energy stored is proportional to the injected light power and the arm length of the interferometer, and thus the spectral sensitivity \tilde{h} is inversely proportional to their square roots.

If, however, the loss per reflection depends on other parameters, these proportional relations break. This can easily be shown by taking the total differentiation of eq. (2),

$$\begin{aligned} d\mathcal{E}_{\text{max}} &\simeq \frac{\ell}{c\mathcal{A}} dP_0 + \frac{P_0}{c\mathcal{A}} d\ell - \frac{P_0 \ell}{c\mathcal{A}^2} d\mathcal{A} \\ &\simeq \frac{\ell}{c\mathcal{A}} \left(1 - \frac{P_0}{\mathcal{A}} \frac{\partial \mathcal{A}}{\partial P_0}\right) dP_0 + \frac{P_0}{c\mathcal{A}} \left(1 - \frac{\ell}{\mathcal{A}} \frac{\partial \mathcal{A}}{\partial \ell}\right) d\ell. \end{aligned} \quad (3)$$

In the linear model where the loss depends neither on the arm length nor the illuminating light power, the partial derivatives $\partial \mathcal{A} / \partial P_0$ and $\partial \mathcal{A} / \partial \ell$ vanish and eq. (3) shows the proportionality.

There will be cases, however, where these partial differentiations are not zero. The former, $\partial\mathcal{A}/\partial P_0$, will be positive when thermally induced distortion of the wavefront becomes significant. On the other hand, the latter will be positive if the loss due to scattering becomes dominant. Let us look at them briefly.

5.1.2: Thermal distortion

An optical component absorbs a small fraction of the light power, either at the surface (coating) or inside the substrate if it is transmitting the light. If the injected light power is significantly increased, this absorbed power heats up the optical component. This heating occurs locally within the beam diameter where most of the light power is concentrated. Since the optical substrate has a finite thermal conductivity, there will be a temperature gradient inside the substrate due to this local heating.

This temperature gradient causes deformation of the optical path, either mechanically by the thermal expansion or optically by the so-called *thermal lensing*. The latter means the gradient of the refraction index due to its dependence on the temperature, and it has an effect similar to a lens as its name implies. It is also expected that there will be another optical effect, the birefringence, due to the stress inside the thermally expanding substrate.

These thermal distortions affect the sensitivity of an interferometer, either by decreasing the amount of signal or by increasing the amount of noise. The former can also be written as the decrease of the stored energy. The latter means that there will be unwanted light components induced by the distortion which can hit the photo-detector to increase the shot noise. This can, however, be eliminated by placing an appropriate mode cleaner (output mode cleaner) before the photo-detector. Thus, the main argument in this chapter is placed on the amount of stored energy when thermally induced distortions are present.

The thermal distortions affect the amount of stored energy through a few mechanisms. One simple example is that the thermal distortions which occur in two arms are not necessarily matched to each other, thus the contrast of the interference at the beamsplitter will be deteriorated. Then the loss inside the power recycling cavity is increased, reducing the amount of energy stored. (This is also the mechanism by which the unwanted light component can hit the photo-detector.)

Another important effect is the stability of cavities. Most of the configurations utilize cavities to store light energy, either in the arms or for power

recycling. When thermal distortions inside a cavity become too strong, the cavity will no longer be ‘stable’, i.e. the circulating light tends to diverge rather than to be confined inside the cavity with a finite aperture. This also reduces the amount of energy stored inside the cavity.

These problems are discussed more in detail in the later part of this chapter. Before doing that, we need to introduce the analytical ‘tools’ to handle these problems. This can be found in the following section, after some more comments on the non-linearities.

5.1.3: The beam size effect

Although it is less known, another non-linearity, $\partial\mathcal{A}/\partial\ell$ in eq. (3), also reduces the amount of stored energy and thus degrades the sensitivity from that expected in the linear model. This, in other words, means that *a longer arm length does not necessarily improve the shot-noise-limited sensitivity of an interferometer.*

When the arm length is increased, the spot size on the mirrors becomes larger, proportionally to the square root of the arm length:

$$w(\ell) \propto \sqrt{\lambda\ell/\pi} \quad (4)$$

(see eq. (11)). Note that this is attributed to a property of the Gaussian beam and cannot be altered.

It was found [57] that the spatial fluctuation (either of mirror surface or of refraction index) which causes the loss due to scattering tends to increase with the beam radius. According to Ref. [57], the power loss due to scattering can be approximated by

$$\mathcal{A}_{\text{scat}} \approx 1 \cdot 10^{-5} \times \left[\frac{w}{5 \text{ mm}} \right]^2 \quad (5)$$

for beam radii bigger than 5 mm. This means, with eq. (4), the longer arm effectively increases the loss per reflection of the fundamental mode.

We must take the mode cleaning effect into account (see 5.2.5) to evaluate the actual loss inside the interferometer. This, however, may only change the factor in eq. (5) but not its exponent. Then the maximum energy stored in an interferometer can be written as

$$\mathcal{E}_{\text{max}} \propto \frac{P_0\ell}{c\mathcal{A}_{\text{scat}}} \propto \frac{P_0\ell}{c} \frac{1}{w^2} \propto \frac{P_0\ell}{c} \frac{\pi}{\lambda\ell} \propto \frac{P_0\pi}{c\lambda}, \quad (6)$$

if the scattering due to spatial fluctuation is the dominating loss in the interferometer. This means that the longer arm does not necessarily improve the shot noise limited sensitivity. In practice, of course, other losses as well as other noise sources must be considered carefully.

A similar problem arises in the choice of the wavelength of the light source. According to eq. (1), \mathcal{E}/λ must be maximized to optimize the sensitivity. The amount of stored energy \mathcal{E} depends on the loss per reflection which in turn depends on the wavelength λ , not only from the quality of the coating at that wavelength but also from the beam radii at that wavelength (see eq. (4)).

All the discussion above will strictly depend on the quality of mirrors (polishing and coatings) and thus we cannot come to a final conclusion of these problems at the moment. What should be emphasized here is the possible existence of these problems. Again, the simple relations in eqs. (1) and (2) are helpful for the understanding of these problems.

5.2: Mode description of scattering

5.2.1: Loss mechanisms

Any light which does not contribute to the signal detection process must be considered as *lost*. One of the mechanisms for losses is the *absorption*, either at the surface (coating) or inside the substrate, or both, of an optical component. In this case the light energy is converted to heat and thus is of no use for signal detection. On the other hand, it is also possible that the energy still remains as light but in a state which does not produce any detectable signal.

This happens, for instance, in a Michelson interferometer when the beams from the two arms are mismatched. Then they will not interfere with the beam from the other arm to produce the detectable signal. One of the possible cause of this is misalignment of the beams from the two arms. This is, however, a sort of problem which one can (and one must) control externally. Even though there will be technical difficulties, this will be able to be overcome with an appropriate alignment system [58].

What is more serious is the mismatch due to the imperfection of the existing optical components. For example, the irregularities of the optical path caused by optical components scattering a part of the injected light into higher order modes. This can also take places either at the surface (coating) or inside the substrate, or both, of an optical component. In either case, it does not

necessarily occur in the same way between the two arms and thus does not interfere at the beamsplitter to produce the detectable signal.

In addition, it is also possible that a part of the injected light is converted into the unwanted polarization component due to the birefringence of optical components. Similar to the scattering, this effect can be caused either at the surface or inside the substrate of an optical components and is not necessarily be matched in the two arms.

In these problems, the accuracy of manufacturing mirrors will limit the matching between the two arms. Furthermore, these effects will be induced by thermal properties of mirrors when injected (and thus absorbed) light power is significantly increased. In other applications such as telescopes, an active control of the mirror surface is considered as a solution for similar problems. In our case, however, it is difficult because of the mechanical requirements on the mirrors.

5.2.2: Reflection at a non-ideal surface

Any deviation of the optical path from the ideal shape scatters a part of the incident light from the original transverse mode (usually the fundamental mode) to higher order modes. Here we describe the scattering as the inter-conversion of transverse modes of light field.

When most of the light intensity is concentrated near its propagation axis, the light field can be treated by the paraxial approximation of the wave equation. This is briefly described in Appendix G and it is shown there that a light field with any spatial distribution can be considered as a superposition of a series of eigen functions.

There is freedom in the choice of the series to which the field is expanded, e.g. Hermite-Gaussian or Laguerre-Gaussian, and also in the choice of the position and the radius of the beam waist. They can, of course, be converted from and to each other. For instance, a fundamental Gaussian beam with a certain beam waist at a certain point can be expanded as a series of Gaussian modes (including higher order modes) with different beam waist at a different point (see G.5 in Appendix G).

In considering the light scattering from a mirror, we assume the following few points for convenience. The incoming beam is a fundamental Gaussian beam which has certain beam parameters (radius and curvature). The mirror surface is considered as a composition of a (perfect) spherical surface with a

certain curvature and the deviation from this spherical surface (see Fig. 5.1). The latter, which we will refer to as the *spatial fluctuation* of the mirror surface, is assumed to be small compared with the wavelength.

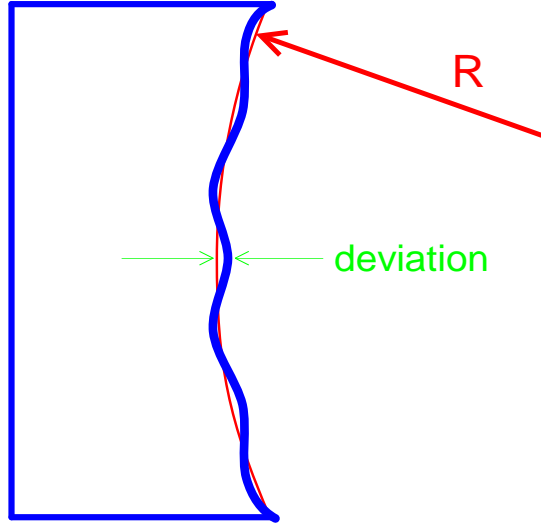


Figure 5.1: A non-ideal surface of a mirror is considered as a superposition of a spherical surface with a certain curvature and the deviation from it. The curvature is not necessarily matched to that of the incident beam. This converts a Gaussian beam to a Gaussian beam with different parameters unless the curvature is matched. On the other hand, the deviation from the ideal sphere converts the fundamental mode to higher order modes.

Even though both originate from the deviation of the mirror surface from the ideal shape, this separation is introduced since the largest component of the deviation is often more or less spherical. By separating this, the rest of the deviation can be treated as small fluctuations. Furthermore, the curvature, which is mostly determined by the fluctuation with spatial wavelength Λ at $\Lambda \simeq \pi w$ [59], converts the beam parameters in a well-defined way (see, for example, Ref. [60]). This will be useful to for analyzing large deviations in detail.

The light scattering of a traversed optical substrate due to the spatial fluctuation of the refractive index can be treated in a similar way. In this case the spatial fluctuation is defined as the deviation from the ideal optical path length (a constant for a non-lensing optical component).

Assuming no transmittance and no absorption, the light field after the reflection from a mirror is represented as

$$\begin{aligned}\Psi(x, y, z) &= \frac{\sqrt{2}}{w} u_{00} \left(\frac{\sqrt{2}x}{w}, \frac{\sqrt{2}y}{w} \right) \exp \left\{ i\psi(z) - ik \frac{x^2+y^2}{2\mathcal{R}_b(z)} \right\} \\ &\quad \times \exp \left\{ i \left[k \frac{x^2+y^2}{2} \left(\frac{1}{\mathcal{R}_b} - \frac{1}{\mathcal{R}_m} \right) + \Phi(x, y) \right] \right\} \\ &= \frac{\sqrt{2}}{w} u_{00} \left(\frac{\sqrt{2}x}{w}, \frac{\sqrt{2}y}{w} \right) \exp \left\{ i\psi(z) - ik \frac{x^2+y^2}{2\mathcal{R}_m(z)} \right\} \exp \{ i\Phi(x, y) \}\end{aligned}\quad (7)$$

$$\text{where } u_{00} \left(\frac{\sqrt{2}x}{w}, \frac{\sqrt{2}y}{w} \right) = \frac{1}{\sqrt{\pi}} \exp \left\{ -\frac{x^2+y^2}{w^2} \right\},$$

using the quadratic approximation of a sphere. Here u_{00} represents the two-dimensional amplitude distribution of the fundamental Gaussian mode, w is the beam radius on the surface, and \mathcal{R}_b and \mathcal{R}_m are the curvature radii of the beam (at the mirror surface) and the mirror, respectively. (As for Φ , see below.)

In the top equation, the upper line of the right hand side represents the incident fundamental mode and the lower line represents the pure phase effects from a mirror with non-ideal surface. These phase effects are separated into the two parts, one is due to the curvature error and the other is the spatial fluctuation of the phase $\Phi(x, y)$ imposed by the fluctuation of the mirror surface (or refraction index). As can be seen in the second equation, the error in the curvature of the mirror converts the fundamental Gaussian beam into a fundamental Gaussian beam with a different curvature.

5.2.3: Mode stability of a cavity

From eq. (7), it is obvious that all the power is preserved in the original mode if (and only if) the curvature of the mirror is matched to that of the beam (at the mirror surface). In fact, it can be shown in a similar way that the power of *any* mode is preserved in that mode if the above condition is satisfied. This leads one to the concept of the mode stability of a cavity. If the light in a cavity always stays in the same mode, the cavity is stable. This requires the curvature of each mirror must be matched to that of the beam.

Using the characteristics of a Gaussian beam described in Appendix G, one can write down the conditions which the mirrors must satisfy. Assuming the beam waist is located at the origin $z=0$, the conditions are

$$\mathcal{R}_j = \mathcal{R}(z) = z_j [1 + (z_R/z_j)^2] \quad \text{or} \quad (z_j)^2 \left(\frac{\mathcal{R}_j}{z_j} - 1 \right) = z_R^2 = \text{const.}, \quad (8)$$

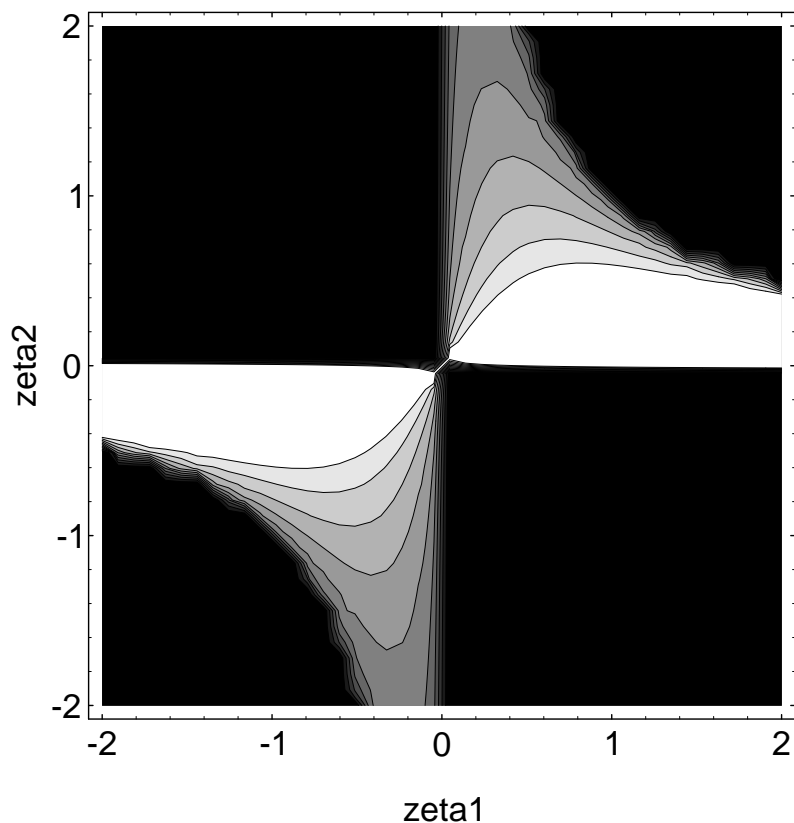


Figure 5.2: The beam radius at one of the mirrors in a Fabry-Perot cavity. The horizontal and vertical axes represent ζ_{\leftarrow} and ζ_{\rightarrow} , respectively (see eq.(10) for their definition). The brightness shows the reciprocal of the beam radius, thus a brighter color represents a smaller beam and a darker color represents a bigger beam. The black part represents that there is no real solution for the beam radius, which means the cavity does not have any stable mode. The white part represents the part in which the beam radius is smaller than the symmetric confocal case ($\zeta_{\leftarrow}=\zeta_{\rightarrow}=0$). The beam size at the other mirror can be obtained by exchanging the two axes.

where \mathcal{R}_j and z_j are the curvature and the distance from the beam waist point of j -th mirror, respectively.

It should be noted that the curvature of a mirror in this notation is defined by the beam propagating in the positive direction of the z -axis. This means a concave mirror facing to the positive direction of the z -axis has a *negative* curvature radius (since it has a convex reflecting surface looking from the other side). Although this is somewhat different from the standard notation, it is necessary in such an application like a three-mirror coupled cavity where the light may hit a mirror at both sides.

For a Fabry-Perot cavity which is composed of two mirrors, the conditions described in eq.(8) can be rewritten, following the ‘standard’ notation, as a

well-known expression [60]

$$0 \leq \zeta_{\triangleleft} \zeta_{\triangleleft} \leq 1 \quad (9)$$

where ζ_{\triangleleft} and ζ_{\triangleleft} are geometric parameters defined as

$$\zeta_{\triangleleft} = 1 - \frac{z_{\triangleleft} - z_{\triangleleft}}{\mathcal{R}_{\triangleleft}} = 1 - \frac{\ell}{\mathcal{R}_{\triangleleft}} \quad \text{and} \quad \zeta_{\triangleleft} = 1 - \frac{z_{\triangleleft} - z_{\triangleleft}}{\mathcal{R}_{\triangleleft}} = 1 - \frac{\ell}{\mathcal{R}_{\triangleleft}}, \quad (10)$$

where $\ell = z_{\triangleleft} - z_{\triangleleft}$ is the separation between the two mirrors. Here subscripts \triangleleft and \triangleleft represent the mirrors at the left and right hand sides, respectively. Since this follows the standard notation, the definition of the curvature radius of a mirror is different from that in eq. (8). Here the curvature radii of mirrors are measured by the beam incident to the reflecting surfaces, which are opposite for the two mirrors facing to each other.

In practice, there are more often the cases in which the beam radius at the waist and its position must be determined from the curvature radii of the mirrors and their separation. Using the above results, these can be solved as follows.

$$\begin{aligned} z_{\text{R}}^2 &= \ell^2 \frac{\zeta_{\triangleleft} \zeta_{\triangleleft} (1 - \zeta_{\triangleleft} \zeta_{\triangleleft})}{(\zeta_{\triangleleft} + \zeta_{\triangleleft} - 2\zeta_{\triangleleft} \zeta_{\triangleleft})^2} = \frac{\ell (\mathcal{R}_{\triangleleft} + \mathcal{R}_{\triangleleft} - \ell) (\mathcal{R}_{\triangleleft} - \ell) (\mathcal{R}_{\triangleleft} - \ell)}{(\mathcal{R}_{\triangleleft} + \mathcal{R}_{\triangleleft} - 2\ell)^2} \\ z_{\triangleleft} &= \ell \frac{-\zeta_{\triangleleft} (1 - \zeta_{\triangleleft})}{\zeta_{\triangleleft} + \zeta_{\triangleleft} - 2\zeta_{\triangleleft} \zeta_{\triangleleft}} = \frac{-\ell (\mathcal{R}_{\triangleleft} - \ell)}{\mathcal{R}_{\triangleleft} + \mathcal{R}_{\triangleleft} - 2\ell} \\ z_{\triangleleft} &= \ell \frac{\zeta_{\triangleleft} (1 - \zeta_{\triangleleft})}{\zeta_{\triangleleft} + \zeta_{\triangleleft} - 2\zeta_{\triangleleft} \zeta_{\triangleleft}} = \frac{\ell (\mathcal{R}_{\triangleleft} - \ell)}{\mathcal{R}_{\triangleleft} + \mathcal{R}_{\triangleleft} - 2\ell} \\ (w_{\text{w}})^2 &= \frac{\ell \lambda}{\pi} \frac{\sqrt{\zeta_{\triangleleft} \zeta_{\triangleleft} (1 - \zeta_{\triangleleft} \zeta_{\triangleleft})}}{|\zeta_{\triangleleft} + \zeta_{\triangleleft} - 2\zeta_{\triangleleft} \zeta_{\triangleleft}|} = \frac{\lambda}{\pi} \frac{\sqrt{\ell (\mathcal{R}_{\triangleleft} - \ell) (\mathcal{R}_{\triangleleft} - \ell) (\mathcal{R}_{\triangleleft} + \mathcal{R}_{\triangleleft} - \ell)}}{|\mathcal{R}_{\triangleleft} + \mathcal{R}_{\triangleleft} - 2\ell|} \quad (11) \\ (w_{\triangleleft})^2 &= \frac{\ell \lambda}{\pi} \sqrt{\frac{\zeta_{\triangleleft}}{\zeta_{\triangleleft} (1 - \zeta_{\triangleleft} \zeta_{\triangleleft})}} = \frac{\lambda |\mathcal{R}_{\triangleleft}|}{\pi} \sqrt{\frac{\ell (\mathcal{R}_{\triangleleft} - \ell)}{(\mathcal{R}_{\triangleleft} - \ell) (\mathcal{R}_{\triangleleft} + \mathcal{R}_{\triangleleft} - \ell)}} \\ (w_{\triangleleft})^2 &= \frac{\ell \lambda}{\pi} \sqrt{\frac{\zeta_{\triangleleft}}{\zeta_{\triangleleft} (1 - \zeta_{\triangleleft} \zeta_{\triangleleft})}} = \frac{\lambda |\mathcal{R}_{\triangleleft}|}{\pi} \sqrt{\frac{\ell (\mathcal{R}_{\triangleleft} - \ell)}{(\mathcal{R}_{\triangleleft} - \ell) (\mathcal{R}_{\triangleleft} + \mathcal{R}_{\triangleleft} - \ell)}} \end{aligned}$$

where w_{w} , w_{\triangleleft} , and w_{\triangleleft} are the beam radii at the waist and at the left and right mirrors, respectively.

From the above equations, it is expected that the beam radius at one of the mirrors becomes infinite when the cavity parameters approach the limitations of eq. (9), i.e. ζ_{\triangleleft} or $\zeta_{\triangleleft} \rightarrow 0$, or $\zeta_{\triangleleft} \zeta_{\triangleleft} \rightarrow 1$. This should be avoided in realistic cases where the mirrors have finite diameters.

Fig. 5.2 shows the beam radius at one of the mirrors, taking ζ_{\triangleleft} and ζ_{\triangleright} as parameters. (The beam radius at the other mirror can be obtained by exchanging the two axes, from their symmetry in eq. (11).) This is in fact identical with the famous stability diagram of a Fabry-Perot cavity, with additional information of the beam radius at one of the mirrors.

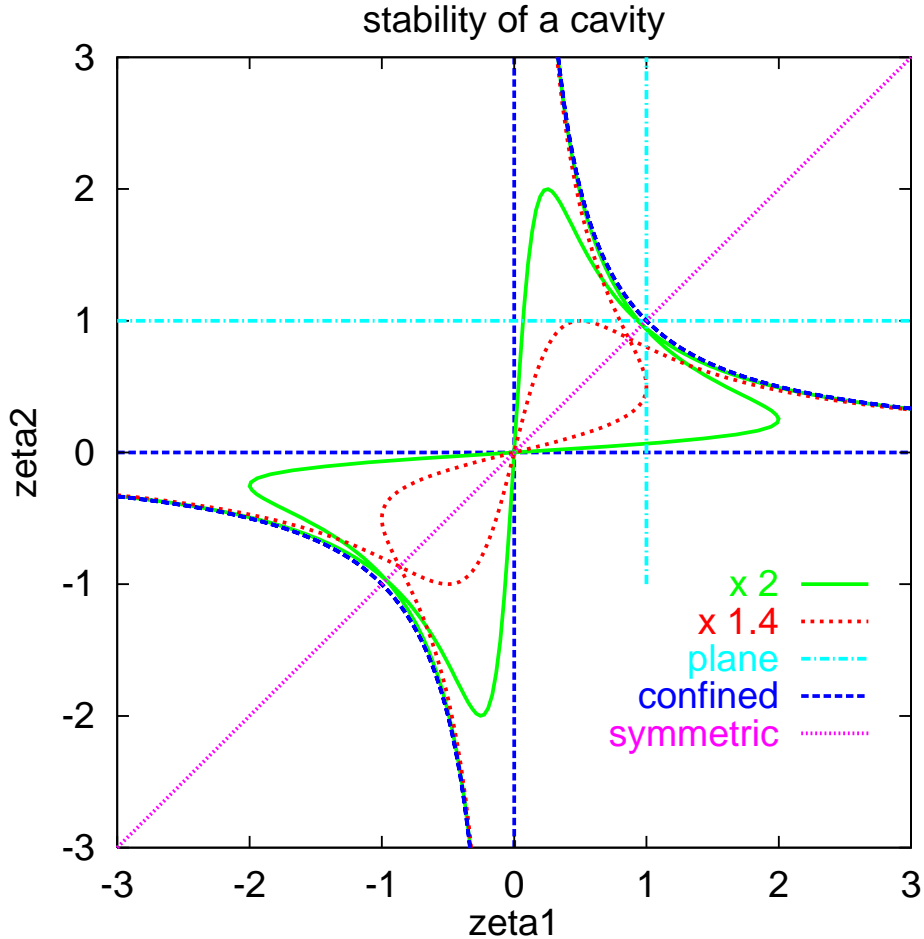


Figure 5.3: A simplified version of Fig. 5.2, but with the constraint from the beam radii at both of the mirrors. The innermost part represents the region in which the beam radii at both mirrors do not exceed $\sqrt{2}$ times those at the symmetric confocal case. The second innermost part is the same but for twice the symmetric confocal case.

Fig. 5.3 shows a somewhat simplified version of Fig. 5.2, but with the constraint from the beam radii at both of the mirrors. If the beam radii should not be bigger than $\sqrt{2}$ times of those in the confocal case $\zeta_{\triangleleft}=\zeta_{\triangleright}=0$ (which makes the product of the two spot sizes smallest), the two ζ -parameters must be within

the innermost part of the diagram. If the tolerance is twice of those in confocal case, the second innermost part is allowed.

5.2.4: Scattering as mode conversion

Now let us consider the second effect in eq. (7), the spatial fluctuation of the phase, assuming the curvature of the mirror is matched to that of the incident beam so that the first effect disappears. The reflected light with distorted phase induced by the mirror surface irregularity can be decomposed into a series of higher order modes with the same beam waist size and its position as the incoming beam. As mentioned earlier, we will assume that the amplitude of the spatial fluctuation is small. Then the exponential function in eq. (7) can be approximated by

$$\exp\{i\Phi(x, y)\} \simeq 1 + i\Phi(x, y) - \frac{1}{2} [\Phi(x, y)]^2 + \dots \quad \text{for } |\Phi| \ll 1. \quad (12)$$

When the Hermite-Gaussian mode expansion is used, the relative amplitude of each higher order mode S_{lm} ($l, m \neq 0$) can be obtained by (see eq. (G36) in Appendix G)

$$\begin{aligned} S_{lm} &= \frac{\sqrt{2}}{w} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \frac{u_l\left(\frac{\sqrt{2}x}{w}\right) u_m\left(\frac{\sqrt{2}y}{w}\right)}{\exp\left\{i[1+l+m]\psi(z) - ik\frac{x^2+y^2}{2\mathcal{R}(z)}\right\}} \Psi(x, y, z) \\ &\simeq \frac{2e^{-i[l+m]\psi(z)}}{w^2} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy u_l\left(\frac{\sqrt{2}x}{w}\right) u_m\left(\frac{\sqrt{2}y}{w}\right) \{1+i\Phi(x, y)\} u_{00}\left(\frac{\sqrt{2}x}{w}, \frac{\sqrt{2}y}{w}\right) \\ &\simeq i \frac{2e^{-i[l+m]\psi(z)}}{w^2} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy u_l\left(\frac{\sqrt{2}x}{w}\right) u_m\left(\frac{\sqrt{2}y}{w}\right) \Phi(x, y) u_0\left(\frac{\sqrt{2}x}{w}\right) u_0\left(\frac{\sqrt{2}y}{w}\right) \\ &\simeq i \frac{2e^{-i[l+m]\psi(z)}}{\pi w^2 \sqrt{2^{l+m} l! m!}} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy H_l\left(\frac{\sqrt{2}x}{w}\right) H_m\left(\frac{\sqrt{2}y}{w}\right) \exp\left\{-\frac{2(x^2+y^2)}{w^2}\right\} \Phi(x, y) \end{aligned} \quad (13)$$

$$\text{where } u_j\left(\frac{\sqrt{2}x}{w}\right) = \frac{H_j(\sqrt{2}x/w)}{\sqrt{2^j j!} \sqrt{\pi}} \exp\left\{-\frac{x^2}{w^2}\right\} = \frac{H_j(\sqrt{2}x/w)}{\sqrt{2^j j!}} u_0\left(\frac{\sqrt{2}x}{w}\right).$$

In a similar way, the relative amplitude of the Laguerre-Gaussian higher order modes can be obtained (see Appendix G). For the relative amplitude of the fundamental mode ($l=m=0$), it is necessary to expand up to the second order for the approximation in order to avoid violating the power conservation.

$$\begin{aligned} S_{00} &= \frac{\sqrt{2}}{w} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy u_{00}\left(\frac{\sqrt{2}x}{w}, \frac{\sqrt{2}y}{w}\right) \exp\left\{-i\psi(z) + ik\frac{x^2+y^2}{2\mathcal{R}(z)}\right\} \Psi(x, y, z) \\ &\simeq \frac{2}{w^2} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \left[u_{00}\left(\frac{\sqrt{2}x}{w}, \frac{\sqrt{2}y}{w}\right)\right]^2 \left\{1+i\Phi(x, y) - \frac{1}{2} [\Phi(x, y)]^2\right\} \\ &\simeq 1 + \frac{2}{\pi w^2} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \exp\left\{-\frac{2(x^2+y^2)}{w^2}\right\} \left\{i\Phi(x, y) - \frac{1}{2} [\Phi(x, y)]^2\right\} \end{aligned} \quad (14)$$

As can be seen in eqs. (13) and (14), the amplitudes of higher order modes will have non-zero components, and correspondingly the amplitude of the fundamental mode will decrease. Thus, the scattering due to the spatial fluctuation of the phase can be considered as the inter-conversion of the power among different transverse modes. This is also true if a beam with non-fundamental mode is injected, then a part of its power will be converted to other transverse modes including the fundamental one.

Considering only a single mirror, the relative power scattered into a certain mode can be obtained by squaring the absolute value of the relative amplitude $|S_{lm}|^2$. The relative power of the fundamental mode represents the ‘fundamental mode reflectivity’, or unity minus the loss due to scattering. Equivalently, summing the relative power of all the higher modes (excepting that of the fundamental mode) gives the loss due to scattering.

$$\begin{aligned} \mathcal{A}_{\text{scat}} &= 1 - |S_{00}|^2 = \sum_{l,m \neq 0} |S_{lm}|^2 \\ &\simeq \frac{2}{\pi w^2} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \exp\left\{-\frac{2(x^2+y^2)}{w^2}\right\} [\Phi(x, y)]^2 \\ &\quad - \left| \frac{2}{\pi w^2} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \exp\left\{-\frac{2(x^2+y^2)}{w^2}\right\} \left\{ i\Phi(x, y) - \frac{1}{2}[\Phi(x, y)]^2 \right\} \right|^2 \end{aligned} \quad (15)$$

The last expression is obtained by using eq. (14).

Let us consider, as an example, a sinusoidal fluctuation in the x -axis whereas no fluctuation is assumed in the y -axis. Then the integration by y in eq. (13) is reduced to δ_{0m} , thus we only need to consider the components S_{l0} . A sinusoidal fluctuation in the x -axis can be written as

$$\Phi(x) = \Phi_0 \cos(2\pi x/\Lambda + \Theta), \quad (16)$$

where Φ_0 is the peak amplitude of the fluctuation, Λ and Θ are the spatial wavelength and the spatial phase (relative to the beam center), respectively.

Inserting eq. (16) into eq. (13) yields

$$\begin{aligned} S_{l0} &= \frac{e^{-il\psi(z)}}{\sqrt{2^l \pi l!}} \frac{\sqrt{2}}{w} \int_{-\infty}^{\infty} dx H_l\left(\frac{\sqrt{2}x}{w}\right) \exp\left\{-\frac{2x^2}{w^2}\right\} \exp\{i\Phi(x)\} \\ &= \frac{e^{-il\psi(z)}}{\sqrt{2^l \pi l!}} \frac{\sqrt{2}}{w} \int_{-\infty}^{\infty} dx H_l\left(\frac{\sqrt{2}x}{w}\right) \exp\left\{-\frac{2x^2}{w^2}\right\} \exp\{i\Phi_0 \cos(2\pi x/\Lambda + \Theta)\} \\ &\simeq ie^{-il\psi(z)} \Phi_0 \cos\left(\Theta + \frac{\pi}{2}l\right) \left(\frac{\Omega}{\sqrt{2}}\right)^l \frac{\exp\{-\Omega^2/4\}}{\sqrt{l!}} \end{aligned} \quad (17)$$

$$\text{where} \quad \Omega = \frac{2\pi}{\Lambda} \frac{w}{\sqrt{2}} \quad : \text{normalized spatial frequency}$$

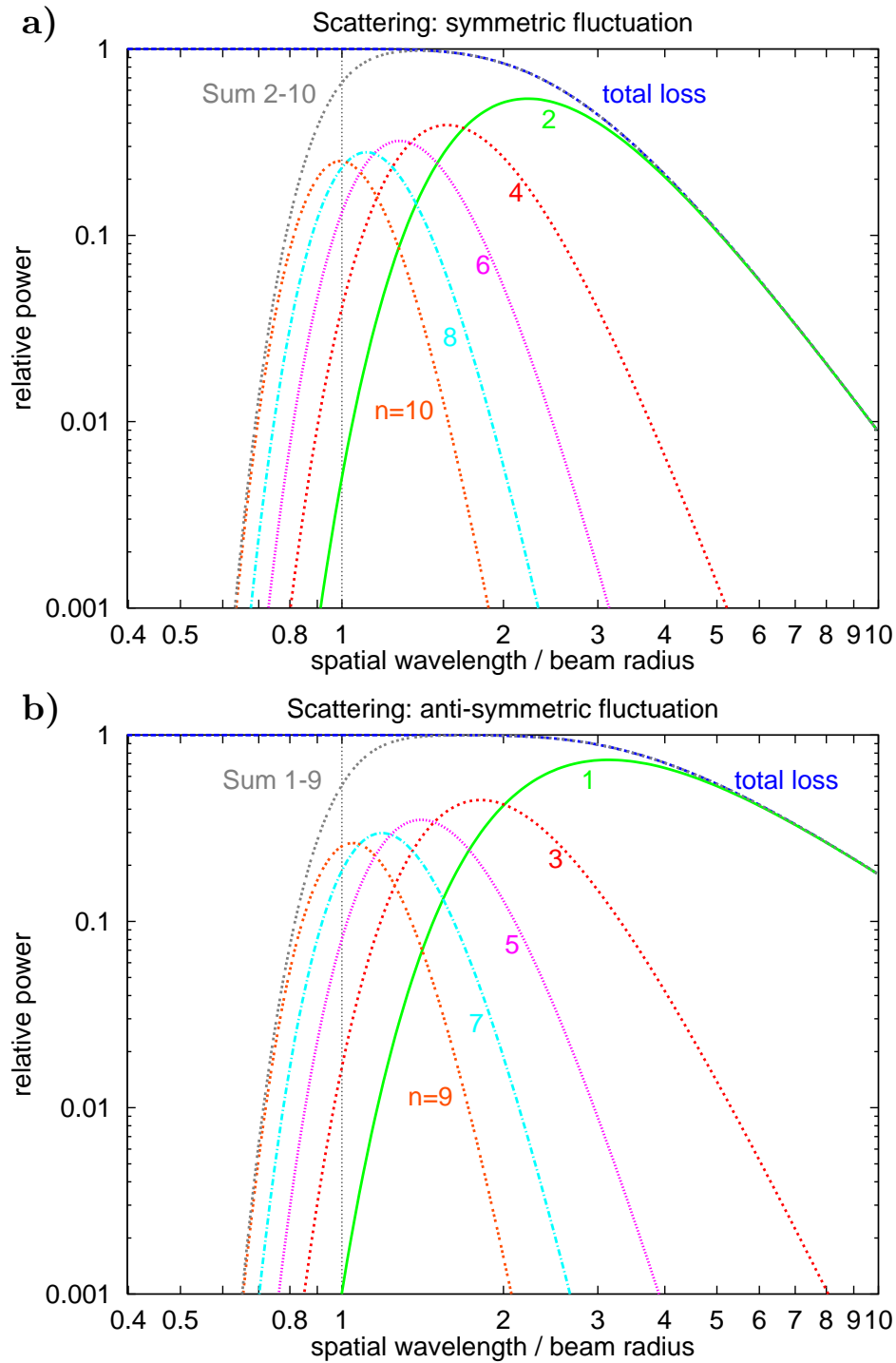


Figure 5.4: The relative power scattered into higher order modes by a one-dimensional sinusoidal fluctuation. Only even modes are excited by **a)** symmetric fluctuation and odd modes by **b)** anti-symmetric fluctuation. The vertical scale is normalized by $\Phi_0^2/2$, where Φ_0 is the zero-to-peak amplitude of the spatial fluctuation. The total loss approaches unity in the normalized scale in the range where the spatial wavelength is shorter than the beam radius.

for a higher order mode and, from eq. (14), the relative amplitude of the fundamental mode is obtained as

$$\begin{aligned} S_{00} &\simeq \frac{\sqrt{2}}{w\sqrt{\pi}} \int_{-\infty}^{\infty} dx \exp\left\{-\frac{2x^2}{w^2}\right\} \left[1 + i\Phi(x) - \frac{1}{2}\Phi^2(x)\right] \\ &\simeq 1 + i\Phi_0 \cos \Theta e^{-\Omega^2/4} - \frac{1}{4}\Phi_0^2 \left(1 + \cos 2\Theta e^{-\Omega^2}\right). \end{aligned} \quad (18)$$

As expected, a symmetric fluctuation ($\Theta=0$) excites only even modes ($l=2j$) and an anti-symmetric fluctuation ($\Theta=\frac{\pi}{2}$) only odd ones ($l=2j+1$). A fluctuation of the spatial wavelength Λ typically excites the mode with the index

$$l_{\text{typ}} \simeq \Omega^2/2 = (\pi w/\Lambda)^2. \quad (19)$$

This is consistent with the typical spatial wavelength $\Lambda_{\text{typ}} \sim 4w/\sqrt{l}$ of the l -th mode which has l nodal lines within the typical diameter¹⁾ $\sim 2w\sqrt{l+1/2}$.

An example of the relative power in each mode scattered by a sinusoidal fluctuation is shown in Fig. 5.4. As can be seen in the figure, the total loss by a sinusoidal fluctuation approaches to a constant value $\mathcal{A}_{\text{scat}}$ when the spatial wavelength becomes shorter than the beam radius ($\Lambda \lesssim w$). The total loss in such a case is determined only by the peak amplitude of fluctuation Φ_0 as

$$\mathcal{A}_{\text{scat}} = \Phi_0^2/2 \quad \text{for } \Lambda \lesssim w \quad (20)$$

(the vertical scale of Fig. 5.4 is normalized by this).

It can also be seen in Fig. 5.4 that the power converted into one given mode decreases with the shorter spatial wavelength. At the same time, however, the number of the modes that are scattered by the sinusoidal fluctuation of a given spatial wavelength increases. This compensates the lesser power in one mode to result in a constant total loss.

It should be mentioned that the first order mode ($l=1$) can be eliminated (to first order) by adjusting the orientation of the mirror properly [59]. Thus, fluctuation with spatial wavelengths much longer than the beam radius ($\Lambda \gtrsim \pi w$) practically do not cause any scattering.

¹⁾ The standard deviation of the fundamental mode is $\sigma=w/\sqrt{2}$, and the characteristic width of l -th mode is given by $\sqrt{2l+1}\sigma$. The nodal lines are present within this characteristic width. See Appendix G for more detail.

5.2.5: Cavity effect (mode cleaning effect)

The above consideration can, however, be applied only if a single mirror is present. If two or more mirrors are aligned to compose a cavity, there are either suppression or enhancement of higher order modes due to the multiple interference inside the cavity.

Let us assume the case of two mirrors composing a Fabry-Perot cavity. Each mirror has an appropriate curvature for a stable mode to exist and the incident beam matches to the mode. Furthermore, one of the mirrors is assumed to be ‘perfect’, i.e. it has no deviation from the ideal shape, and the other has only small spatial fluctuation from the ideal shape. By restricting ourselves to small fluctuations, we can neglect the higher order effects, i.e. the effects of scattering back into the fundamental mode from higher order modes.

In a cavity, the light scattered into one of the low order modes ($l, m \lesssim 10$) has a radius w smaller than the size of the mirrors. Then this light component will be reflected back and forth inside the cavity, forming a multiple interference inside the cavity. As a result, the amplitude of that mode inside the cavity depends on the tuning condition of the cavity for that particular mode.

In other words, the amount of the power converted into a transverse mode depends on the boundary condition which is determined by the mirrors. This is actually the same mechanism as the case of the frequency (longitudinal mode) conversion induced by gravitational waves. In both cases the perturbation (spatial fluctuation or gravitational waves) converts a certain amount of *amplitude* into a different mode (transverse or longitudinal), but the *power* converted into that mode is also affected by the boundary condition.

As can be found in Appendix G, the phase shift along the propagation, the so-called Guoy phase shift, depends on the number of nodal lines. This makes it possible to distinguish the fundamental mode from higher order modes. By adjusting the cavity so that it is resonant with the fundamental mode but not with higher order modes, the amount of power converted into the higher order modes can be suppressed. This practically reduces the loss due to scattering.

The amplitude *inside* a Fabry-Perot cavity is modified (either enhanced or suppressed) by a factor

$$G(\phi) = \frac{1}{1 - \rho_1 \rho_2 \exp\{-i\phi\}}, \quad (21)$$

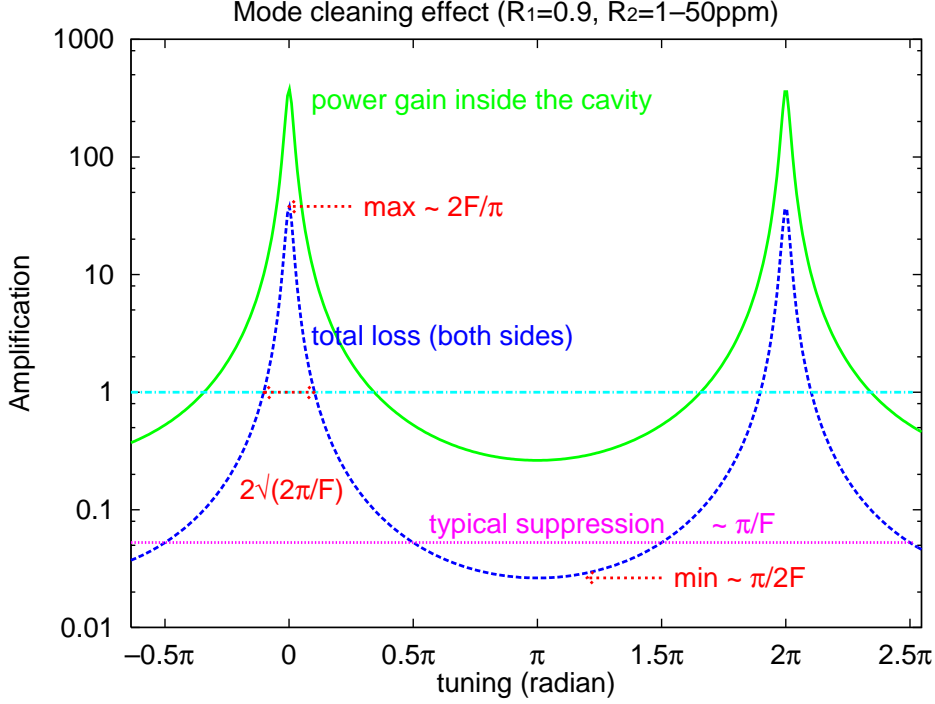


Figure 5.5: Mode suppression or enhancement factor in a Fabry-Perot cavity with respect to the tuning. The power transmittances of the mirrors are 0.1 and $5 \cdot 10^{-5}$, respectively, which corresponds to a finesse of ~ 60 . It has a maximum of $\sim \frac{2}{\pi} \mathcal{F}$ when resonant (~ 0 or 2π radian) and a minimum of $\sim 1/(\frac{2}{\pi} \mathcal{F})$ when anti-resonant ($\sim \pm \pi$ radian). The typical suppression factor is $\sim \pi/\mathcal{F}$.

where ρ_1 and ρ_2 are the amplitude reflectivities of the mirrors, respectively, and ϕ is the tuning of the cavity in one round trip at that ‘mode’.

In this case both the transverse and longitudinal mode must be taken into account. Then the tuning can be written as

$$\begin{aligned} \phi_{lm}(\omega) &= \omega t_r - (1+l+m)\psi_r \quad (\text{mod } 2\pi) \\ &= \omega_g t_r - (l+m)\psi_r + \phi_0, \end{aligned} \quad (22)$$

where t_r is the retarded time for a round trip inside the cavity, ψ_r is the Guoy phase shift of the fundamental mode in a round trip, l and m are the number of nodal lines in the two axes, respectively. The second line applies when there is some modulation inside the cavity, like the signal cavity in an interferometric gravitational-wave detector; ω_g is the angular frequency of the modulation that produces the sidebands (negative for the lower sideband), and $\phi_0 = t_r \omega_0 - \psi_r$ is the tuning of the fundamental mode at the carrier frequency.

The relative loss at a mode is given by the power which leaks out of the cavity through either of the mirrors,

$$\begin{aligned} P_{lm} &= |a_{lm}|^2 \times \left[|\tau_1 G(\phi_{lm})|^2 + |\tau_2 G(\phi_{lm})|^2 \right] \\ &= |a_{lm}|^2 \times \frac{(\tau_1)^2 + (\tau_2)^2}{1 + (\rho_1 \rho_2)^2 - 2\rho_1 \rho_2 \cos \phi_{lm}}. \end{aligned} \quad (23)$$

Thus, the second factor can be considered as the effect of the mode suppression or enhancement.

Here a problem arises: How can one determine the reflectivity and transmittance of a mirror without knowing the amount of loss due to scattering, which is affected by the mode suppression or enhancement? To get the exact answer for this problem, we must consider the second or even higher order contribution of mode conversion, i.e. the effects from such light that is converted from a high order mode to the other modes. We will, however, neglect them here and use the reflectivity and transmittance without the cavity effect. Since we assume the amplitude of fluctuation is small, the higher order effects of mode conversion are small and this will be a good enough approximation.

Fig. 5.5 shows an example of the mode suppression or enhancement factor with respect to the tuning. It can be seen that the relative loss is enhanced if the cavity is resonant with that mode and it is suppressed if the cavity is off-resonant. The border between enhancement and suppression is given by

$$|\phi| = \arccos(\rho_1 \rho_2) \simeq \sqrt{2\pi/\mathcal{F}} \quad (\text{for } \mathcal{F} \gg 1). \quad (24)$$

Thus, in the case of reasonably high finesse, most of the tuning range corresponds to suppression rather than enhancement.

The maximum enhancement or suppression factor ($\phi=0$ or π) is roughly given by $\sim \frac{2}{\pi}\mathcal{F}$ or $\sim 1/(\frac{2}{\pi}\mathcal{F})$ where \mathcal{F} is the finesse of the cavity. (Note that this is defined in power, not amplitude. The amount of signal enhancement is, as is discussed in 3.4.1, determined by the amplitude.) The off-resonant modes are suppressed typically by a factor of one half of the maximum case $\sim \pi/\mathcal{F}$, and those with $|\phi| \lesssim \frac{\pi}{2} \pmod{2\pi}$ are suppressed more strongly.

For our purpose of reducing the loss due to scattering into higher order modes, it is necessary to set the tunings of higher order modes so that they will be suppressed effectively. It is impossible, however, to suppress all the higher order modes. Since the tuning difference between two successive modes

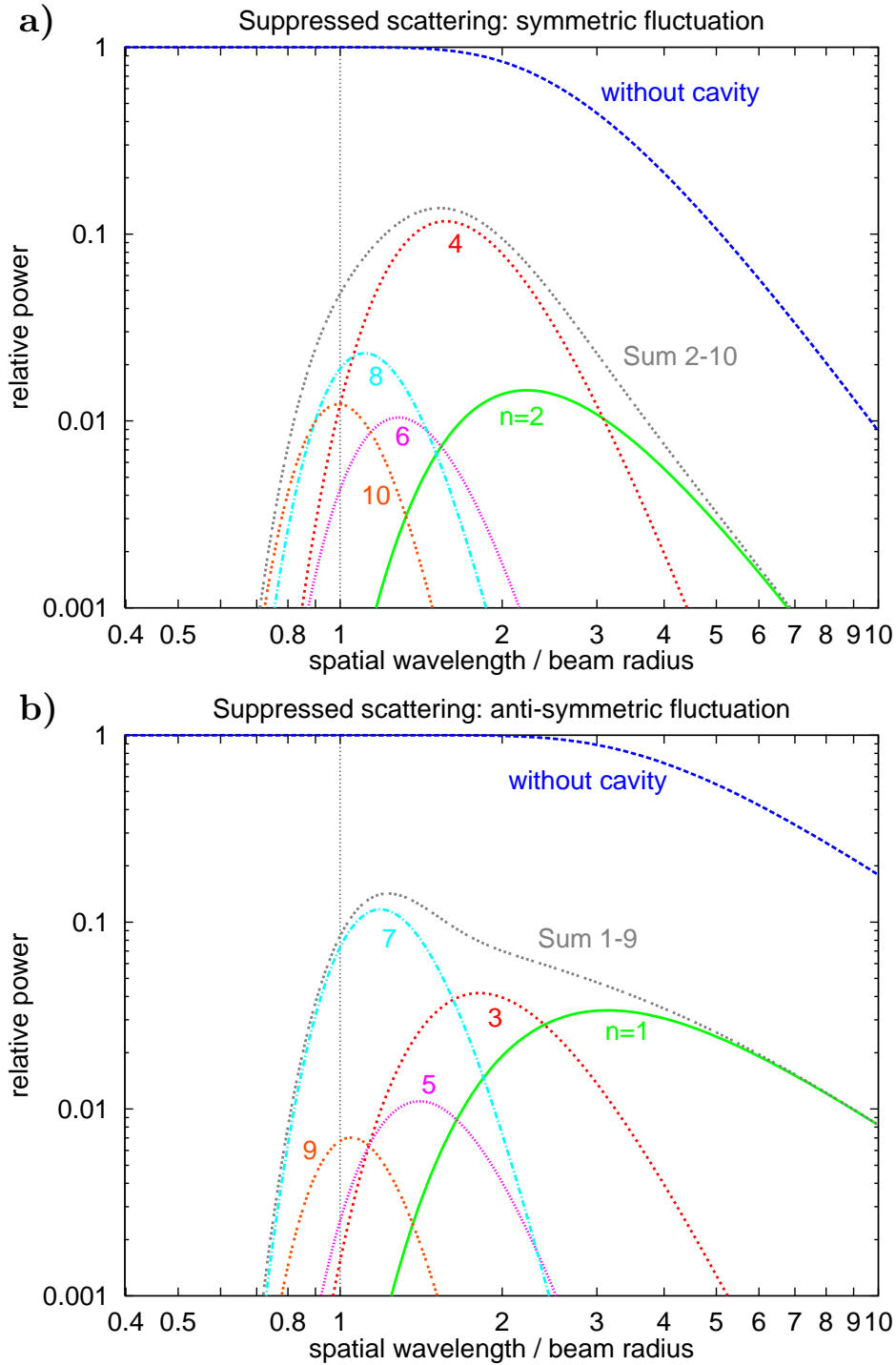


Figure 5.6: The scattering loss due to a sinusoidal fluctuation, with the mode cleaning effect of the cavity taken into account. The power transmittances of two mirrors are the same as in Fig. 5.5, 0.1 and $5 \cdot 10^{-5}$, respectively. The Guoy phase shift per round trip in a cavity of the fundamental mode is assumed to be $\psi_r \cong 0.274 \times 2\pi$. It can be seen that most of the modes can be suppressed, though by different factors, compared with those without the cavity (see Fig. 5.4).

is a constant, one among $[\sqrt{\pi\mathcal{F}/2}]$ modes²⁾ must be enhanced. Thus, it is important to have a high finesse for the mode cleaning purpose both to suppress effectively and to suppress as many modes as possible.

It is also known [57] that, fortunately, the most of the power goes into relatively low order modes. This makes it possible to adjust the cavity to suppress the relatively strong modes effectively. On the other hand, the number of the modes involved increases with the nodal number $(l+m)$. It must be avoided to have strong enhancement even at very high order modes.

It is ψ_r , the Guoy phase shift in a round trip, that determines the tuning difference of between two successive modes. It is given by

$$\psi_r = \arctan(ct_r/z_R) \quad (25)$$

where z_R is the Rayleigh range, which in turn is determined by the curvature radii of the mirrors and their distance, as can be found in eq. (11).

To avoid the strong enhancement of higher order modes, this must be chosen so that $\psi_r/2\pi$ will not be a rational number. Furthermore, it should be chosen to suppress relatively low order modes effectively. For example, $\psi_r/2\pi \sim 0.25\text{--}0.3$ or $1.7\text{--}1.75$ seems to be a good choice to suppress the second to the fifth modes (note that the first order mode can be eliminated by choosing the orientation of the mirror appropriately).

Fig. 5.6 shows an example of mode cleaning effect, again taking a sinusoidal fluctuation at one of the mirrors. Thus, the result can be directly compared with Fig. 5.4. It is seen that most of the modes and thus the total loss can be suppressed by a typical factor $\sim\pi/\mathcal{F}$, even by a cavity with relatively low finesse ($\mathcal{F}\sim 60$). With higher finesse cavities, it is easier to suppress them more effectively, though one must be careful to avoid accordingly strong enhancement.

²⁾ Here $[x]$ means the smallest integer bigger than x .

5.3: Thermal distortion

5.3.1: Thermal expansion

Now we return to the problems of thermally-induced distortions mentioned in 5.1.2. First of all, let us consider one of the simplest cases, which is the thermal expansion due to the absorption at the surface of a mirror. As is discussed in 5.2.2, the deformation due to the thermal expansion is assumed to be more or less spherical. The deviation from the ideal sphere is treated as the spatial fluctuation whose amplitude is small compared with the wavelength.

The spatial fluctuation increases the loss due to scattering, as is discussed in 5.2.4. This loss may, however, be reduced by the mode cleaning effect if it occurs inside a cavity, as is discussed in 5.2.5. Furthermore, it can in principle be compensated by increasing the injected light power, if we have a strong enough light source.

Another thing to be considered is the change in the curvature of the sphere. As can be seen in Fig. 5.3, the curvature of each mirror of a Fabry-Perot cavity must satisfy a certain condition. The change in the curvature of one or both of the mirrors may cause this condition to be violated, resulting in an unstable cavity, which cannot be compensated by means of injecting power. Thus, the major emphasis here is placed on the curvature change due to thermal expansion and its implication on the cavity stability, which will be referred to as the *thermal stability* of a cavity.

Due to the intensity distribution of a Gaussian beam, the thermal expansion also depends on the distance from the center of the beam. The amount of deformation is characterized by the *sagitta* ς , which is defined as the curvature depth of the mirror surface measured across the beam diameter. It is given as

$$\varsigma \simeq \frac{w^2}{2\mathcal{R}} \quad (26)$$

by the curvature radius \mathcal{R} of the mirror and the beam radius w at the mirror using a quadratic approximation of a sphere. The sagitta is positive for a converging mirror and negative for a diverging mirror.

Since the intensity outside the beam diameter is small compared with that within the diameter, it is assumed that the temperature gradient occurs within the hemisphere of the same radius as the beam spot on the mirror. Then the

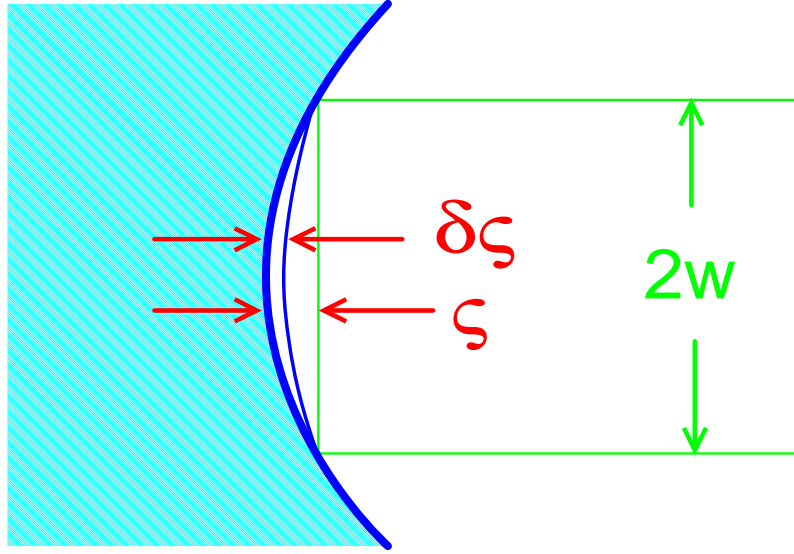


Figure 5.7: The local curvature of a mirror is characterized by the sagitta ζ , which represents the curvature depth of the mirror surface measured across the beam diameter. The thermal expansion due to the local heating changes the curvature of the mirror which is represented by the change in sagitta $\delta\zeta$.

change in sagitta, which represents the local curvature change (see Fig. 5.7) due to thermally induced deformation, can be estimated [54] as

$$\delta\zeta \simeq \frac{\alpha/\kappa}{4\pi} P_A \quad (27)$$

for a fixed beam radius. Here P_A is the total absorbed light power, α and κ are the thermal expansion coefficient and the thermal conductivity of the substrate material, respectively.

It is interesting to note that $\delta\zeta$ depends only on the absorbed power and the thermal properties of the substrate α/κ , but not on the beam radius. When the *eigen mode* of a cavity is considered, however, this is no longer true since the change in the curvature varies the beam radii on the mirrors. From eq. (26), it is expected that this in turn changes the sagitta. Due to this ‘secondary’ effect, the dependence of the sagitta on the thermal distortion is non-linear for the eigen mode of a cavity.

The thermal expansion due to the absorption inside the substrate will also occur. This can, however, be neglected because the substrate is assumed to be much thicker than the beam radius (from mechanical requirements) and thus the end effect is expected to be small. It is also assumed that the light power transmitted through the substrate is much less than that hitting the surface and thus a weaker effect is expected from this cause. Furthermore, the thermal

lensing, to be discussed next, is found to have a bigger effect than the thermal expansion, as can be seen in Table 5.1.

5.3.2: Thermal lensing

If the thermal lensing in a transmitting optics is concerned, what matters is the ‘optical’ path length change due to the gradient of the refraction index. This effect is characterized by a parameter $\beta = \partial n / \partial \mathcal{T}$, the temperature dependence of the refraction index. This has a similar effect with α , the thermal expansion coefficient, in the previous problem.

In this case, the sagitta must be defined by the spherical curvature of the wavefront, or the path length difference between the beam center and the point of the beam radius. The change in sagitta by $\delta\zeta$ has a similar effect with a lens of a focal length $z_F \simeq w^2 / 2\delta\zeta$ for single passage (note that light often passes through a thermal lens twice in a round trip of a cavity).

The temperature gradient due to the absorption at the mirror surface will cause a thermal lensing effect that can be described by replacing α in eq. (27) by β [54] as

$$\delta\zeta \simeq \frac{\beta/\kappa}{4\pi} P_A . \quad (28)$$

Again, this depends only on the total absorbed power and the thermal properties of the substrate β/κ , but not on the beam radius.

The thermal lensing due to the absorption inside the substrate also has a significance, since its effect is proportional to the length through which the light is transmitted. This is a difference from the thermal expansion which appears only at the end of the substrate. The thermal lensing inside the substrate can, again, be written in a similar form to eqs. (27) and (28) [56] as

$$\delta\zeta \simeq 1.3 \frac{\beta/\kappa}{4\pi} \frac{dP_A}{dz} z_d , \quad (29)$$

where dP_A/dx is the absorbed power per unit length and z_d is the thickness of the substrate. (The reason for writing it in this form rather than a simple total absorbed power is that the absorption per unit length is thought to be a constant and thus the thermal lensing effect in the substrate is expected to be proportional to the substrate thickness—an important issue in choosing parameters.)

The additional factor 1.3 from the case of the surface absorption arises from the fact that the area through which the heat escapes by conduction is less than

that in the surface case. This results in a larger temperature gradient and thus a stronger lensing effect. Another difference is that there is a temperature gradient along the optical axis in the surface case. This also affects the total lensing strength.

5.3.3: Thermal properties of substrate materials

Table 5.1 reviews the relevant thermal properties of the materials which are of interest as the substrates of optical components. As can be seen in the table, the thermal lensing always has a stronger effect than the thermal expansion for these materials.

In the choice of the substrate material for our optical components, not only optical properties but also mechanical properties must be taken into account. Though their values are shown for comparison, both of typical low-expansion materials, Zerodur and ULE, must be excluded as candidates because of their relatively large internal damping. It is also true that they are not so much of interest when thermal lensing is concerned.

Among the rest, fused silica is thought to be the most promising material, because of the mechanical properties and moreover because of the availability for the required size as the substrate for long arm interferometers. It has, however, relatively strong thermal lensing effect, as can be seen in the table, due to its relatively poor thermal conductivity.

For non-transmitting optics such as delay line mirrors or far end mirrors of arm cavities, silicon is a good candidate. This also has good mechanical properties. This may, however, require development of polishing technology since it has rarely been used as substrates for high-quality mirrors so far.

Although sapphire has good thermal and mechanical properties and thus can be used for transmitting optics, its inherent birefringence may introduce another problem [61]. Obviously diamond would be the optimum, both from the thermal and mechanical properties. Both of these materials may, however, be difficult to obtain in a big enough piece for our purpose.

From the above consideration, fused silica seems to be a reasonable choice for the substrate of transmitting optics, at least for the moment. For non-transmitting optics, silicon can also be used. Thus, in the following discussion, we estimate the sensitivities of interferometers with the two cases in mind: one is those made of fused silica only and the other is a combination of fused silica and silicon.

Table 5.1: Thermal properties

material	mass density (g/cm ³)	sound velocity (km/s)	expansion α/κ (nm/W)	lensing β/κ (nm/W)
diamond	3.51	17.5	1.3	10
sapphire	3.98	9.3	280	600
silicon	2.33	8.4	12.8	—
fused silica	2.20	5.7	330	8700
ULE	2.20	5.5	± 23	8500
Zerodur	2.53	6.0	± 14	

Another parameter which is relevant to the thermal distortion problem is the fraction of absorption. The best values reported so far are of the order of ~ 1 ppm for a coated surface and ~ 1 ppm/cm for a bulk of fused silica. These values may be improved with the development of techniques, especially for coatings. At the moment, however, we will use these values for the estimation of the thermal distortion effects.

5.3.4: The effect of thermal distortion on the interference

As is mentioned in 5.1.2, thermal distortions deteriorate the interference between the two arms of an interferometer, unless they occur identically in the two arms. In practice, it is natural to assume that there will be mismatching between the two arms to some extent. Here we will consider its effect.

Following the assumption discussed in 5.2.2, the thermal lensing (or expansion—we will take lensing as the example since it is expected to have larger effects) is separated into a perfect lens with a certain focal length and the deviation from it. When a fundamental Gaussian beam passes through this distorted path, the former converts it to a Gaussian beam with different beam parameters and the latter converts a part of the injected power into different spatial modes. Since the latter is assumed to be small, we will consider here mainly the former effect.

According to the ABCD law [62], a thin lens with focal length z_F converts the parameters of a fundamental Gaussian beam as

$$z' = \frac{Az + B}{Cz + D}, \quad z = z + iz_R \quad (30)$$

where $\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1/z_F & 1 \end{pmatrix}$ for a thin lens with focal length z_F

where z and z' are the complex beam parameters of the original beam and the converted beam, respectively. The origin ($z=0$) is assumed to be located at the waist point of the original beam. This can be rewritten as

$$z' = z \frac{1 - (z^2 + z_R^2)/(zz_F)}{(1 - z/z_F)^2 + (z_R/z_F)^2}, \quad z'_R = \frac{z_R}{(1 - z/z_F)^2 + (z_R/z_F)^2}, \quad (31)$$

by taking the real and the imaginary parts of eq. (30).

If the two arms of an interferometer have lenses with different focal lengths z'_F and z''_F , its interference visibility can be obtained by inserting eq. (31) into eq. (G56) in Appendix G, as

$$\begin{aligned} \mathcal{C} = |a_{00}|^2 &= \frac{4z'_R z''_R}{(z' - z'')^2 + (z'_R + z''_R)^2} \\ &= \frac{4(z'_F z''_F z_R)^2}{4(z'_F z''_F z_R)^2 + (z^2 + z_R^2)^2 (z'_F - z''_F)^2} \\ &= \left[1 + \left(\frac{z^2 + z_R^2}{2z} \right)^2 \left(\frac{1}{z'_F} - \frac{1}{z''_F} \right)^2 \right]^{-1}. \end{aligned} \quad (32)$$

The loss due to this deteriorated visibility is

$$\begin{aligned} \mathcal{A}_C &= \frac{1 - \mathcal{C}}{2} \simeq \frac{1}{2} \left(\frac{z^2 + z_R^2}{2z} \right)^2 \left(\frac{1}{z'_F} - \frac{1}{z''_F} \right)^2 \\ &\simeq \frac{1}{2} \left(\frac{z^2 + z_R^2}{z} \right)^2 \frac{(\delta\zeta' - \delta\zeta'')^2}{w^4} \simeq \frac{1}{2} \left(\frac{k\Delta\zeta}{2} \right)^2, \end{aligned} \quad (33)$$

where w and $k = 2\pi/\lambda$ are the beam radius and the wave number, respectively. Here $\Delta\zeta = \delta\zeta' - \delta\zeta''$ and $z_F \simeq w^2/2\delta\zeta$ were used.

The above result means that there will be a loss of $\sim 5 \cdot 10^{-4}$ if the thermal lensing in the two arms differ in sagitta by $\lambda/100$. Assuming 10% error between the two arms, this corresponds to a change in sagitta of $\lambda/10$ in one of the arms. Supposing if fused silica is used as the substrate material, this amount of deformation will be caused by an absorption of 150 mW at a surface, or of

120 mW inside the substrate. If there is no compensation at all, however, the mismatch between the two arms will be more significant [54].

Fortunately, there is a remedy for this problem. As is discussed in 5.2.5, the loss due to scattering can be reduced if it occurs inside a cavity. When a mirror is present at the interfering output, as in signal recycling or in resonant sideband extraction, the interference occurs inside the cavity that is composed of the signal (recycling/extraction) mirror and the arms. Then the modes that are not resonant with this cavity will be suppressed, resulting in a reduction of the loss. In addition, this suppression can be made quite efficient if the modes expected to be produced are known in advance, as in the thermal distortion case [56].

With the mode cleaning effect taken into account, the power loss due to the imperfect visibility can be approximated as follows:

$$\begin{aligned} \mathcal{A}_{\text{int}} &\simeq \frac{(\tau_s)^2}{2} \frac{1}{2} \left(\frac{k}{2} 2\Delta\zeta \right)^2 \simeq \frac{(\tau_s)^2}{4} \left(\frac{k\mathcal{Y}}{2} 2\delta\zeta \right)^2 \\ &\simeq \frac{(\tau_s)^2}{4} \left\{ \frac{k\mathcal{Y}}{4\pi} \left(\sum_j \frac{\alpha_j + \beta_j}{\kappa_j} P_{Cj} + 1.3 \sum_m \frac{\beta_m}{\kappa_m} P_{Bm} \right) \right\}^2 \\ \mathcal{Y} &= \frac{\Delta\zeta}{\delta\zeta} = \frac{\delta\zeta_1 - \delta\zeta_2}{\delta\zeta}, \end{aligned} \quad (34)$$

where \mathcal{Y} is the fraction of the mismatch between the thermal distortions in the two arms, $(\tau_s)^2$ is the power transmittance of the signal mirror, and P_C and P_B are the absorbed power at the surface (coating) and inside the substrate (bulk), respectively. The subscripts j and m represent the different points where absorptions take place, and thermal properties at each point can be different. It is assumed that the light passes a thermal lens twice and that the power transmittance of the signal mirror dominates the losses inside the signal cavity.

5.3.5: Thermal distortion in a cavity

To consider the thermal stability problem, let us start from a simple Fabry-Perot cavity case. When there is no other optics inside the cavity (as in the case of the arm cavities), only the thermal expansion of the mirrors must be considered. The small deviation from the ideal sphere may increase the loss due to scattering, but it will be suppressed by the mode cleaning effect. Thus, in this case the curvature of the sphere is most important.

From the discussion in 5.2.3, it is expected that a big change in the curvature of one or both of the mirrors makes the initially stable cavity unstable. If

the two mirrors have the same thermal properties, the maximum change allowed is from -1 to 0 or 0 to 1 in terms of ζ -parameters (see eqs. (9) and (10) and Fig. 5.3). If one of the mirrors has a lower expansion coefficient (as in the case of a combination of fused silica and silicon), this tolerable range may be even bigger by choosing an appropriate initial geometry.

It should be noted, however, that the beam radius at one of the mirrors will be too large when the parameters approach a boundary of the stable region, as can be seen in Fig. 5.3. If we want the beam always to be smaller than $\sqrt{2}$ times of that in the confocal case, the ζ -parameters are not allowed to change by more than unity for any initial geometry. Furthermore, the assumption of small deviation may not be satisfied if thermal deformation becomes too large. Thus, we will use $1/2$ in ζ -parameter as the upper limit, taking a safety factor into account.

The above limitation corresponds to a change in sagitta by $\delta\zeta \sim \lambda/4\pi$ for a near-confocal geometry. Though the change of sagitta for the eigen mode of a cavity is non-linear, its effect is small in this case since we use only the region where the change in the beam radius is small. Thus, it is possible to estimate the critical absorbed power³⁾ as

$$\overline{P_E} \simeq \frac{\lambda}{4\pi} \frac{4\pi}{\alpha/\kappa} \simeq \frac{\lambda}{\alpha/\kappa} \approx 3.2 \text{ W} \times \left[\frac{\lambda}{1.064 \mu\text{m}} \right] \left[\frac{\alpha/\kappa}{0.33 \mu\text{m}/\text{W}} \right]. \quad (35)$$

The numerical result quoted is that for fused silica, and it will be 64 W for silicon.

The coupling of the injected beam to the cavity will change when the eigen mode of the cavity changes by thermal distortion. This can in principle be compensated by adjusting the input beam actively. Within the region with relatively small change, however, the coupling will not be so much deteriorated ($\gtrsim 0.9$) even without any compensation. This requires just 10% more injecting power in order to store the same amount of energy.

When transmitting optics are present inside the cavity (as the beamsplitter in a power recycling cavity), the thermal lensing inside that optical component must be taken into account. In this case, a lens effectively changes the curvature of a mirror and the distance between the mirrors simultaneously. This has, however, relatively small effect as long as the focal length is much longer than

³⁾ Note that the definition of the critical power here is identical with Ref. [53], but is different from Ref. [51] by a factor of 2.

the cavity length. This condition is satisfied, again, in the region of small change.

Thus, we can use the same threshold as in the expansion case, $\delta\zeta \sim \lambda/4\pi$, and the critical absorbed power can be defined in a similar way:

$$\overline{P}_C \simeq \frac{\lambda}{4\pi} \frac{4\pi}{\beta/\kappa} \simeq \frac{\lambda}{\beta/\kappa} \approx 120 \text{ mW} \times \left[\frac{\lambda}{1.064 \mu\text{m}} \right] \left[\frac{\beta/\kappa}{8.7 \mu\text{m}/\text{W}} \right] \quad (36)$$

for surface (coating) absorption and

$$\overline{P}_B \simeq \frac{\lambda}{4\pi} \frac{4\pi}{1.3\beta/\kappa} \simeq \frac{\lambda}{1.3\beta/\kappa} \approx 94 \text{ mW} \times \left[\frac{\lambda}{1.064 \mu\text{m}} \right] \left[\frac{\beta/\kappa}{8.7 \mu\text{m}/\text{W}} \right] \quad (37)$$

for substrate (bulk) absorption. The numerical results quoted are those for fused silica.

5.4: Sensitivity with thermal distortion

5.4.1: Sample parameters of an interferometer

In the following discussion, we assume an interferometer with arm lengths of 3 km which is illuminated by a laser with 1.064 μm wavelength. Each coated surface and each transmitting optical component are assumed to have minimum losses of 25 ppm and 3 ppm/cm, respectively, due to absorption and scattering (to high order modes that cannot be suppressed).

The beam radius will be 32 mm when a near-confocal geometry is adopted. The actual geometry will not be far from this. Taking a safety factor into account, the mirror diameter will be ~ 250 mm. From the mechanical requirements, a mirror must have an aspect ratio close to unity, thus its thickness must be $\gtrsim 100$ mm. Though exact numbers depend on many considerations (such as manufacturing facilities), we will take 150 mm as a typical value.

The injected beam is separated into two arms by a beamsplitter, and the returning beams from the two arms interferes with each other at the same beamsplitter. In order to retain the symmetry of the two arms, a compensation plate with similar properties to the beamsplitter will be used. The beamsplitter (and the compensation plate) has similar dimensions as the mirrors. Note that the path length inside it is 1.13 times the physical thickness due to the refraction, when the beam is injected at 45 degrees.

The interferometer has either a delay line or a Fabry-Perot cavity in each arm to increase the storage time in the arms. To maximize the amount of

energy stored, power recycling is incorporated and optimized. Signal recycling is an option either for a delay line based or a Fabry-Perot based configuration. For a Fabry-Perot based configuration, resonant sideband extraction is another possibility. In either case, the detector bandwidth must be set for the desired observation purpose.

To prevent the non-interfering beam to hit the photo-detector and to contribute to the shot noise, an output mode cleaner must be used (unless the generic mode cleaning effect of the configuration is strong enough). To avoid the unwanted coupling to the interferometer response, optical isolators must be placed between the interferometer and the mode cleaner.

The interfering output is detected after beating with a local oscillator light (by, for example, using external modulation), and then, if necessary, demodulated to reproduce the original gravitational wave signal. It is assumed that the local oscillator light is strong enough so that it dominates the shot noise at the photo-detector. This is necessary to avoid the degradation of the signal-to-noise ratio by, for example, the remaining higher order modes due to the distortions.

5.4.2: Delay lines in the arms

Supposing that delay lines are present in the arms, the light power hitting the mirrors is the same as that transmitted through the beamsplitter. Then, from the values estimated above, it is expected that the thermal lensing in the beamsplitter (and the compensation plate) will be the limiting factor. The maximum light power which can be transmitted through these substrate will be ~ 5.5 kW from the values quoted above.

The amount of energy stored in the optical system can be written, when it is limited by thermal distortion, as

$$\begin{aligned} \mathcal{E}_{\text{DL}} &\simeq 2 \times \frac{\overline{P}_{\text{B}}}{\mathcal{A}_{\text{BS}}} \times \frac{N\ell}{c} \\ &\approx 1.1 \text{ J} \times \left[\frac{\lambda}{1.064 \mu\text{m}} \right] \left[\frac{\beta/\kappa}{8.7 \mu\text{m}/\text{W}} \right]^{-1} \left[\frac{\mathcal{A}_{\text{BS}}}{17 \text{ ppm}} \right]^{-1} \left[\frac{N}{10} \right] \left[\frac{\ell}{3 \text{ km}} \right], \end{aligned} \quad (38)$$

where \mathcal{A}_{BS} is the total absorption inside the beamsplitter, N is the number of transits in the delay line, and ℓ is the arm length of the interferometer. Here 2 in the top equation stands for the two arms of an interferometer.

As can be seen in eq. (38), increasing the number of bounces increases the amount of energy, and thus improves the sensitivity. It may, however, be limited

by the required mirror size, as is discussed in 3.2.1. This also means that it is difficult to match the detector bandwidth to the desired observation purpose, especially for low frequency cases.

To optimize the detector bandwidth, it is necessary to combine with signal recycling scheme. The power transmittance of the signal recycling mirror is determined from the required detector bandwidth as

$$(\tau_{\text{SR}})^2 \simeq \frac{4\pi\Delta f_{3\text{dB}}}{c/N\ell} \approx 0.25 \times \left[\frac{\Delta f_{3\text{dB}}}{200 \text{ Hz}} \right] \left[\frac{N}{10} \right] \left[\frac{\ell}{3 \text{ km}} \right], \quad (39)$$

assuming a broad-band response, i.e. without detuning. With this configuration, the spectral sensitivity will be

$$\begin{aligned} \tilde{h}_{\text{DL}} &\simeq \sqrt{\frac{\hbar\lambda}{c} \frac{\Delta f_{3\text{dB}}}{\mathcal{E}}} \\ &\approx 8.2 \cdot 10^{-24}/\sqrt{\text{Hz}} \times \left[\frac{\lambda}{1.064 \mu\text{m}} \right]^{\frac{1}{2}} \left[\frac{\Delta f_{3\text{dB}}}{200 \text{ Hz}} \right]^{\frac{1}{2}} \left[\frac{\mathcal{E}}{1.1 \text{ J}} \right]^{-\frac{1}{2}} \\ &\approx 8.2 \cdot 10^{-24}/\sqrt{\text{Hz}} \times \left[\frac{\Delta f_{3\text{dB}}}{200 \text{ Hz}} \right]^{\frac{1}{2}} \left[\frac{\beta/\kappa}{8.7 \mu\text{m/W}} \right]^{-1} \left[\frac{\mathcal{A}_{\text{BS}}}{17 \text{ ppm}} \right]^{\frac{1}{2}} \left[\frac{N\ell}{30 \text{ km}} \right]^{-\frac{1}{2}} \end{aligned} \quad (40)$$

for a broad-band response.

When detuning is used to shift the peak frequency of the response, the bandwidth is broader by a factor of 2 but the sensitivity will be worse by a factor of 2 (assuming the same reflectivity for the signal recycling mirror as the broad-band case). In this case, however, its full bandwidth is useful whereas the low frequency part will be ‘wasted’ in the broad-band responses.

From eq. (34), the thermally induced loss of this interferometer is expected to be around

$$\mathcal{A}_{\text{th}} \simeq \frac{(\tau_{\text{SR}})^2}{4} (\mathcal{R}k \delta\zeta)^2 \approx 160 \text{ ppm} \times \left[\frac{(\tau_{\text{SR}})^2}{0.25} \right] \left[\frac{\mathcal{R}}{0.1} \right]^2, \quad (41)$$

which is reasonably small as long as moderate matching of the two arms is achieved. Without any compensation, however, this loss will be significantly large. The required illuminating light power to store the expected energy is given by

$$P_{\text{in}} \simeq \mathcal{E} \frac{\mathcal{A}_{\text{tot}}}{N\ell/c} \simeq 2 \overline{P}_{\text{B}} \frac{\mathcal{A}_{\text{tot}}}{\mathcal{A}_{\text{BS}}} \approx 5.5 \text{ W} \times \left[\frac{\overline{P}_{\text{B}}}{94 \text{ mW}} \right] \left[\frac{\mathcal{A}_{\text{tot}}}{500 \text{ ppm}} \right] \left[\frac{\mathcal{A}_{\text{BS}}}{17 \text{ ppm}} \right]^{-1} \quad (42)$$

where $\mathcal{A}_{\text{tot}} \simeq N\mathcal{A}_{\text{min}} + \mathcal{A}_{\text{th}}$ is the total loss per one round trip inside the power recycling cavity. It is a natural consequence that the critical power and the fraction of the absorption at the critical part among the total loss determine the required illuminating light power.

5.4.3: Fabry-Perot cavities in the arms

When Fabry-Perot cavities are used in the arms, the number of the optical components where thermal effects take place increases, as can be seen in Fig. 5.8. The power inside the arm cavities will be extremely high and their thermal stability must be considered. The stability of the power recycling cavity must also be taken into account, since it has stronger thermal effects due to the optics inside it. Finally, the loss due to imperfect interference may contribute.

From the thermal stability of the arm cavities, the light power hitting on the mirror surface must be less than ~ 3.2 MW, which is probably good enough for our requirement. However, the power absorbed at the surface of the coupling mirrors cause thermal lensing inside the power recycling cavity. From this view point, the light power allowed to hit a coupling mirror is ~ 120 kW. This results in the amount of energy that can be stored as

$$\begin{aligned}\mathcal{E}_{\text{AP}} &\simeq 2 \times \frac{\overline{P}_{\text{C}}}{\mathcal{A}_{\text{surf}}} \times \frac{2\ell}{c} \\ &\approx 4.8 \text{ J} \times \left[\frac{\lambda}{1.064 \mu\text{m}} \right] \left[\frac{\beta/\kappa}{8.7 \mu\text{m}/\text{W}} \right]^{-1} \left[\frac{\mathcal{A}_{\text{surf}}}{1 \text{ ppm}} \right]^{-1} \left[\frac{\ell}{3 \text{ km}} \right],\end{aligned}\quad (43)$$

where $\mathcal{A}_{\text{surf}}$ is the absorption at the surface of the coupling mirror.

Another limitation arises from the thermal lensing due to the absorption inside the beamsplitter and the coupling mirror. The presence of the coupling mirrors increases the absorption and the associated thermal lensing inside the power recycling cavity. The maximum light power transmitted through these optics will be ~ 2.9 kW. Then the amount of energy stored is

$$\begin{aligned}\mathcal{E}_{\text{RP}} &\simeq 2 \times \frac{\overline{P}_{\text{B}}}{\mathcal{A}_{\text{IC}} + \mathcal{A}_{\text{BS}}} \times \frac{2\mathcal{F}}{\pi} \times \frac{2\ell}{c} \\ &\approx 9.4 \text{ J} \times \left[\frac{\lambda}{1.064 \mu\text{m}} \right] \left[\frac{\beta/\kappa}{8.7 \mu\text{m}/\text{W}} \right]^{-1} \left[\frac{\mathcal{A}_{\text{BS}} + \mathcal{A}_{\text{IC}}}{32 \text{ ppm}} \right]^{-1} \left[\frac{\mathcal{F}}{125} \right] \left[\frac{\ell}{3 \text{ km}} \right],\end{aligned}\quad (44)$$

where \mathcal{A}_{IC} is the absorption in the substrate of the coupling mirror.

Of the above two conditions, the lower one will limit the stored energy, thus the amount of stored energy never exceeds the value quoted in eq. (43). For a relatively narrow-band case, the power in the arm cavity is likely to be the

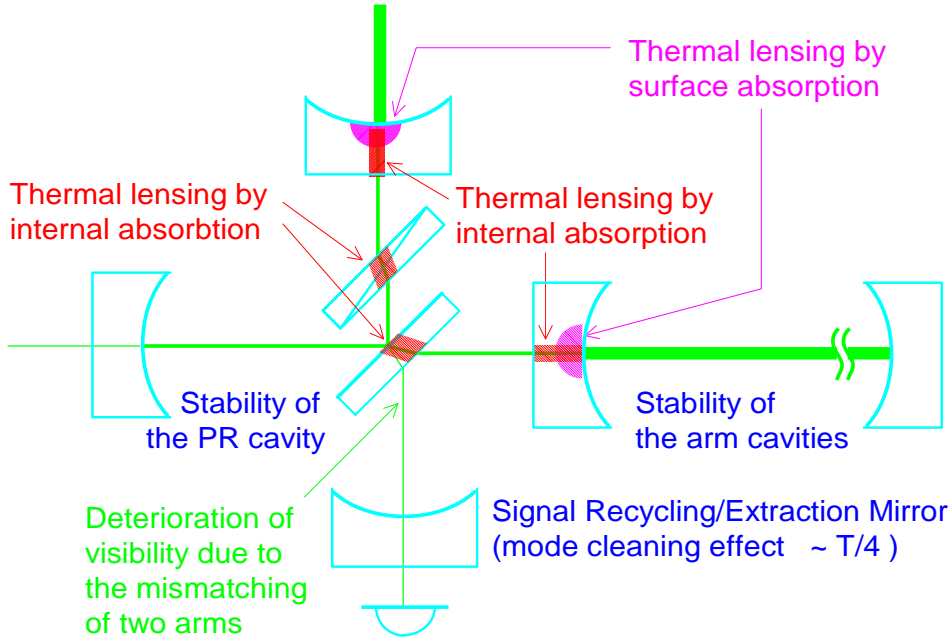


Figure 5.8: The major points which must be taken into account in concerning the sensitivity with thermally induced distortions. The possible places and the causes of thermal distortions are shown, with comments on associated problems. When the coupling mirror of the arm cavity absorbs a fraction of the light inside the cavity, it will cause thermal lensing inside its substrate, which affects the stability of the power recycling cavity. The coupling mirrors and the beamsplitter (and the compensation plate, if exist) will cause thermal lensing due to the absorption inside their substrates. If signal recycling or signal extraction mirror presents, its mode cleaning effect must be taken into account in considering the distortion-induced mismatch of the two beams.

limitation and for a broad-band case the light power inside the power recycling cavity will be the dominant factor. The threshold is

$$\mathcal{F} \approx 64 \times \left[\frac{\mathcal{A}_{BS} + \mathcal{A}_{IC}}{32 \text{ ppm}} \right] \left[\frac{\mathcal{A}_{\text{surf}}}{1 \text{ ppm}} \right]^{-1} \text{ or equivalently} \quad (45)$$

$$\Delta f_{3\text{dB}} \approx 390 \text{ Hz} \times \left[\frac{\ell}{3 \text{ km}} \right]^{-1} \left[\frac{\mathcal{A}_{BS} + \mathcal{A}_{IC}}{32 \text{ ppm}} \right]^{-1} \left[\frac{\mathcal{A}_{\text{surf}}}{1 \text{ ppm}} \right].$$

In practice, both effects contribute to the distortion, especially in cases that are close to this threshold ($\Delta f_{3\text{dB}} \gtrsim 200 \text{ Hz}$ is already close enough to be of significance).

Assuming the light power in the arms is the limiting factor, the sensitivity expected from the above numbers is

$$\begin{aligned}\tilde{h}_{\text{FP}} &\approx 3.9 \cdot 10^{-24}/\sqrt{\text{Hz}} \times \left[\frac{\lambda}{1.064 \mu\text{m}} \right]^{\frac{1}{2}} \left[\frac{\Delta f_{3\text{dB}}}{200 \text{ Hz}} \right]^{\frac{1}{2}} \left[\frac{\mathcal{E}}{4.8 \text{ J}} \right]^{-\frac{1}{2}} \\ &\approx 3.9 \cdot 10^{-24}/\sqrt{\text{Hz}} \times \left[\frac{\Delta f_{3\text{dB}}}{200 \text{ Hz}} \right]^{\frac{1}{2}} \left[\frac{\beta/\kappa}{8.7 \mu\text{m/W}} \right]^{-1} \left[\frac{\mathcal{A}_{\text{surf}}}{1 \text{ ppm}} \right]^{\frac{1}{2}} \left[\frac{\ell}{3 \text{ km}} \right]^{-\frac{1}{2}}\end{aligned}\quad (46)$$

for $\Delta f_{3\text{dB}} \lesssim 200 \text{ Hz}$.

For broad-band response, the sensitivity will be worse than that expected from eq. (46) with additional dependence on the bandwidth $\Delta f_{3\text{dB}}$. This can, however, be improved by using resonant sideband extraction, as is discussed in 5.4.4.

When the power recycling cavity is at the limit of thermal stability, either from the light power in the arm cavities or that in the power recycling cavity, the loss due to the imperfect visibility will be

$$\mathcal{A}_{\text{th}} \simeq \frac{1}{2} \left(\gamma k \delta_{\text{StH}} \right)^2 \simeq 1.25 \cdot 10^{-3} \times \left[\frac{\gamma}{0.1} \right]^2. \quad (47)$$

This should be compared with the loss in the arms

$$\mathcal{A}_{\text{arm}} \simeq \frac{2}{\pi} \mathcal{F} \mathcal{A}_{\text{min}} \approx 4 \cdot 10^{-3} \times \left[\frac{\mathcal{F}}{125} \right] \left[\frac{\mathcal{A}_{\text{min}}}{25 \text{ ppm}} \right]. \quad (48)$$

For broad-band cases with

$$\begin{aligned}\mathcal{F} &\approx 40 \times \left[\frac{\mathcal{A}_{\text{min}}}{25 \text{ ppm}} \right]^{-1} \left[\frac{\gamma}{0.1} \right]^2 \quad \text{or equivalently} \\ \Delta f_{3\text{dB}} &\approx 625 \text{ Hz} \times \left[\frac{\ell}{3 \text{ km}} \right]^{-1} \left[\frac{\mathcal{A}_{\text{min}}}{25 \text{ ppm}} \right] \left[\frac{\gamma}{0.1} \right]^{-2},\end{aligned}\quad (49)$$

the light power injected is wasted rather than stored in the arm cavity.

For relatively narrow-band cases, the illuminating light power required to store the expected energy is given by

$$P_{\text{in}} \simeq \mathcal{E} \frac{\mathcal{A}_{\text{eq}}}{2\ell/c} \simeq 2 \overline{P_C} \frac{\mathcal{A}_{\text{eq}}}{\mathcal{A}_{\text{surf}}} \approx 16 \text{ W} \times \left[\frac{\overline{P_C}}{120 \text{ mW}} \right] \left[\frac{\mathcal{A}_{\text{eq}}}{66 \text{ ppm}} \right] \left[\frac{\mathcal{A}_{\text{surf}}}{1 \text{ ppm}} \right]^{-1} \quad (50)$$

where $\mathcal{A}_{\text{eq}} \simeq 2\mathcal{A}_{\text{min}} + \mathcal{A}_{\text{th}}/(\frac{2}{\pi}\mathcal{F})$ is the equivalent loss per one round trip in the arm cavity. Again, the required power is determined by the critical power and the fraction of the absorption there among the other losses.

5.4.4: Resonant sideband extraction

As is discussed above, the sensitivity of broad-band type response is likely to be limited by the light power inside the power recycling cavity. In addition, thermally induced loss may reduce the efficiency of storing light energy, i.e. more light power is required than that expected in the linear model. To reduce this loss, a mirror at the interfering output will be useful. For a Fabry-Perot based configuration, two possible options are signal recycling and resonant sideband extraction.

To install signal recycling, the arm cavities must have broader linewidths (i.e. lower finesse) than that required from the desired detector bandwidth. On the other hand, it is necessary to have narrower linewidths (higher finesse) for the arm cavities in resonant sideband extraction. Since the arm cavities with low finesse introduced these problems, it is a natural consequence to choose resonant sideband extraction.

The additional mirror, signal extraction mirror, works to broaden the detector bandwidth from that determined by the arm cavity. At the same time, it also has an effect to suppress the loss due to the thermal lensing inside the power recycling cavity. Furthermore, the light power which is transmitted through the optics can be significantly reduced by having high finesse for the arm cavities. This makes the thermal lensing due to the absorption inside the substrate completely negligible.

The surface absorption, however, still remains as a cause of thermal lensing which can limit the sensitivity of an interferometer. When this is the limiting factor, the amount of energy that can be stored in the arms is given by eq. (43), and the expected sensitivity is that described in eq. (46), however without any restriction on the bandwidth. Thus, for broad-band purposes with $\Delta f_{3\text{dB}} \gtrsim 390 \text{ Hz}$, an improvement of

$$\frac{\overline{P_C}}{\overline{P_B}} \frac{\mathcal{A}_{\text{IC}} + \mathcal{A}_{\text{BS}}}{\mathcal{A}_{\text{surf}}} \frac{\pi}{2\mathcal{F}} \approx 1.3 \times \left[\frac{\mathcal{A}_{\text{IC}} + \mathcal{A}_{\text{BS}}}{32 \text{ ppm}} \right] \left[\frac{\mathcal{A}_{\text{surf}}}{1 \text{ ppm}} \right]^{-1} \left[\frac{\Delta f_{3\text{dB}}}{500 \text{ Hz}} \right] \left[\frac{\ell}{3 \text{ km}} \right] \quad (51)$$

in terms of energy (and by the reciprocal of its square root in spectral sensitivity) is expected by using resonant sideband extraction.

The required illuminating light power to store the thermally limited energy is given in eq. (50). The loss due to thermal distortion will be reduced by the mode cleaning effect of the signal extraction cavity. Furthermore, the loss due to thermal distortion will have less significance by increasing the finesse of the

arm cavities. As a consequence, the equivalent loss per one round trip will be dominated by the loss in the arms, i.e. $\mathcal{A}_{\text{eq}} \simeq 2\mathcal{A}_{\text{min}} \approx 50$ ppm. Thus, the light power required to achieve the sensitivity limited by thermal distortion will be less by a factor of ~ 0.76 , or equivalently ~ 12 W of light power is required.

5.4.5: Possible remedies

From the above consideration, we may conclude that thermally induced distortion can limit the sensitivity from that expected in the linear model (20 J in terms of stored energy) by a factor ~ 2 . However, there are still several possibilities remaining to improve the sensitivity a little more than those in the discussion above. Here we briefly review them (see Fig. 5.9).

First of all, it should be emphasized that the limiting factor is thermal lensing at the coupling mirrors of the arm cavities due to the absorption at the surface. This naturally suggests the possible remedies: reducing thermal lensing, reduce the absorption at the coating, and finally store more energy with these limitations.

To reduce thermal lensing, one straightforward way is the use of materials with lower β/κ . This, however, strictly depends on the availability of the material. What is required is not only the big enough bulk but also the polishing techniques and so on. Thermal lensing due to surface absorption may also be reduced by having a (thick enough) coating of high thermal conductivity, which distributes the heat homogeneously and thus reduces the temperature gradient.

Another less obvious way is to control the thermal lensing actively. This can be done, either by pre- or post-figuring using extra light (possibly at wavelengths that are absorbed by the substrates). The former means that the optics are designed with thermal distortion taking into account and extra light is used to preset the distorted condition. The latter means that the extra light is used to cancel the thermal distortion induced by the ‘main’ beam.

The reduction of the absorption at a coating depends on the development of the technology. It is not hopeless, however, since the present value is limited by the real process rather than the properties of the coating material. Theoretically there is room for improvements by a factor of $\sim 10^3$. The absorption inside the substrate is, however, close to the theoretical limit for fused silica and further improvements seem to be difficult as long as it is used for the substrates.

A straightforward way of storing more energy with these limitations is the use of longer cavity length. This does not necessarily mean a longer geometrical

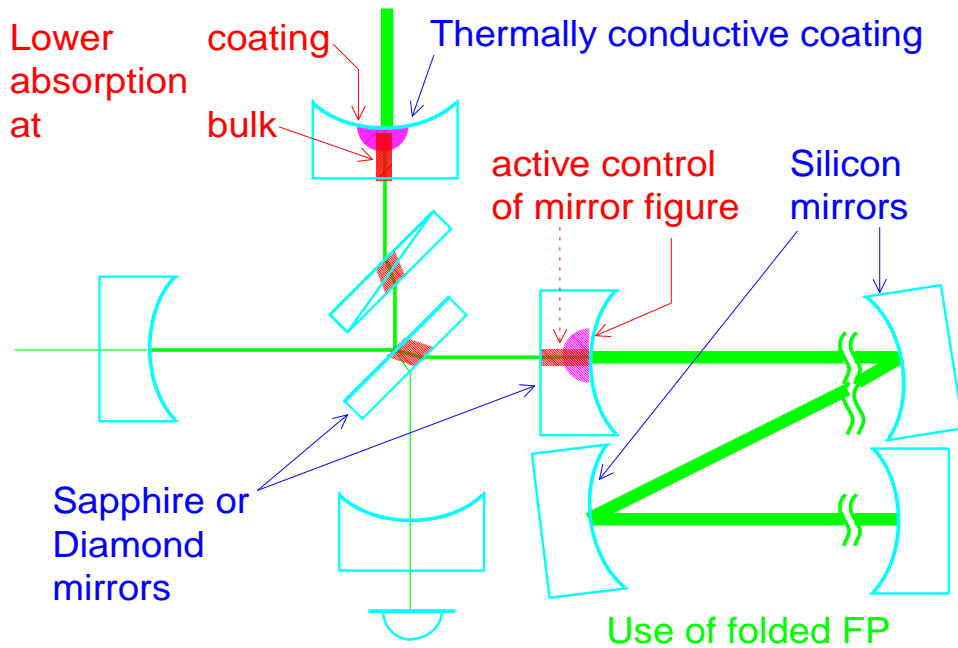


Figure 5.9: Possible schemes to improve the thermal distortion limited sensitivity. One straightforward way is to store more energy under the same limitation, by using folded Fabry-Perot cavities in the arms. It is also possible to reduce the distortion by controlling the thermally induced distortion actively, by using extra light. Reducing the absorption of the optics (at the surface or inside the substrate) decreases the cause of the problem. Improving the thermal conductivity (mirrors made of silicon or sapphire or diamond, or coating by conductive material) decreases the temperature gradient and associated lensing.

arm length, but elongating the cavity length by using *folded Fabry-Perot cavities* would suffice. Then ℓ in eq. (43), (45) and (46) must be considered as the cavity length rather than the geometrical arm length, which must be replaced by $N\ell/2$ where N is the number of transits. Even though N cannot be so large because of the limitation on mirror size, this scheme will be useful when thermal lensing becomes the limiting factor.

It is worth mentioning that resonant sideband extraction has some advantages when one of such remedies is adopted. First of all, the thermal lensing effect due to the absorption inside the substrates can be made negligible. The heat conducting coating can help the thermal lensing due to surface effect, but not the bulk effect. Active control of thermal distortion will be easier if it is due to surface absorption than that due to the absorption inside the substrates.

When the absorption at a coating is reduced, the finesse of arm cavities must be increased, according to eq. (45), so that again the surface effect dominates the limitation. To maintain the detector bandwidth, resonant sideband extraction is useful. Similarly, the linewidth of the arm cavities becomes narrower when a longer arm length or folded cavities are used (assuming the same finesse). Again, the use of resonant sideband extraction can be a solution for this problem.

Summary and prospects

We have discussed in Chapter 3 that the sensitivity of an interferometric gravitational-wave detector can be given by its (noise) bandwidth Δf_{NB} and the amount of energy \mathcal{E} stored in the optical system, as

$$\tilde{h}_0 \gtrsim \sqrt{\frac{2\hbar\lambda}{\pi c} \frac{\Delta f_{\text{NB}}}{\mathcal{E}}},$$

not depending on the configuration employed. According to this sensitivity theorem, the sensitivities expected in future advanced detectors require 20 J of $1\ \mu\text{m}$ light to be stored. The physical background of this theorem can be understood qualitatively, though its full explanation is left to be done.

In Chapter 4, the normalized response of resonant sideband extraction is analyzed in detail assuming the linear model. Its behavior is well understood, which makes it possible to choose appropriate parameters for the desired response. When appropriate parameters are chosen, resonant sideband extraction shows similar or better sensitivities compared with other configurations in many cases including broad-band type responses and low-frequency responses (with accordingly narrow bandwidths). For a narrow-band response at a relatively high peak frequency, however, signal recycling seems to be a better choice, as long as a reasonably long storage time in the arm can be achieved.

It is shown in Chapter 2 that the detection scheme affects the normalized response of an optical configuration, though its dependence is relatively small. The phase of the local oscillator relative to the signal sidebands determines the amount of the signal which can be detected for given upper and lower sidebands. If these two are symmetric as in a broad-band response, the signal is pure phase modulation and the amount of the signal which can be detected is proportional to the amplitude of one of the sidebands.

The sensitivities in the presence of thermally induced distortions of the beam were estimated in Chapter 5. As expected, resonant sideband extraction is advantageous when thermally induced distortion becomes significant, especially for broad-band purposes. It is also shown that the use of the sensitivity theorem described above makes the discussion easier and more straightforward in considering such a non-linear problem.

Assuming fused silica as the substrates of optical components and coating losses of state-of-art level, the amount of energy that can be stored in an optical system is likely to be limited by thermally induced distortions. Unless either lower coating loss or better substrate material will be available, somewhat more complicated optical configurations will be necessary. Again in such cases, an application of resonant sideband extraction with some modification will be useful.

There are, however, things left to be done before the use of resonant sideband extraction. One of the biggest problems is controlling the position of each optical component. As is discussed in Chapter 2, all the relative positions of the optical components must be kept appropriately (within a certain tolerance) in order to make the interferometer work at its best sensitivity. To achieve this with the help of a feedback control system, it is necessary to measure the relative positions of optical components with moderately good sensitivities.

The increasing number of optical components makes this problem more difficult. In this aspect, resonant sideband extraction is less favorable due to its more complicated optical layout. The difficulty arises not only from the information required to control all of the optical components but also from the necessity of separating the pieces of information contained in one of the detected signals. Although this does not degrade the principal advantages, detailed analysis must be carried out in order to solve this problem.

Another problem left to be solved is the detection of the differential phase of the Michelson interferometer, which contains the possible gravitational wave signals. A good efficiency of detecting the signal-induced sidebands is required and there are several schemes proposed for this purpose, as described in Chapter 2. Among them, external modulation can always be used, however at the expense of a relatively complicated optical layout. In order to apply other schemes, e.g. pre-modulation, to resonant sideband extraction, a detailed analysis will be required because of the presence of the signal extraction mirror. In addition, the spurious signal caused by the motion of the signal extraction mirror must also be considered carefully.

Appendices

Appendix A: An additional factor of two

The additional factor of two in eq.(3.2) is a subtle, but a non-negligible point. In short, this arises from two factors, both due to the virtue of a Michelson interferometer. The explanation is, however, not so straightforward and thus the author postponed it. Here it is discussed, though only briefly.

What is being discussed in 3.2.3 and the successive sections is the effect of a gravitational wave on the light propagating perpendicularly to it. One of the arms of an interferometer produces the signal sidebands according to the normalized response $G(\omega)$. If the incident light has power of P_0 , its amplitude is given by $a_0 = \sqrt{P_0/Z_0}$ and the amplitude of a single sideband is expected to be $G(\omega) a_0$.

In fact, however, the light from the source is first divided into the two arms by a beamsplitter. Assuming that the beamsplitter has a power reflectivity and a transmittance of $\rho^2 = \tau^2 = 0.5$, the ‘effective’ input amplitude in each arm is $a_0/\sqrt{2}$ (we ignore the effects of power recycling or signal recycling since they are included in the normalized response). Then the amplitude of a single sideband produced in *one* arm is given by $G(\omega) a_0/\sqrt{2}$.

Assuming a gravitational wave with optimum propagation axis and polarization (as is assumed throughout this paper), the other arm has the same effect but in the opposite polarity. The light from the two arms are recombined at the beamsplitter so that the sidebands from the two arms are added *linearly* at the detecting output. Thus the total sideband produced in an interferometer is $a_s(\omega) = \sqrt{2} G(\omega) a_0$. This is $\sqrt{2}$ more than the amplitude expected from only one arm with the same (total) light power.

This is a natural consequence of the quadrupole nature of a gravitational wave; the correlation between the two arms improves the sensitivity by a factor of $\sqrt{2}$ from that expected for one arm. In terms of scattering cross-sections, the effective cross-section of two orthogonal arms is twice that of one arm with the same length.

Another factor of $\sqrt{2}$ is a little more complicated to explain, although it comes from a similar origin. Just as the ‘input’ (not optical but for gravitational

waves) of a Michelson interferometer has a correlation between the two arms, the ‘output’ from an interferometer has a correlation.

As is discussed in 2.4.1, the signal of the differential phase appears equally in the photo-currents of the two detectors in differential method. By taking the difference of the two photo-currents, the signal is summed *linearly* whereas the noise is added quadratically. This, again, results in an improvement of $\sqrt{2}$ on the signal-to-noise ratio.

The above discussion can explain the reason of the factor in the differential method. Many interferometers are, however, operated at a dark fringe in order to separate the carrier and the signal sidebands (see 2.4.2). If, for example, the internal modulation scheme is used, only one of the interfering outputs is detected by a photo-detector. Then, the correlation between two photo-currents cannot explain the factor of $\sqrt{2}$.

In such a case, the interferometer must be modulated near a dark fringe. Then the photo-current, which determines the amount of shot noise, is also modulated at twice the frequency modulating the interferometer. The amount of signal is proportional to the local oscillator amplitude, whereas the averaged photo-current is half the square of the local oscillator amplitude, as can be seen in eq. (2.35). Thus, a factor of $\sqrt{2}$ can be gained compared with the case of a local oscillator at DC range.

This can also be understood as the correlation of the two slopes of a Michelson interferometer, utilized in the time domain. In this case the phase is ‘measured’ twice, once at each side of a dark fringe, during one modulation period. When the photo-current is demodulated, the signals at these two slopes are summed linearly and the noise contributions are added quadratically. Thus, this gives a similar effect to the previous case.

Although what are discussed here are relatively simple detection schemes only, similar effects are expected in more complicated cases. Thus, the author assumes that the relation given in eq. (3.2) is a universal one which is *always* satisfied in any detection scheme.

Appendix B: Basic assumptions

In order to plot the normalized responses shown in Chapter 3, it is necessary to determine the associated parameters. Many of them are derived from others according to the requirements from the desired response. There are, however, a few parameters which must be given *a priori*, from which others are determined. Among them are the arm length of the interferometer, the minimum loss per reflection upon a mirror, the loss inside the substrate of a transmitting optics, the loss at an anti-reflection surface with non-zero reflectivity, and the loss due to the imperfect visibility.

The assumptions the author used in this paper in plotting the normalized response are as follows:

$$\begin{aligned} \ell &= 3 \text{ km} & \mathcal{A}_{\min} &= 25 \cdot 10^{-6} = 25 \text{ ppm} \\ \mathcal{C}_0 &= 0.99 & \mathcal{A}_{\text{sub}} &= \mathcal{A}_{\text{AR}} = 0. \end{aligned} \tag{B1}$$

As can be seen in the list, the losses inside a substrate and at an anti-reflection surface are assumed to be zero, since they are usually less significant than those due to the imperfect visibility. If one expects a much better visibility (or a much worse substrate or anti-reflection coating), they must also be taken into account. These are, however, merely assumptions and their validity must be discussed.

In some cases, the author discussed ‘ideal’ interferometers. In such cases, the visibility \mathcal{C}_0 is assumed to be unity. The minimum loss per reflection is, however, assumed to be finite since otherwise the amount of energy stored (or the power recycling gain) will be infinity (see eq. (3.5)), which is impractical and meaningless.

Using the above values, a normalized response $|G(\omega)|$ typically has a value of the order of $\sim 10^{13}$ (see eq. (B4)). Such a big number is, however, less convenient to handle. In addition, the upper and lower sidebands must be summed, according to eq. (3.51), before making the plots. Thus, the author used the *normalized sensitivity* as the vertical scale of the figures, whose definition is

$$\begin{aligned} G'(\omega) &= \frac{4}{\omega_0 t_{3\text{km}}} \times \mathbf{G}(\omega) = \frac{4}{\omega_0 t_{3\text{km}}} \times \left[G(\omega) + G^*(-\omega) \right] \\ &\approx 1.1 \cdot 10^{-10} \times \left[\frac{\lambda}{1 \mu\text{m}} \right] \times \left[G(\omega) + G^*(-\omega) \right] \end{aligned} \tag{B2}$$

where $t_{3\text{km}} = 2 \times (3 \text{ km})/c \approx 20 \mu\text{s}$.

As can be seen in eq. (3.51), the sum of the two sidebands' responses reduces to twice the single sideband in broad-band cases.

Using the normalized sensitivity defined above, the noise-equivalent (linear) spectral amplitude of gravitational waves is given by

$$\begin{aligned}\tilde{h}(\omega) &= \frac{1}{2\mathbf{G}(\omega)} \sqrt{\frac{2\hbar\omega_0}{P_0}} = \frac{2}{G'(\omega)} \frac{\sqrt{2\hbar\omega_0/P_0}}{\omega_0 t_{3\text{km}}} \\ &\approx 1.1 \cdot 10^{-20}/\sqrt{\text{Hz}} \times \left[\frac{G'(\omega)}{1} \right]^{-1} \left[\frac{P_0}{10 \text{ W}} \right]^{-\frac{1}{2}} \left[\frac{\lambda}{1 \mu\text{m}} \right]^{\frac{1}{2}}.\end{aligned}\quad (\text{B3})$$

Note that this depends on the power and the wavelength of the light source. The dependence on the arm length is, however, absorbed in the normalized sensitivity so that one can compare the normalized sensitivities of interferometers with different arm lengths.

From the sensitivity theorem, the normalized response at the peak frequency ω_p is limited by

$$\begin{aligned}|G(\omega_p)| &\leq \frac{1}{2} \frac{\omega_0}{4} \sqrt{\frac{2\ell/c}{\Delta f_{\text{NB}}}} \sqrt{\frac{1-\mathcal{A}_{\text{min}}}{\mathcal{A}_{\text{min}}}} \\ &\lesssim 2.1 \cdot 10^{13} \times \left[\frac{\lambda}{1 \mu\text{m}} \right]^{-1} \left[\frac{\Delta f_{\text{NB}}}{100 \text{ Hz}} \right]^{-\frac{1}{2}} \left[\frac{\ell}{3 \text{ km}} \right]^{\frac{1}{2}} \left[\frac{\mathcal{A}_{\text{min}}}{25 \text{ ppm}} \right]^{-\frac{1}{2}}.\end{aligned}\quad (\text{B4})$$

Similarly, the normalized sensitivity is limited by

$$\begin{aligned}|G'(\omega_p)| &\leq \frac{1}{2 t_{3\text{km}}} \sqrt{\frac{2\ell/c}{\Delta f_{\text{NB}}}} \sqrt{\frac{1-\mathcal{A}_{\text{min}}}{\mathcal{A}_{\text{min}}}} \\ &\lesssim 100 \times \sqrt{N_{\text{eff}}} \times \left[\frac{\ell}{3 \text{ km}} \right] \left[\frac{\mathcal{A}_{\text{min}}}{25 \text{ ppm}} \right]^{-\frac{1}{2}} \\ &\lesssim 2.2 \cdot 10^3 \times \left[\frac{\Delta f_{\text{NB}}}{100 \text{ Hz}} \right]^{-\frac{1}{2}} \left[\frac{\ell}{3 \text{ km}} \right]^{-\frac{1}{2}} \left[\frac{\mathcal{A}_{\text{min}}}{25 \text{ ppm}} \right]^{-\frac{1}{2}},\end{aligned}\quad (\text{B5})$$

where $N_{\text{eff}} = (c/2\ell)/\Delta f_{\text{NB}}$ is the ‘effective number of transits’ in the arm. This is the ‘yardstick’ of the values of the normalized sensitivity.

The assumptions eq. (B1) are not necessarily fair all the way. A typical example is the non-linearity discussed in Chapter 5. Even without such effects, however, the losses may depend on other conditions. For example, the minimum loss per reflection may depend on the reflectivity of the mirror, because the higher reflectivity requires more layers of coating which may result in a bigger loss.

Appendices

What is more serious is the loss due to the imperfect visibility; it may depend, for example, on the number of reflections in the multi-reflecting arms. It is expected that there will be more distortion of the wave front with more reflections in the arms, and thus the visibility deteriorates. The situation will be more complicated when Fabry-Perot cavities are used in the arms—the visibility may depend on the finesse of the arm cavities, but in a manner which is not yet understood. Furthermore, the visibility will improve if a signal recycling or extraction mirror is present in front of the photo-detector, as is discussed in 5.2.5.

In plotting the normalized responses, however, all these effects are neglected and a constant visibility is assumed. Thus, the comparison here may not be fair for some of the configurations, such as a dual recycled interferometer with a few bounces in the arms. The better comparison can be made if all the configurations are compared using proper parameters, preferably measured under appropriate conditions.

There is, however, no systematic data on these points, especially for large beams which will be used in a full scale interferometric gravitational-wave detector. This is the reason why the author had to use relatively ‘simple’ assumptions. With the more realistic data, some of the conclusions may alter.

Appendix C: Cancellation of phase delay

The behavior of a three-mirror coupled cavity described in 4.3 can be partially understood as the rotation of the oblique line in Fig. 4.10. An alternative way of understanding the behavior is to investigate the phase shift inside the three-mirror coupled cavity. This allows one to get a qualitative understanding of resonant sideband extraction without referring to Fig. 4.10.

In general, the resonance feature is determined by the phase rather than the amplitude. The reason why a cavity has a finite linewidth is the *phase delay*, i.e. the difference in phase shift between different frequency components, incurred while travelling a finite distance. The reflectivities of the mirrors determine the sharpness of the multiple interference inside the cavity, and thus the tolerable phase shift per round trip to maintain the resonance condition. When light of a particular frequency is resonant with a cavity, that of a different frequency will incur a different phase shift while travelling the cavity length and thus may not be resonant with the cavity.

In our application, what is of interest is the linewidth of the ‘tentative cavity’ used in the brief discussion in 4.1.3. The phase difference $\Delta\phi_a$ between different frequency components (with difference of $\Delta\omega_m$) while travelling the arm cavity length ℓ_a is given by

$$\Delta\phi_a = (2\ell_a/c) \times \Delta\omega_m \quad \text{or} \quad \frac{d\phi_a}{d\omega_m} = \frac{2\ell_a}{c}. \quad (\text{C1})$$

This coefficient is determined by the arm length and is always positive. Since the phase ϕ_a appears negative in $e^{-i\phi_a}$, this means a phase *delay*.

The phase shift during the ‘round trip’ of the tentative cavity is the sum of the phase delay while travelling the arm and the phase shift upon the reflection from the ‘compound mirror’. The latter is determined by the reflectivities of the mirrors and the length of the compound mirror, as

$$\frac{d}{d\omega_m} \angle \rho_{\text{CM}}(\omega_m) = \frac{d\phi_s}{d\omega_m} \times \frac{d}{d\phi_s} \angle \rho_{\text{CM}}(\phi_s) = \frac{2\ell_s}{c} \times \frac{d}{d\phi_s} \angle \rho_{\text{CM}}(\phi_s). \quad (\text{C2})$$

Appendices

Here the second factor is determined only by the reflectivities of the mirrors. If the compound mirror is under-coupled, this factor can be positive, i.e.

$$\begin{aligned}
\frac{d}{d\phi_s} \angle \rho_{\text{CM}}(\phi_s) &= \frac{d}{d\phi_s} \left[\arctan \frac{\frac{1-(\rho_c)^2}{\rho_c} \sin 2\phi_s}{\frac{1+(\rho_s)^2}{\rho_s} - \frac{1+(\rho_c)^2}{\rho_c} \cos 2\phi_s} \right] \\
&= \frac{2 \cdot \frac{1-(\rho_c)^2}{\rho_c} \left[\frac{1+(\rho_s)^2}{\rho_s} \cos 2\phi_s - \frac{1+(\rho_c)^2}{\rho_c} \right]}{\left[\frac{1+(\rho_s)^2}{\rho_s} \right]^2 + \left[\frac{1+(\rho_c)^2}{\rho_c} \right]^2 - 2 \left[\frac{1+(\rho_s)^2}{\rho_s} \right] \left[\frac{1+(\rho_c)^2}{\rho_c} \right] \cos 2\phi_s - \frac{1}{4} \sin^2 2\phi_s} \quad (\text{C3}) \\
&\simeq \frac{2 \cdot \frac{1-(\rho_c)^2}{\rho_c}}{\frac{1+(\rho_s)^2}{\rho_s} - \frac{1+(\rho_c)^2}{\rho_c}} = \frac{2\rho_s[1 - (\rho_c)^2]}{(\rho_c - \rho_s)(1 - \rho_s\rho_c)} > 0 \quad (\text{for } \phi_s \simeq 0),
\end{aligned}$$

within a limited region around the resonance, as can be seen in Fig. 4.8. This has the effect of cancelling the phase delay incurred while travelling the arm cavity.

On the other hand, the first factor in eq. (C2) is determined by the length of the signal extraction cavity. Thus, it is possible to compensate the phase delay incurred while travelling the arm cavity by adjusting this factor. This is, however, possible only within a limited frequency range, and the range in turn gets narrower by elongating the signal extraction cavity. Furthermore, the tolerable phase delay in a round trip also varies since the signal extraction cavity has a frequency-dependent reflectivity.

Provided that the signal extraction cavity linewidth is broad enough, only the cancellation around the resonance is significant. This reduces the sharpness of the curve at the peak of the resonance. Elongating the signal extraction cavity makes the cancellation stronger, and thus the top of the curve becomes a ‘plateau’ rather than a ‘peak’ (see Fig. 4.15). At the same time, however, the linewidth of the signal extraction cavity gets narrower and thus the initial assumption may become invalid.

When the signal extraction cavity linewidth becomes comparable (not necessarily equal) to that of the arm cavity, a resonance peak appears at each end of the signal extraction cavity linewidth (see Fig. 4.15). This is because the signal extraction cavity has higher reflectivities at slightly off-resonant frequencies than the exact resonance frequency, which improves the sensitivity at these frequencies. This is only true within the linewidth of the signal extraction cavity where the phase cancellation applies.

Further elongation makes the linewidth of the signal extraction cavity even narrower, thus the two resonance peaks move even closer to each other. At the same time each peak becomes higher, since the reflectivity becomes higher at the frequency where the amount of phase shift upon the reflection matches to the phase delay along the arm. In this case the phase may actually be ‘over-compensated’ at the carrier frequency and thus a ‘notch’, instead of a peak, will be formed at the carrier frequency (see Fig. 4.15 in 4.3.4).

Due to these complexities, it is not so easy to find the appropriate length of the signal extraction cavity. This effect is, however, relatively small—the increase in the bandwidth measured at the -3 dB point is typically no more than a factor of two. Furthermore, this effect has a significance only with a quite long signal extraction cavity (typically $\gtrsim 1/10$ of the arm cavity length). Thus, the cavity length may be limited by practical conditions rather than suitability.

Appendix D: Derivation of narrow-band parameters

The frequency response of resonant sideband extraction is dominated by the effect of the signal cavity given in eq.(4.15). Since the absolute value of its numerator is constant, the resonance feature is mainly determined by the denominator. Assuming the signal extraction cavity length to be infinitesimal, the denominator can be written using the light frequency ω as

$$D(\omega, \psi) = 1 - \rho_s \rho_c e^{-i\psi} - \rho_c \rho_r e^{-i\omega t_a} + A \rho_s \rho_r e^{-i(\omega t_a + \psi)}. \quad (D1)$$

Here $A = (1 - \mathcal{A}_c)$ and ψ is the amount of detuning of the signal extraction cavity at the carrier frequency measured in radian.

To choose the optimum tuning for a given target frequency ω_z , one need to find ψ_0 which minimizes the absolute value of $D(\psi) = D(\omega_z, \psi)$. Differentiating the square of the absolute value of a complex-value function (by its real argument) gives

$$\begin{aligned} \frac{\partial}{\partial \psi} |D|^2 &= \frac{\partial}{\partial \psi} (DD^*) = D \frac{\partial}{\partial \psi} D^* + D^* \frac{\partial}{\partial \psi} D = 2 \Re \left\{ D \frac{\partial}{\partial \psi} D^* \right\} \\ &= 2 \left(\Re \{ D \} \Re \left\{ \frac{\partial}{\partial \psi} D \right\} + \Im \{ D \} \Im \left\{ \frac{\partial}{\partial \psi} D \right\} \right). \end{aligned} \quad (D2)$$

Using this, the condition required in the present problem can be written as

$$\begin{aligned} 0 &= \Re \{ D \} \Re \left\{ \frac{\partial}{\partial \psi} D \right\} + \Im \{ D \} \Im \left\{ \frac{\partial}{\partial \psi} D \right\} \\ &= \rho_s \left\{ \rho_c \sin \psi_0 - A \rho_r \sin(\omega_z t_a + \psi_0) \right. \\ &\quad \left. + (\rho_c)^2 \rho_r (\sin \omega_z t_a \cos \psi_0 - \cos \omega_z t_a \sin \psi_0) \right. \\ &\quad \left. + A \rho_c (\rho_r)^2 [\sin(\omega_z t_a + \psi_0) \cos \omega_z t_a - \cos(\omega_z t_a + \psi_0) \sin \omega_z t_a] \right\} \\ &= \rho_s \rho_r \sin \psi_0 \times \\ &\quad \left\{ (\rho_c / \rho_r) [1 + A (\rho_r)^2] - [A + (\rho_c)^2] \cos \omega_z t_a - [A - (\rho_c)^2] \sin \omega_z t_a \cot \psi_0 \right\}. \end{aligned} \quad (D3)$$

One of the solutions of this equation is

$$\psi_0 = \arctan \frac{[A - (\rho_c)^2] \sin \omega_z t_a}{(\rho_c / \rho_r) [1 + A (\rho_r)^2] - [A + (\rho_c)^2] \cos \omega_z t_a}, \quad (D4)$$

which is presented in eq.(4.29). Note that this does not depend on ρ_s , the effective reflectivity of the signal extraction mirror.

If the signal extraction cavity has a finite length, its effect must be taken into account. The tuning at the target frequency must be ψ_0 , and detuning at

the carrier frequency must be less than this by the phase shift inside the signal extraction cavity. Subtracting this effect results in

$$\psi_\chi = \psi_0 - \omega_z t_s = \psi_0 - \chi \omega_z t_a, \quad (\text{D5})$$

as is shown in eq. (4.30). Again, remembering the tile pattern and the oblique line shown in Fig. 4.11 is helpful for the understanding of this problem.

Once the amount of detuning is determined as described above, it is convenient to consider the locus of $D(\omega) = D(\omega, \psi_0)$ in the complex plane along the change of ω . At the peak frequency, $|D(\omega)|$ must be the minimum, i.e. the locus must be closest to the origin $(0, 0)$. When detuning is optimized, the target frequency ω_z can be assumed to be close to the peak frequency. This will be true in most of the narrow-band purposes with $Q = \omega_p / (2\pi \Delta f_{3\text{dB}}) \gtrsim 1$.

Under such conditions, $D(\omega)$ can be approximated around $D(\omega_z)$ by taking up to the first order terms of ω ,

$$\begin{aligned} D(\omega) &\simeq D_1(\omega) = D(\omega_z) + (\omega - \omega_z) \frac{\partial}{\partial \omega} D(\omega_z) \\ &= D(\omega_z) - it_a (\omega - \omega_z) \dot{D}(\omega_z) \end{aligned} \quad (\text{D6})$$

where

$$\begin{aligned} \dot{D}(\omega) &= \frac{-1}{it_a} \frac{\partial}{\partial \omega} D(\omega) = -\rho_c \rho_r e^{-i\omega t_a} + A \rho_s \rho_r e^{-i(\omega t_a + \psi)} \\ &= -\rho_r e^{-i\omega t_a} (\rho_c - A \rho_s e^{-i\psi}). \end{aligned} \quad (\text{D7})$$

(This definition of $\dot{D}(\omega)$ is to keep it dimensionless.)

The locus of $D_1(\omega)$ in the complex plane is a straight line which passes through $D(\omega_z)$, and it must be compared with that of $D(\omega)$ to confirm the validity of the approximation. If the locus of $D_1(\omega)$ is close enough to that of $D(\omega)$, ω_p which minimizes $|D_1(\omega)|$ can be considered as being an approximation of ω'_p which minimizes $|D(\omega)|$.

If $|D_1(\omega)|$ has a minimum at ω_p , the vector to point $D_1(\omega_p)$ must be orthogonal to the *locus* of $D_1(\omega)$ (see Fig. D.1), i.e.

$$D_1(\omega_p) \perp \frac{\partial}{\partial \omega} D_1(\omega). \quad (\text{D8})$$

Since the first term of $D_1(\omega)$ in eq. (D6) is just an offset, this means

$$D_1(\omega_p) \perp it_a \dot{D}(\omega_z) \quad \text{or} \quad D_1(\omega_p) \parallel \dot{D}(\omega_z). \quad (\text{D9})$$

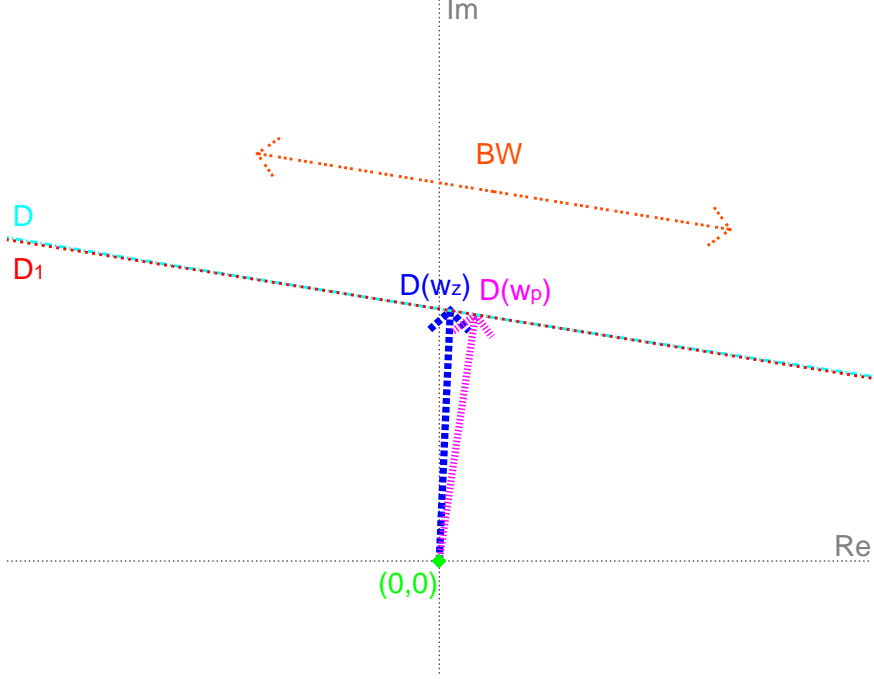


Figure D.1: The loci of $D(\omega)$ and its approximation $D_1(\omega)$ (see eqs. (D1) and (D7) in the text for their definitions), although they are scarcely distinguishable. The point of the peak frequency $D(\omega_p)$ is not necessarily the same as that of the target frequency $D(\omega_z)$. The 3 dB bandwidth is obtained as the range within which the length of the vector remains $\leq \sqrt{2}$ times of the minimum at $D(\omega_p)$.

This can be rewritten as

$$\begin{aligned} 0 &= \Im \left\{ \frac{D_1(\omega_p)}{\dot{D}(\omega_z)} \right\} = \Im \left\{ \frac{D(\omega_z) - it_a(\omega_p - \omega_z) \dot{D}(\omega_z)}{\dot{D}(\omega_z)} \right\} \\ &= \Im \left\{ \frac{D(\omega_z)}{\dot{D}(\omega_z)} \right\} - t_a(\omega_p - \omega_z). \end{aligned} \quad (\text{D10})$$

Thus, the first relation in eq. (4.31) can be obtained

$$\omega_p = \omega_z + \frac{1}{t_a} \Im \{ E(\omega_z, \psi_0) \} \quad (\text{D11})$$

where

$$\begin{aligned} E(\omega, \psi) &= \frac{D(\omega, \psi)}{\dot{D}(\omega, \psi)} = \frac{1 - \rho_s \rho_c e^{-i\psi} - \rho_c \rho_r e^{-i\omega t_a} + A \rho_s \rho_r e^{-i(\omega t_a + \psi)}}{-\rho_c \rho_r e^{-i\omega t_a} + A \rho_s \rho_r e^{-i(\omega t_a + \psi)}} \\ &= 1 - \frac{1 - \rho_s \rho_c e^{-i\psi}}{\rho_r e^{-i\omega t_a} (\rho_c - A \rho_s \rho_r e^{-i\psi})}, \end{aligned} \quad (\text{D12})$$

using the definitions in eqs. (D1) and (D7). This is what is shown in eq. (4.32).

Appendix D: Derivation of narrow-band parameters

The 3 dB bandwidth is measured at the two points whose distances from the origin are 3 dB $\approx \sqrt{2}$ times longer than that of the minimum (see Fig. D.1). Since the locus of $D_1(\omega)$ is orthogonal to $D(\omega_p)$, the above condition can be written as

$$(1 \pm i) \times D_1(\omega_p) = D_1(\omega_{\pm}) = D_1(\omega_p) - it_a(\omega_{\pm} - \omega_p) \dot{D}(\omega_z). \quad (\text{D13})$$

Thus we obtain

$$-it_a \Delta\omega = -it_a(\omega_+ - \omega_-) = 2iD_1(\omega_p)/\dot{D}(\omega_z). \quad (\text{D14})$$

In the above equation, $D_1(\omega_p)$ can be written, from the relation in eq. (D9), as

$$\begin{aligned} D_1(\omega_p) &= D(\omega_z) - it_a(\omega_p - \omega_z) \dot{D}(\omega_z) \\ &= \left\{ \frac{D(\omega_z)}{\dot{D}(\omega_z)} - it_a(\omega_p - \omega_z) \right\} \dot{D}(\omega_z) = \Re \left\{ \frac{D(\omega_z)}{\dot{D}(\omega_z)} \right\} \dot{D}(\omega_z). \end{aligned} \quad (\text{D15})$$

Using this result, the second relation in eq. (4.31)

$$\Delta f_{3\text{dB}} = \frac{|\Delta\omega|}{2\pi} = \frac{1}{\pi t_a} \left| \Re \{ E(\omega_z, \psi_0) \} \right|. \quad (\text{D16})$$

is obtained.

Appendix E: Et cetera

If the interaction between a gravitational wave and light can be thought of as being a ‘scattering’, there are a few things worth commenting on. The discussions here are actually not of practical use. They might, however, be interesting issues when one compares interferometric gravitational-wave detectors with other type of detectors.

In a broad-band detector, a gravitational wave produces the upper and lower sidebands. The upper sideband has more energy than the carrier which is received (or absorbed) from the gravitational wave, whereas the lower sideband has less energy, which is lost as (or emitted into) the gravitational wave. The difference of energy from the carrier is the same amount in both cases, thus no energy is lost from the gravitational wave according to conservation of energy. Thus, it can be considered as an *elastic scattering*.

On the other hand, in the single sideband case a part of the gravitational wave energy may be absorbed by a photon, causing an *inelastic scattering* by the detector. Furthermore, if the detector is optimized for the lower sideband, a part of energy is *lost* from a photon. This should be thought of as being a gravitational wave ‘emitted’ by a photon. This can be understood in close analogies to stimulated emission and absorption of a photon by an atom. One question (which the author cannot answer) is: *What will happen to conservation of spin in this case?* Does a photon change its spin by absorbing or emitting a graviton (which has a spin of two)? Does this cause any effect on the detection process?

From eq. (3.58), the effective *scattering cross-section* of the detector can be estimated as $\sim 1.6 \cdot 10^{-25} \text{ m}^2$ —which is an extremely small value considering the physical size of a detector. This is, as mentioned in 3.5.3, due to the weakness of gravitational interactions. Another disadvantage of an optical detector is the smallness (or tininess) of the energy per photon. The product of the last two factors in eq. (3.58)

$$\alpha_G \times \left(\frac{2\pi\hbar c/\lambda}{M_p c^2} \right)^2 \approx 1.0 \cdot 10^{-56} \times \left[\frac{\lambda}{1 \mu\text{m}} \right]^{-2} \quad (\text{E1})$$

indicates the weakness of a typical probability of gravitational interaction between photons.

This disadvantage of using photons as interacting media is partially compensated by the fact that photons are *bosons*. This makes it possible to store

in a single state as many photons as needed—in our case, 10^{20} photons are required to achieve the desired sensitivity. An interferometer using particles with heavier (rest) masses, e.g. an electron interferometer or even a neutron interferometer, may be advantageous in terms of the probability of interaction, or the scattering cross-section. These particles are, however, *fermions* and each particle must occupy a different state, which makes it difficult to predict the behavior of the detector.

Another advantage of using light as the interacting media is the readiness to reach the quantum-noise limited sensitivity. For ‘optical’ frequencies, $\hbar\omega_0 \gg k_B\mathcal{T}$, i.e. a photon has much more energy than the typical thermal energy. This makes it possible to forget about ‘thermally excited’ photons completely, which is a big difference from resonant mass detectors. Although this type of detector has a much bigger scattering cross-section to gravitational waves than an optical detector, its sensitivity is likely to be limited by thermally excited motions which cannot be distinguished from that excited by gravitational waves.

In the above paragraph, the term ‘thermally excited’ was put in quotation marks because it is used for the narrowest meaning which is a less common definition. What the author meant here is the photons excited purely by thermal origin—in other words, the photons due to the Planck radiation. A more commonly used meaning of ‘thermal noise’ in an optical detector is the conversion of a photon from one state to another due to the coupling with the thermally excited motion of the optical components. Since we have so many photons in a single state, any small perturbation may convert a few (or more) of them into other states.

Thermal noise in the latter sense is another serious problem in interferometric gravitational-wave detectors, though it is rarely mentioned in this paper. This can reduce the number of photons which are useful as interacting media, and, moreover, this may cause spurious signals by converting a photon into the mode we detect. Then it may not be possible to distinguish it from that due to gravitational waves, just as in the resonant mass detector case.

The coupling of thermally excited motion of a mirror with the photons are investigated in Refs. [50] and [51]. According to these results, the coupling of relatively high order vibration modes with the fundamental optical mode (which we use as the interacting medium) seems to be a serious problem. Fortunately, this coupling can be suppressed by using the mode cleaning effect of cavities described in 5.2.5. To achieve a good suppression factor, however, one needs

to use relatively high finesse cavities in order to suppress the coupling with the most sensitive optics, i.e. the main test mass mirrors.

In this respect, resonant sideband extraction is advantageous not only because it utilizes high finesse cavities in the arms but also because the signal extraction cavity effectively has a *mode selective reflectivity*¹⁾. If the relevant parameters are chosen properly, it is resonant practically only with the (optical) fundamental mode. Then it suppresses all the optical higher order modes which are caused by the surface irregularity (static scattering) or by thermally excited motions of the mirrors (*dynamic scattering*).

In a quantum mechanical view, the probability of transition from one state to another depends on the mode density around the final state, as can be seen in eq. (3.59). Cavities, especially multi-mirror coupled cavities like the signal cavity of resonant sideband extraction, in general have an effect of changing the mode density inside it. Thus, it is possible to reduce the mode density around the relevant higher order modes, i.e. the modes that are expected to couple strongly with the spatial fluctuations.

The relevant parameters must, however, be chosen carefully to utilize this effect since this can work also in the other way, i.e. it can enhance one of the unwanted higher order modes. The resonance feature of scattered light is determined by tuning defined in eq. (5.22). In the ‘dynamic’ case, both mode number and the frequency of the vibration is important. To avoid the spurious signal inside the observation frequency range, the tuning at that frequency and in the mode expected to be produced must be off-resonant.

¹⁾ This concept was initially proposed by Meers [49] for the use in signal recycling scheme.

Appendix F: Frequency response and the local oscillator

As is discussed in 2.3.4, the frequency response of a narrow-band configuration depends on the phase of the local oscillator that detects the signal sidebands. In Chapters 3 and 4, the author assumed the detection of the ‘phase modulation component’, since it is the best scheme for broad-band responses, it is the scheme used in present prototypes, and altering it to other schemes does not make big differences. In order to show the last point, we consider here the effect of the phase of the local oscillator taking dual recycling and resonant sideband extraction as examples (they are the only ‘realistic’ narrow-band configurations).

For a broad-band response, the signal is a pure phase modulation and the phase of the local oscillator affects the sensitivity sinusoidally, as shown in eq. (2.33). On the other hand, if the signal is a pure single sideband, the local oscillator phase does not affect the normalized response at all. Thus, it is expected that the effect of the local oscillator phase will be most significant when the upper and lower sidebands have different but comparable (real) amplitudes. In this appendix, we take such ‘extreme’ cases as examples in order to show the effects clearly. Although there will be similar effects in other cases, they may not be as significant as those shown here.

Fig. F.1 shows the (real) amplitude of the signal sideband in the interferometer using dual recycling with detuning, before summing the upper and lower ones. Zero-frequency corresponds to the carrier frequency, and the positive and the negative frequencies correspond to the upper and lower sidebands, respectively. Even if the signal recycling cavity is tuned to the upper sideband, the lower sideband can have a smaller but non-negligible amplitude, as can be seen in the figure. This affects the normalized response when the two sidebands are summed, as is illustrated in Fig. F.2. The peak frequency as well as the bandwidth varies with the local oscillator phase, but only by a small fraction. There is stronger dependence at out-of-resonance frequencies.

It can be seen that the sensitivities at lower frequencies and higher frequencies are not compatible, i.e. choosing the local oscillator phase which maximizes one of them makes the sensitivity in the other region worse. This can be understood by considering that the phase of the upper sideband varies by π radian with the frequency around the resonance, whereas that of the lower sideband does not change so much. There may be a local oscillator phase which gives the

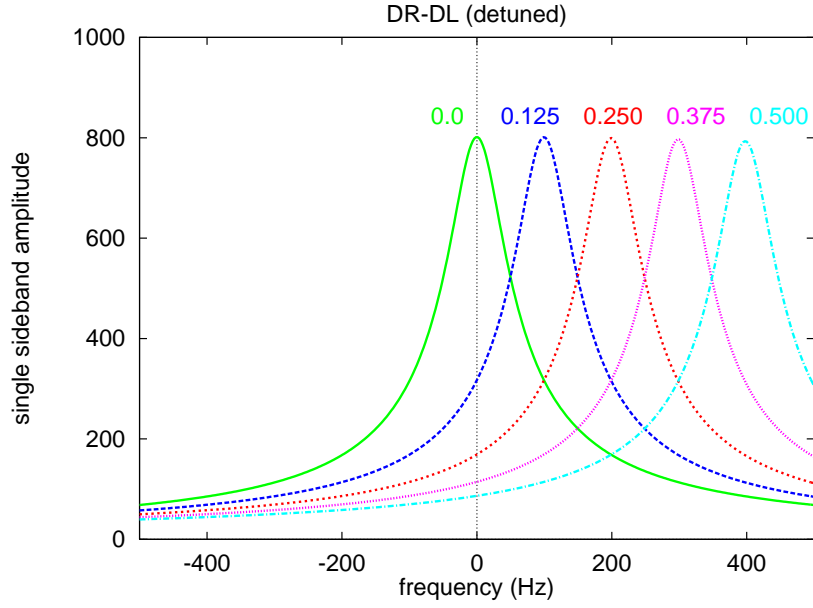


Figure F.1: The amplitude of a single sideband in a dual recycled interferometer with various amounts of detuning (keys indicate them measured in radian). The shape of the response does not change; only the peak frequency moves with detuning. The horizontal axis represents the frequency offset from the carrier; the positive and negative ones corresponds to the upper and lower sidebands, respectively. The parameters assumed are $N=20$ and $\rho_{\text{SR}}^2 \approx 0.95$.

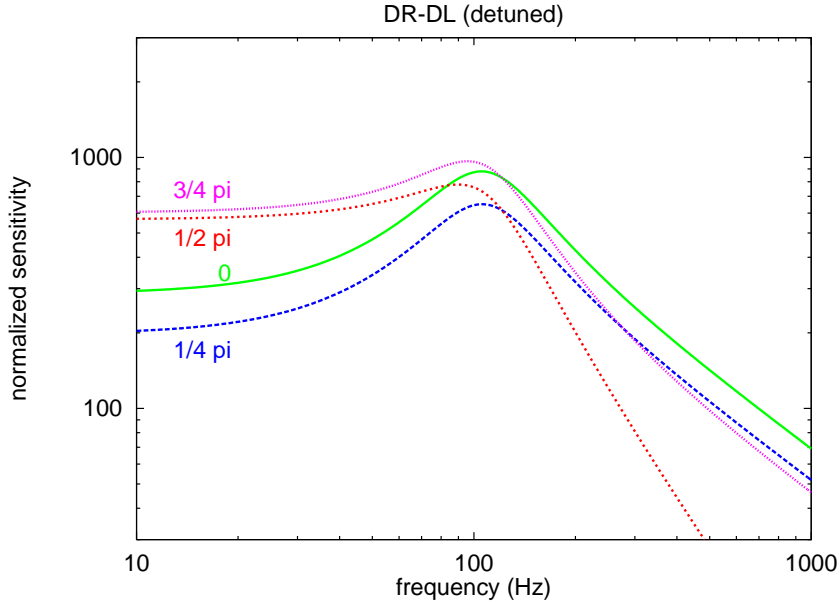


Figure F.2: The normalized responses of a dual recycled interferometer with different local oscillator phases. Keys indicate the local oscillator phase; 0 corresponds to detecting the phase modulation component and $\frac{\pi}{2}$ corresponds to detecting the amplitude modulation component of the signal carried by the upper and lower sidebands. The detuning of the signal recycling cavity is 0.125 radian and other parameters are the same as Fig. F.1.

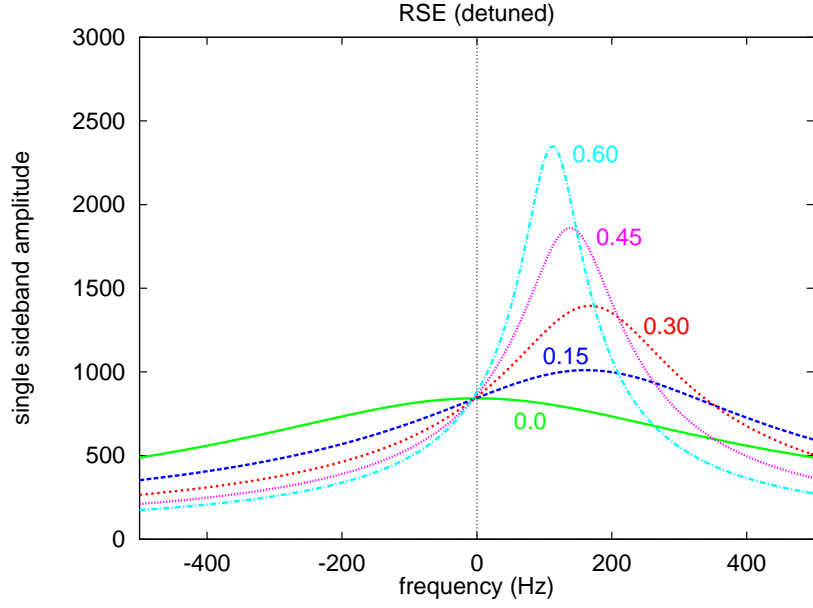


Figure F.3: The amplitude of a single sideband in an interferometer with resonant sideband extraction, with various amounts of detuning (keys indicate them measured in radian). The peak frequency as well as the bandwidth varies with detuning. The parameters assumed are $\mathcal{F}_{\text{arm}} \approx 1240$, and $\rho_{\text{SX}}^2 \approx 0.8$.

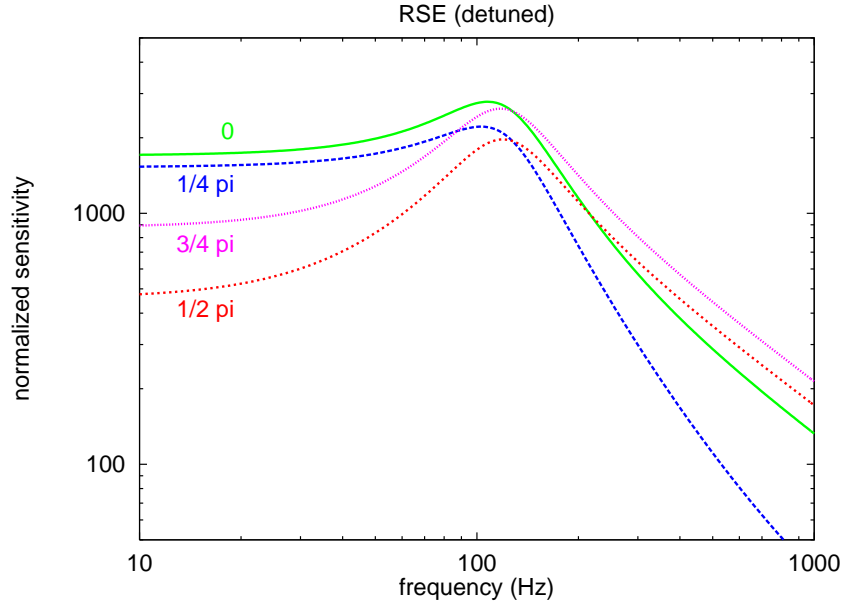


Figure F.4: The normalized responses of an interferometer with resonant sideband extraction, with different local oscillator phases. Keys indicate the local oscillator phase as explained in Fig. F.2. The detuning of the signal extraction mirror is 0.6 radian and other parameters are the same as Fig. F.3.

Appendices

best compromise for all the frequencies, but it will depend on the parameters used as well as the observation purpose.

In the application of a narrow-band response, however, only the sensitivity around the resonance peak will probably be of interest. Furthermore, in such cases the bandwidth of the signal recycling cavity will be narrower than in the example shown here, in order to achieve the higher peak sensitivity. Then the lower sideband will have smaller amplitude and thus weaker dependence on the local oscillator phase is expected.

In a similar way to the dual recycling case, the normalized response of an interferometer using resonant sideband extraction with detuning depends on the local oscillator phase. The amplitude of the single sideband in this case is shown in Fig. F.3 and the normalized response with various local oscillator phases are illustrated in Fig. F.4, each corresponding to Fig. F.1 and Fig. F.2 in the previous case.

The qualitative behavior is quite similar to the previous case, although the local oscillator that optimizes, for example, low frequencies is different. Again, there may be best local oscillator phase, but it will depend on the parameters as well as the observation purpose.

One must, however, be careful that all the present experiments and most of the analyses are carried out assuming the phase modulation component detection. If one alters the detection scheme, some of the well-known behavior may be also altered. Although a detailed analysis is left to be done, what the author can imagine as an example is the effects from the fluctuation of the light source, either in the frequency or the amplitude.

The present knowledge (see, for example, Ref. [25]) is derived assuming that the phase modulation component of a broad-band response is detected. In a narrow-band interferometer, especially when the detection scheme is altered, the response of the interferometer to the fluctuation of the light source may change from that in the previous experience.

Appendix G: Gaussian beam

G.1: Paraxial equation

The spatial distribution of the amplitude of electro-magnetic waves, often called *modes*, can be obtained as solutions of the paraxial wave equation²⁾

$$\left[\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) - i2k \frac{\partial}{\partial z} \right] \Psi(x, y, z) = 0, \quad (\text{G1})$$

which corresponds to an approximation of the wave equation

$$\left[\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] \mathbf{E}(x, y, z, t) = 0 \quad (\text{G2})$$

with the assumptions of 1) single polarization, 2) the wave propagating along the z -axis, and 3) the intensity is concentrated near the propagation axis (in this case the z -axis). Eq. (1) can be obtained by inserting

$$\mathbf{E}(x, y, z, t) = \mathbf{u} \Re \left\{ a \Psi(x, y, z) \exp\{i(\omega t - kz)\} \right\} \quad (a \in \mathbf{C}) \quad (\text{G3})$$

into eq. (2) and by ignoring the second derivative of Ψ with respect to z (i.e. $|\frac{\partial^2}{\partial z^2} \Psi| \ll |2k \frac{\partial}{\partial z} \Psi|$). Here \mathbf{u} is the unit vector to represent the polarization and a is a constant factor that represents the magnitude of the amplitude.

It should be noted in eq. (3) that the definition of $\Psi(x, y, z)$ does not contain the waving factor $\exp\{i(\omega t - kz)\}$. Thus $\Psi(x_p, y_p, z_p)$ must be understood as the complex amplitude of the oscillation at the point (x_p, y_p, z_p) , not the oscillation itself. The whole $\Psi(x, y, z)$ can be considered as the spatial distribution of the complex amplitude of the wave propagating along the z -axis.

The intensity, the spatial distribution of the power, is obtained by squaring the absolute value of the amplitude, i.e. $|a \Psi|^2$. The total power carried by the electro-magnetic wave is obtained by integrating the intensity over the surface perpendicular to its propagation axis (i.e. $z=\text{const.}$),

$$P = \frac{|a|^2}{Z_0} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy |\Psi(x, y, z)|^2 \quad (\text{G4})$$

$$\text{where} \quad Z_0 = \sqrt{\mu_0/\varepsilon_0} = 1/(c\varepsilon_0) \approx 376.7 \Omega.$$

²⁾ The following analysis is based on [57, 59], with some notations replaced.

Appendices

Here the proportional constant Z_0 is called the *characteristic impedance of the vacuum*³⁾. The fact that both electric and magnetic fields carry the same amount of power is also taken into account. It should be noticed that the total power does not depend on z from the conservation law.

Since the constant factor a in eq. (4) represents the magnitude of the amplitude, it will be convenient if Ψ is defined so that it will be dimensionless after the spatial integration. Under this condition, the dimensions of a and Ψ will be (V) and (1/m), respectively. It will especially be useful if Ψ is normalized so that the integration will be unity. We will return to this point later.

G.2: Fundamental Gaussian mode

In rectangular coordinates, Ψ can be separated into a product of the solutions of the x - and y -axes components, i.e.

$$\Psi(x, y, z) = \Psi_x(x, z) \times \Psi_y(y, z), \quad (\text{G5})$$

where $\Psi_x(x, z)$ must satisfy a differential equation

$$\left[\frac{\partial^2}{\partial x^2} - i2k \frac{\partial}{\partial z} \right] \Psi_x(x, z) = 0, \quad (\text{G6})$$

and so must $\Psi_y(y, z)$, replacing x by y .

One of the simplest solution of eq. (6) can be obtained by assuming a solution with the form of

$$\Psi_x(x, z) = A(z) \exp\left\{ -\frac{ik}{2} \frac{x^2}{B(z)} \right\} \quad (A(z), B(z) \in \mathbf{C}). \quad (\text{G7})$$

Inserting this into eq. (6) requires two functions $A(z)$ and $B(z)$ to satisfy

$$\frac{dB(z)}{dz} = 1 \quad \text{and} \quad \frac{dA(z)}{dz} = -\frac{A(z)}{2B(z)}, \quad (\text{G8})$$

which can be solved as

$$B(z) = z + B_x \quad \text{and} \quad A(z) = \sqrt{A_x/B(z)} = \sqrt{A_x/(z + B_x)}. \quad (\text{G9})$$

³⁾ This can be obtained by considering that $\frac{1}{2}\epsilon_0|E|^2$ represents the electric energy density and its two-dimensional integration means $d\mathcal{E}/dz$. Multiplying this by $dz/dt=c$ results in $P=d\mathcal{E}/dt$, the power. If $a\Psi(x, y, z)$ in eq. (4) represents the amplitude of the magnetic field oscillation, then $1/Z_0$ will be replaced by Z_0 .

Then the solution of eq. (6) is written as

$$\Psi_x(x, z) = \sqrt{\frac{A_x}{z + B_x}} \exp\left\{-\frac{ik}{2} \frac{x^2}{z + B_x}\right\}. \quad (\text{G10})$$

Similarly, a solution of $\Psi_y(y, z)$ can be obtained by replacing x in eq. (10) by y .

By combining the two solutions for the x - and y -axes, we obtain one of the solutions of eq. (1) as

$$\Psi(x, y, z) = \sqrt{\frac{A_x A_y}{(z + B_x)(z + B_y)}} \exp\left\{-\frac{ik}{2} \left(\frac{x^2}{z + B_x} + \frac{y^2}{z + B_y}\right)\right\}. \quad (\text{G11})$$

Although B_x and B_y , A_x and A_y are in general independent of each other, we assume for a while that they are identical (this corresponds to cylindrical symmetry), i.e.

$$\begin{aligned} A_x &= A_y = iA_0 \\ B_x &= B_y = -z_w + iz_R \quad (z_w, z_R \in \mathbf{R}). \end{aligned} \quad (\text{G12})$$

(Here ‘ i ’ in the upper equation is introduced merely to make the following expressions simpler.) Then eq. (10) can be rewritten as

$$\begin{aligned} \Psi(x, y, z) &= \frac{iA_0}{z - z_w + iz_R} \exp\left\{-\frac{ik}{2} \frac{x^2 + y^2}{z - z_w + iz_R}\right\} \\ &= \frac{A_0 \exp\left\{i \arctan \frac{z - z_w}{z_R}\right\}}{\sqrt{(z - z_w)^2 + (z_R)^2}} \exp\left\{-\frac{x^2 + y^2}{2} \times \frac{ik(z - z_w) + kz_R}{(z - z_w)^2 + (z_R)^2}\right\} \\ &= \frac{A_0}{z_R} \frac{w_w}{w(\Delta z)} \exp\left\{-\frac{x^2 + y^2}{w^2(\Delta z)}\right\} \exp\left\{i\psi(\Delta z) - ik \frac{x^2 + y^2}{2\mathcal{R}(\Delta z)}\right\} \end{aligned} \quad (\text{G13})$$

where

$$\begin{cases} w(z) = w_w \sqrt{1 + (z/z_R)^2} & w_w = \sqrt{2z_R/k} = \sqrt{z_R \lambda / \pi} \\ \mathcal{R}(z) = z [1 + (z_R/z)^2] & \Delta z = z - z_w \\ \psi(z) = \arctan(z/z_R) & \left[e^{i\psi(z)} = i \sqrt{\frac{z - iz_R}{z + iz_R}} = i \frac{z - iz_R}{|z + iz_R|} \right]. \end{cases} \quad (\text{G14})$$

This represents the *fundamental Gaussian mode*.

The last expression in eq. (13) is helpful in understanding the physical meanings of the parameters: $w(\Delta z)$ is the beam radius at z , measured as the distance from the center ($x=y=0$) to the point where the amplitude is $1/e$ of that on the axis, or equivalently to the point where the intensity is $1/e^2$. The beam radius has a minimum value w_w at $z=z_w$, the *beam waist* point.

Appendices

The beam radius gets larger as the distance from the waist point increases. Its divergence $\vartheta(z-z_w)$ at z (the half angle, from center to $1/e$ amplitude point) is given by

$$\vartheta(z) = \frac{dw(z)}{dz} = \frac{u_w/z_R}{\sqrt{1+(z_R/z)^2}} \xrightarrow{z \rightarrow \infty} \vartheta_\infty = \frac{u_w}{z_R} = \sqrt{\frac{\lambda}{\pi z_R}}. \quad (\text{G15})$$

This can be understood from the Uncertainty Principle. The beam has uncertainties⁴⁾ Δx of the position in the x -axis determined by the beam size at the waist, and Δp_x of the momentum in that axis which determines the divergence. The product of the two uncertainties has a minimum value of $\hbar/2$:

$$\Delta x \Delta p_x = \left(\frac{u_w}{2}\right) \left(\frac{\vartheta_\infty}{2} \frac{2\pi\hbar}{\lambda}\right) = \sqrt{\frac{z_R\lambda}{\pi}} \sqrt{\frac{\lambda}{\pi z_R}} \frac{\pi\hbar}{2\lambda} = \frac{\hbar}{2}, \quad (\text{G16})$$

in accordance with the Heisenberg's Uncertainty Principle.

As the beam diverges, the amplitude at the center correspondingly decreases, i.e. proportionally to $u_w/w(\Delta z)$, from the value at the waist A_0/z_R . The amplitude at $r = \sqrt{x^2 + y^2}$ away from the axis is smaller by a Gaussian factor $\exp\{-\frac{r^2}{w^2(\Delta z)}\}$ with a standard deviation of $w(\Delta z)/\sqrt{2}$.

The wavefront, which is defined as the surface with a constant phase, has a curvature radius of $\mathcal{R}(\Delta z)$. The curvature radius is defined to be positive for diverging beams and negative for converging ones. The absolute value of the curvature radius $|\mathcal{R}|$ is the minimum of $2z_R$ at $|\Delta z| = z_R$; it becomes infinity, i.e. a plane wave, at $z=z_w$, and $\mathcal{R}(z) \simeq z$ when $|z| \gg z_R$. There is also a Guoy phase shift $\psi(\Delta z)$ due to the divergence.

As can be seen in eqs. (14) and (15), z_R is the most useful parameter from which all the others can be derived. This, the so-called *Rayleigh range*, is measured as the distance along the propagation axis from the waist to the point where the beam radius is $\sqrt{2}$ times of that at the waist. The region where $|\Delta z| \ll z_R$ is called the *near field*, in which the beam radius is approximately equal to that at the waist, whereas in the *far field* where $|\Delta z| \gg z_R$ the divergence $\vartheta(\Delta z)$ is almost a constant and the curvature radius $\mathcal{R}(\Delta z)$ is roughly equal to the distance Δz from the waist. The Guoy phase shift $\psi(\Delta z)$ varies rapidly with Δz in the near field, but slowly in the far field.

If the conditions in eq. (12) are not satisfied, the x - and y -axes have different waist sizes at different points. This results in an elliptic beam shape whose

⁴⁾ Note that the uncertainty is given by a radius of $u_w/2$, not $u_w/\sqrt{2}$. This is because what represents the position of a photon is $|\Psi|^2$, not $|\Psi|$.

eccentricity varies with the propagation. This may, however, be better to be considered as a superposition of higher order modes which are described below.

G.3: Hermite-Gaussian mode

From the previous result eq. (10), we can try a little modification in order to find more complicated solutions of eq. (6). For simplicity, we shift the origin of the z -coordinate to be z_w , then each Δz can be replaced by z . Keeping the exponential term to be the same as before, the new trial can be made with a form of

$$\Psi_x(x, z) = A(z) B\left(\frac{\sqrt{2}x}{w(z)}\right) \exp\left\{-\frac{ik}{2} \frac{x^2}{z+iz_R}\right\}, \quad (\text{G17})$$

where $A(z)$ and $B(\xi)$ are different from those in the previous result. The new additional factor $B(\xi)$ is a function of a *combination* of x and z , from the real part of the exponent of eq. (10).

By inserting eq. (17) into eq. (6), we obtain a differential equation which $B(\xi)$ must satisfy,

$$0 = \ddot{B}(\xi) - 2\xi\dot{B}(\xi) + \left\{-1 + \frac{1}{iz_R} \left[z + 2(z^2 + z_R^2) \frac{\dot{A}(z)}{A(z)}\right]\right\} B(\xi) \quad (\text{G18})$$

where $\xi = \sqrt{2}x/w(z) = x\sqrt{kz_R/(z^2 + z_R^2)}$.

If the inside of the big curly brackets ($\{\dots\}$) is equal to $2m$ where m is a natural number (or zero), then this reduces to

$$\ddot{H}_m(\xi) - 2\xi\dot{H}_m(\xi) + 2mH_m(\xi) = 0 \quad (m \in \mathbf{N}_0), \quad (\text{G19})$$

whose solutions, known as the Hermite polynomials, are given by

$$H_m(\xi) = (-1)^m e^{\xi^2} \frac{d^m}{d\xi^m} e^{-\xi^2} \quad \text{or} \quad \begin{aligned} H_0(\xi) &= 1 \\ H_1(\xi) &= 2\xi \\ H_2(\xi) &= 4\xi^2 - 2 \\ H_3(\xi) &= 8\xi^3 - 12\xi \\ H_4(\xi) &= 16\xi^4 - 48\xi^2 + 12 \\ H_5(\xi) &= 32\xi^5 - 160\xi^3 + 120\xi \\ &\vdots \end{aligned} \quad (\text{G20})$$

ignoring the arbitrary constant factor.

For the Hermite polynomials to be the solution of eq. (6), the condition mentioned above

$$2m = -1 + \frac{1}{iz_R} \left[z + 2(z^2 + z_R^2) \frac{\dot{A}(z)}{A(z)}\right] \quad (\text{G21})$$

Appendices

must be satisfied. This results in a form of $A(z)$ as

$$A(z) = \sqrt{A_x} \frac{\exp\left\{i \frac{2m+1}{2} \arctan \frac{z}{z_R}\right\}}{\sqrt[4]{z^2 + z_R^2}}. \quad (\text{G22})$$

In the end, the solution of eq. (6) can be written as

$$\begin{aligned} \Psi_x(x, z) &= \sqrt{A_x} \frac{\exp\left\{i \frac{2m+1}{2} \arctan \frac{z}{z_R}\right\}}{\sqrt[4]{z^2 + z_R^2}} H_m\left(x \sqrt{\frac{kz_R}{z^2 + z_R^2}}\right) \exp\left\{-\frac{ik}{2} \frac{x^2}{z + iz_R}\right\} \\ &= \sqrt{\frac{A_x}{z_R} \frac{u_w}{w(z)}} H_m\left(\frac{\sqrt{2}x}{w(z)}\right) \exp\left\{-\frac{x^2}{w^2(z)}\right\} \exp\left\{i\left[\frac{1}{2} + m\right]\psi(z) - ik \frac{x^2}{2\mathcal{R}(z)}\right\}, \end{aligned} \quad (\text{G23})$$

using the definitions in eq. (14). Note that choosing $m=0$ coincides with the simplest solution in eq. (10) except for the constant factor.

One may notice the similarity of this to the solution of the harmonic oscillator, whose normalized eigen functions are given by

$$u_m(x/\sigma) = \frac{H_m(x/\sigma)}{\sqrt{2^m m!} \sqrt{\pi}} \exp\left\{-\frac{(x/\sigma)^2}{2}\right\} = u_0(x/\sigma) \frac{H_m(x/\sigma)}{\sqrt{2^m m!}} \quad (m \in \mathbf{N}_0), \quad (\text{G24})$$

where σ is the standard deviation of x in the lowest mode ($m=0$). By letting $\sigma = w(z)/\sqrt{2}$, eq. (23) can be rewritten as

$$\Psi_x(x, z) = \sqrt{\frac{A_x u_w 2^m m! \sqrt{\pi}}{z_R w(z)}} u_m\left(\frac{\sqrt{2}x}{w(z)}\right) \exp\left\{i\left[\frac{1}{2} + m\right]\psi(z) - ik \frac{x^2}{2\mathcal{R}(z)}\right\}. \quad (\text{G25})$$

Since the last factor becomes a pure phase effect, it can be understood that the Gaussian beam behaves⁵⁾ as a *harmonic oscillator* on a wave front, i.e. a surface of constant phase.

As is well known, $u_m(x/\sigma)$ has a typical width of $\sim \sigma \sqrt{1+2m}$ and m zero-crossing points within this width. These appear as the typical width of the beam in that axis and the number of the *nodal lines*, respectively.

It is also known that a set of eigen functions of a Harmonic oscillator $\{u_m(x/\sigma), m \in \mathbf{N}_0\}$ composes a *complete* set, i.e. an arbitrary function $\Psi(x)$ can be represented as a superposition of these functions, as

$$\Psi(x) = \sum_{m=0}^{\infty} C_m u_m(x/\sigma), \quad (\text{G26})$$

⁵⁾ More exactly, we were trying to find such solutions.

by choosing appropriate coefficients $\{C_m, m \in \mathbf{N}_0\}$. To find these coefficients, another feature of u_m , *ortho-normality*⁶⁾

$$\int_{-\infty}^{\infty} u_l(x/\sigma) u_m(x/\sigma) d\left(\frac{x}{\sigma}\right) = \frac{1}{\sigma} \int_{-\infty}^{\infty} u_l(x/\sigma) u_m(x/\sigma) dx = \delta_{lm} \quad (\text{G27})$$

can be used.

$$\begin{aligned} \int_{-\infty}^{\infty} u_m(x/\sigma) \Psi(x) dx &= \int_{-\infty}^{\infty} u_m(x/\sigma) \sum_{l=0}^{\infty} C_l u_l(x/\sigma) dx \\ &= \sum_{l=0}^{\infty} C_l \int_{-\infty}^{\infty} u_m(x/\sigma) u_l(x/\sigma) dx = \sum_{l=0}^{\infty} \sigma \delta_{lm} C_l = \sigma C_m \end{aligned} \quad (\text{G28})$$

Since there is no restriction on the choice of σ , there is an infinitive number of sets of $\{u_m(x/\sigma)\}$ to represent a given function $\Psi(x)$, though their coefficients sets are different from each other. They can, of course, be converted into each other by using eq. (28):

$$C_{lm} = \frac{1}{\sigma} \int_{-\infty}^{\infty} u_l(x/\sigma) u_m(x/\sigma') dx. \quad (\text{G29})$$

This matrix C_{lm} with infinity dimension converts the coefficients $\{C_l\}$ for σ to $\{C'_m\}$ for σ' .

As in the fundamental mode case, eq. (25) is a solution for the y -axis as well when all x are replaced by y . So the solution of eq. (1) can be obtained as a product of the two:

$$\begin{aligned} \Psi_{lm}(x, y, z) &= \frac{\sqrt{2}}{w(z)} u_l\left(\frac{\sqrt{2}x}{w(z)}\right) u_m\left(\frac{\sqrt{2}y}{w(z)}\right) \exp\left\{i[1+l+m]\psi(z) - ik\frac{x^2+y^2}{2\mathcal{R}(z)}\right\} \\ &= \frac{\sqrt{2z_R/\lambda}}{z + iz_R} \frac{H_l\left(\frac{\sqrt{2}x}{w(z)}\right)}{\sqrt{2^l l!}} \frac{H_m\left(\frac{\sqrt{2}y}{w(z)}\right)}{\sqrt{2^m m!}} \exp\left\{-\frac{ik}{2} \frac{x^2 + y^2}{z + iz_R}\right\} e^{i[l+m]\psi(z)}. \end{aligned} \quad (\text{G30})$$

This represents the lm -th *Hermite-Gaussian mode*. This is normalized⁷⁾ so that it will satisfy

$$\langle \Psi_{l'm'} | \Psi_{lm} \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \Psi_{l'm'}^*(x, y, z) \Psi_{lm}(x, y, z) = \delta_{l'l} \delta_{m'm}, \quad (\text{G31})$$

⁶⁾ According to the definition, the functions u_m are dimensionless and must be integrated without any dimension so that the result will be again dimensionless.

⁷⁾ The meaning of normalization here is somewhat different from that for the u_m which are dimensionless by themselves, whereas Ψ has a dimension of $(1/m)$. Thus, ‘normalized’ may be considered as ‘convenient to handle’.

Appendices

by choosing arbitrary constants A_x and A_y to be

$$A_x A_y = \frac{z_R/\lambda}{2^{(l+m-1)} l! m!}. \quad (\text{G32})$$

Using the above definition, the power carried by a mode is obtained by squaring the absolute value of the *normalized amplitude* a_{lm} .

$$\begin{aligned} P_{lm} &= \frac{|a_{lm}|^2}{Z_0} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy |\Psi_{lm}(x, y, z)|^2 = \frac{|a_{lm}|^2}{Z_0} \langle \Psi_{lm} | \Psi_{lm} \rangle \\ &= \frac{|a_{lm}|^2}{Z_0} \left\{ \frac{\sqrt{2}}{w(z)} \right\}^2 \int_{-\infty}^{\infty} \left\{ u_l \left(\frac{\sqrt{2}x}{w(z)} \right) \right\}^2 dx \int_{-\infty}^{\infty} \left\{ u_m \left(\frac{\sqrt{2}y}{w(z)} \right) \right\}^2 dy \\ &= |a_{lm}|^2 / Z_0 \end{aligned} \quad (\text{G33})$$

It is a_{00} , the normalized amplitude of the fundamental mode, that which is called normalized amplitude in Chapters 2 through 4 where only the fundamental mode is considered.

Due to the completeness of u_m , any solution of eq. (1) can be represented by a superposition of Hermite-Gaussian modes,

$$\Psi(x, y, z) = \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} a_{lm} \Psi_{lm}(x, y, z). \quad (\text{G34})$$

Furthermore, thanks to its ortho-normality, the total power carried by this wave can be obtained by summing the power of each mode:

$$\begin{aligned} P &= \frac{1}{Z_0} \langle \Psi | \Psi \rangle = \frac{1}{Z_0} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \left| \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} a_{lm} \Psi_{lm}(x, y, z) \right|^2 \\ &= \frac{1}{Z_0} \sum_{l=0}^{\infty} \sum_{l'=0}^{\infty} \sum_{m=0}^{\infty} \sum_{m'=0}^{\infty} a_{l'm'}^* a_{lm} \langle \Psi_{l'm'} | \Psi_{lm} \rangle \\ &= \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} |a_{lm}|^2 / Z_0 = \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} P_{lm}. \end{aligned} \quad (\text{G35})$$

This property is useful in analyzing the power loss due to scattering.

To find the normalized complex amplitude of a certain mode in an arbitrary spatial distribution, the ortho-normality can be used in a similar way to eq. (28)⁸⁾.

$$\begin{aligned} \langle \Psi_{lm} | \Psi \rangle &= \sum_{l', m'=0}^{\infty} \langle \Psi_{lm} | a_{l'm'} \Psi_{l'm'} \rangle = \sum_{l', m'=0}^{\infty} a_{l'm'} \delta_{l'l} \delta_{m'm} = a_{lm} \\ &= \frac{\sqrt{2}}{w(z)} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \frac{u_l \left(\frac{\sqrt{2}x}{w(z)} \right) u_m \left(\frac{\sqrt{2}y}{w(z)} \right)}{\exp \left\{ i [1+l+m] \psi(z) - ik \frac{x^2+y^2}{2\mathcal{R}(z)} \right\}} \Psi(x, y, z) \end{aligned} \quad (\text{G36})$$

⁸⁾ Here a rational expression is used merely to fit the long equation into the limited page width, nothing more. Note that the complex conjugate of Ψ_{lm} was used inside the integral.

The differences among the modes with different numbers of nodes appear as the characteristic radius $\sim w(z)\sqrt{1+2m}$ for the m -th mode (in either axis), and the Guoy phase shift $[1+l+m]\psi(z)$ for the lm -th mode. These features are used in the iris (diaphragm) and the mode selector [63], respectively, to select the desired mode which is usually the fundamental mode.

G.4: Laguerre-Gaussian mode

Another complete set of eigen functions of eq. (1) can be obtained by using cylindrical coordinates

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta, \end{cases} \quad (\text{G37})$$

which is useful when axial symmetry can be assumed. Then eq. (1) can be rewritten as

$$\left[\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} \right) - i2k \frac{\partial}{\partial z} \right] \Psi(r, \theta, z) = 0. \quad (\text{G38})$$

A solution of this differential equation can be obtained by trying a form similar to eq. (17),

$$\Psi(r, \theta, z) = A(z) B\left(\frac{2r^2}{w^2(z)}\right) C(\theta) \exp\left\{-\frac{ik}{2} \frac{r^2}{z+iz_R}\right\}. \quad (\text{G39})$$

Among the factors, $C(\theta)$ can be easily separated and solved as $C(\theta) = \exp\{il\theta\}$ where l is an integer. Then the remaining factors must satisfy

$$\chi \ddot{B}(\chi) + (1-\chi) \dot{B}(\chi) - \left\{ \frac{l^2}{4\chi} + \frac{1}{2} - \frac{1}{2iz_R} \left[z + (z^2 + z_R^2) \frac{\dot{A}(z)}{A(z)} \right] \right\} B(\chi) = 0, \quad (\text{G40})$$

where $\chi = 2r^2/w^2(z)$.

Assuming $B(\chi)$ to be $B(\chi) = \chi^j D(\chi)$, then eq. (40) is reduced to

$$\chi \ddot{D}(\chi) + (2j+1-\chi) \dot{D}(\chi) + m D(\chi) = 0, \quad (\text{G41})$$

whose solutions are known as Laguerre polynomials shown in eq. (46), if j and $A(z)$ satisfy the following conditions.

$$j = |l|/2 \quad \text{and} \quad 2m = -1 - 2j + \frac{1}{iz_R} \left[z + (z^2 + z_R^2) \frac{\dot{A}(z)}{A(z)} \right] \quad (\text{G42})$$

The latter equation can be solved as

$$A(z) = A_0 \frac{\exp\{i[1+|l|+2m]\psi(z)\}}{\sqrt{z^2 + z_R^2}}. \quad (\text{G43})$$

Appendices

In the end, the solution of eq. (38), which represents a *Laguerre-Gaussian mode*, can be written in a similar way to eq. (30) as⁹⁾

$$\Psi_m^{(l)}(r, \theta, z) = \frac{\sqrt{2}}{w(z)} \check{u}_m^{(l)}\left(\frac{\sqrt{2}r}{w(z)}, \theta\right) \exp\left\{i[1+|l|+2m]\psi(z) - ik\frac{r^2}{2\mathcal{R}(z)}\right\} \quad (G44)$$

$$(l \in \mathbf{Z}, m \in \mathbf{N}_0),$$

where $\psi(z)$, $w(z)$, and $\mathcal{R}(z)$ are identical to those in Hermite-Gaussian modes, defined in eq. (14). Again, this is normalized so that the power carried by that mode is given by squaring the absolute value of an additional factor, the normalized amplitude.

The normalized eigen functions $\check{u}_m^{(l)}(\xi, \theta)$ are defined as

$$\check{u}_m^{(l)}(\xi, \theta) = u_m^{(|l|)}(\xi) \exp\{il\theta\} \quad (l \in \mathbf{Z}, m \in \mathbf{N}_0)$$

$$u_m^{(l)}(\xi) = \sqrt{\frac{m!}{\pi(m+l)!}} \xi^l L_m^{(l)}(\xi^2) \exp(-\xi^2/2) \quad (l, m \in \mathbf{N}_0) \quad (G45)$$

$$= u_0^{(0)}(\xi) \sqrt{\frac{m!}{(m+l)!}} \xi^l L_m^{(l)}(\xi^2),$$

where the indices m and $|l|$ represent the numbers of radial and angular nodes, respectively. Note that this is *not* a real function but a complex one (unless $l=0$). In eq. (45), $L_m^{(l)}(\chi)$ is the generalized Laguerre polynomial defined as follows.

$$L_m^{(\mu)}(\chi) = \frac{e^\chi \chi^{-\mu}}{m!} \frac{d^m}{d\chi^m} e^{-\chi} \chi^{m+\mu} \quad (\mu > -1, m \in \mathbf{N}_0)$$

$$\begin{aligned} L_0^{(\mu)}(\chi) &= 1 \\ L_1^{(\mu)}(\chi) &= (1+\mu) - \chi \\ L_2^{(\mu)}(\chi) &= \frac{1}{2}\{(2+\mu)(1+\mu) - 2(2+\mu)\chi + \chi^2\} \\ L_3^{(\mu)}(\chi) &= \frac{1}{6}\{(3+\mu)(2+\mu)(1+\mu) - 3(3+\mu)(2+\mu)\chi + 3(3+\mu)\chi^2 - \chi^3\} \\ &\vdots \end{aligned} \quad (G46)$$

Among the normalized eigen functions, those with $l=0$ are cylindrically symmetric, i.e. have no dependence on θ .

$$\check{u}_m^{(0)}(\xi, \theta) = u_m^{(0)}(\xi) = \frac{1}{\sqrt{\pi}} L_m^{(0)}(\xi^2) \exp(-\xi^2/2) \quad (m \in \mathbf{N}_0) \quad (G47)$$

⁹⁾ Here the parenthesized superscript, in e.g. $u_m^{(l)}$, for the angular index is adopted to avoid a possible confusion with an exponent. One may find those without parentheses in literature.

Furthermore, the lowest mode coincides with the counterpart in the Hermite-Gaussian mode set, i.e. the fundamental mode.

$$u_0^{(0)}(r/\sigma) = \check{u}_0^{(0)}(r/\sigma, \theta) = u_0(x/\sigma) u_0(y/\sigma) = u_{00}(x/\sigma, y/\sigma) \quad (\text{G48})$$

As in the Hermite-Gaussian case, the normalized eigen functions $u_m^{(l)}$ satisfy the ortho-normal relation

$$\int_0^\infty d\xi \int_0^{2\pi} \xi d\theta \left\{ \check{u}_m^{(l)}(\xi, \theta) \right\}^* \left\{ \check{u}_{m'}^{(l')}(\xi, \theta) \right\} = \delta_{ll'} \delta_{mm'}. \quad (\text{G49})$$

Note that one of the eigen functions $\check{u}_m^{(l)}(\xi, \theta)$ in the integral must be taken as its complex conjugate, since it is a complex function. Except for this point, all the discussion for Hermite-Gaussian modes (from eq. (33) through eq. (36)) are applicable also to Laguerre-Gaussian modes.

The biggest difference of Laguerre-Gaussian modes from Hermite-Gaussian modes is that the eigen functions are complex functions. This means that the wave front is no longer given by a spherical surface with the curvature of $\mathcal{R}(z)$ unless $l=0$. The constant phase surface of Laguerre-Gaussian mode with $l \neq 0$ will be a spiral whose pitch is determined by l . This causes a singularity at the center ($x=y=0$), but the amplitude there is zero (since it is the crossing point of nodal lines) and the phase cannot be defined there.

Another consequence of the complex eigen function is that the angular nodal lines move with the time. To see this effect, we need to return to eq. (3). Inserting eqs. (44) and (45) into eq. (3), the phase part is represented by

$$\angle \mathbf{E} = [1 + |l| + 2m] \psi(z) - \frac{kr^2}{2\mathcal{R}(z)} + l\theta + (\omega t - kz). \quad (\text{G50})$$

In a $z=\text{const.}$ plane, the constant phase surface (such as nodal lines) rotate with an angular frequency of $-\omega/l$. This can be understood as the cross section of a propagating spiral.

It is possible to compose a real eigen function by appropriately combining those with $l > 0$ and $l < 0$, i.e.

$$\begin{aligned} \dot{u}_m^{(l)}(\xi, \theta) &= \frac{1}{\sqrt{2}} \left[\check{u}_m^{(l)}(\xi, \theta) + \check{u}_m^{(-l)}(\xi, \theta) \right] = \sqrt{2} u_m^{(l)}(\xi) \cos(l\theta) \\ \dot{u}_m^{(l)}(\xi, \theta) &= \frac{-i}{\sqrt{2}} \left[\check{u}_m^{(l)}(\xi, \theta) - \check{u}_m^{(-l)}(\xi, \theta) \right] = \sqrt{2} u_m^{(l)}(\xi) \sin(l\theta) \end{aligned} \quad (\text{G51})$$

$(l \in \mathbf{N}, m \in \mathbf{N}_0)$

Appendices

Note that l in these definitions are restricted to positive integers ($l > 0$). Either $\hat{u}_m^{(l)}$ or $\hat{v}_m^{(l)}$ satisfy the ortho-normal relation eq. (49) and they are orthogonal to each other. The union of these two sets and the set of cylindrically symmetric ones $\{u_m^{(0)}, \hat{u}_m^{(l)}, \hat{v}_m^{(l)}\}$ is complete, but each set is not complete by itself.

G.5: Gaussian beam with different parameters

Let us consider, as an example, to expand a fundamental Gaussian beam with different parameters, however with the same axis of propagation, to a series of eigen modes. This example is useful in considering, for instance, the visibility of an interference of two fundamental Gaussian beams but with different parameters (curvature and beam radius).

As can be seen in eq. (14), these parameters are determined by the distance from the waist point $z - z_w$ and the Rayleigh range z_R . It is convenient to combine these two into a single parameter as

$$z(z) = (z - z_w) + iz_R, \quad (\text{G52})$$

which is called the *complex beam parameter*. Using this, the parameters in eq. (14) can be written as

$$\frac{1}{z(z)} = \frac{1}{\mathcal{R}(z)} - i \frac{w_w^2}{z_R w^2(z)} \quad \exp\{i\psi(z)\} = i \sqrt{\frac{z^*(z)}{z(z)}}, \quad (\text{G53})$$

and a Hermite-Gaussian mode eq. (30) as

$$\begin{aligned} \Psi_{lm}(x, y, z) = & \\ & \frac{\sqrt{2z_R/\lambda}}{z(z)} \frac{H_l\left(\frac{\sqrt{2}x}{w(\Delta z)}\right)}{\sqrt{2^l l!}} \frac{H_m\left(\frac{\sqrt{2}y}{w(\Delta z)}\right)}{\sqrt{2^m m!}} \exp\left\{-\frac{ik}{2} \frac{x^2 + y^2}{z(z)}\right\} e^{i[l+m]\psi(\Delta z)}. \end{aligned} \quad (\text{G54})$$

A fundamental Gaussian beam Ψ' with a complex beam parameter z' can be expanded to a series of Hermite-Gaussian mode with a different parameter z described above. From eq. (36), this can be written as

$$\begin{aligned} a_{lm} = \langle \Psi_{lm} | \Psi' \rangle &= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \Psi_{lm}^*(x, y, z) \Psi'(x, y, z) \\ &= \frac{2\sqrt{z_R z'_R}}{\lambda z^*(z) z'(z)} \frac{\exp\{-i[l+m]\psi(\Delta z)\}}{\sqrt{2^{l+m} l! m!}} \times \\ & \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy H_l\left(\frac{\sqrt{2}x}{w(\Delta z)}\right) H_m\left(\frac{\sqrt{2}y}{w(\Delta z)}\right) \exp\left\{-\frac{ik}{2} \left(\frac{1}{z'(z)} - \frac{1}{z^*(z)}\right) (x^2 + y^2)\right\}. \end{aligned} \quad (\text{G55})$$

For example, the amplitude of the fundamental mode component ($l=m=0$) can be obtained as

$$\begin{aligned}
 a_{00} &= \langle \Psi_{00} | \Psi' \rangle = \frac{2\sqrt{z_R z'_R}}{\lambda z^*(z) z'(z)} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \exp\{-D(z)(x^2 + y^2)\} \\
 &= \frac{2\sqrt{z_R z'_R}}{\lambda z^*(z) z'(z)} \left\{ \int_{-\infty}^{\infty} \exp[-D(z)x^2] dx \right\} \left\{ \int_{-\infty}^{\infty} \exp[-D(z)y^2] dy \right\} \\
 &= \frac{2\sqrt{z_R z'_R}}{\lambda z^*(z) z'(z)} \left\{ \sqrt{\frac{\pi}{D(z)}} \right\}^2 = \frac{i2\sqrt{z'_R z_R}}{z'(z) - z^*(z)} = \frac{i2\sqrt{z_R z'_R}}{(z_w - z'_w) + i(z_R + z'_R)} \\
 &\quad \text{where} \quad D(z) = \frac{ik}{2} \left(\frac{1}{z'(z)} - \frac{1}{z^*(z)} \right).
 \end{aligned} \tag{G56}$$

This will be unity if and only if $z'_w = z_w$ and $z'_R = z_R$, i.e. the two beams have exactly the same complex parameter.

If the distances from the waist point are the same ($z'_w = z_w$), the amplitude of the fundamental component is determined by the Rayleigh range or, equivalently, the beam radius at the waist.

$$a_{00}|_{\Delta w} = \frac{2\sqrt{z_R z'_R}}{z_R + z'_R} = \frac{2w_w w'_w}{w_w^2 + w'^2_w} \simeq 1 - \frac{1}{2} \left(\frac{\Delta w}{w} \right)^2. \tag{G57}$$

The relative power that does not contribute to the fundamental mode due to the mismatch of the beam radius can be considered as a ‘loss’, which can be written as

$$\begin{aligned}
 \mathcal{A}_{\Delta w} &= 1 - |a_{00}|_{\Delta w}|^2 = \left(\frac{w_w^2 - w'^2_w}{w_w^2 + w'^2_w} \right)^2 \\
 &\simeq \left(\frac{\Delta w_w}{w_w} \right)^2 \quad \text{for} \quad \Delta w_w = w_w - w'_w \ll w_w.
 \end{aligned} \tag{G58}$$

Similarly, if the beam radii at the waist (not at the interfering point) are the same, i.e. $z'_R = z_R = \pi w_w^2 / \lambda$, but the distances from the waist are different, then the amplitude of the fundamental component and the loss due to mismatch are given as

$$\begin{aligned}
 a_{00}|_{\Delta z_w} &= \frac{i2z_R}{(z_w - z'_w) + i2z_R} = \left[1 + \frac{z_w - z'_w}{i2z_R} \right]^{-1} \\
 \mathcal{A}_{\Delta z_w} &= 1 - |a_{00}|_{\Delta z_w}|^2 = \left(\frac{z_w - z'_w}{2z_R} \right)^2 / \left[1 + \left(\frac{z_w - z'_w}{2z_R} \right)^2 \right] \\
 &\simeq \left(\frac{\Delta z_w}{2z_R} \right)^2 \quad \text{for} \quad \Delta z_w = z_w - z'_w \ll z_R,
 \end{aligned} \tag{G59}$$

Appendices

respectively.

In a similar way to eq. (56), the amplitudes of higher order components can be rewritten as

$$\begin{aligned}
a_{lm} &= \langle \Psi_{lm} | \Psi' \rangle \\
&= \frac{2\sqrt{z_R z'_R}}{\lambda z^*(z) z'(z)} \frac{\exp\{-i[l+m]\psi(\Delta z)\}}{\sqrt{2^{l+m} l! m!}} \times \\
&\quad \left\{ \int_{-\infty}^{\infty} \exp\{-D(z) x^2\} H_l\left(\frac{\sqrt{2}x}{w(\Delta z)}\right) dx \right\} \left\{ \int_{-\infty}^{\infty} \exp\{-D(z) y^2\} H_m\left(\frac{\sqrt{2}y}{w(\Delta z)}\right) dy \right\},
\end{aligned} \tag{G60}$$

where $D(z)$ is the same as in eq. (56). Since the exponential term in each integral is an even function, the integral with a Hermite polynomial with an odd index must vanish. Thus, the non-zero components are those with even numbers for both indices.

From the definition of Hermite polynomials in eq. (20), some of the low order terms can be obtained as follows.

$$\begin{aligned}
2E a_{00} &= \sqrt{2} a_{02} = \sqrt{2} a_{20} \\
2E^2 a_{00} &= a_{22} = \sqrt{2/3} a_{04} = \sqrt{2/3} a_{40} \\
2E^3 a_{00} &= \sqrt{1/5} a_{06} = \sqrt{1/5} a_{60} = \sqrt{1/3} a_{24} = \sqrt{1/3} a_{42} \\
2E^4 a_{00} &= \sqrt{\frac{2}{35}} a_{08} = \sqrt{\frac{2}{35}} a_{80} = \sqrt{\frac{1}{10}} a_{26} = \sqrt{\frac{1}{10}} a_{62} = \frac{1}{3} a_{44} \\
&\vdots
\end{aligned} \tag{G61}$$

$$\begin{aligned}
\text{where } E &= \frac{\exp\{-i2\psi(\Delta z)\}}{2} \times \left\{ \frac{2}{w^2(\Delta z) D(z)} - 1 \right\} \\
&= -\frac{1}{2} \frac{z}{z^*} \left\{ \frac{-i2z_R z^* z'}{|z|^2 (z^* - z')} - 1 \right\} = \frac{1}{2} \frac{z - z'}{z^* - z'} = \frac{1}{2} \frac{(z_w - z'_w) - i(z_R - z'_R)}{(z_w - z'_w) + i(z_R + z'_R)}
\end{aligned}$$

Note that E is a constant, i.e. not depending on z . This implies the fact that the amplitude of each mode does not change along the propagation, as expected.

Appendix H: Higher order local oscillators

When modulation-induced sidebands are used as the local oscillators, one must be careful that there may be other components than those used as the local oscillators. For example, if the sidebands are produced by a phase modulation¹⁰⁾, each component can be written, from eq. (2.20), as

$$a_l = J_l(m_a) e^{i(\omega_0 \pm l\omega_a)t} \quad (l \in \mathbf{N}_0) \quad (\text{H1})$$

where m_a and ω_a are the index and the frequency of the modulation induced to produce the sidebands, respectively.

What is usually used as the local oscillator are the first order sidebands, i.e. those with $l=1$ in eq. (H1). In the case of external modulation, these first order sidebands must be as large as possible in order to maximize the amount of output signal. Then the required modulation index will be $m_a \approx 1.84$ radian. With such a high modulation index, the higher order sidebands with $l \geq 2$ in eq. (H1) will have non-negligible amplitudes. The remaining carrier component ($l=0$) also has a relatively large amplitude. Both of them may affect the amount of the signal as well as that of the noise.

In the case of internal or pre-modulation schemes, the optimum modulation index is determined as a compromise of the local oscillator amplitude and the remaining carrier component (see Ref. [25]), and often not a small value (~ 0.5). Thus, higher order sidebands affect the signal and noise, although their effects are less significant than the external modulation case. (In this case the carrier component has no effect since that from two arms cancel each other and does not hit the photo-detector.)

These higher order sidebands (and the remaining carrier component) behave as local oscillators, with, however, different frequency from the ‘main’ local oscillators ($l=1$). Thus, these convert the optical information to the photo-current but at different frequencies from those of the main local oscillators, i.e. $l\omega_a$ instead of ω_a . They will contribute to the output signal when these are also (electrically) demodulated, i.e. multiplied by their frequencies $l\omega_a$.

Note, however, that only the odd components ($l = 2j+1$) behave as the phase detector simultaneously with the main ($l=1$) local oscillators, whereas the

¹⁰⁾ This is what is commonly done because of its easiness and efficiency. Since amplitude modulation requires the change in instantaneous power, it may result in the loss of the maximum available power.

Appendices

even components ($l = 2j$) detect the amplitude modulation. This is due to the relative phase of $\angle i = \frac{\pi}{2}$ between the odd components and the even components (see eq. (2.20) in 2.2.1).

What requires a more serious consideration is the noise. Even if these higher order components are not multiplied by their frequencies, these produce noise at the signal frequency range, because they interfere with the *vacuum fluctuation* which causes the shot noise (see the discussion in 3.5.2). In the end, the fluctuation in the detected power due to the shot noise is determined only by the total amount of the detected power, including these components which do not contribute to the signal. Thus, it is desirable to make only the local oscillator contribute to the signal in order to optimize the signal-to-noise ratio.

For example, the noise at frequency ω_n in the output signal is demodulated from $\omega_a \pm \omega_n$ in the photo-current (assuming demodulation only at ω_a). The $\omega_a + \omega_n$ component in the photo-current is a result of the interferences of the light at frequencies of $\omega_0 \pm \omega_n$ and $\omega_0 \mp \omega_a$. Similarly, the $\omega_a - \omega_n$ component in the photo-current is resulting from $\omega_0 \mp \omega_n$ and $\omega_0 \pm \omega_a$. Thus, these are the components related to the output noise at ω_n .

If higher order (optical) local oscillators with $\omega_0 \pm l\omega_a$ are present, however, the interference of, for example, $\omega_0 \pm (l-1)\omega_a + \omega_n$ and $\omega_0 \pm l\omega_a$ also produces the $\omega_a \pm \omega_n$ component in the photo-current, and thus contributes to the output noise at ω_n . Even though there may not be any interesting signal at $\omega_0 \pm (l-1)\omega_a + \omega_n$, there are the vacuum fluctuations which contribute to the noise.

Furthermore, since these different frequency components interfere with the same vacuum fluctuation, the fluctuation in the photo-current has correlations among the different frequency components. For example, the vacuum fluctuation at $\omega_0 + \omega_n$ interferes with all the components with $\omega_0 \pm l\omega_a$. This results in the fluctuation in the detected power with correlation among the frequency components with $\omega_n \pm l\omega_a$.

When the photo-current is demodulated at ω_a , only $\omega_a \pm \omega_n$ components contribute to the noise at ω_n . When it is also demodulated at $l\omega_a$, however, other components also affect the noise in the demodulated output. This does not necessarily mean that the amount of noise increases but it may decrease when the correlated noise contributions are subtracted properly. Further discussion on this topic can be found in Refs. [27] and [31].

References

- [1] A. Abramovici, W.E. Althouse, R.W.P. Drever, Y. Gürsel, S. Kawamura, F.J. Raab, D. Shoemaker, L. Sievers, R.E. Spero, K.S. Thorne, R.E. Vogt, R. Weiss, S. Whitcomb, and M.E. Zucker, *Science* **256** (1992) p.325.
- [2] C. Bradaschia, R. Del Fabbro, A. Di Virgilio, A. Giazotto, H. Kautzkey, V. Montelatici, D. Passuello, A. Brilliet, O. Cregut, P. Hello, C.N. Man, P.T. Manh, A. Marraud, D. Shoemaker, J.-Y. Vinet, F. Barone, L. Di Fiore, L. Milano, G. Russo, S. Solimeno, J.M. Aguirregabiria, H. Bel, J.-P. Duruisseau, G. Le Denmat, P. Tournenc, M. Capozzi, M. Longo, M. Lops, I. Pinto, G. Rotoli, T. Damour, S. Bonazzola, J.A. Marck, Y. Gourghoulon, L.E. Holloway, F. Fuligni, V. Iafolla, and G. Natale, *Nucl. Instr. Meth. Phys. Res.* **A289** (1990) p.518.
- [3] K. Danzmann, H. Lück, A. Rüdiger, R. Schilling, M. Schrempel, W. Winkler, J. Hough, G.P. Newton, N.A. Robertson, A.M. Campbell, J.E. Logan, D.I. Robertson, K.A. Strain, J.R.J. Bennett, V. Kose, M. Kühne, B.F. Schutz, D. Nicholson, J. Shuttleworth, H. Welling, P. Aufmuth, R. Rinkleff, A. Tünnermann, and B. Willke, *Internal report MPQ 190* (1994).
- [4] K. Kawabe, *Report on Japanese interferometer project*, in *Gravitational Wave Astrophysics*, ed. P.R. Saulson (Report on Snowmass Meeting in 1994, in press).
- [5] D. Blair, talk at *Aspen Winter conference* (Aspen, 1995).
- [6] C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (W.H. Freeman and Co., San Francisco, 1973).
- [7] A. Einstein, *Preuss. Akad. Wiss. Berlin* (1916) p.688.
- [8] K.S. Thorne, in *300 Years of Gravitation*, eds. S.W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1987) p.330.
- [9] J.H. Taylor and J.M. Weisberg, *Astrophys. J.* **253** (1982) p.908.
- [10] J.H. Taylor and J.M. Weisberg, *Astrophys. J.* **345** (1989) p.434.
- [11] J. Weber, *Phys. Rev.* **117** (1960) p.306.
- [12] J. Weber, *Phys. Rev. Lett.* **22** (1969) p.1320.
- [13] D.G. Blair, in *The Detection of Gravitational Waves*, ed. D.G. Blair (Cambridge University Press, Cambridge, 1991) p.73.
- [14] B.F. Schutz ed., *Gravitational Wave Data Analysis* (Kluwer Academic Publishers, Dordrecht, 1989).
- [15] C.M. Caves, *Phys. Rev.* **D23** (1981) p.1693.

References

- [16] N.A. Robertson, in *The Detection of Gravitational Waves*, ed. D.G. Blair (Cambridge University Press, Cambridge, 1991) p.353.
- [17] P.R. Saulson, *Phys. Rev. D***42** (1990) p.2437.
- [18] A. Rüdiger and colleagues of GEO, in *Relativistic Gravity Research With Emphasis on Experiments and Observations*, eds. J. Ehlers and G. Schäfer (Springer-Verlag, Berlin, 1992) p.239.
- [19] J. Hough, H. Ward, G.A. Kerr, N.L. Mackenzie, B.J. Meers, G.P. Newton, D.I. Robertson, N.A. Robertson, and R. Schilling in *The Detection of Gravitational Waves*, ed. D.G. Blair (Cambridge University Press, Cambridge, 1991) p.329.
- [20] P.R. Saulson, *Fundamentals of Interferometric Gravitational Wave Detectors* (World Scientific, Singapore, 1994).
- [21] R.L. Forward, *Phys. Rev. D***17** (1978) p.379.
- [22] R. Weiss, *MIT Quarterly Progress Report (Research Laboratory of Electronics)* **105** (1972) p.54.
- [23] R.W.P. Drever, talk at *10th Int. Conf. on General Relativity and Gravitation* (Padua, 1983).
- [24] C.N. Man, D. Shoemaker, M. Pham Tu, and D. Dewey, *Phys. Lett. A***148** (1990) p.8.
- [25] D. Shoemaker, R. Schilling, L. Schnupp, W. Winkler, K. Maischberger, and A. Rüdiger, *Phys. Rev. D***38** (1988) p.423.
- [26] S. Kawamura, J. Mizuno, J. Hirao, N. Kawashima, and R. Schilling, *ISAS report* **637** (1989).
- [27] T.M. Niebauer, R. Schilling, K. Danzmann, A. Rüdiger, and W. Winkler, *Phys. Rev. A***43** (1990) p.5022.
- [28] L. Schnupp, talk at a *European Collaboration Meeting on Interferometric Detection of Gravitational Waves* (Sorrento, 1988).
- [29] R. Takahashi, J. Mizuno, S. Miyoki, and N. Kawashima, *Phys. Lett. A***187** (1994) p.157.
- [30] M.W. Regehr, Ph.D Thesis, California Institute of Technology, Pasadena (1995), unpublished.
- [31] B.J. Meers and K.A. Strain, *Phys. Rev. A***44** (1991) p.4693.
- [32] W. Winkler, in *The Detection of Gravitational Waves*, ed. D.G. Blair (Cambridge University Press, Cambridge, 1991) p.269.
- [33] J.-Y. Vinet, B. Meers, C.N. Man, and A. Brillet, *Phys. Rev. D***38** (1988) p.433.
- [34] B.J. Meers, *Phys. Lett. A***142** (1989) p.465.

- [35] J. Hough, B.J. Meers, G.P. Newton, N.A. Robertson, H. Ward, G. Leuchs, T.M. Niebauer, A. Rüdiger, R. Schilling, L. Schnupp, H. Walther, W. Winkler, B.F. Schutz, J. Ehlers, P. Kafka, G. Schäfer, M.W. Hamilton, I. Schütz, H. Welling, J.R.J. Bennett, I.F. Corbett, B.W.H. Edwards, R.J.S. Greenhalgh, and V. Kose, *Internal report MPQ 147* (1989).
- [36] D. Herriott, H. Kogelnik, and R. Kompfner, *Applied Optics* **3** (1964) p.523.
- [37] D.R. Herriott and H.J. Schulte, *Applied Optics* **4** (1965) p.883.
- [38] R.W.P. Drever, J. Hough, A.J. Munley, S.-A. Lee, R. Spero, S.E. Whitcomb, H. Ward, G.M. Ford, M. Hereld, N.A. Robertson, I. Kerr, J.R. Pugh, G.P. Newton, B. Meers, E.D. Brook III, and Y. Gürsel, in *Quantum Optics, Experimental Gravitation, and Measurement Theory*, eds. P. Meystre and M.O. Scully (Plenum Press, New York, 1983) p.503.
- [39] R.W.P. Drever, in *Gravitational Radiation*, eds. N. Deruelle and T. Piran (North-Holland, Amsterdam, 1983) p.321.
- [40] E. Morrison, J. Hough, B.J. Meers, G.P. Newton, D.I. Robertson, K.A. Strain, P.J. Veitch, and H. Ward, in *Proc. of The Sixth Marcel Grossmann Meeting on General Relativity*, eds. H. Sato and T. Nakamura (World Scientific, Singapore, 1992) p.1505.
- [41] M.E. Zucker, in *Proc. of The Sixth Marcel Grossmann Meeting on General Relativity*, eds. H. Sato and T. Nakamura (World Scientific, Singapore, 1992) p.224.
- [42] R. Schilling, L. Schnupp, W. Winkler, K. Maischberger, and A. Rüdiger, *J. Phys. E: Sci. Instrum.* **14** (1981) p.65.
- [43] P. Horowitz and W. Hill, *The Art of Electronics* second ed. (Cambridge University Press, Cambridge, 1989).
- [44] B.J. Meers, *Phys. Rev. D* **38** (1988) p.2317.
- [45] R. Schilling, private communication (1994).
- [46] J.-Y. Vinet, *J. Physique (Paris)* **47** (1986) p.639.
- [47] Ling-An Wu, H.J. Kimble, J.L. Hall, and Huifa Wu, *Phys. Rev. Lett.* **57** (1986) p.2520.
- [48] D.G. Blair, in *The Detection of Gravitational Waves*, ed. D.G. Blair (Cambridge University Press, Cambridge, 1991) p.16.
- [49] P.A.M. Dirac, *Principles of Quantum Mechanics* fourth ed. (Oxford University Press, London, 1958).
- [50] A. Gillespie and F. Raab, in preparation (1994).
- [51] F. Bondu and J.-Y. Vinet, *Phys. Lett. A* **198** (1995) p.74.
- [52] B.J. Meers and K.A. Strain, *Phys. Rev. D* **43** (1991) p.3117.

References

- [53] J. Mizuno, K.A. Strain, P.G. Nelson, J.M. Chen, R. Schilling, A. Rüdiger, W. Winkler, and K. Danzmann, *Phys. Lett.* **A175** (1993) p.273.
- [54] W. Winkler, K. Danzmann, A. Rüdiger, and R. Schilling, *Phys. Rev.* **A44** (1991) p.7022.
- [55] P. Hello and J.-Y. Vinet, *Phys. Lett.* **A178** (1993) p.351.
- [56] K.A. Strain, K. Danzmann, J. Mizuno, P.G. Nelson, A. Rüdiger, R. Schilling, and W. Winkler, *Phys. Lett.* **A194** (1994) p.124.
- [57] R. Schilling, talk at *The Seventh Marcel Grossmann Meeting on General Relativity* (Stanford, 1994).
- [58] E. Morrison, B.J. Meers, D.I. Robertson, and H. Ward, *Applied Optics* **33** (1994) p.5037.
- [59] W. Winkler, R. Schilling, K. Danzmann, J. Mizuno, A. Rüdiger, and K.A. Strain, *Applied Optics* **33** (1994) p.7547.
- [60] A.E. Siegman, *Lasers* (University Science Books, Mill Valley, 1986).
- [61] W. Winkler, A. Rüdiger, R. Schilling, K.A. Strain, and K. Danzmann, *Optics Communications* **112** (1994) p.245.
- [62] A. Yariv, *Quantum Electronics* third ed. (John Wiley and Sons, New York, 1989).
- [63] A. Rüdiger, R. Schilling, L. Schnupp, W. Winkler, H. Billing, and K. Maischberger, *Optica Acta* **28** (1981) p.641.