2012

Advances in quantum optical metrology and the establishment of an invisible quantum tripwire

Steven Blane McCracken
Louisiana State University and Agricultural and Mechanical College

Recommended Citation

McCracken, Steven Blane, "Advances in quantum optical metrology and the establishment of an invisible quantum tripwire" (2012). *LSU Doctoral Dissertations*. 2313.
https://digitalcommons.lsu.edu/gradschool_dissertations/2313

ADVANCES IN QUANTUM OPTICAL METROLOGY AND THE ESTABLISHMENT OF AN INVISIBLE QUANTUM TRIPWIRE

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana Statef University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Physics and Astronomy

by
Steven Blane McCracken
B.S., Francis Marion University, 2004
August 2012

Table of Contents

Abstract

This thesis presents a summary of the foundation and background of the field of quantum optics, and an analysis of some recent discoveries in various fields of which I have aided in furthering investigative research and advancement through publications. Such topics include numerical optimization of generalized quantum states used in phase sensitive quantum metrology, an analysis of object detection through the use of quantum interferometry in the presence of lossy conditions, and the use of the latter technique to propose an invisible quantum tripwire.

First is a collaborative effort to numerically optimize quantum optical states for quantum metrological applications. We optimize two-mode, entangled, number states of light in the presence of loss in order to maximize the Fisher information, which is equivalent to minimizing the phase uncertainty. We find that in the limit of zero loss the optimal state is the so-called N00N state, for small loss, the optimal state gradually deviates from the N00N state, and in the limit of large loss the optimal state converges to a generalized two-mode coherent state, with a finite total number of photons. The results provide a general protocol for optimizing the performance of a quantum optical interferometer in the presence of photon loss.

The next topic is statistical hypothesis testing of interaction free measurement and a quantitative limit on the obtainable error. Previous analyses have been based solely on detection probabilities known only in the infinite photon limit. Our analysis assumes a finite number of photons, and an investigation of reliability in the presence of photon loss and phase fluctuations. We use symmetric hypothesis testing and the Chernoff bound to provide error estimation after N independent, single photon trials.

Finally, we present a quantum optical interrogation technique capable of detecting an intrusion with very low probability of the quantum ”tripwire” being revealed to the intruder. The tripwire exploits a curious nonlinear behavior of the quantum Zeno effect we discovered, which occurs in a lossy system. We also employ statistical hypothesis testing, allowing us to calculate a confidence level of interaction-free measurement after a given number of trials.

Chapter 1
A Brief History

In the earliest time of electrostatic exploration, the relationship between electricity and magnetism was not well known. Furthermore, the propagation of such fields was little more than speculation – rendering the fundamental properties of light a mystery. The turn of the nineteenth century showed a great deal of interest in electrostatic research and the corresponding mathematical foundations with scientists such as Poisson, Ampere, Biot, Savart, and Gauss. However, it was not until the research of Michael Faraday (and Joseph Henry, independently) that the principles of electricity and magnetism became explicitly knit together through the principles of induction and the discovery of the electromotive force. These crude, yet revolutionary, principles were later refined and recast in the form of mathematical elegance known simply as Maxwell's equations.

Expanding on Faraday's visualization of electric field lines, James Clerk Maxwell developed his theories on an fictional physical grid he dubbed the molecular vortex. This theoretical visual depiction consists of a uniform lattice of electromagnetic vortices and smaller theoretical molecular particles between the vortices capable of carrying charge through the lattice in the event adjacent vortex rotations were properly aligned [1].

$$\nabla \times E = -\frac{\partial B}{\partial t} \tag{1.1a}$$

$$\nabla \times B = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial E}{\partial t} \tag{1.1b}$$

$$\nabla \cdot E = \frac{\rho}{\epsilon_0} \tag{1.1c}$$

$$\nabla \cdot B = 0 \tag{1.1d}$$

Maxwell's equations exhibit an explicit, quantitative relationship between electric and magnetic fields with the implication for complementary kinetic and potential energy conversions between them. Furthermore, this equation set suggests the existence of the propagation of electromagnetic energy and is the basis for the derivation of the electromagnetic nonhomogeneous wave equations [1]:

$$\left(\nabla^2 - \mu\epsilon\frac{\partial^2}{\partial t^2}\right)\mathbf{E} = \mu\frac{\partial \mathbf{J}}{\partial t} + \nabla\frac{\rho}{\epsilon} \tag{1.2a}$$

$$\left(\nabla^2 - \mu\epsilon\frac{\partial^2}{\partial t^2}\right)\mathbf{B} = -\mu\nabla \times \mathbf{J}. \tag{1.2b}$$

These suggests that the occurrence of a spatial or temporal alteration in the current $\mathbf{J}$ or charge ρ densities will result in the spatial propagation of a coupled electromagnetic wave. The equation set further suggests that the wave – once in the absence of the initial disturbance – has symmetric electric and magnetic properties and propagates with the velocity $c = 1/\sqrt{\mu\epsilon}$. While it was unproven at the time of discovery, this velocity was (correctly) speculated to be the propagation velocity of visible light in free space.

Following Maxwell's theoretical discovery of the existence of freestanding electromagnetic wave propagation, experimentalists inevitably strove to confirm the postulate. By the late nineteenth century, Heinrich Hertz's had succeeded in experimentally validating Maxwell's work and its implications electromagnetic propagation. His experimental setup consisted of an oscillating high voltage spark gap source and a wire loop antenna which, similarly, had an open segment to act as a spark gap. The high voltage of the transmitter ionized the air of the spark gap and created a continuous oscillation of current between the electrodes which produced an azimuthally symmetric propagating electromagnetic wave. The wire loop was effective at "detecting" the transmitted waves up to hundreds of meters – as indicated by the production of spark in the gap of the receiving loop. Hertz's experiments marked the first experimental evidence of the wireless transmission of radio waves and would lead directly to the advent of telecommunication. Perhaps more relevant is the fact that these transmitted waves obeyed all of the same observable properties as visible light (including reflection and refraction), thus strengthening beyond doubt that light is indeed a high frequency electromagnetic wave that can be explained and predicted by Maxwell's equations.[2]

Chapter 2
Optical Elements

Before the advent of quantum mechanics light could only be analyzed as a three dimensional continuous wave propagating through free space. Interestingly, even after the discovery of quantum phenomena, this classical wave analysis is still the only wave to describe many of light's properties. Such examples of this would be color (the frequency of the oscillating electromagnetic field), wavefront characteristics (just as water ripples propagate, reflect, and refract), relative wave phases and phase shifts (which is the basis for various interference effects), and – the only characteristic listed here specifically inherent to transverse waves – polarization.

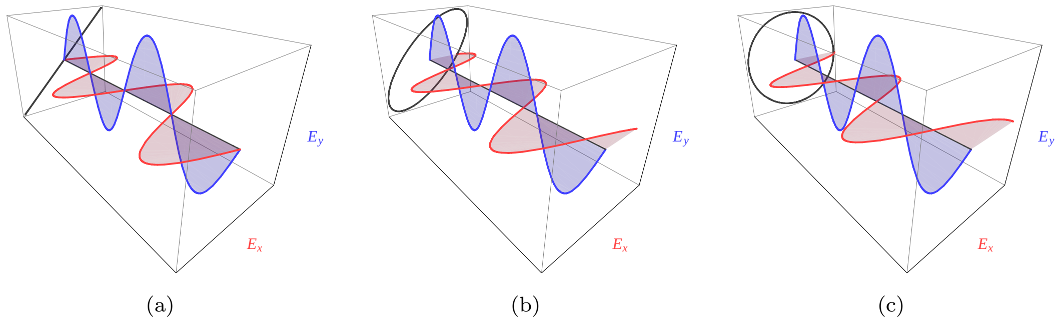
2.1 Polarizers

Polarization is a subtle but universal property of light waves and can yield some rather complex properties depending on the medium with which it interacts. In its most basic sense, polarization is conventionally described by the angle of the plane of oscillation of the electric field component. Specifically, the transverse oscillations are directly coupled, and therefore a single description of one is sufficient. Most "natural" sources of electromagnetic radiation are unpolarized (i.e., there is no predefined affinity for the angle of the electric field oscillation). That is to say, the electric field vector $\mathbf{E}(z,t) = \mathbf{E}_x(z,t) + \mathbf{E}_y(z,t)$ is perfectly not biases to either of the two of the orthogonal oscillating electric components given by

$$\mathbf{E}_x(z,t) = \hat{i}E_{0x}\cos(kz - \omega t) \tag{2.1}$$

$$\mathbf{E}_y(z,t) = \hat{j}E_{0x}\cos(kz - \omega t + \epsilon) \tag{2.2}$$

with ϵ representing an arbitrary phase difference between the components. Once a specific polarization state has been arranged, however, this phase factor becomes crucial in determining the polarization of the field. For $\epsilon = \pi n$ increments, the field will be linearly polarized (see figure (a)). It should be noted that the field will only exhibit identical polarization for $\epsilon = 2\pi n$, otherwise the linear polarizations will manifest as a physical 90°rotation of the plane of polarization. In reference to figures (a), this means that the projected polarization would span from the top left corner to the bottom right instead.



(a) (b) (c)

FIGURE 2.1. A representation of the additive properties of the orthogonal oscillating components of electric field vector and its projected polarization. Figure (a) represents no phase shift between the component fields ($\epsilon = 0$) and the resulting linear polarization. In figure (b), the relative phase has been shifted by $\epsilon = \pi/4$ yielding an elliptical polarization, while fully circular polarization results from a phase different of exactly $\epsilon = \pi/2$.

As seen in figure (c), circular polarization is a phenomenon that occurs when the components of the electric field vector possess a relative phase difference of $\epsilon = \pm\pi/2$. Whether the axis is shifted ahead by 2π or behind by -2π determines whether the light is right or left circularly polarized. Obviously, any phase difference of $0 < \epsilon < \pi/2$ will result in various degrees of elliptical polarization. There are numerous methods of producing polarized light - both natural and manufactured; circular and linear. Though it is not within the scope of this thesis to dwell in detail on each method, several common and practical methods are worth a brief discussion.

2.1.1 Scattering

Polarization, or dichroism, is the result of selective absorption by an anisotropic dichroic material. Therefore, it can be readily plausible that polarization is the direct effect of atomic interaction where the light field absorbed and may or may not be reemitted in a constructive resonance with the incident field (which hinges specifically on the properties of the dichroic material with which is it interacting). With this is mind, natural scattering is a reasonable fundamental for understanding the principles behind more complex dichroism.

A concretely common, yet enlightening, example of polarization due to scattering is our own sun's interaction with the earth's atmosphere. Atmospheric scattering occurs when incident (unpolarized) sunlight interacts with a air molecules which absorb specific frequencies of the light-energy. Although the incident light is unpolarized, given the assumption that all light rays are approximately parallel, all possible orientations of the electric field vectors are transverse to the direction of propagation of the light rays. Combined with the knowledge that the electric field is the driving force for the molecular resonance and absorption, it stands to reason that the absorbing molecules are resonating transversely to the angle of incident sunlight. This directly implies that there can be no molecular resonance in the direction of propagation of the light. Therefore, analysis of any perpendicularly scattered light will reveal polarization in the direction of the electric field oscillation, but never a component along the direction of the ray's propagation [3].

2.1.2 Wire-Grid and Polaroid

These are perhaps the most abundant and practical of polarizers. Although often misinterpreted by those not privy to a developed knowledge of electromagnetic interactions, the wire-grid filter involves a rather straightforward mode of operation. Often explained as a "gate," a common interpretations is that the slits allow a single polarization through the grid. While this naivety holds true of a resultant linearly polarized beam, the method of operation is in fact exactly opposite.

It is necessary to recall that dichroic materials *absorb* polarization states through electromagnetic interaction. As the name states, the grid is composed of dense parallelly aligned conducting strands. The incident electric field aligned parallel to the grid induces a alternating current in the grid and is thereby absorbed. A fraction of the energy is also converted into heat from the induced current. Any remaining transmitted field parallel to the grid tends to be cancelled by the phase lagged, reradiated field from the current induced in the grid. Clearly, any field aligned perpendicular to the grid passes directly through with little absorption. A polaroid is the molecular analog to the wire-grid. Iodine-soaked stretched vinyl forms long chains of close-knit molecular conductance channels. The obvious benefits of polaroid sheets are weight, size, durability, and cost to name a few of the many.

Polarization is of specific importance to the research topics presented in this thesis as they provide a basis of orthogonality for interferometry and detection. According to Malus's Law,

$$I(\theta) = I(0)\cos^2\theta \tag{2.3}$$

where $I(0)$ is the light intensity after an initial linear dichroic interface aligned to $\theta = 0$, and has a maximum possible intensity of

$$I_{0,max} = \frac{c\epsilon_0}{2}E_0^2. \tag{2.4}$$

This unfortunately suggests that the use of any polarizer to perform an experiment will reduce the intensity of the available light by at least 1/2, which can be especially troubling when working with very weak intensities of light (in some cases even down to the single photon level). Another innate property of this principle is the potential for indiscernibility of the beam's polarization. Depending on the specific experiment, this could be a desired effect or lead to complicated probabilistic detection schemes.

2.1.3 Dichroic Crystals

The final method of polarization discussed here in brevity, will be revisited in later sections for its more widely utilized, conjoining effects. Dichroic crystals are such that the composition of their lattice structure lends a uniform anisotropy along preferred directions. In similar fashion to the other dichroic materials discussed here, the molecules composing these crystals have an affinity to vibrate more readily at a specific angle to the lattice. This affinity is the result of weaker bonding when compared to other incident angles of the same lattice, which implies a reduced natural frequency and absorption energy. Therefore, at all frequencies, light polarized in the plane of these weakly bonded molecules will be more readily absorbed than other planes of polarization. The plane of polarization for which the crystal is non-absorbing is known as the optic axis. Obviously, unpolarized light incident on such a dichroic crystal will emerge polarized in the direction of the optic axis. The unique shape of such crystals is the origin of the anisotropy and the related optic axis. The structure is such that the lattice has a uniaxial symmetric (as apposed to cubic or other more symmetric crystalline structures). Given the intrinsic periodicity of all crystals, the optic axis is more correctly referred to as a principal plane. This also yields other interesting characteristics that will be discussed in their relevant sections.

2.2 Wave Plates

A wave plate is an optical element capable of manipulating the polarization of pre-polarized light through methods similar to the above mentioned dichroic crystals. Wave plates are manufactured by physically altering a uniaxial crystal such that the planes of incidence are normal and parallel to the optic axis. As discussed previously, when the electric field is normally incident on the optic axis with its field vector oscillating parallel to the optic axis, it passes readily through the crystal. However, if the alignment is such that the linear polarization is perpendicular to the principal plane, the crystal appear refractively isotropic. The field of the perpendicular polarization continually absorbed reradiated through the crystal at a relatively reduced velocity. This velocity reduce is directly associated with a relative shift in the crystal's index of refraction between the parallel and perpendicular alignments. Typically, alignment parallel to the optic axis has a smaller index of refraction (at a given frequency) and therefore a greater velocity ($v = c/n$) [3]. While the converse is possible, all discussions in this work will be restricted to this so called negative uniaxial crystal (e.g., calcite). Due to it's tendency toward higher velocity light propagation, the optic axis is known as the fast axis. Likewise, the plane perpendicular to the principal plane is dubbed the slow axis. The angular dependent refractive index directly leads to angular dependent optical distances through the same physical path in the crystal. This distance is given by

$$\Lambda = d(|n_o - n_e|) \tag{2.5}$$

where n_o represents the lower refractive index of the optic axis, n_e the relatively higher refractive index of the orthogonal alignment, and d is the thickness of the crystal. The subscripts are represented prematurely here for continuity, but will be discussed in further depth in later sections. From this description of optical path length, it follows that the phase difference between the two orthogonal components of polarization is

$$\Delta\phi = k_0\Lambda = \frac{2\pi}{\lambda_0}d(|n_o - n_e|). \tag{2.6}$$

For a single, polarized, monochromatic wave (or equivalently, a superposition of in phase orthogonal polarization waves), it is clear that if the optical paths of the polarization components is a 2π multiple of the spatial frequency (wave number), the emerging light will be identical in polarization to the incident beam.

2.2.1 Full-Wave Plate

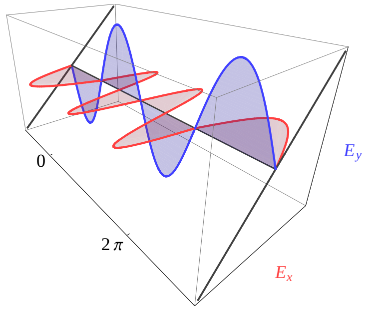


FIGURE 2.2. Polarized light passing through a full-wave plate with a thickness from 0 to 2π. The component of the electric field vector along the fast axis (optic axis) is transmitted with negligible phase shift, while the component parallel to the slow axis has an optical distance of λ_0. Note that the projected diagonal polarization state is the same entering and exiting the crystal.

The name "full-wave plate" directly suggests that the slow axis is retarded by a full-wavelength before exiting the optical medium. Due to the fact that the index of refraction is a function of frequency, the optical length of the plate is also a function of frequency. This results in identical polarization wave emission (as illustrated in figure 2.2) strictly for monochrome waves, and therefore acts as an effective frequency bandwidth filter.

The apparatus must not be tuned, however, in such a way as to select the desired frequency directly, since a $2\pi\lambda_0$ shift would clearly result in no detectable change. Instead, the crystal should be placed between two orthogonal polarizers. Natural light becomes linearly polarized by the initially dichroism and passes through the crystal. As the light leaves the crystal, it is already effectively frequency separated by its polarization characteristics. The wavelength λ_0 whose optical distance through $n_o d$ and $n_e d$ is a 2π multiple of the unshifted wavelength is the single wavelength that will emerge from the crystal unaffected (without regard to higher order harmonics). All other wavelengths will be slightly out of phase, and therefore have gained some degree of ellipticity. At this point, a polarizer in alignment with the initial would do little but narrow the frequency bandwidth of the incident natural light without specific frequency selection. Therefore, an orthogonal polarizer is used to detect the frequency which resulted in the least eccentric superposition (i.e., $\lambda/4 = d(|n_o - n_e|)$ yields circular polarization - the maximum possible linear orthogonality) and therefore the highest absolute probability of being the complement to the eliminated frequency.

2.2.2 Half-Wave Plates

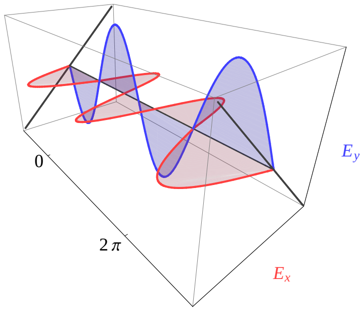


FIGURE 2.3. Polarized light passing through a half-wave plate with a thickness from 0 to 2π. The component of the electric field vector along the fast axis (optic axis) is transmitted with negligible phase shift, while the component parallel to the slow axis has an optical distance of $\lambda_0/2$. Note that while the beams entering and exiting the crystal are both diagonal polarization, it has been rotated by 90°upon exit.

Half-wave plates rotate a beam's polarization while maintaining its original phase relationship and are of particular important to later chapters of this thesis. The relative optical length between the orthogonal polarization components is defined as $\lambda/2$. According to equation (2.6), this condition can be expressed as

$$d(|n+o-n_e|) = (2m+1)\lambda_0/2. \tag{2.7}$$

Figure 2.3 represents the angle of linear polarization 45° between the fast and slow axes. This ensures that electric field vector projects on the fast and slow axes with equal amplitudes, which ensures maximum polarization rotation. Specifically, if the wave were completely aligned with the fast (or slow) axis, it would transmit through the crystal with a minimum (or maximum) optical distance, but with no relative phase shift. Therefore, any beam polarized along either of these axes will emerge in precisely the same state. For a negative uniaxial crystal, the optic axis yields the maximum transmission velocity and can be taken as a $\theta = 0°$ reference angle. When the polarization is incident at $\theta = 45° = \pi/4$ angle from the fast axis, the electric field component along the slow axis lags by $\lambda/2 = \pi$ radians, thereby effectively rotating the field by $2\theta = \pi/2 = 90°$. Conceptually, the incident polarization is "reflected" about the optic axis creating the effect of 2θ radians of state rotation. Therefore, any angle of alignment $0 <= \theta <= \pi/4$ may be realized to produce plane polarization rotations $0 <= 2\theta <= \pi/2$. On a final note, as is true with the full-wave plate, the half-wave plate has a subtle frequency dependence that may affect the precision of its performance.

2.2.3 Quarter-Wave Plates

Finally, in the same manner as the previously discussed wave plates, a quarter-wave plate retards the polarization component along the slow axis by $\lambda/4$. Again,

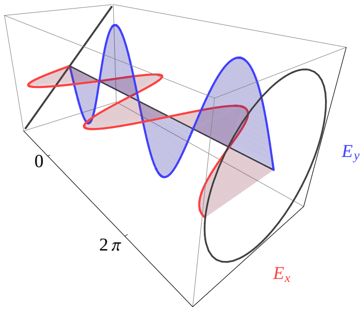


FIGURE 2.4. Polarized light passing through a quarter-wave plate with a thickness from 0 to 2π. The component of the electric field vector along the fast axis (optic axis) is transmitted with negligible phase shift, while the component parallel to the slow axis has an optical distance of $\lambda_0/4$. Note that the polarization state of the beam has been converted from linear to circular polarization. Conversely, this crystal will also convert circular to linear polarization due to the equivalent phase shift.

according to equation (2.6), this condition is given by

$$d(|n_o - n_e|) = (4m + 1)\lambda_0/4. \tag{2.8}$$

As with the half-wave plate, the $(4m+1)$ coefficient represents the arbitrary additional phase of 2π. As illustrated by figure 2.4, a quarter-wave plate is capable of converting linearly polarized light aligned at $\theta = \pi/4 = 45°$ between the principal axes of the uniaxial crystal into circularly polarized light (which is essentially a perfectly even distribution of linear polarization at all angles). Conversely, if circular light is incident on a quarter-wave plate of arbitrary orientation, the emerging beam will be linearly polarized at a 45° to the principal axes regardless of their orientation. Also, just as the half-wave plate is cable of rotating polarization at angles less than 90°, the quarter-wave plate can create various degrees of elliptical polarizations.

2.3 Beam Splitters

In a basic, classical sense, beam splitter is literally the splitting of a single incident beam of light into two emergent beams of lesser intensities (equal intensity or otherwise, within the restriction that the sum of their intensities must represent the intensity of the original incident beam). For the initial introduction of such an optical element, this section will maintain focus on the continuous wave classical analysis of beam splitting. While there are many methods of producing this effect, two specific methods are essential to material to be discussed later chapters: beam splitting due simply to a reflective probability ($R < 1$), and beam splitting based on the beam's angle of linear polarization.

2.3.1 Fresnel Equations

While a full derivation of the optical properties necessary to analyze the behavior of light incident on various types of boundaries (and with various boundary conditions) is far too broad to lend much relevance to the topics here, a condensed derivation of basic reflectance and transmittance properties will be presented in the form of the Fresnel equations. These arrive from considering a scenario of an electromagnetic wave incident on a boundary surface with the possibility of reflecting off of and(or) transmitting through the boundary. The basic relationship between the pertinent variables describing an electromagnetic field and its propagation are the electric field $\mathbf{E}$, magnetic field $\mathbf{B}$, and the wave vector $\mathbf{k}$. These are related by the following:

$$\hat{\mathbf{k}}\times\mathbf{E} = v\mathbf{B} \tag{2.9a}$$

$$\hat{\mathbf{k}}\cdot\mathbf{E} = 0 \tag{2.9b}$$

$$\mathbf{E}_i + \mathbf{E}_r = \mathbf{E}_t. \tag{2.9c}$$

where the last equation is arranged to preserve continuity of E across the boundary.

The right-handed system formed by $\hat{\mathbf{E}} \times \hat{\mathbf{B}} = \hat{\mathbf{k}}$ suggests separate perpendicular and parallel boundary conditions at the interface between the two media. Since the electric field in parallel to the surface at a flat angle of incidence ($\theta_i = \pi/w$), its parallel component should be directly continuous while the magnetic fields perpendicular component (pointing directly into the medium at the same angle of incidence) should be continuous directly. For a magnetic field parallel to the surface, $\mathbf{B}/\mu$ should be taken as continuous. Considering the case B is normal to the interface, the boundary condition for the parallel component becomes

$$-B_i/\mu_i \cos\theta_i + B_r/\mu_i \cos\theta_r = -B_t/\mu_t \cos\theta_t. \tag{2.10}$$

The reflection(transmission) amplitudes of the boundary are defined as the ratio of the reflected(transmitted) electric field to the incident electric field. Therefore, algebraically converting B to E above with a bit a manipulation yields the set of reflection and transmission coefficients for the scenario of an electric field parallel to the boundary. A bit of manipulation will show the perpendicular reflection and transmission amplitudes to be

$$r_\perp = \left(\tfrac{E_r}{E_i}\right)_\perp = \frac{n_i\cos\theta_i - n_t\cos\theta_t}{n_i\cos\theta_i + n_t\cos\theta_t} \tag{2.11}$$

$$t_\perp = \left(\tfrac{E_t}{E_i}\right)_\perp = \frac{2n_i\cos\theta_i}{n_i\cos\theta_i + n_t\cos\theta_t}. \tag{2.12}$$

In a similar manner, the problem can be analyzed while assuming the complementary situation: a magnetic field parallel to the interface layer. In this case, the tangential electric field is continuous, and the boundary condition is represented as

$$E_i\cos\theta_i - E_r\cos\theta_r = E_t\cos\theta_t \tag{2.13}$$

and the parallel boundary condition for the magnetic field are the same as above except with $+E_r$. Using the defined ratio, the parallel components of the reflection and transmission amplitudes are found to be

$$r_\parallel = \frac{n_t\cos\theta_i - n_i\cos\theta_t}{n_i\cos\theta_t + n_t\cos\theta_i} \tag{2.14}$$

$$t_\parallel = \frac{2n_i\cos\theta_i}{n_i\cos\theta_t + n_t\cos\theta_i}. \tag{2.15}$$

These relationships will be useful in deriving and discussing the physical implications of the various beam splitting techniques to be discussed.

2.3.2 Frustrated Total Internal Reflection

Total internal refraction is the possible result of light incident on a boundary whose index of refraction is lower than the medium from which the light is incident. As light passes from a medium of higher index of refraction to a medium of lower refractive index, it bends away from the boundary's normal line in accordance with Snell's Law ($n_i \sin\theta_i = n_t \sin\theta_t$). If the angle of incidence is greater than the critical angle (where the refracted angle is exactly 90°) for the specific media interface ($\theta_i >= \theta_c$), the light is reflected off the interface back into the material of greater refractive index.

Exactly at the critical point, the light is refracted and the boundary interface as a surface wave travelling exactly parallel to the surface but with a nonzero wave vector extending outward from the surface. Assuming the plane of interaction is the x-y plane and the boundary surface is the x-z plane, when $\theta_i = \theta_c$ total internal reflection will occur. However, in order for the boundary condition to hold true, the y component of the transmitted wave vector (in the surface wave) must be nonzero. In a general sense, the transmitted electric field is given by

$$\mathbf{E_t} = E_t e^{\mathbf{k_t}\cdot\mathbf{r}-\omega t} \tag{2.16}$$

where $\mathbf{k_t}\cdot\mathbf{r} = k_{tx}x + k_{ty}y = k_t(x\sin\theta_t + y\cos\theta_t)$. Using Snell's law, the wave numbers can be rewritten as

$$k_{ty} = \pm i k_t \left(\frac{n_i^2 \sin^2\theta_i}{n_t^2} - 1\right)^{1/2} \equiv \pm i\beta \tag{2.17}$$

$$k_{tx} = \frac{n_i k_t}{n_t} \sin\theta_i \tag{2.18}$$

where k_{ty} has been arranged suggestively in expectance of $n_i > n_t$. Plugging this back into our original expression for the transmitted electric field, we find that

$$\mathbf{E_t} = E_t e^{-\beta y} e^{i(n_i k_t x \sin\theta_i / n_t - \omega t)} \tag{2.19}$$

where $e^{+\beta y}$ has been neglected as it doesn't satisfy boundary conditions at infinity. Inspection of equation (2.19) shows that while the x-component represents the oscillation and propagation of the previously mentioned surface wave, the y-component (perpendicular to the boundary layer) now have the characteristic of an exponential decay but is, in fact, nonzero *above* the boundary layer. This component of the field is known as the evanescent wave.

The process of frustrated total internal reflection relies exactly on the existence of this perpendicular evanescent wave for successful operation. A beam splitter of this type uses two triangular shaped "prisms" of high refractive index, arranged in cubic fashion, with a small gap of lower refractive index between the two halves. A simple example of this would be two pieces of glass separated by a small air gap. Although, in practice, an air gap would be quite difficult to align and maintain proper separation, so commonly a low index bonding agent is employed.

The operation of such a beam splitter is literally two-fold. First, the relative indices of refraction and prism angle must be such that light can be normally incident on a side of the cube, while maintaining total internal reflection at the boundary layer. Also, this angle must be close enough to the critical angle - and the low index gap thin enough - to allow for the near-field evanescent wave to successfully traverse the gap into the next prism where it is propagated out normal to the opposite side. As with all other optical elements discussed, careful consideration must be made to their frequency dependence.

2.3.3 Birefringent Crystals

The last optical element of discussion is a beam splitter produced from a special arrangement of the previously discussed anisotropic, uniaxial crystal. In previous sections, this crystal was discussed in terms of its ability to produce a polarized output and its axially variable refractive index. Recall that for a negative uniaxial crystal, the optic index (fast axis) is aligned on the principal plane and has a lower index of refraction compared to the complementary slow axis. Recall also that linearly polarized light in an alignment centered between the two principal axes project equal amplitudes and displays an appreciably retarded propagation velocity along the slow axis relative to the fast axis.

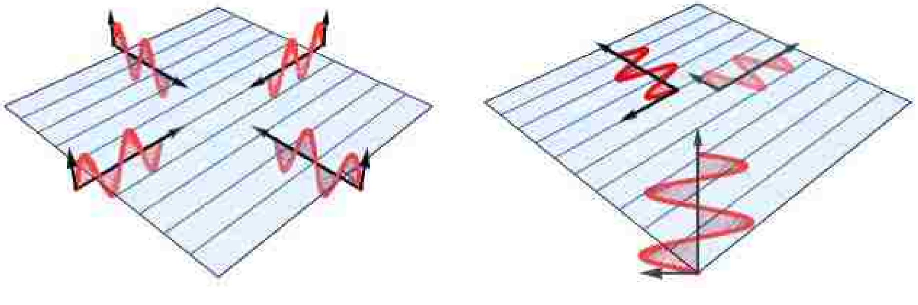

(a) (b)

FIGURE 2.5. A representation of the two orthogonal components of polarized light as they are incident on a uniaxial crystal with the principal plane shown and the optic axis represented by the horizontal lines. Figure (a) represents polarization oriented along the slow axis of the crystal, and figure (b) represents polarization aligned parallel to the fast axis.

The final relevant characteristic of such uniaxial crystals is their refractive properties. As seen in figure 2.5(a), when the electric field is aligned perpendicular to the principal plane, the crystal appears isotropic in absorption since there is no component of the electric field parallel to the optic axis. In this case, the light would be absorbed and reradiated evenly (i.e., isotropically) through the crystal and its path and polarization would remain unchanged. This emerging beam is referred to as the ordinary wave, or "o-wave."

Figure 2.5(b), however, is parallel to the plane while passing through the crystal. The figure represent the incident angle typical of natural uniaxial crystals. That is, the flat side of the crystal on which the light is incident is typically at an angle to the optic axis, even when the polarization is in the principal plane. In this case, the wave vector's orthogonal components experience different optical paths through the crystal. The component whose electric field vector is perpendicular to the optic axis travels at precisely the speed of the orientation depicted in figure 2.5(a). However, the component whose electric field vector is *parallel* to the optic axis experiences a different index of refraction and therefore a different optical velocity.

The key feature here is the anisotropy of the crystal between having the field vector parallel or perpendicular to the optic axis. For the negative uniaxial crystal to which our discussions have been limited, the component of the wave vector parallel to the optic axis travels optically *faster* through the crystalline lattice than does the orthogonal component of the wave vector for the same principally parallel polarization state. Having the velocity vector elongated along the path with parallel electric field (upper left wave in figure 2.5(b), causes the resultant wave vector to refract along this direction. This resultant, refracted wave is known as the extraordinary wave, or "e-wave."

The end result is a doubly refracted beam from this type of birefringent crystal. The polarization component which is everywhere orthogonal to the principal plane (o-wave) emerges with the same linear polarization (perpendicular to the principal plane) and in the same direction. The polarization component which is in the principal plane (e-wave) and at an angle to the optic axis emerges with the same polarization, but refracted in the direction of the optic axis due to a component of its wave vector travelling at a relatively faster speed along the fast axis. The result is a polarization “split” beam.

Known as a Wollaston prism, the typical polarizing beam splitter makes clever use of uniaxial symmetry phenomena while maintaining a physical arrangement similar to that of the frustrated total internal refracting beam splitter. It consists of two triangular shaped uniaxial crystals fitted together to make a cube; but unlike the FTIR beam splitter, there is no gap between the halves. Instead, the optic axes of the two pieces are aligned perpendicular to one another. As an example, let’s arbitrarily choose the horizontal optic axis to be the side of incidence. Just as with most of the other elements we’ve seen by now, when a linearly polarized beam is incident at a 45°angle there is maximum projection onto both the fast and slow axes, and the beam is equally separated into horizontal and vertical components.

For our example, let’s continue considering a negative uniaxial crystal (such as calcite) in which the optic axis is the fast axis with a lower relative index of refraction. So, the horizontal component is on the fast axis with relatively lower index of refraction and relatively higher velocity. Once the beam encounters the section of the crystal where the optic axis is suddenly vertical, it behaves as if it’s interacting with a boundary layer with an index shift. The horizontal, fast light sees the transition to the slow axis as an increase in relative index and refracts towards the normal of the boundary. The vertical, slow component, however, see the shift to the optic axis as a decrease in index and refracts away from the boundary. Upon exiting the crystal, both beams refract away from the normal and separate even farther, and the result is two independent beams of orthogonal polarization with feasible spatial separation.

Chapter 3
Quantization

3.1 Necessity for Quantization

At the turn of the century classical thermodynamics and statistical physics were becoming well established alongside the fundamentals of electricity and magnetism. It was at this time and with the new advances in physics that progress was taking shape on a century old fundamental question: what is the quantitative relationship between thermal and electromagnetic radiation. It had been noticed that a universal characteristic of all objects is to emit various wavelengths and intensities of light when heated to a certain temperature. Moreover, it was shown that this irradiative characteristic and associated temperature are irrespective of the physical composition of the heated. While several attempts were made to predict this behavior with the established classical methods, it was clear that the theory remained incomplete, and out of this necessity spawned the origins of quantum mechanics.

3.1.1 Black Body Radiation

The term black body refers to an object of any composition that neither reflects nor scatters any incident light, and is therefore perfectly black in appearance. The question then becomes, “How is a perfectly black, absorbing object to emit anything - especially light?” In 1859, Kirchoff answered this question by proving the equivalence of absorptive power, emitted power, and radiative power (per unit area per unit frequency) to be

$$e_f = J(f,T)A_f \tag{3.1}$$

where $J(f,T)$ is a universal function of power radiation (per unit area per unit frequency) for all objects. As suggested by its name, a perfect black body absorbs all incident light, and there for $A_f = 1$ and this relation becomes

$$e_f = J(f,T) \tag{3.2}$$

which directly links the temperature of a glowing hot object to it's radiative frequency and emitted power [2]. In practice, a perfect black body is formed from a cavity with a hole in the side such that any light that does manage to enter the cavity is scattered within and not re-emitted. This guarantees that any light see from the hole is solely the result of heating the body to radiative temperatures. In 1879, from experimental data, Stefan established the law of black body power emittance as a direct function of temperature by integrating over all frequencies of the emitted power. Namely,

$$e_t = \int_0^\infty e_f df = \sigma T^4 \tag{3.3}$$

shows that e_t is now the total emitted power for all frequencies, and σ is a constant now known as the Stefan-Boltzmann constant [2]. Similarly, when the black body radiation energy emitted per unit area per unit time integrated over a specific frequency range (λ to $\lambda + d\lambda$), radiation energy density of the cavity is given by

$$w(\lambda, T) = \frac{4E(\lambda, T)}{c}. \tag{3.4}$$

As a function of frequency, this relates to equation 3.2 as

$$u(f,T) = \frac{4J(f,T)}{c}. \tag{3.5}$$

In 1894, using the only available tools - classical thermodynamics - Wilhelm Wien showed that this energy density must have the form

$$w(\lambda T) = \frac{\gamma(\lambda, T)}{\lambda^5} \tag{3.6}$$

where γ is simply some function of the multiplied variables λT. It was known from experimental data that this function must have a maximum (at λ_{max}), so differentiation was applied as follows [4]:

$$u(f,T) = w(\lambda,T)\left|\frac{d\lambda}{df}\right| = w\left(\frac{c}{f},T\right)\frac{c}{f^2}. \tag{3.7}$$

Based on Maxwell's particle velocity distribution and experimental data, Wien's final formula for the spectral energy density was

$$u(f,T) = Af^3e^{-\beta f/T} \tag{3.8}$$

which was strongly supported by data up to the experimental capability of roughly $4\mu m$. Later results revealed grave inaccuracies at higher wavelengths, strengthening the necessity for a new theory to explain the behavior. [2]

3.1.2 Rayleigh-Jeans and Planck

Both the Rayleigh-Jeans method (classically based with continuous variables) and Planck method (quantum based with discrete allowed energy values) stem from a consideration of the density of modes of the cavity. Specifically, both begin with a product of the number of oscillators per unit volume (a count of the allowed degrees of freedom in the cavity with respect to the particle or waves present) times the average energy per oscillator,

$$u(f,T) = \bar{E}N(f)df. \tag{3.9}$$

3.1.2.1 Density of Modes

The calculation for the density of modes in the cavity is irrespective of the shape of the cavity, so the derivation assumes a cube to simplify boundary conditions. In accordance with Maxwell's equations, the electric field of the cavity contains a sinusoidal time dependent term and a spatial, time independent term that satisfies

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + k^2\right)E(x,y,z) = 0 \tag{3.10}$$

with

$$k = \frac{2\pi}{\lambda} = \frac{2\pi f}{c} \tag{3.11}$$

Assuming an orthogonal coordinate separable solution for the energy, $E = u(x)v(y)w(z)$, the problem simplifies to

$$\frac{d^2u}{dx_i^2} + k^2u = 0 \tag{3.12a}$$

$$\frac{d^2v}{dy_i^2} + k^2u = 0 \tag{3.12b}$$

$$\frac{d^2w}{dz_i^2} + k^2u = 0 \tag{3.12c}$$

$$k^2 = k_x^2 + k_y^2 + k_z^2$$

The above set of equations describes the motion of a simple harmonic oscillator in three independent, orthogonal dimensions. The solution to each dimension is known to be

$$u(x) = B\cos(k_x x) + C\sin(k_x x) \tag{3.13}$$

The boundary conditions of such a cubical cavity of length L require that the electric field vanish at $x = 0$ and $x = L$, which determines $u(x)$ is a function of $\sin k_x x$ only. Therefore, the energy is given by

$$E(x, y, z) = A\sin(k_x x)\sin(k_y y)\sin(k_z z) \tag{3.14}$$

$$k^2 = \frac{\pi^2}{L^2}(n_x^2 + n_y^2 + n_z^2)$$

Here, k describes the magnitude of a vector in reciprocal space with one standing wave per volume element $(\pi/L)^3$. The number of possible modes is described by the volume of a spherical shell in the positive quadrant (1/8 of the shell's total volume). Therefore, the density of modes for a region $k + dk$ is given by

$$\frac{N(k)dk}{V} = \left(\frac{k^2\pi/2dk}{(\pi/L)^3}\right)2 = \frac{Vk^2dk}{\pi^2} \tag{3.15}$$

The factor of 2 is included since we are counting degrees of freedom of electric fields, and there are also two orthogonal polarizations for each oscillator that the cavity allows. Using the fact that $k = \omega/c = 2\pi f/c$, we obtain [2]

$$\frac{N(f)df}{V} = \frac{\omega^2}{\pi^2 c^3} d\omega = \frac{8\pi f^2}{c^3} df \tag{3.16}$$

or using $f = c/\lambda$,

$$N(\lambda) d\lambda = \frac{8\pi}{\lambda^4} d\lambda \tag{3.17}$$

3.1.2.2 Classical Method

It is at this point that the techniques adopted by Rayleigh and Jeans differ from those employed by Planck. Rayleigh and Jeans were considering the density of modes of the classical electromagnetic fields in the cavity directly, while Planck was focusing on the origination of the waves: the electric oscillations of the particles composing the cavity walls. Let us first focus on the classically derived continuous Rayleigh-Jeans case.

Rayleigh and Jeans took the assumption that the electromagnetic waves in resonance within the cavity were an accurate representation of the temperature of the cavity itself, given it was in thermal equilibrium. They adopted a classical thermodynamic principle to describe the average energy of the oscillators that are in an energy state (E) above some threshold energy level (E_0). This energy probability is described by the Boltzmann distribution

$$P(E) = P_0 e^{-(E-E_0)/k_B T} \tag{3.18}$$

where P_0 is the probability of the system having minimal energy. The normalized average energy of the continuous system is then given by

$$\bar{E} = \frac{\int_0^\infty E P(E) dE}{\int_0^\infty P(E) dE} = \frac{\int_0^\infty E e^{-E/k_B T} dE}{\int_0^\infty e^{-E/k_B T} dE} = k_B T \tag{3.19}$$

Now the spectral energy density from equation (3.9) is

$$u(f,T)df = \frac{8\pi f^2}{c^3} k_B T df \tag{3.20a}$$

$$w(\lambda,T)d\lambda = \frac{8\pi}{\lambda^4} k_B T d\lambda \tag{3.20b}$$

This expression clearly doesn't fit the data at shorter wavelengths, but is a reasonable estimate at longer wavelengths where Wien's expression previously failed.

3.1.2.3 Quantum Method

Planck, having known from experimental data that the curve was T dependent at long wavelengths and that Wien's formula worked well for shorter wavelengths, set out to interpolate the two. As mentioned earlier, he analyzed the system in a similar manner to Rayleigh and Jeans, but took a quantized approach as to the values allowed by the oscillating charges in the cavity walls that produce the measured electric fields. As is standard practice, he adopted the same expression for the density of modes for the cavity and the Boltzmann probability distribution, but discretized the spectral energy density calculation. As such, the formula for the average energy is given by

$$\bar{E} = \frac{\sum_{n=0}^{\infty}(nhf)P_0 e^{-nhf/k_B T}}{\sum_{n=0}^{\infty} P_0 e^{-nhf/k_B T}} \tag{3.21}$$

Recognizing this as a modified geometric series, the energy density can be rewritten,

$$\bar{E} = hf(1 - e^{-hf/k_B T}) \sum_{n=0}^{\infty} n e^{-nhf/k_B T} \tag{3.22}$$

The remaining sum is merely the derivative of a geometric series, and is evaluated similarly yielding the following:

$$\bar{E} = \frac{hf e^{-hf/k_B T}}{1 - e^{-hf/k_B T}} = \frac{hf}{e^{hf/k_B T} - 1} \tag{3.23}$$

Plugging this into equation (3.9), the final quantum based Planck distribution is

$$\begin{aligned} u(f,T)df &= \frac{8\pi f^3}{c^3}\left(\frac{h}{e^{hf/k_BT}-1}\right)df \\ w(\lambda,T)d\lambda &= \frac{8\pi c}{\lambda^5}\left(\frac{h}{e^{hc/\lambda k_bT}-1}\right)d\lambda \end{aligned} \tag{3.24}$$

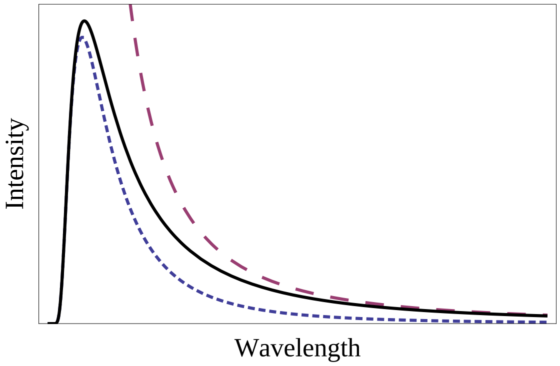


FIGURE 3.1. An illustration of the trend of the trend of black body curves in arbitrary units (all constants have been set to 1). The dotted blue line represents Wien's result which is a good approximation a high frequencies. The dashed red line represents the Rayleigh-Jeans result, which is a good approximation at low frequencies. The solid black line represents the Planck distribution. Only the Planck distribution correctly describes the behavior of black body radiation at all frequencies by including the effects of quantum mechanics.

Just as should be expected, Planck's distribution can be simplified in limiting cases to recover the Wien and Rayleigh-Jeans equations. In the high frequency limit, $hf/k_BT \gg 1$, and the exponential term can be approximated as

$$\frac{1}{e^{hf/k_BT}-1} \approx e^{-hf/k_BT} \tag{3.25}$$

and the energy density becomes

$$u(f,T) \approx \frac{8\pi hf^3}{c^3}e^{-hf/k_BT} \tag{3.26}$$

This is precisely the Wien equation with $A = \frac{8\pi h}{c^3}$ and $\beta = h/k_B$. This is verified graphically in figure 3.1.2.3 where the dashed blue line represents the Wien equation and it's close comparison the Planck's distribution for small wavelengths.

Similarly, in the low frequency regime, $hf/k_BT \ll 1$, the exponential can be approximated by a Taylor expansion to first order,

$$\frac{1}{e^{hf/k_BT}-1} = \frac{1}{1+\frac{hf}{k_BT}+\cdots-1} \approx \frac{k_BT}{hf} \tag{3.27}$$

in which case the energy density is approximated by

$$u(f,T) \approx \frac{8\pi f^2}{c^3}k_BT \tag{3.28}$$

This too can be visually verified from the figure 3.1.2.3 where the dashed red line is the Rayleigh-Jeans equation. In the limit of large wavelength (low frequency regime), this curve asymptotically approaches the Planck distribution and the expected T/f trend observed experimentally [2].

3.2 Electromagnetic Field Quantization

Thus far, most discussions have related to the classical behavior of light as a wave; however, as seen from the derivation of Planck's distribution, it is often quite imperative to develop a fully quantized picture of the situation under analysis. To do so we need to establish the origin of quantization: bounded peiodicity.

3.2.1 EM Field Operators

Just as was done in section 3.1.2.1, we should establish an expansion of all physically allowed electric field resonance modes within the boundary. The one dimensional electric field can be separated into the product of a time dependent component and a oscillating spatial component. Following the procedure of section 3.1.2.1, the solution for a single mode oscillating electric field propagating in the $\hat{z}$ direction is found to be

$$E_x(z,t) = \sqrt{\frac{2\omega^2}{V\epsilon_0}}q(t)\sin(kz) \tag{3.29}$$

Or, generalizing to a multimode expansion, this can be expressed as

$$E_x(z,t) = \sum_j \sqrt{\frac{2\omega_j^2}{V\epsilon_0}} q_j(t)\sin(k_j z) \tag{3.30}$$

where V is the volume of the cavity, $q(t)$ is the time dependent separable term, and $\omega = cn\pi/L$ with $n = 1, 2, ...$ Using Maxwell's equations

$$\nabla \times \mathbf{E} = \frac{-\partial \mathbf{B}}{\partial t} \tag{3.31}$$

$$\nabla \times \mathbf{B} = \mu_0\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \tag{3.32}$$

an expression for the multimode magnetic field is found to be

$$B_y(z,t) = \sum_j \frac{\mu_0\epsilon_0}{k_j}\sqrt{\frac{2\omega_j^2}{V\epsilon_0}} \dot{q}_j(t)\cos(k_j z) \tag{3.33}$$

The energy for this one dimensional, classical representation of the field is given by the following Hamiltonian:

$$H = \frac{1}{2}\int dV \left[\epsilon_0 E_x^2(z,t) + \frac{1}{\mu_0}B_x^2(z,t)\right] \tag{3.34}$$

Plugging in the expressions for the electric field (3.30) and the magnetic field (3.33), the Hamilton simplifies to the familiar expression for a harmonic oscillator as a function of position and momentum.

$$H = \frac{1}{2}\sum_j (p_j^2 + \omega_j^2 q_j^2) \tag{3.35}$$

Here, the fact that the canonical momentum is the time derivative of position ($\dot{q} = p$) has been utilized. From this expression, it is clear that the energy for an oscillating electromagnetic field (in the jth state) is identical to that of the mechanical simple harmonic oscillator with unit mass.

Having identified the basic canonical variables for the classical system, it should be recognized that they are directly analogous to their quantized operator counterparts, $\hat{q}$ and $\hat{p}$. The difference here is that the quantized operators are expanded into N dimensional Hilbert space via a matrix formalism. The canonical commutation relation in n-dimensional operator form is given by

$$\begin{aligned}[\hat{q}_i, \hat{p}_j] &= i\hbar\delta_{ij}\hat{\mathbb{I}} \\ [q_i, q_j] &= [p_i, p_j] = 0\end{aligned} \tag{3.36}$$

It follows directly that the analogous operator expressions for the electric and magnetic fields are [6, 5]

$$\hat{E}_x(z,t) = \sum_j \sqrt{\frac{2\omega_j^2}{V\epsilon_0}}\hat{q}_j(t)\sin(k_j z) \tag{3.37a}$$

$$\hat{B}_y(z,t) = \sum_j \frac{\mu_0\epsilon_0}{k_j}\sqrt{\frac{2\omega_j^2}{V\epsilon_0}}\hat{p}_j(t)\cos(k_j z) \tag{3.37b}$$

and the resulting Hamiltonian operator is now

$$\hat{H} = \frac{1}{2}\sum_j(\hat{p}_j^2 + \omega_j^2\hat{q}_j^2) \tag{3.38}$$

3.2.2 Creation and Annihilation Operators

When performing calculation on quantized states, it is more convenient and conventional to introduce the canonical transformation from $\hat{p}$ and $\hat{q}$ to the creation and annihilation operators $\hat{a}$ and $\hat{a}^\dagger$, respectively. Redefining the position and momentum operators as a sum and difference of these new canonical operators gives

$$\hat{q}_j = (\hat{a}_j + \hat{a}_j^\dagger)\sqrt{\frac{\hbar}{2\omega_j}} \tag{3.39a}$$

$$\hat{p}_j = (\hat{a}_j - \hat{a}_j^\dagger)\sqrt{\frac{\hbar}{2\omega_j}} \tag{3.39b}$$

From this, it is straightforward to obtain the creation and annihilation operators directly.

$$\hat{a}_j = \frac{\omega_j\hat{q}_j + i\hat{p}_j}{\sqrt{2\hbar\omega_j}} \tag{3.40a}$$

$$\hat{a}_j^\dagger = \frac{\omega_j\hat{q}_j - i\hat{p}_j}{\sqrt{2\hbar\omega_j}} \tag{3.40b}$$

It is important to note that while the position and momentum operator are Hermitian observables, the creation and annihilation operators are non-Hermition and as such do not directly represent measured observable quantities. The choice for the definition of these operators tends to eliminate constants when calculating, and simplifies their commutation relation to

$$[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij} \tag{3.41a}$$

$$[\hat{a}_i, \hat{a}_j] = [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0 \tag{3.41b}$$

Note that, unlike the position and momentum operators, the results of these commutations are plain numbers and not matrices. From this new set of canonical operators, the electric and magnetic field operators become [6, 5]

$$\hat{E}_x(z,t) = \sum_j \mathcal{E}_j(\hat{a}_j + \hat{a}_j^\dagger)\sin(k_j z) \tag{3.42a}$$

$$\hat{B}_y(z,t) = -i\sum_j \mathcal{B}_j(\hat{a}_j - \hat{a}_j^\dagger)\cos(k_j z) \tag{3.42b}$$

where

$$\mathcal{E}_j = \sqrt{\frac{\hbar\omega_j}{\epsilon_0 V}} \tag{3.43a}$$

$$\mathcal{B}_j = \frac{\mu_0}{k_j}\sqrt{\frac{\epsilon_0\hbar\omega_j^3}{V}} \tag{3.43b}$$

In terms of these new operators, the Hamiltonian is now

$$\hat{H} = \sum_j \hbar\omega_j\left(\hat{a}_j^\dagger\hat{a}_j + \frac{1}{2}\right) \tag{3.44}$$

which is again analogous with a mechanical simple harmonic oscillator.

3.2.3 Time Dependence

In section 3.2.1, it was stated that the spatial and temporal components of the electric and magnetic fields are separable. All derivations and discussions since have been pertaining to the spatial portion of the field and the associated operators. In order to add time dependence to the established operators, we refer to the general form of Heisenberg's equation

$$\frac{d\hat{O}}{dt} = \frac{i}{\hbar}[\hat{H}, \hat{O}] \tag{3.45}$$

where $\hat{O}$ is any arbitrary (time independent) operator. When applied to the annihilation operator and using the commutation relation defined in section 3.2.2, we get

$$\begin{aligned} \frac{d\hat{a}}{dt} &= \frac{i}{\hbar}[\hat{H}, \hat{a}] \\ &= i\omega(\hat{a}^\dagger \hat{a}\hat{a} - \hat{a}\hat{a}^\dagger \hat{a}) \\ &= i\omega[\hat{a}^\dagger, a]\hat{a} = i\omega\hat{a} \\ \frac{d\hat{a}}{dt} &= -i\omega\hat{a} \end{aligned} \tag{3.46}$$

where the commutation relation, $[\hat{a}^\dagger, \hat{a}] = -1$, has been used. Given that differentiation of a variable that returns $-i\omega$ times the original variable, the variable itself must contain an exponential. Hence, the time dependence of the annihilation operator is contained within an exponential. From this, the creation operator can also be found by taking the complex conjugate.

$$\hat{a}(t) = \hat{a}(0)e^{-i\omega t} \tag{3.47a}$$

$$\hat{a}^\dagger(t) = \hat{a}(0)e^{i\omega t} \tag{3.47b}$$

3.2.4 Free Space Generalization

Derivations thus far have been confined to a one dimensional analysis with hard boundaries where the field must vanish. These restrictions can be lifted to generalize the bound cavity to an unbound area of free space. In this approach, the vector potential is expressed as a superposition of classical free space plane waves given by [6]

$$\mathbf{A}(\mathbf{r},t)=\sum_{\mathbf{k}}\hat{\epsilon}_{\mathbf{k}}\left[\alpha_{\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{r}}+\alpha_{\mathbf{k}}^{*}(t)e^{-i\mathbf{k}\cdot\mathbf{r}}\right] \tag{3.48}$$

To satisfy the periodic boundary condition in a region of free space, the wave vector must be multiples of 2π

$$\mathbf{k}=\frac{2\pi}{L}(m_x,m_y,m_z) \tag{3.49}$$

where m_i is an integer. Assuming no local charge density or current density reduces Maxwell's divergence formulae to

$$\nabla\cdot\mathbf{E}=0 \tag{3.50a}$$

$$\nabla\cdot\mathbf{B}=0 \tag{3.50b}$$

which specifies transversality condition for the vector potential as

$$\mathbf{k}\cdot\hat{\epsilon}_{\mathbf{k}}=0 \tag{3.51a}$$

$$\epsilon_{\mathbf{k}1}\times\epsilon_{\mathbf{k}2}=\hat{\mathbf{k}} \tag{3.51b}$$

where $\epsilon_{\mathbf{k}1}$ and $\epsilon_{\mathbf{k}2}$ represent the orthogonal polarization unit vectors. Since the vector potential must satisfy the wave equation,

$$\frac{d^2\alpha_{\mathbf{k}}}{dt^2}+\omega_k^2\alpha_{\mathbf{k}}=0 \tag{3.52}$$

the time dependence and spatial dependence are separable as

$$\alpha_{\mathbf{k}}(t)=\alpha_{\mathbf{k}}e^{-i\omega_k t} \tag{3.53}$$

Now, the electric field can be calculated from the vector potential.

$$\mathbf{E}(\mathbf{r},t)=-\frac{\partial\mathbf{A}(\mathbf{r},t)}{\partial t}$$

$$\mathbf{E}(\mathbf{r},t)=i\sum_{\mathbf{k}}\omega_k\hat{\epsilon}_{\mathbf{k}}\left[\alpha_{\mathbf{k}}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}-\alpha_{\mathbf{k}}^{*}e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}\right] \tag{3.54}$$

Similarly, the magnetic field can also be calculated from the vector potential as

$$\mathbf{B}(\mathbf{r},t) = \nabla \times \mathbf{A}(\mathbf{r},t) \tag{3.55}$$

$$\mathbf{B}(\mathbf{r},t) = \frac{i}{c}\sum_{\mathbf{k}} \omega_k(\hat{\mathbf{k}} \times \hat{\epsilon}_{\mathbf{k}}) \left[\alpha_{\mathbf{k}} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} - \alpha^*_{\mathbf{k}} e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}\right] \tag{3.56}$$

Plugging these expressions into the Hamiltonian yields

$$H = \frac{1}{2}\int dV \left[\epsilon_0 \mathbf{E}\cdot\mathbf{E} + \frac{1}{\mu_0}\mathbf{B}\cdot\mathbf{B}\right] \tag{3.57}$$

$$= 2\epsilon_0 V \sum_{\mathbf{k}} \omega_k^2 \alpha_{\mathbf{k}} \alpha^*_{\mathbf{k}} \tag{3.58}$$

Choosing the amplitude $\alpha_{\mathbf{k}}$

$$\alpha_{\mathbf{k}} = \frac{1}{2\omega_k\sqrt{\epsilon_0 V}}[\omega_k q_{\mathbf{k}} + i p_{\mathbf{k}}] \tag{3.59}$$

reduces the Hamiltonian to equations (3.38)

$$H = \frac{1}{2}\sum_{\mathbf{k}} \left(p^2_{\mathbf{k}} + \omega_k^2 q_{\mathbf{k}}\right) \tag{3.60}$$

which is the known correct value for the energy of an oscillator. This is also the form of the Hamiltonian upon substitute the definition for the creation and annihilation operators. Therefore, juxtaposing our definition of $\hat{\alpha}_{\mathbf{k}}$ (in quantized operator form) and the definition for the annihilation operator,

$$\hat{\alpha}_{\mathbf{k}} = \frac{\omega_k \hat{q}_{\mathbf{k}} + i\hat{p}_{\mathbf{k}}}{2\omega_k\sqrt{\epsilon_0 V}} \tag{3.61}$$

$$\hat{a}_{\mathbf{k}} = \frac{\omega_k \hat{q}_{\mathbf{k}} + i\hat{p}_{\mathbf{k}}}{\sqrt{2\hbar\omega_k}} \tag{3.62}$$

it is immediately clear that the value for the amplitude of the vector potential is

$$\hat{\alpha}_{\mathbf{k}} = \sqrt{\frac{\hbar}{2\omega_k \epsilon_0 V}}\hat{a}_{\mathbf{k}}. \tag{3.63}$$

Hence, the free space electric and magnetic fields are [6, 5]

$$\hat{\mathbf{E}}(\mathbf{r},t) = i\sum_{\mathbf{k}} \sqrt{\frac{\hbar\omega_k}{2\epsilon_0 V}} \left[\hat{a}_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{r}} - \hat{a}^\dagger_{\mathbf{k}}(t) e^{-i\mathbf{k}\cdot\mathbf{r}}\right] \hat{\epsilon}_{\mathbf{k}} \tag{3.64a}$$

$$\hat{\mathbf{B}}(\mathbf{r},t) = \frac{i}{c}\sum_{\mathbf{k}} \sqrt{\frac{\hbar\omega_k}{2\epsilon_0 V}} \left[\hat{a}_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{r}} - \hat{a}^\dagger_{\mathbf{k}}(t) e^{-i\mathbf{k}\cdot\mathbf{r}}\right] (\hat{\mathbf{k}} \times \hat{\epsilon}_{\mathbf{k}}) \tag{3.64b}$$

Chapter 4
Quantum Optics

4.1 Quantum Optical States

A quantum state, by definition, refers to any combination of orthogonal components of quantized value. Nevertheless, there exists an infinite array of such states – each with its own specific set of characteristics and properties. In the interest of topical focus, this section aims to detail two specific classes of states: quantum-like photon number states and classical-like coherent states. These are arguable the most commonly encountered quantum optical states and exemplify the broad contrasts that can exist among quantum state characteristics.

4.1.1 Photon Number States

Fock states (or "number states") are quantum states containing a discrete and definite photon number. As mentioned in previous sections, the conventional choice of canonical operators in optical physics are the creation and annihilation operators (or typically known as raising and lower operators in atomic physics). This set of operators has earned their names from their mathematical ability to describe the energy transfer between atomic orbitals and the ambient radiation field. If a high intensity photon interacts with a lower energy atom, the photon can be "annihilated" from the radiation field while simultaneously raising the energy level of the atom. Likewise, high energy atoms have to ability to undergo an emission process in which its energy level is lowered while "creating" an ambient photon with equivalent energy. Henceforth, all mention to these operators will be in reference to the photonic energy of the radiation field.

A number state in quantum optics is labeled as a vector, or "ket" in Dirac notation, labeled with a number representing the number of photons present. From this, the function of the creation ($\hat{a}^{\dagger}$) and annihilation ($\hat{a}$) operators is to increase or decrease this number by a single photon, respectively. Specifically [7],

$$\hat{a_k}^{\dagger}|n_k\rangle = \sqrt{n_k+1}|n_k+1\rangle \tag{4.1a}$$

$$\hat{a_k}|n_k\rangle = \sqrt{n_k}|n_k-1\rangle \tag{4.1b}$$

$$\hat{a_k}|0\rangle = 0|0\rangle$$

where the subscript is included to explicitly demonstrator that these are single mode operators of wave vector $\mathbf{k}$ and polarization $\hat{\epsilon_{\mathbf{k}}}$. The commutation relation for this set of operators is the following bose commutation:

$$[\hat{a}, \hat{a}^{\dagger}] = 1.$$

Applying the creation operation in succession can create any arbitrary number state from the vacuum state as

$$|n_k\rangle = \frac{(\hat{a}_k^{\dagger})^{n_k}}{\sqrt{n_k!}}|0\rangle \tag{4.2}$$

Other properties of these states include orthonormality

$$\langle n_k|m_{k'}\rangle = \delta_{mn}\delta_{kk'} \tag{4.3}$$

and completeness

$$\sum_{n_k=0}^{\infty}|n_k\rangle\langle n_k| = 1. \tag{4.4}$$

In tandem, these operators for the single mode number operator for Fock states.

$$\hat{n}_k|m\rangle_{k'} = \hat{a}_k^{\dagger}\hat{a}_k|m\rangle_{k'} = m\delta_{kk'}|m\rangle_{k'} \tag{4.5}$$

While it is explicitly required in many circumstances, the mode label k is typically implied and will not appear in the general discussions that follow. According to the Hamiltonian (3.44), the energy eigenvalue equation is

$$\hat{H}|n\rangle = \hbar\omega\left(\hat{a}^{\dagger}\hat{a} + \frac{1}{2}\right)|n\rangle = E_n|n\rangle \tag{4.6}$$

Applying the creation operator to both sides of this expression from the left hand side gives the quantized energy value added to the field via the creation of a single photon [6].

$$\hbar\omega\left(\hat{a}^{\dagger}\hat{a}^{\dagger}\hat{a} + \frac{1}{2}\hat{a}^{\dagger}\right)|n\rangle = D_n\hat{a}^{\dagger}|n\rangle$$

Using the commutation relation (3.41a), the operators on the left hand side must be anti-normally ordered to allow factoring of the creation operator to the right.

$$\hbar\omega\left(\hat{a}^{\dagger}\hat{a}\hat{a}^{\dagger} - \hat{a} + \frac{1}{2}\hat{a}^{\dagger}\right)|n\rangle = (E_n + \hbar\omega)\hat{a}^{\dagger}|n\rangle$$

$$\hbar\omega\left(\hat{a}^{\dagger}\hat{a} + \frac{1}{2}\right)(\hat{a}^{\dagger}|n\rangle) = (E_n + \hbar\omega)(\hat{a}^{\dagger}|n\rangle) \tag{4.7}$$

Similarly, applying the annihilation operator to the energy eigenvalue equation yields

$$\hat{H}(\hat{a}|n\rangle) = (E_0 - \hbar\omega)(\hat{a}|n\rangle) \tag{4.8}$$

This shows that $(\hat{a}|n\rangle)$ is an eigenstate of the Hamiltonian with energy eigenvalue $E_0 - \hbar\omega$. Special care must be take in the case of the vacuum state as equation (4.8) breaks down for $\hat{a}|0\rangle = 0$ without resolving a minimum energy eigenvalue $(E_0 > 0)$. Therefore, the eigenvalue equation for the ground state is given by [6]

$$\hat{H}|0\rangle = \hbar\omega\left(\hat{n} + \frac{1}{2}\right)|0\rangle = \frac{1}{2}\hbar\omega|0\rangle \tag{4.9}$$

where the minimum energy eigenvalue $E_0 = \hbar\omega/2$ is known as the zero-point energy.

4.1.2 Coherent States

In quantum optics, the coherent state is analogous to a classical state. While still technically a quantum state containing discrete number state components, it contains an infinite number of Fock states and is weighted smoothly about a mean photon number. Purely classical fields are described simultaneously by a definite amplitude and a definite phase. Due to the Heisenberg uncertainty principle, this cannot be true in a quantum description of the field. On the other hand, purely quantum mechanical number states have no direct phase association, and therefore are considered to have a uniform phase distribution from 0 to 2π radians. Coherent states, however, are described by a mean photon number possess a specific phase association while maintaining a minimum uncertainty between the two quantities as defined by the product of the uncertainty in amplitude (related to the mean photon number) and the uncertainty in phase [7].

Furthermore, the behavior of these states can be modeled as a "displaced" minimum energy state simple harmonic oscillator (zero-point energy state of eigenvalue $\hbar\omega/2$), whose displacement in turn oscillators about the potential well of a harmonic oscillator [5].

The coherent state can be realized out of the classical scenario where oscillating current density $\mathbf{J}$ provides the radiation source. The vector potential given by equation (3.48) can be modified as $\mathbf{A}(\mathbf{r},t) \to -i\mathbf{A}(\mathbf{r},t)$ which then produces real expressions for the electric and magnetic fields given by equations (3.64a) and (3.64b). The free space Hamiltonian that couples the current source to the vector potential is given by

$$H(t) = \int \mathbf{J}(\mathbf{r},t)\cdot \mathbf{A}(\mathbf{r},t) d^3r \tag{4.10}$$

For this redefinition of the vector potential, the time dependent Schrodinger equation can be integrated as follows

$$\frac{d}{dt}|\psi(t)\rangle = -\frac{i}{\hbar}H|\psi(t)\rangle$$

$$|\psi(t)\rangle = \exp\left[-\frac{i}{\hbar}\int_0^t dt' H(t')\right]|\psi(0)\rangle.$$

Relabeling the ground state as the vacuum state ($|\psi(0)\rangle = |0\rangle$) and performing the integration over a single mode $\mathbf{k}$ gives

$$|\psi(t)\rangle_{\mathbf{k}} = \exp(\alpha_{\mathbf{k}}\hat{a}^{\dagger}_{\mathbf{k}} - \alpha^*_{\mathbf{k}}\hat{a}_{\mathbf{k}})|0\rangle_{\mathbf{k}} \equiv |\alpha_{\mathbf{k}}\rangle \tag{4.11}$$

where $|\alpha\rangle$ is the conventional labeling for a coherent state and $\alpha_{\mathbf{k}}$ is given by [5]

$$\alpha_{\mathbf{k}} = \frac{1}{\hbar\omega_k}\mathcal{E}_{\mathbf{k}}\int_0^t dt'\int d\mathbf{r}\left[\hat{\epsilon}_{\mathbf{k}}\cdot \mathbf{J}_{\omega}(\mathbf{r},t)e^{i(\omega_k t' - \mathbf{k}\cdot\mathbf{r})}\right]. \tag{4.12}$$

As mentioned earlier, a coherent is equivalent to the displacement of the ground state energy wave function $|\psi(0)\rangle = |0\rangle$. Apparently, according to (4.11), the definition of the displacement operator is

$$\hat{D}(\alpha) = \exp(\alpha\hat{a}^{\dagger} - \alpha^*\hat{a}) \tag{4.13}$$

Under the conditions that a commutator commutes with both of its elements (i.e., $[A,[A,B]] = [B,[A,B]] = 0$), the Campbell-Baker-Hausdorff formula is

$$e^{A+B} = e^A e^B e^{-[A,B]/2}. \tag{4.14}$$

Applying this to the displacement operator,

$$D(\alpha) = e^{-|\alpha|^2/2}e^{\alpha a^{\dagger}}e^{-\alpha^* a} \tag{4.15}$$

where the operator "hat" notation has been assumed. Properties of the displacement operator include [7]

$$D^{\dagger}(\alpha) = D^{-1}(\alpha) = D(-\alpha) \tag{4.16a}$$

$$D^{\dagger}(\alpha)aD(\alpha) = a + \alpha \tag{4.16b}$$

$$D^{\dagger}(\alpha)a^{\dagger}D(\alpha) = a^{\dagger} + \alpha^{*} \tag{4.16c}$$

Returning to equation (4.11), it is now clear that the coherent state can be expressed in terms of the vacuum state as

$$|\alpha\rangle = D(\alpha)|0\rangle. \tag{4.17}$$

Using the above properties, it can be shown that coherent states are eigenstates of the annihilation operator.

$$\begin{aligned}
\hat{a}|\alpha\rangle &= \hat{a}\hat{D}(\alpha)|0\rangle \\
&= \hat{D}(\alpha)\hat{D}^{\dagger}(\alpha)\hat{a}\hat{D}(\alpha)|0\rangle \\
&= \hat{D}(\alpha)(\hat{a} + \alpha)|0\rangle = \alpha\hat{D}(\alpha)|0\rangle \\
\hat{a}|\alpha\rangle &= \alpha|\alpha\rangle \qquad (4.18)
\end{aligned}$$

In a general sense, any arbitrary single-mode state can be expressed as a finite (or infinite) sum of discrete number states (since the infinite set of such states is orthonormal as discussed above). Specifically, this is given by

$$|\psi\rangle = \sum_{n} c_n|n\rangle \tag{4.19}$$

Clearly, the coherent state should be no exception. By projecting equation (4.18) onto a number state we get

$$(\langle n|\hat{a})|\alpha\rangle = \langle n|(\hat{a}|\alpha\rangle) \tag{4.20}$$

$$\sqrt{n+1}\langle n+1|\alpha\rangle = \alpha\langle c|\alpha\rangle. \tag{4.21}$$

From equation (4.2), it follows that

$$\langle n|\alpha\rangle = \frac{\alpha^{n}}{\sqrt{n!}}\langle 0|\alpha\rangle. \tag{4.22}$$

As was postulated in equation (4.19), and using the fact that the number states form a complete set (equation (4.4)), the coherent state can be expanded in terms of number states with the above state coefficients. Specifically,

$$|\alpha\rangle = \sum_n |n\rangle\langle n|\alpha\rangle = \langle 0|\alpha\rangle \sum_n \frac{\alpha^n}{\sqrt{n!}}|n\rangle. \tag{4.23}$$

To ensure normalization of such an expansion, we take the mod-square of the state vector as follows:

$$\begin{aligned}
|\langle\alpha|\alpha\rangle|^2 &= |\langle 0|\alpha\rangle|^2 \sum_n \frac{|\alpha|^{2n}}{n!}|n\rangle \\
&= \left|\langle 0|\hat{D}(\alpha)|0\rangle\right|^2 e^{|\alpha|^2} \\
&= \left|e^{-|\alpha|^2/2}\langle 0|e^{\alpha\hat{a}^\dagger}e^{-\alpha^*\hat{a}}|0\rangle\right|^2 e^{|\alpha|^2} \\
|\langle\alpha|\alpha\rangle|^2 &= 1.
\end{aligned} \tag{4.24}$$

Therefore, in terms of a number state expansion, the coherent state is defined as

$$|\alpha\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}|n\rangle. \tag{4.25}$$

As mentioned at the beginning of this section, the coherent state is a well formed distribution of states described by an mean photon number. Specifically, this distribution is the Poissonian distribution

$$P(n) = |\langle n|\alpha\rangle|^2 = \frac{|\alpha|^{2n}e^{-|\alpha|^2}}{n!} \tag{4.26}$$

where $\langle\hat{n}\rangle = \langle\alpha|\hat{a}^\dagger\hat{a}|n\rangle = |\alpha|^2$ is the mean photon number [7].

4.2 Quantum Interference of Photons

Quantum mechanical photonic interaction – specifically, the use of beam splitters to create quantum effects – has become the basis for methods of interferometry and quantum measurement. When dealing with discrete photons and other quantum states of light, the beam splitter behaves quite differently from the classical description of section 2.3. Where section 2.3 described the effects in terms of continuous wave electromagnetic fields and associated wave vectors, the establishment of more formal governing constraints on a beam splitter requires an analysis on a quantum level.

4.2.1 Beam Splitter Derivation

(a) (b)

FIGURE 4.1. Representation of beam splitter transformation. The annihilation operators transform from $\hat{a}$ to $\hat{b}$. Figure (a) shows a source incident on the top mode of the beam splitter and transforms the jones vector in unprimed notation, while figure (b) illustrates incidence from the bottom and transforms to primed notation. Here, r and t are completely general coefficients.

A lossless beam splitter is typically described by an SU(2) Lie algebraic operator. This derivation will adopt the jones vector notation of figure 4.1. For initial discussions, the input column vector will be unimodal and represent the orthogonal inputs generic amplitudes (i.e., the dimensionless unity implies the input mode of the source regardless of the specific source or state type). In its most general form, the matrix can is constructed of the two column vectors of figure 4.1, which is given by

$$\begin{bmatrix}\hat{b}_1\\ \hat{b}_2\end{bmatrix} = \begin{bmatrix}r & t'\\ t & r'\end{bmatrix}\begin{bmatrix}\hat{a}_1\\ \hat{a}_2\end{bmatrix}. \tag{4.27}$$

Each of the beam splitter coefficients is assumed to have an intrinsic phase association, such that [8]

$$\begin{aligned} r &= |r|e^{i\phi_r} & \qquad r' &= |r'|e^{i\phi_{r'}} \\ t &= |t|e^{i\phi_t} & \qquad t' &= |t'|e^{i\phi_{t'}} \end{aligned} \tag{4.28}$$

By extracting the two equations for $\hat{b}_1$ and $\hat{b}_2$ from equation (4.27) and using the boson commutation relation applied to the output annihilation operator,

$$[\hat{b}_i, \hat{b}_j^\dagger] = \hat{b}_i\hat{b}_j^\dagger - \hat{b}_j^\dagger\hat{b}_i = \delta_{ij},$$

the following set of restraints can be derived [6]:

$$\begin{aligned} |r|^2 + |t|^2 &= 1 \\ |r'|^2 + |t'|^2 &= 1 \end{aligned} \tag{4.29a}$$

$$\begin{aligned} rt'^* + tr'^* &= 0 \\ r't^* + t'r^* &= 0 \end{aligned} \tag{4.29b}$$

$$\begin{aligned} |r'| &= |r| \\ |t'| &= |t| \end{aligned} \tag{4.29c}$$

Alternatively, these can be derived from matrix multiplication. Knowing that the beam splitter must be a unitary operator ($\hat{U}_{BS}^\dagger = \hat{U}_{BS}^{-1}$), multiplying it by its Hermitian conjugate gives [9]

$$\hat{U}_{BS}^\dagger\hat{U}_{BS} = \begin{bmatrix} |r|^2 + |t|^2 & r't^* + t'r^* \\ rt'^* + tr'^* & |r'|^2 + |t'|^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{4.30}$$

where the product, being unitary, has be set equal to the identity matrix. Equating the corresponding elements recovers equation sets (4.29a) and (4.29b). Moreover, unitarity demands the following must also hold true:

$$\hat{U}_{BS}\hat{U}_{BS}^\dagger = \begin{bmatrix} |r|^2 + |t'|^2 & rt^* + t'r'^* \\ tr^* + r't'^* & |r'|^2 + |t|^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{4.31}$$

Now, equating the elements of (4.30) and (4.31) recovers the final constraint of equation set (4.29c).

From the four possible off-diagonal terms of the unitary products above (two of which are explicitly given in (4.29b)), two unique expressions for phase relations arrive after substituting equations (4.28).

$$\begin{gathered} \phi_r - \phi_t + \phi_{r'} - \phi_{t'} = \pm\pi \\ -\phi_r + \phi_t - \phi_{r'} + \phi_{t'} = \pm\pi \end{gathered} \tag{4.32}$$

In fact, these are equivalent given $e^{i\pi} = e^{-i\pi} = -1$, and it is obvious that the relative phase shift of the reflected and transmitted components is $\pi/2$. A balanced beam splitter has the same relative phase when incident from either side of the beam splitter. Specifically,

$$\phi_r - \phi_t = \phi_{r'} - \phi_{t'} = \pi/2 \tag{4.33}$$

such that

$$\hat{U}_{BS} = \begin{bmatrix} r & it \\ it & r \end{bmatrix}. \tag{4.34}$$

since according to the constraints all phase information must exist in the off-diagonal terms. It is exactly these phased off-diagonal terms that allow for quantum effects. Without loss of generality, the reflection and transmission probabilities can be defined as trigonometric functions.

$$|r|^2 = |r'|^2 = R = \cos^2\theta \tag{4.35a}$$

$$|t|^2 = |t'|^2 = T = \sin^2\theta \tag{4.35b}$$

Therefore, the final form for the balanced beam splitter is

$$\hat{U}_{BS}(\theta) = \begin{bmatrix} \cos\theta & i\sin\theta \\ i\sin\theta_i & \cos\theta \end{bmatrix}. \tag{4.36}$$

While the balanced beam splitter is appealing due to its symmetric, it can become tedious to track the imaginary components throughout calculations. Due to this, an asymmetric representation is often preferred. In this description, the sum of the phase differences on both sides of the beam splitter remains 90°as in equation (4.32); however, the two sides no longer have mathematically equivalent phase shifts ($\phi_r - \phi_t \neq \phi_{r'} - \phi_{t'}$). Instead, all π radians of the phase difference is placed into one matrix element resulting in the elimination of imaginary components. Two mathematically equivalent examples would be

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \text{ and } \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

It should be noted that the physical existence of such a device isn't feasible without the coupling of additional elements. For example, the addition of a $\pi/2$ phase delay at one of the output modes would effectively emulate this descriptor.

4.2.2 Photon Bunching

As seen in the previous section, when analyzed from a fully quantum mechanical perspective, state description – and even the optical elements themselves – harbor seemingly "hidden" phase information that simply isn't accessible to classical theory. The implications and applications of these phenomena are quite vast; therefore, it is the goal of this section to provide a specific example of quantum mechanical interference via a single 50:50 beam splitter with various input state conditions.

As a first example, I present the effects of a unimodal beam with finite photon number incident on a single beam splitter. Photon number resolving detection is assumed at both output modes of the beam splitter. This calculation requires a method of tracking the probability of the individual photons to be reflected or transmitted. Because the beam splitter in use is 50% probable to reflect or transmit, the probability function for a single photon is single $p = 1/2$. In the assumption that each photon is unique, the probability for the ensemble would simply be the product of the individual probabilities ($(p/2)^n$). However, the photons are indistinguishable. Of equal importance is the fact that the photon can take one of only two paths upon reaching the beam splitter (i.e., results in a binary output). Due to this, the ensemble probability is described by the binomial probability distribution function

$$P(n, k) = \mathbb{C}_k^n p^k (1 - p)^{n-k}. \tag{4.37}$$

where n is the total number of photons, and k is the number of photons measured at a single output port. Upon plugging in the probability function $p = 1/2$, this distribution becomes

$$P(k, n - k|n, 0) = \mathbb{C}_k^n \left(\frac{1}{2}\right)^n. \tag{4.38}$$

This is the exact form of the distribution for a finite, definite number of photons in the input state. The domain of this distribution ($[0, n]$) defines a nonzero value for all possible output combinations. Instinctively, this must define quantum mechanical probabilities since from a classical perspective, it is expected that the beam will be evenly divided between the outputs ($k_r = k_t = N/2$).

Analysis of the situation in the limit of large photon number must be accomplished through approximations. Namely, Stirling's approximation of $n!$ for large n is

$$n! \approx e^{-n} n^n \sqrt{2\pi n}. \tag{4.39}$$

According to the statistics for a binomial distribution, the mean value for the distribution is given by $k = np = n/2$. Therefore, an offset parameter (ϵ) can be defined to allow for analysis infinitesimally close to the PDF's mean value. This offset takes the form

$$k = (1+\epsilon)n/2, \tag{4.40}$$

and can be applied – in combination with Stirling's approximation – to the factorial terms in the binomial coefficients to yield [11]

$$\mathbb{C}_k^n \approx \frac{2^{n+1}}{\sqrt{2\pi n}\left(1+\frac{n}{2}\epsilon^2\right)}. \tag{4.41}$$

This change of variable now shifts the distribution centered upon $\epsilon = 0$ ($k = n/2$), and has a domain of $\epsilon \in [-1, 1]$ (corresponding to $k \in [0, n]$).

The maximum of a binomial distribution (and likewise the probability distribution in question) occurs at its mean value of $k = n/2$. For this value, the approximated probability distribution becomes

$$P_{max}(n,k) = P\left(\frac{n}{2}, \frac{n}{2}\middle| n, 0\right) = \frac{2}{\sqrt{2\pi n}}. \tag{4.42}$$

From equation (4.41), it can be seen that for $\epsilon = \sqrt{2/n}$ the distribution drops to exactly half its value at $\epsilon = 0$. That is,

$$P(\epsilon = \sqrt{2/n}) = \frac{1}{2}P(\epsilon = 0) \tag{4.43}$$

Evidently, in the infinite limit, $n \to \infty$ and $\epsilon \to 0$ while remaining at half-max probability. Hence, the distribution approach singular value at $n/2$ with probability $n/2$. This is exactly what we expect for the infinite photon classical limit. This single valued probability is explicitly given by $P(n/2, n/2|n, 0) = 1$.

Now, let's consider the effects of two incident beams. Given that in the fully quantum mechanical description, individual photons *always* have a nonzero probability of existing in either mode, there is an associated probability of mode exchange of the photons. Inherent in this process are the effects of interference due to the off diagonal phase values of the quantum state descriptor.

The probability amplitude of finding n_1 photons out the output of the beam splitter corresponding to the transmission mode of an input beam with N_1 photons is given by [11]

$$A(n_1, N_1 + N_2 - n_1 | N_1, N_2) = \frac{(-1)^{N_1}}{2^{(N_1+N_2)/2}} \sum_k (-1)^k \sqrt{\mathbb{C}^{N_1}_{k} \mathbb{C}^{N_2}_{n_1-k} \mathbb{C}^{n_1}_{k} \mathbb{C}^{n_2}_{N_1-k}}. \quad (4.44)$$

Full analysis of the mathematical components of the amplitude would be mathematically tedious while offering little physical insight. Instead, see figure 4.2.2 for the results of the fully quantum mechanical calculation of a state containing exactly $n = 2$ photons [11]. In table 4.2.2, the results from the previous discussion of single mode input are shown for states $|2, 0\rangle$ and $|0, 2\rangle$. The phenomenon of present interest is the fact that the identical, dual mode Fock state $|1, 1\rangle$ produces identically 0% probability of emitting a single photon into either of the beam splitter output.

TABLE 4.1. Results for beam splitter output probabilities with all possible arrangements of a two photon input state.

Input	Output $(N_1, N_2\vert$		
$\vert N_1, N_2)$	$(0, 2\vert$	$(1, 1\vert$	$(2, 0\vert$
$\vert 2, 0)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
$\vert 1, 1)$	$\frac{1}{2}$	0	$\frac{1}{2}$
$\vert 0, 2)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

4.2.3 Entanglement

Entanglement refers to a correlation between two elementary particles – regardless of their spatial separation – due to a their physically individual characteristics being described by a single wave function. Once the wave function for one of the particles is collapsed (measured), the correlated properties of the secondary particle are immediately influenced accordingly. The implications of the phenomenon lead to extraordinarily large and diverse areas of research, but this thesis will focus primarily on the use of entanglement for metrology and sensing applications. Also, as there are superfluous means of obtaining entangled photon pairs, it should be mentioned that the topics discussed herein focus on the entangled description of the state *within* the sensing interferometer, rendering the specific generation methods far less relevant.

As will be detailed in later chapters, one of the most significant advantages of perfect entanglement is its capability of super-resolving phase sensitivity in metrological applications. Also of great importance to the topics of this work is the ability of a beam splitter to create single photon entanglement with the vacuum and ultimately result in the photon's self-interference. This is a highly nonclassical effect that requires extremely reliable methods of implementation. We suggest later methods of coping with the extreme fragility of entangled pairs to allow for a more robust sensing technique. For now, I shall focus on the specific mathematical framework of entanglement and elementary methods of obtaining the entangled pairs.

As mentioned, entanglement simultaneously describes correlated properties of two, otherwise independent, particles. Mathematically, this is equivalent to a state vector composed of a summation of the independent terms in such a way that the expression is non-separable. In essence, if the wave function is capable of being factored into a product of the two independent wave functions, there exists no entanglement. A simple example of an entangled state is

$$|\psi_{ent}\rangle = |n\rangle_a|m\rangle_b + |m\rangle_a|n\rangle_b,$$

where an example of an unentangled/separable state might be

$$|\psi_{sep}\rangle = |n\rangle_a|m\rangle_b + |n\rangle_a|k\rangle_b = |n\rangle_a\left(|m\rangle_b + |k\rangle_b\right) = |\psi_a\rangle|\psi_b\rangle.$$

Thus far, this chapter has introduced quantum-like photon number states, classical-like coherent states, and fully quantum mechanical operator for the unitary beam splitter. These concepts may now be combined as a comparison of quantum and classical states to entangled and separable states.

Let us begin with the finite number quantum Fock states with two mode generalization. Clearly, the set of states exhibit symmetric properties about the number distribution $|n/2, n/2\rangle$ as seen in table 4.2.2. For this reason, discussion will simply be restricted to no more than a single photon Fock state per input mode. Also adopted here is the shorthand notation for a beam splitter transformation where the annihilation operators directly transformed in terms of the reflection and transmission amplitudes given generally by

$$\begin{aligned}\hat{b}_0 &= t\hat{a}_0 + e^{i\phi_1} r\hat{a}_1 \\ \hat{b}_1 &= t\hat{a}_1 + e^{i\phi_0} r\hat{a}_0.\end{aligned}$$

Note that the mode labeling is reversed from the unitary operator constructed earlier for the purpose of adhering to convention and the evidence that supports this phase arrangement [11]. Physically, there is no difference. Therefore, balanced the beam splitters will be assumed symmetric with 50:50 reflectivity such that their transformation becomes

$$\hat{b}_0 = \frac{1}{\sqrt{2}}(\hat{a}_0 + i\hat{a}_1) \tag{4.45a}$$

$$\hat{b}_1 = \frac{1}{\sqrt{2}}(\hat{a}_1 + i\hat{a}_0). \tag{4.45b}$$

For a single photon incident at one mode of the beam splitter, and vacuum at the other, this transformation is

$$\begin{aligned}|0\rangle_0|1\rangle_1 = \hat{a}_1^\dagger|0\rangle_0|0\rangle_1 \xrightarrow{BS} & \frac{1}{\sqrt{2}}\left(i\hat{b}_0^\dagger + \hat{b}_1^\dagger\right)|0\rangle_0|0\rangle_1 \\ = & \frac{1}{\sqrt{2}}\left(i|1\rangle_0|0\rangle_1 + |0\rangle_0|1\rangle_1\right) \qquad (4.46)\end{aligned}$$

Clearly this operation results in an entangled state, since (4.46) is non-separable. This scenario was discussed in earlier sections as the simple probability of a single photon incident on a beam splitter to have a 50% probability of reflecting/transmitting. As long as the output mode of the photon is not measured, there exists the a complete lack of knowledge about which path the photon has taken and the photon remains entangled with the vacuum. As soon as a measurement is made, however, the "which path information" is determined with certainty and the entanglement collapses. Note that if the two components of this entangled state were again incident on a beam splitter (as if a simulated dual mode input), single photon interference would result even from the probability amplitude that the photon *may* exist in either mode. This is the nature of entanglement. However, if there were any attempts to discern the photon's mode (even without disturbing the path of the photon), the entanglement would collapse and classical detection probabilities would ensue.

As a second example, consider the dual Fock input state of table 4.2.2. The transformation for this state is

$$\begin{aligned}|1\rangle_0|1\rangle_1 = \hat{a}_0^\dagger\hat{a}_1^\dagger|0\rangle_0|0\rangle_1 \xrightarrow{BS} & \frac{1}{2}\left(\hat{b}_0^\dagger + i\hat{b}_1^\dagger\right)\left(i\hat{b}_0^\dagger + \hat{b}_1^\dagger\right) \qquad (4.47) \\ = & \frac{i}{\sqrt{2}}\left(\hat{b}_0^\dagger\hat{b}_0^\dagger + \hat{b}_1^\dagger\hat{b}_1^\dagger\right)|0\rangle_0|0\rangle_1 \qquad (4.48)\end{aligned}$$

The result is another maximally entangled state similar to the single photon input. As expected, the complex square of the coefficients confirms the counterintuitive gap in probability of the corresponding state in table 4.2.2.

As a final example, the single mode coherent state shall be transformed. Recall from equations (4.17) and 4.15 that a coherent state is equivalent to displaced vacuum and that the displacement operator contains the operators to be transformed. It follows that [6]

$$\begin{aligned}|0\rangle_0|\alpha\rangle_1 \xrightarrow{BS} &\exp\left[\frac{\alpha}{\sqrt{2}}\left(i\hat{b}_0^\dagger + \hat{b}_1^\dagger\right) - \frac{\alpha^*}{\sqrt{2}}\left(-i\hat{b}_0 + \hat{b}_1\right)\right]|0\rangle_0|0\rangle_1 \\ =&\exp\left[\left(\frac{i\alpha}{\sqrt{2}}\right)\hat{b}_0^\dagger - \left(\frac{-i\alpha^*}{\sqrt{2}}\right)\hat{b}_0\right]\exp\left[\left(\frac{\alpha}{\sqrt{2}}\right)\hat{b}_1^\dagger - \left(\frac{\alpha^*}{\sqrt{2}}\right)\hat{b}_1\right]|0\rangle_0|0\rangle_1 \\ =&\left|\frac{i\alpha}{\sqrt{2}}\right\rangle_0\left|\frac{\alpha}{\sqrt{2}}\right\rangle_1 . \qquad (4.49)\end{aligned}$$

Per earlier discussions, this state certainly *is* separable and is therefore *not* entangled. A complex square of this separable state would show that simply half the light intensity is in one mode and half is in the other. This is to be expected since we are dealing with the classically limited coherent state with infinite photon number. Specifically, an expansion of this output state using equation (4.23) gives

$$|out\rangle = \left|\frac{i\alpha}{\sqrt{2}}\right\rangle_0\left|\frac{\alpha}{\sqrt{2}}\right\rangle_1 = e^{-|\alpha|^2/2}\sum_{N=0}^{\infty}\frac{\alpha^N}{\sqrt{N!}}|\psi_N\rangle \qquad (4.50)$$

where

$$|\psi_N\rangle = \frac{1}{2^{N/2}}\sum_{n=0}^{N} i^n\left[\frac{N!}{n!(N-n)!}\right]^{1/2}|n\rangle_0|N-n\rangle_1. \qquad (4.51)$$

This implies that the output is, in fact, of coherent state form, but not summed directly over the Fock state number basis as with the standard definition. Instead, do to the probability of mode mixing, the coherent state is summed over a binomial distribution of photon number between the two output modes [6].

Chapter 5

Optimized Quantum States for Lossy Metrology

5.1 Theory of Quantum Metrology

Quantum metrology refers to the sensitive measurement of deviations in a quantum interferometer that result in a minute phase shift between the arms. The phase sensitivity of a quantum interferometer is often determined from the Fisher information of the device setup.

Quantum states of light play an important role in applications including metrology, imaging, sensing, and quantum information processing [13]. In quantum interferometry, entangled states of light, such as the maximally path-entangled N00N states, replace conventional laser light to achieve a sensitivity below the shot-noise limit, even reaching the Heisenberg limit, and a resolution well below the Rayleigh diffraction limit [12]. For an overview of quantum metrology applications see, for example, Ref. [13]. However, for real-world applications, diffraction, scattering, and absorption of quantum states of light need to be taken into account. Recently it has been shown that many quantum-enhanced metrology schemes using N00N states perform poorly when a considerable amount of loss is present [14, 15, 16]. However, our team has also discovered a new class of entangled number states, which are more resilient to loss [17]. These so-called $M\&M'$ states still outperform classical light sources under a moderate 3 dB of loss.

We systematize the numerical search for optimal quantum states in a two-mode interferometer in the presence of loss. We employ the Fisher information to obtain the phase sensitivity of the interferometer. An exhaustive review and application of the Fisher information concept to the sensitivity of a March-Zehnder interferometer, particularly in the zero loss case, has been presented in the recent work by Durkin and Dowling [38]. The chief utility of the Fisher information approach is that it provides a bound on the phase sensitivity, even in the absence of a fully specified detection scheme, and is now widely adopted in studies of interferometer sensitivity. Such numerical optimization has been previously carried out in the absence of loss, and with loss over a restricted class of input states [18, 39]. Here, we provide a completely general optimization scheme that is applied to the two-mode interferometer, but also has application to the optimization of linear optical systems for quantum linear optical information processing [40, 41].

Using this scheme, we first recover the well-known fact that N00N states are optimal in the absence of loss [13]. For large loss, the optimal states belong to a class of two-mode coherent states with finite total photon number. The optimization procedure yields the optimal Fisher information – and hence the minimal phase uncertainty – for every level of loss. The validity of our numerical optimization is verified using several methods, including genetic algorithms and simulated annealing, and the close agreement among these methods provides evidence that we are indeed finding the global optimum.

5.2 Phase Measurement and Fisher Information

The goal of quantum metrology is to obtain maximally sensitive phase measurements, which ideally approach the theoretical limit of the Cramer-Rao bound. This is a bound of the variance of the phase fluctuation, and as such places a physical limit on the certainty of any phase measurement. Specifically, the Cramer-Rao bound is given by

$$\delta\phi^2 \geq \frac{1}{F(\phi)} \tag{5.1}$$

where the function $F(\phi)$ is the Fisher information. Fisher information, which is in turn a measure of the amount of certainty that can be extracted from a system, is given by

$$\begin{aligned} F(\phi) &= \sum_{n_a,n_b} P_{n_a,n_b}(\phi)\left(\frac{\partial}{\partial\phi}\ln\left[P_{n_a,n_b}(\phi)\right]\right)^2 \\ &= \sum_{n_a,n_b}\frac{\left(\partial_\phi P_{n_a,n_b}(\phi)\right)^2}{P_{n_a,n_b}(\phi)}. \end{aligned} \tag{5.2}$$

P_{n_a,n_b} is the probability of concurrently detecting n_a, n_b photons in modes a, b as depicted in figure 5.4 [18]. This probability can be determined by

$$P_{n_a,n_b}(\phi) = Tr\left[\hat{A}_{n_a,n_b}\hat{\rho}_{input}\hat{A}^{\dagger}_{n_a,n_b}\right] \tag{5.3}$$

where

$$\hat{A}_{n_a,n_b} = |0\rangle_a|0\rangle_{ba}\langle n_a|_b\langle n_b| \tag{5.4}$$

is the operator for single photon detection.

5.3 Initialization and Phase Introduction

As seen from the definition of the Fisher information, the determination of the phase sensitivity of the interferometric sensor requires a photon number resolving detection scheme. Since measurement is dependent on precise photon numbers, discrete photon loss in the interferometer was modeled by placing a beam splitter in each mode following the introduction of the single mode phase shift to be measured [19]. We assumed the general N photon pure state corresponding to the two modes a, b in figure 5.1.

$$|\psi\rangle = \sum_{k=0}^{N} c_k|N-k\rangle_a|k\rangle_b \tag{5.5}$$

To calculate the photon loss, this pure state must be written as an $(N+1)^2$ dimensional density matrix

$$\rho_{in} = |\psi\rangle\langle\psi| = \sum_{i,j=0}^{N} c_i c_j^*|N-i,i\rangle\langle N-j,j|. \tag{5.6}$$

It is assumed that the phase fluctuations to be measured precede any possible photon loss. Therefore, the phase operator imparting the observable phase acts on the generalized input state vector as

$$e^{i\phi\hat{n}_a}|\psi\rangle = \sum_{k=0}^{N} c_k e^{i\phi(N-k)}|N-k\rangle_a|k\rangle_b$$

which produces the density matrix

$$\hat{\rho}_\phi = \sum_{i,j=0}^{N} c_i c_j^* e^{i\phi(j-i)}|N-i,i\rangle\langle N-j,j|.$$

5.4 Model for Discrete Photon Loss

Photon loss is typically modeled by a beam splitter that routes photons out of the interferometer [19]. In implementing this model, we first enlarge the Hilbert space to include modes representing the scattered photons and then, after the scattering, trace out these modes. Here we extend the beam splitter model for photon loss to two propagating modes that represent the two paths in the optical interferometer (figure 5.1).

Once the state is propagated through the beam splitters, all additional input and output modes of these beam splitters must also be included in the system description. This is accomplished an expansion of the density matrix to $(N+2)(N+1)/2$ dimensions to allow for the additional beam splitter modes. Once discrete photons have been "routed out" of the propagation modes, they must be removed for the system description prior to the detection stage. This process is simulated via a trace of the previously introduced beam splitters modes, which effectively reverts the system back to the dual mode description of the input. However, the environmental tracing is a nonunitary process that results in an irreversible alteration to the state density matrix. The result is an $(N+2)(N+1)/2$ dimensional reduced density matrix of block diagonal form [20]. In this block diagonal form, each subspace represents $L\epsilon\{0, 1, ..., N\}$ total photons lost. After calculating the reduced density matrix, Eqn (5.1) is used to calculate the phase variance.

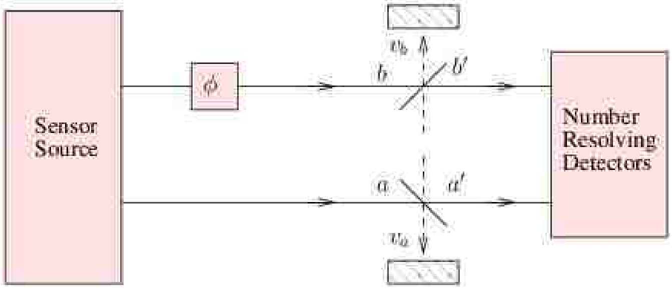


FIGURE 5.1. Abstract interferometer condensing the input state plus the first beam splitter into the first box, followed by two propagating modes with discrete photon loss modeled by additional beam splitters. The box on the right includes a beam splitter and the photon-number resolving detectors.

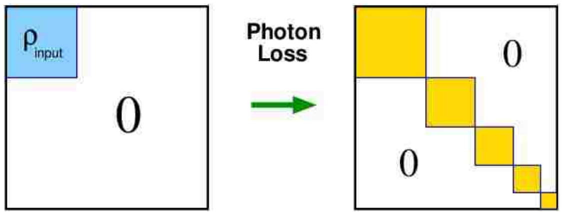


FIGURE 5.2. Illustration of the evolution of the density matrix as it undergoes discrete photon loss. The resulting sub-block state density matrices are sequentially reduced in dimension by 1x1 representing the loss of an addition single photon.

For an interferometer with two input ports A and B as depicted in figure 5.1, an arbitrary pure-state input with N photons can be written as $|\psi\rangle_{\text{input}} = \sum_{k=0}^{N} c_k |N-k, k\rangle_{A,B}$ where c_k are the input amplitudes to be optimized. The ket $|N-k, k\rangle_{A,B}$ is a basis state in which $N-k$ and k photons are in mode-A (detection arm) and mode-B (control arm), respectively. Such a quantum state resides in an $N+1$ dimensional Hilbert space (depicted as ρ_{input} in figure 5.2) and is represent by the generalized density matrix construction

$$\rho = |\psi\rangle\langle\psi| = \sum_{k,m=0}^{N} c_{k,m} |N-k, k\rangle\langle N-m, m|. \tag{5.7}$$

The transformation of the quantum state by any passive lossless optical elements, such as beam splitters, phase shifters, and mirrors, can then be described by an $(N+1)\times(N+1)$ unitary matrix.

When the propagation channels suffer from photon loss , we need to consider the total density matrix that includes all the scattered photon modes ($(N+1)(N+2)/2$ square dimensionality). Then we obtain the reduced density matrix for the two interferometer modes by tracing out these additional modes [20]. This reduced density matrix is associated with a much larger Hilbert space of dimension $(N+1)(N+2)/2$, which includes all states with a total of $N, N-1, ..., 0$ photons in the two interferometer modes. As evident from the right side of figure 5.2 (after tracing the photon loss modes), the resulting reduced density matrix is in block diagonal form. The upper left block is identical to the original $(N+1)\times(N+1)$ density matrix ρ_{input}. Subsequent blocks represent the density matrix of the state after having lost a single photon. This continues until the last block is simply a 1×1 matrix representing the vacuum state. Specifically, the reduced density matrix is of the form

$$\rho_L = \sum_{i,j=0}^{N-L} \rho_{N,L,i,j} |N-L-i, i\rangle\langle N-L-j, j| \tag{5.8}$$

where $\rho_{N,L,i,j}$ is the sub-block density matrix for the state having lost $L \in 0, ..., N$ photons. Therefore, each block is represented by

$$\rho_{N,L,i,j} = \sum_{m=0}^{L} \left(c_{i+m}c^*_{i+m}\right) A_{i+m} B^*_{j+m} \tag{5.9}$$

where

$$A_k = \begin{bmatrix} N-k \\ N-L-i \end{bmatrix}^{1/2} t^{N-L-i} r^{L+i-k} \begin{bmatrix} k \\ i \end{bmatrix}^{1/2} t'^{i} r'^{k-i} \tag{5.10a}$$

$$B_k = \begin{bmatrix} N-k' \\ N-L-j \end{bmatrix}^{1/2} t^{N-L-j} r^{L+j-k'} \begin{bmatrix} k' \\ j \end{bmatrix}^{1/2} t'^{j} r'^{k'-j}. \tag{5.10b}$$

where represent binomial distributions of the reflectance and transmittance (r,t) and (t',r') in the upper path and the lower path, respectively. Note that for a given number of photons N, the L value labels the block of the reduced density matrix, and i, j specify the matrix element inside that block. The quantum state of light ends up in a mixed state associated with an $N+1-L$ dimensional Hilbert space, if a total of L photons are lost.

5.5 Numerical Approach to State Optimization

The initial input state was randomized, and systematically optimized to improve the phase sensitivity. Specifically, numerical maximization of the resulting Fisher information was performed by means of a genetic algorithm and simulated annealing to ensure accuracy of results. Each was allowed enough iterations to ensure the result was a global maximum at discrete intervals of loss for specified total photon numbers.

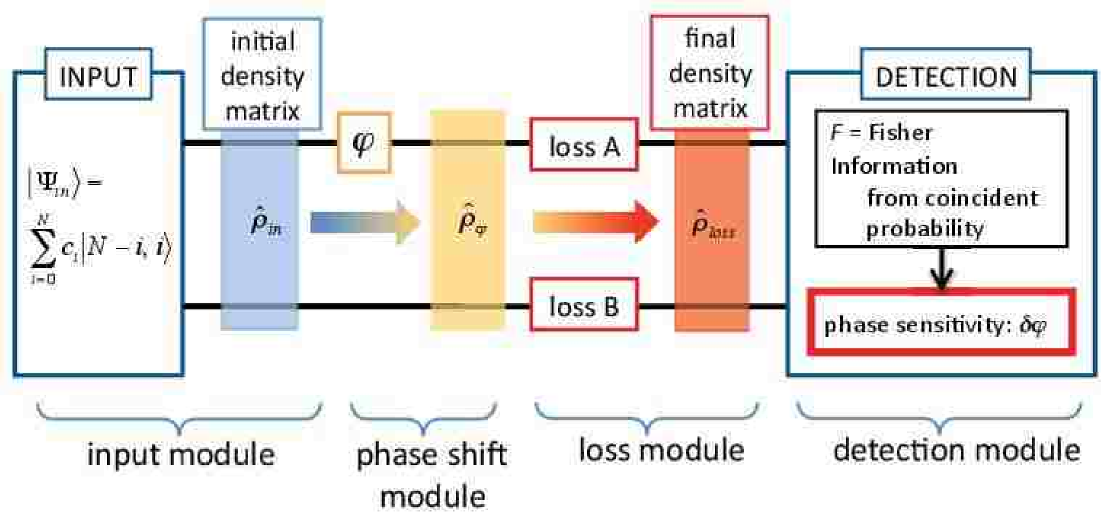


FIGURE 5.3. Flowchart representation of the numerical implementation of the forward prob for detecting the introduction of an arbitrary phase with maximum resolution after the effects of discrete loss are applied to both propagation modes.

The optimal quantum input state is now obtained numerically. First, a forward problem solver is developed using a density matrix approach and is depicted in figure 5.3 by the shaded regions of the mode propagations. An input state is written as a density matrix, ρ_{in} (blue region of figure 5.3). The single mode phase shift during photon propagation is taken into account by operating with $e^{i\phi\widehat{n}_A}$ on this density matrix as detailed in previous sections. Next, photon losses are applied using equation (5.7) and (5.10), producing a reduced density matrix of dimension $(N+2)(N+1)/2$. In the last step of the forward problem solver, the minimum detectable phase sensitivity $\delta\phi$ is computed from the final density matrix. The phase detection is modeled by a final 50-50 beam splitter followed by two number resolving photodetectors. The joint probability of simultaneously detecting m_1 photons at the first photodetector and m_2 photons at the second photodetector is computed as $P_m = \sum_{i=1}^{(N+2)(N+1)/2} \hat{U}_{\text{bs}_{m,i}} \sum_{j=1}^{(N+2)(N+1)/2} \hat{\rho}_{\text{out}_{i,j}} \hat{U}^{\dagger}_{\text{bs}_{j,m}}$, where the label m represents a pair of numbers (m_1, m_2). Here, $\hat{U}_{\text{bs}}$ is a unitary transformation representing a 50-50 beam splitter and $\hat{\rho}_{\text{out}}$ is the final density matrix obtained after loss. Then, phase sensitivities are estimated from the Fisher information, F, for a single measurement, $\delta\phi = 1/\sqrt{F}$, where $F = \sum_{m=1}^{(N+2)(N+1)/2} P_m \left(\partial \ln P_m/\partial\phi\right)^2$ [18]. We note that, in all of our calculations, we assume a large flux of entangled states is to be used, and we normalize our results by this flux.

We optimize the system to find the minimum detectable phase sensitivity, given fixed losses in the detection and control arms. For this, a genetic global optimization algorithm is applied to the forward problem solver. The parameters to be optimized are the complex coefficients c_k; the optimal sensitivity is necessarily ϕ-independent since a change in ϕ can be absorbed into the relative phases of c_k. During the numerical computation of F, we observe that the landscape of F in the optimization

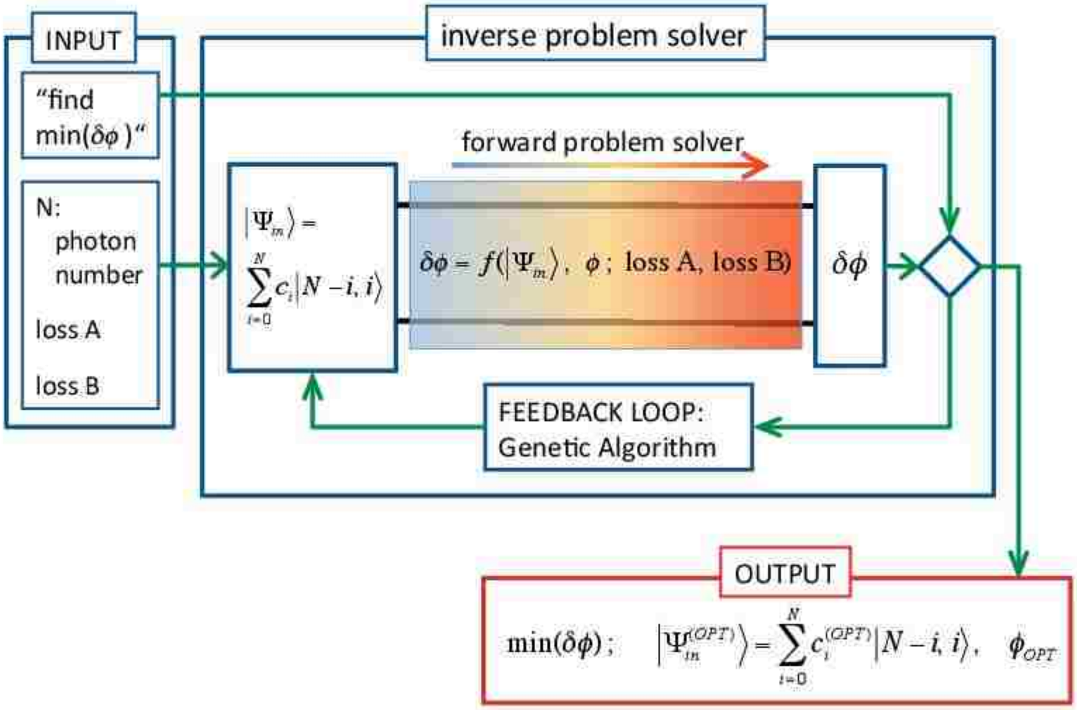


FIGURE 5.4. Flowchart for the numerical optimization of the forward problem. The randomized initial state is run through the forward problem modules, and the phase sensitivity of this state is determined. By use of a genetic algorithm and simulated annealing, the output state is altered and the process repeats. Eventually, the code will converge onto a solution state whose phase sensitivity meets predefined tolerance levels.

parameter space possesses several local maxima contrary to the convex $\tilde{F}_Q$ used by Doner *et al.* [39].

5.6 Results

Our initial results indicated that maximum phase sensitivity is attained by maintaining a lossless control arm (fixed phase mode), and decreases as the loss in the target arm increases. Also, an analytical limiting analysis supports the data in that phase sensitivity in increased by adjusting the entangling source to transmit an increased number of photons into the mode containing more significant loss. Furthermore, the results indicate that while phase sensitivity greatly increases with total photon number for low loss, there is little benefit for increased photons in the high loss regime. In fact, normalization of the Fisher information by the photon number yields N-independent optimum states asymptotically approaching the sensitivity of a classical coherent state. These finite number two mode coherent states are known as generalized Perelomov coherent states (GPCS), and defined as $|GPCS\rangle = (N!)^{-1/2}[\hat{a}^\dagger e^{i\beta}\cos\alpha - \hat{b}^\dagger e^{-i\beta}\sin\alpha]^N|0\rangle$ for real parameters α, β [21].

Now we compare a classical interferometer with the optimized quantum state approach. Consider a single-mode coherent state $|\alpha\rangle = \mathrm{e}^{-\frac{|\alpha|^2}{2}}\sum_{k=0}^{\infty}\frac{\alpha^k}{\sqrt{k!}}|k\rangle$ as the input state to a Mach-Zehnder interferometer (MZI) [20]. The first beam splitter partitions the state between two modes. In the first mode, the state acquires a phase shift ϕ and undergoes a loss of intensity by a factor of $|t|^2$. Then, the two beams are redirected to the second beam splitter, and photons are detected in each output port. The Fisher information, normalized to the average number of photons $\bar{n} = |\alpha|^2$, is $F/\bar{n} = (4|t|^2\sin^2\theta)/(1+|t|^2\tan^2\theta)$, where θ describes the angle of σ_y rotation by the first beam splitter. This equation provides the Cramer-Rao asymptotic accuracy of measurement of the unknown phase shift ϕ using a classical scheme. We single out two cases. First, we assume that the first beam splitter has a fixed 50-50 ratio. Then in the limit of large loss $t \to 0$ we have

$$F/\bar{n} = \frac{2|t|^2}{1+|t|^2} \overset{|t|\to 0}{\to} 2|t|^2\,. \tag{5.11}$$

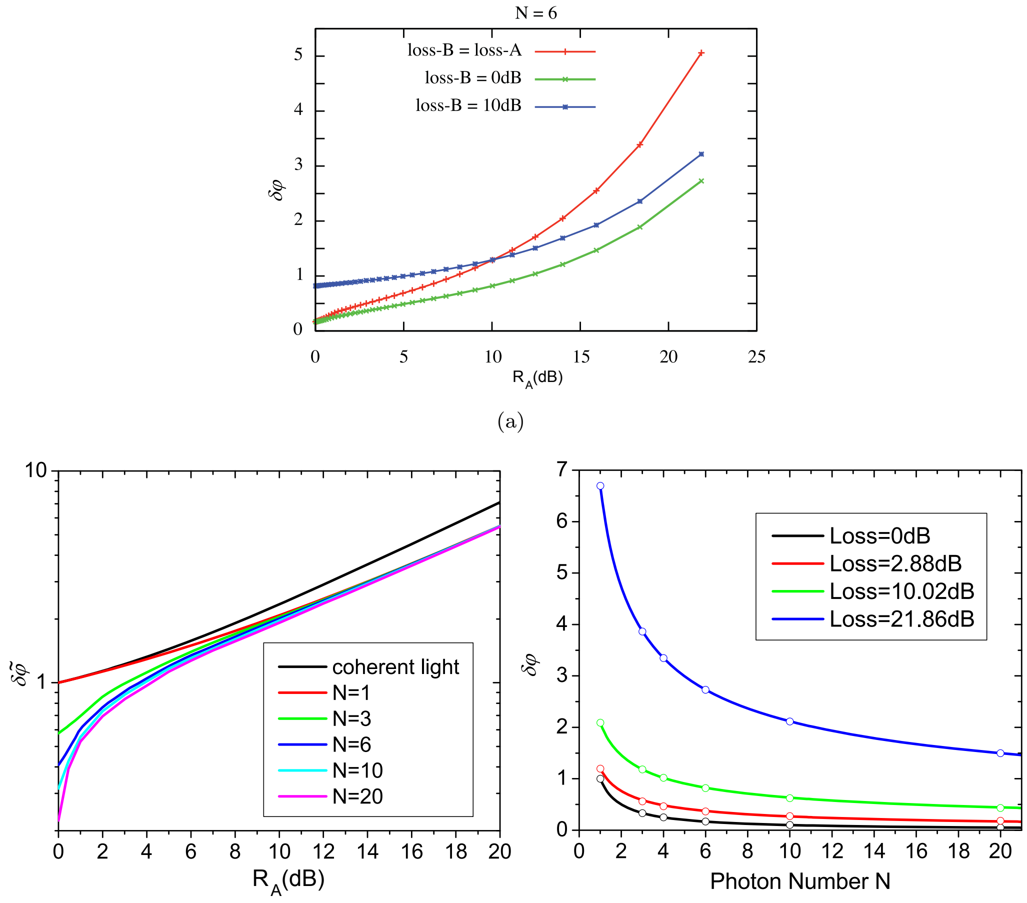
Fisher information can be increased by optimizing the first beam splitter to compensate for the loss. In this case we obtain

$$F/\bar{n} = \frac{4|t|^2}{(1+|t|)^2} \overset{|t|\to 0}{\to} 4|t|^2\,. \tag{5.12}$$

The results of numerical optimization of $\delta\phi$ are presented in figure 5.5. We denote the losses in dB in the detection and control arms as R_A and R_B, respectively. First, in figure 5.5(a) we investigate the overall influence of loss in the control arm. One set of simulations is conducted with equal losses in the detection and control arms ($R_A = R_B$). We also consider fixed 10 dB loss and fixed 0 dB loss in the control arm ($R_B = 10$ dB, $R_B = 0$ dB) as loss in the detection arm is varied. In all cases, $N = 6$ is assumed. We consistently find that an increase in R_B results in higher $\delta\phi$. Thus, one can expect the best phase sensitivity to be achieved with the smallest possible loss in the control arm.

In figure 5.5(b), the numerically optimized phase sensitivity is presented as a function of R_A for $N = 1$, 3, 6, 10, and 20. The classical (coherent light) baseline of Eq. (5.11) is also shown. To compare the quantum results for different N with each other and with the classical case, we rescale the phase sensitivity of the pure quantum state by normalizing the Fisher information similarly to Eqs. (5.11) and (5.12): $\delta\tilde{\phi} = 1/\sqrt{F/N} = \delta\phi\sqrt{N}$. Since $\delta\tilde{\phi}$ is obtained from the Fisher information per single photon, it is also the measure of the synergically-enhanced phase sensitivity per single photon with N photons acting together. For pure quantum states, figure5.5(b) shows that larger N produces smaller $\delta\tilde{\phi}$ for any given amount of loss. The N-dependence of $\delta\tilde{\phi}$ is greatest at $R_A = 0$ dB, and weakest in the limit of extremely high loss, where the lines merge together. Coherent light does not show enhancement with N at any level of loss, and at every level of loss, coherent light exhibits worse performance in phase sensitivity compared to entangled quantum states.

One interesting observation is that $\delta\tilde{\phi}$ in the extremely high loss region ($R_A >\sim 16$ dB) becomes N-independent even when using optimally entangled quantum states. In other words, the optimal phase sensitivity $\delta\phi$ given by the optimal quantum state becomes proportional to $N^{-1/2}$ in this high-loss regime, i.e., its scaling with N in this regime is the same as for coherent light governed by the shot noise limit. However, despite the same scaling with N, the phase sensitivity $\delta\phi$ is still



(a)

(b)

(c)

FIGURE 5.5. Minimum detectable phase sensitivities calculated from the normalized Fisher information: (a) As a function of detector arm loss (R_A) with three different losses in the control arm (R_B) for $N = 6$. (b) As a function of R_A for $R_B = 0$ in log scale. (c) As a function of input photon number N for $R_B = 0$ with fixed R_A. Lines represent the results of curve fitting using a functional form $1/N^x$. In the absence of loss, the Heisenberg limit, $x = 1$, is obtained. For high loss, x tends toward 0.5, approaching the shot noise scaling.

better with entangled quantum states than with coherent light. This can be explained by the optimal preparation of the initial state. As we will see later, the probability amplitudes c_k are distributed asymmetrically to generate the smallest possible $\delta\phi$ for nonzero loss, while coherent light always enters the system through 50-50 beam splitter, i.e., it is symmetrical between the control and detection arms, Eq. (5.11). With coherent light, a similar improvement can be achieved by adjusting (or optimizing) the first beam splitter, resulting in Eq. (5.12). In the latter case, $\delta\tilde{\phi}$ of coherent light becomes identical to that of the pure quantum state with $N = 1$. However, we emphasize that, when losses are not too high, the phase sensitivity of $N > 1$ pure quantum states is always better than that of coherent light, even with an optimized first beam splitter.

Figure 5.5(c) shows the optimal phase sensitivity as a function of photon number N for given fixed R_A values, with $R_B = 0$. Here $\delta\phi$ is shown for each N, with circles indicating $N = 1$, 3, 4, 6, 10, and 20. The lines are drawn by curve-fitting to a power law, $\delta\phi \sim 1/N^x$. From this, we find that $\delta\phi$ is well represented by $1/N$, $1/N^{0.68}$, $1/N^{0.517}$, and $1/N^{0.5}$ for $R_A = 0$ dB, 2.88 dB, 10.02 dB, and 21.86 dB, respectively. This result provides an overall view of how the phase sensitivity changes from the Heisenberg limit to the shot noise limit with increased loss, i.e., it follows a power law with $1/N^x$ where $1/2 \leqslant x \leqslant 1$.

To characterize the optimal state, we use three classes of well-defined states: N00N, $M\&M'$, and a two-mode SU(2) coherent state, often called the Generalized Perelomov Coherent State (GPCS). GPCS is defined as $|GPCS\rangle = (N!)^{-1/2}[\hat{a}_1^\dagger e^{i\beta}\cos\alpha - \hat{a}_2^\dagger e^{-i\beta}\sin\alpha]^N|0\rangle$, where $\hat{a}_1^\dagger$ and $\hat{a}_2^\dagger$ are creation operators in the two modes, and α and β are two real parameters [21]. In particular, $M\&M'$ is the first class of path entangled states shown analytically to have robustness to photon loss. It is interesting to see how the true optimal state may differ from the $M\&M'$ state in lossy environments. Characteristics of the optimal state are presented in figure 5.6 for a fixed $R_B = 0$ dB. The similarity of the optimal states with each of these three benchmark states is measured by the squared overlap between the optimal state and the benchmark state. The results show that the optimal state is closest to the N00N state for low loss. As loss increases, the N00N state portion gradually decreases and the optimal state becomes closer to GPCS than to N00N at around 5 dB of loss. The degree of similarity between the optimal state and the $M\&M'$ state is rather low for loss smaller than 6 dB. For every value of loss, GPCS is higher than $M\&M'$. Figure 5.6(b) shows how the input amplitudes of the optimal state are arranged for different loss levels. In the lossless case, we have the N00N state. As loss increases, the optimal state is reshuffled and acquires an asymmetric shape. This serves as critical information for achieving a highly sensitive interferometric system. Based on the results shown, it is obvious that generating such optimal input states should be the first consideration in the development of an interferometric sensor using entangled photons.

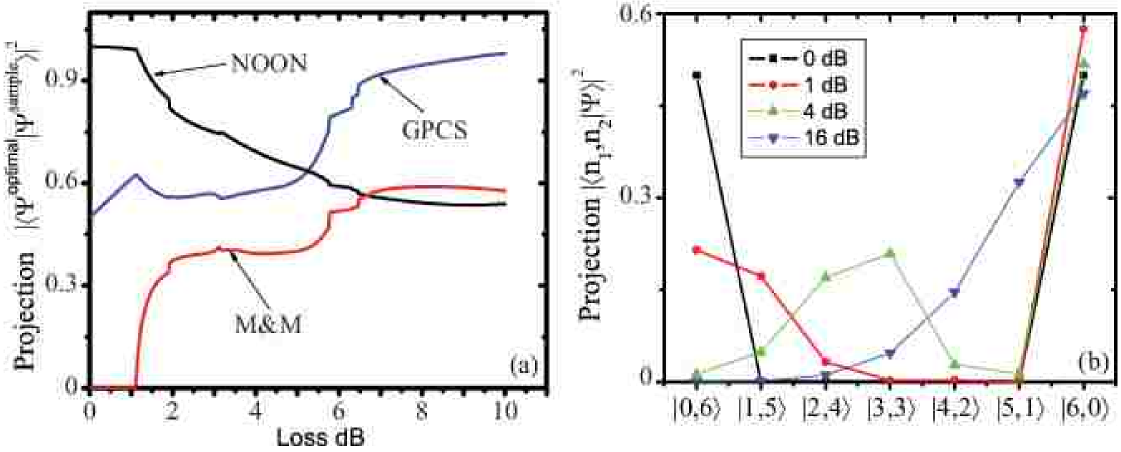


FIGURE 5.6. (Color online) (a) Projection of the optimal state on N00N, $M\&M'$, and GPCS as a function of R_A, for $N = 4$. (b) The optimal input state composition. The vertical axis shows coefficients $|c_k|^2$ of the optimally entangled input state. Both figures clearly demonstrate that the optimal state changes from N00N-type to GPCS-type as loss increases, with the crossover occurring at approximately 5 dB loss.

Chapter 6

Interaction Free Measurement (IFM)

6.1 Concept

The field of quantum mechanics is full of predictions and paradoxes that contradict with our intuition. One such counter-intuitive prediction was made about the possibility of using a single particle to detect the presence of an object (classical or quantum) without direct interaction [30]. This "interaction-free" measurement inspired studies in quantum interrogation with the ultimately promise of undetectable remote sensing and imaging with single photons [31, 25].

It should be explicitly mentioned, however, that this paradoxical property is fundamentally based on the ability of a single quantum particle (e.g., a photon) to interfere with itself through entanglement within the independent modes of the interferometer. Due to the superposition of the particle's possible existence in *either* of the modes, it is clear that the governing wave function describes the particle simultaneously in *both* configurations. Thus, the rather audacious statement of "interaction-free" should be qualified with the reality that the objects presence in fact collapses the entanglement of the wave function, and has thereby actually influenced (i.e., interacted) with the quantum state [33]. However, once the entangled wave function is collapsed by the presence of an object, there is a distinct probability that the quantum particle actually has not (and will not, since the collapse has now occurred) physically interacted with the object.

Interaction Free Measurement (IFM) originated with the Elitzur-Vaidman "Bomb" gedanken experiment that showed it was possible to detected a single-photon, hair-triggered bomb in an interferometer — without setting it off — by exploiting single particle interference combined with the presence of quantum “which-path” information [30]. The original bomb protocol had a success probability of only 25%. (In another 50% of the runs the bomb was detonated, and in the remaining 25% no information about the bomb was obtained.) The protocol was improved upon by Kwiat, et al., who combined lossless IFM with a multi-pass quantum Zeno effect [22].

The concept of interaction free measurement in the single photon regime seems to have been extensively studied in previous papers. It is the focus of these papers, however, to restrict analysis to detection/absorption probabilities and define an efficiency as the ratio of detection probability to the probability of detection or absorption [22] with arduous attempts at improving experimental efficacy. Also, various theoretical detection schemes have been proposed as a means of improving said efficiency [34]. While these are feasible perspective, it gives little consideration to other interesting characteristics when relaxing the theoretical assumptions to include phase, semi-transparence, and loss. Furthermore, in the event of complete single photon loss, there is necessity for an investigation of the amount of resources (e.g., number of photons, number of trials, etc) required to obtain sufficient measurement confidence and a practical description of efficiency in terms of these resources.

6.2 Mach-Zehnder Interferometer

The most basic MZI configuration involves two beam splitters and guarantees one of the two output modes to be “dark” in terms of photon detection due to self-interference of the single photon. However, if an object is introduced into one of the modes (as seen in figure 6.1), the entanglement is destroyed giving full which path information for the photon. Therefore, it will no longer exhibit the quantum interference effect at the final beam splitter and has a 25% chance of exiting the MZI in the dark port. This directly implies that the presence of the object has been discerned without the photon interacting with the object – a process known as interaction free measurement (IFM).

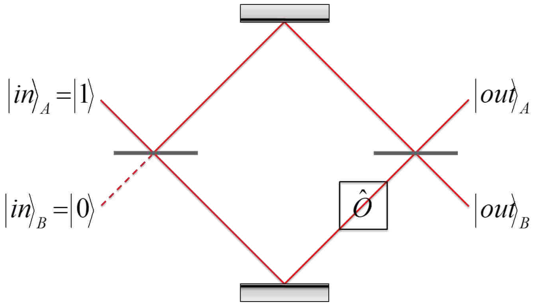


FIGURE 6.1. MZI with an object introduced into one of the modes. There is a 50% chance of the photon striking the object, and a 25% chance that the photon will be detected at the dark port yielding a successful IFM.

In general, an interaction free measurement can be performed by using a Mach-Zehnder interferometer (MZI) configured for a dark port arrangement at the output. As seen in figure 6.1, such a device consists of a single photon input into a single spatial mode incident on a set of two consecutive beam splitters ($\hat{U}_1$ and $\hat{U}_2$). The unitary evolution of the beam splitter on the state is described by

$$\hat{U}(\theta) = \begin{bmatrix} \sqrt{R} & i\sqrt{T} \\ i\sqrt{T} & \sqrt{R} \end{bmatrix}. \tag{6.1}$$

Without loss of generality, the reflectance and transmittance coefficients can be defined as trigonometric functions (while preserving R+T=1), such that the two parameters become a function of a single variable. Specifically, the unitary beam splitter transformation takes the form

$$\hat{U} = \hat{R}(\theta) = \begin{bmatrix} \cos\theta & i\sin\theta \\ i\sin\theta & \cos\theta \end{bmatrix}. \tag{6.2}$$

which is also the form of a rotation matrix $\hat{R}(\theta)$ about a single axis by an angle θ.

Prior to an object's introduction into the system, the interferometer results in a single spatial mode output due to single photon interference. The presence of an object, however, breaks the interference resulting in a nonzero probability of detecting the photon at the previously dark port. Mathematically, we define the presence of an object as a two-dimensional piecewise operator $\hat{O}$,

$$\hat{O} = \begin{bmatrix} 1 & 0 \\ 0 & \beta \end{bmatrix}, \quad \beta = \begin{cases} 1, & \text{no object present} \\ 0, & \text{object present} \end{cases}. \tag{6.3}$$

For reasons that will be dually apparent later, we define the input mode of the interferometer to be spatially opposite to the mode containing the object (reference figure 6.1); or in the absence of an object, this mode is considered to still have a nonzero loss probability coefficient. In practice, this could be implemented by maintaining a lossless control arm in a high quality, carefully constructed fiber loop.

To further generalize this setup, it is possible to consider the effects of a non-unitary loss probability in either or both arms of the interferometer. For sensing applications, we assume a lossless control arm and define any arbitrary environmental loss and/or phase modulation to be present in the lower spatial mode only.

$$\hat{L} = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi}\sqrt{1-L} \end{bmatrix} \tag{6.4}$$

Generalizing the rotation angle of the beam splitters and assuming a single incident photon, the output of this two mode interferometer is given by

$$\begin{aligned} |output\rangle &= \hat{U}_2 \hat{O} \hat{L} \hat{U}_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta_1 \cos\theta_2 - e^{i\phi}\sqrt{1-L}\sin\theta_1 \sin\theta_2 \\ i\left(\cos\theta_1 \sin\theta_2 + e^{i\phi}\sqrt{1-L}\sin\theta_1\cos\theta_2\right) \end{bmatrix}. \end{aligned} \tag{6.5}$$

To maintain a dark port arrangement for this interferometer requires $\theta_1 + \theta_2 = \pi/2$ (implying $R_1 = T_2$ and $T_1 = R_2$) [30], balanced path lengths ($\phi = 0$), no loss (alternatively, equivalent loss in both arms), and no object present in either mode. Assuming these conditions guarantees that the photon will always be detected at the bright port. Once an object is introduced into one of the modes, however, there becomes some nonzero probability of detecting the photon at the dark port. The obvious problem is that there is also a nonzero probability of the photon hitting the object and spoiling the interaction-free aspect of the device. [22] defines the efficiency of a successful interaction-free measurement as

$$\eta = \frac{P_{ifm}}{P_{ifm} + P_{abs}} = \frac{R_1 R_2}{R_1 R_2 + T_1} = \frac{1 - T_1}{2 - T_1} \tag{6.6}$$

where P_{ifm} is the probability of a successful interaction free measurement and P_{abs} is the probability of photon absorption by the object. From (6.6), it can be seen that as $T_1 \to 0$ (correspondingly, $R_2 \to 0$) such a scheme is only possible of achieving 50% efficiency. It is important to note, however that the condition for maximum efficiency ($T_1 = 0$) implies that the photon will never take the path containing the object, and the efficiency decreases as the photon becomes more probable of interacting with the object.

6.3 Quantum Zeno Effect

An improved setup for interaction-free measurement, proposed by Kwiat et al., increased the number of beam splitters to an arbitrarily large number of “passes” through the single pass setup discussed in the previous section [22].

This method requires that the reflectivity of the central beam splitter be reduced (thereby increasing the conjugate transmissivity) in order to create a process of “coherent evolution” from the input mode (upper mode as shown in figure 6.2) of the interferometer to the output (i.e., lower) mode after N beam splitters, effec-

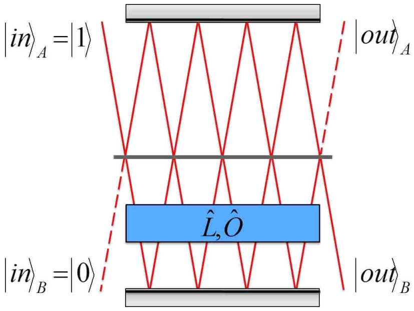


FIGURE 6.2. In a practical implementation, multiple passes through a MZI can be emulated by a dual cavity resonator composed of asymmetric Fabry-Perot interferometers (N=5 shown here). Theoretically, outer walls of the resonator are perfectly reflecting, while the inner barrier acts as a variable beam splitter. In this arrangement, the object (or loss) must be uniformly distributed along the lower cavity for the analysis presented.

tively rebalancing the dark port arrangement. As implied in figure 6.2, the incident angle and length of the cavity determine the number of effective conjoined MZI passes are present and must be "in tune" with the reflectivity of the central beam splitter.

Recalling the definition for the beam splitter (6.2) and its equivalence to a uniaxial rotation operator, it follows that a succession products of rotations (about the same axis) is equivalent to a single rotation about the accumulated single angle. That is, our unitary beam splitter obeys the following rotational property [32]:

$$\hat{R}(\theta)\hat{R}(\phi) = \begin{bmatrix} \cos\theta & i\sin\theta \\ i\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\phi & i\sin\phi \\ i\sin\phi & \cos\phi \end{bmatrix} \tag{6.7}$$

$$= \begin{bmatrix} \cos(\theta+\phi) & i\sin(\theta+\phi) \\ i\sin(\theta+\phi) & \cos(\theta+\phi) \end{bmatrix} \tag{6.8}$$

$$= \hat{R}(\theta+\phi) \tag{6.9}$$

To achieve a fully orthogonal transformation on a two-level state, a 90°rotation ($\hat{R}(\pi/2)$) must be applied which results in

$$\hat{U} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} i = \begin{bmatrix} 0 \\ 1 \end{bmatrix} i. \tag{6.10}$$

By inspection, the parameter θ can be redefined as $\theta = \pi/(2N)$ such that after N passes through the beam splitter,

$$\hat{U}^N = \hat{R}^N(\theta) = \hat{R}(N\theta) = \hat{R}(\pi/2) \tag{6.11}$$

which is exactly equivalent to the orthogonal unitary transformation above.

As illustrated in figure 6.2, this is most easily expressed in the polarization basis were the total polarization rotation (after N beam splitters) should be orthogonal to the input polarization. Specifically,

$$\sum_{n=0}^{N} \theta_n = \frac{\pi}{2} \tag{6.12}$$

where θ_n is the beam splitter rotation angle as given in (6.2). Since this is a one-dimensional rotation, the product of any number of individual rotations is equal to a single rotation operation over the total (summed) angle. Therefore, to maintain orthogonal evolution, the reflection and transmission angles of each of the n beam splitters is $\theta_n = \pi/2N$.

Specifically, if there is no object present in the interferometer, than the photon is allowed to coherently evolve from the top spatial mode (original polarization) to the bottom spatial mode (orthogonal polarization) unhindered due to a lack of “which path” information between the two modes. The presence of an object, or any amount of loss, in a single mode makes an effective hard measurement on the mode containing the object providing information about the photon’s path and destroying the bimodal entanglement. The reduction in polarization rotation decreases the probability of the photon taking the lossy/object mode; therefore, any measurement made by the loss/object will find the photon in the lossless path with high probability. This collapses the bimodal superposition probability of the coherent evolution process and preserves the state of the input.

The ultimate theoretical advantage of this scheme lies in its efficiency. Recalling the definition of efficiency from (6.6), we can redefine P_{ifm} and P_{abs} in terms of the total number of beam splitters N as

$$P_{ifm} = \cos^{2N}\theta_n \tag{6.13}$$

$$P_{abs} = \sin^2\theta_n \sum_{k=0}^{N-2} \cos^{2k}\theta_n = 1 - (\cos^2\theta_n)^{N-1}. \tag{6.14}$$

However, in the limit of large N,

$$\lim_{N\to\infty}\left(\cos^{2N}\theta_n\right) \approx \lim_{N\to\infty}\left(1 - \frac{\pi^2}{4N}\right) = 1. \tag{6.15}$$

Therefore, in this limit, $P_{ifm} \to 1$ and $P_{abs} \to 0$; thus the efficiency asymptotically increases to unity.

6.4 Polarization Based IFM

It has been shown that by increasing the number of beam splitters (equivalent to redirecting the output modes of figure 6.1 back to the input modes for subsequent passes through the interferometer; or more practically, utilizing a dual Fabry-Perot interferometric setup) improves IFM probability while decreasing photon-object interaction probability [22]. This effect results from the fact that the reflectance of the beam splitter must be reduced to maintain the dark port condition in the absence of an object.

The scheme containing $N > 2$ beam splitters is more naturally implemented in the orthogonal polarization bases $|V\rangle$ and $|H\rangle$ through the use of polarizing beam splitters. In this configuration, a half-wave plate is used to rotate the state's polarization by a small angle N times such that $\sum_k^N \theta_k = \pi/2$ (refer to figure 6.3). This orthogonal total rotation angle ensures the dark port arrangement of the lossless interferometer [23]. Now the reflectance of the birefringent polarizing beam splitters is a function of the angle of linear polarization of the electric field of the

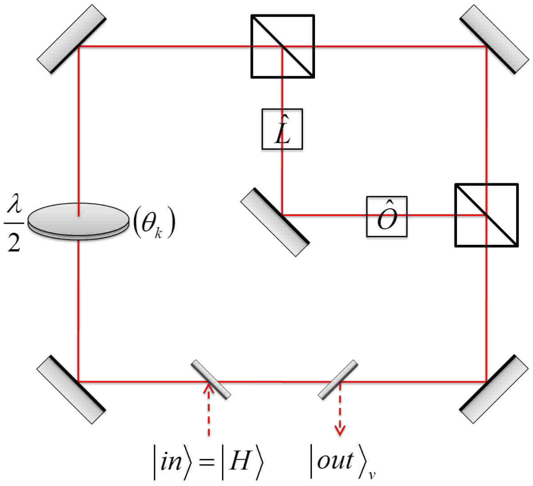


FIGURE 6.3. Schematic for an MZI operating in the polarization basis. The photon is cycled through the device effectively creating a series of adjacent MZIs while the HWP rotates the state by a small angle at the start of every cycle. Adhering to convention, the reflectivity and transmissivity of the two beam splitters are redefined to ensure a gradual coherent evolution from the lossless control mode to the target mode. In the presence of an object or loss, this evolutionary process is spoiled and the photon remains in the lossless path.

photon (i.e., if the polarization angle $\theta = 0$ corresponds to state $|H\rangle$, then $R = 0$; if the polarization angle is $\theta = \pi/2$ corresponds to state $|V\rangle$, then $R = 1$). This implies that the unitary transformation corresponding to a single cycle through the lossless interferometer is now determined by the half-wave plate instead of the beam splitter(s) directly. However, since the state is literally being rotated in polarization by angle θ, the rotation operator as equation (6.2) applies. Note that the interferometer can now be used for sensing, where PBS_1 entangles the state between a carefully maintained (lossless) control arm and a (possibly lossy) target arm. In this scenario, a delay line on the control arm must be implemented to maintain temporal coincidence of the state on PBS_2. After k of N cycles, a measure of it's polarization gives some information about the target arm of the sensor: an orthogonal projection implies a clear and lossless path, whereas $\langle input|output\rangle = 1$ implies some obstruction of the target arm – or phase shift resulting from retarded wave packet velocity, also signifying the presence of some transparent obstruction. Assuming that the input state is horizontally polarized ($|in\rangle = |H\rangle$), the respective probabilities of detecting the photon horizontally or vertically polarized after k cycles is

$$P_{H,k} = |\langle H|\hat{M}^k|H\rangle|^2 \tag{6.16a}$$

$$P_{V,k} = |\langle V|\hat{M}^k|H\rangle|^2 \tag{6.16b}$$

where $\hat{M}$ is a transformation operator effectively simulates one full cycle through the interferometer (including half-wave plate rotation, mirror phase effects, loss interaction, and object measurement). In the absence of loss or an object, the state should perform a fully orthogonal rotation; therefore, this transfer matrix must obey the following relations:

$$P_{H,N} = |\langle H|\hat{M}^N|H\rangle|^2 = |\langle H|V\rangle|^2 = 0$$

$$P_{V,N} = |\langle V|\hat{M}^N|H\rangle|^2 = |\langle V|V\rangle|^2 = 1$$

6.5 Detection Methods

Thus far, the mathematical and physical description of the problem has been idealized by presuming lossless modes of the interferometer with respect to the ultimate goal of determining the presence of an object in the path of the interferometer. In the absence of an object, a detection at the output would result in a polarization angle of exactly $\theta = \pi/2$ relative to the input. If an object is present, however, the quantum Zeno effect would prevent the state from rotating the full amount. The concern would then be for the introduction of environmental loss, or a transparent or semi-transparent object. As will be detailed in the following chapters, the answer to these is essentially the production of an attenuated, non-orthogonal state vector at the output port.

Due to the non-orthogonality of the output state given appreciable loss $0 < L < 1$, any projective measurement of its polarization could – in the worst case – be completely random and yield no information about the detection. Because of this, there have been numerous proposals for state discrimination at the detection stage [34]. Such processes can become quite complicated, requiring ancillary resources while still resulting in various probabilities of correct discrimination (P_d). The scheme introduced by [34] is on in which the polarization state is purposefully rotated $\theta < \pi/2$ to take advantage of specific established measurement processes. Some such processes include an Ivanovic-Dieks-Peres measurement which requires ancillary photons and have a success probability of $P_d = 1 - \cos(k\theta)$. Another randomly chooses between measurement bases ("guess and measure" strategy) with success of $P_d = (1 - \cos^2(k\theta))/2$. Other measurement schemes, such as the Helstrom measurement, allow for the possibility of a misidentification, which can obfuscate results. As will be presented in later chapters, it is in part the objective of this thesis to present a theoretical alternative for handling such lossy scenarios in a practical manner.

Chapter 7

Analytical Analysis of Lossy IFM

7.1 Motivation

This chapter deals directly with what is arguably the most intriguing, yet most challenging, aspect of a lossy Mach-Zehnder interferometer: analytical descriptions for the evolution of the photon's wave function's as it propagates through the modes of the interferometer. While, as implied by the title, these derivations were performed with the specific goal of further understanding the characteristics of lossy interaction free measurement, all calculations are based on the assumption that an object is *not* present in the interferometer and are therefore universally applicable to generalized two mode interferometric schemes with single mode loss.

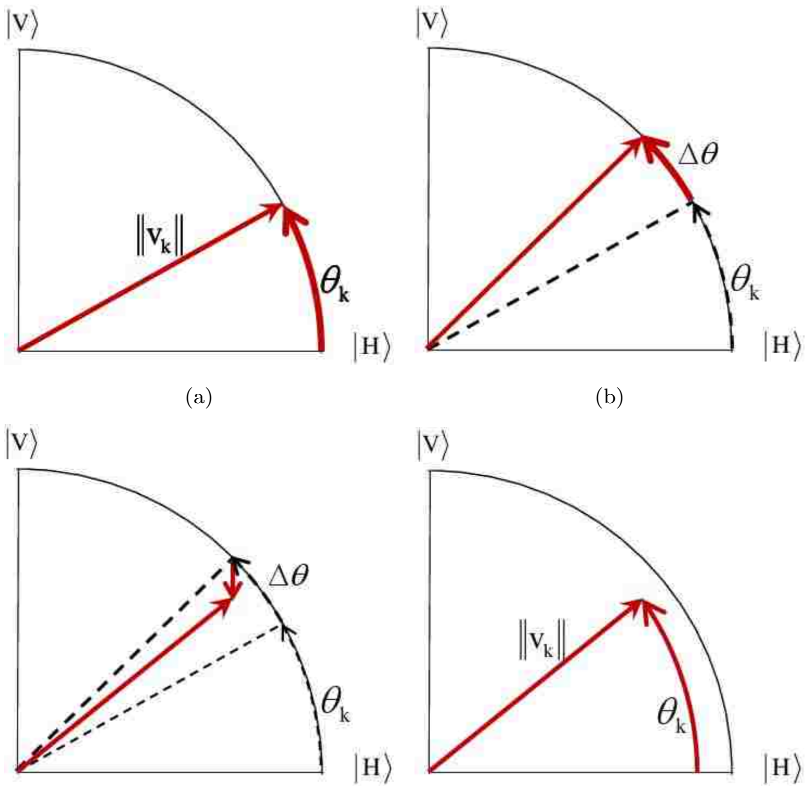
The stipulation for the inception of this scheme being that, in the absence of losses, the interference pattern should be configured such that destructive interference occurs at one of the output ports. Further assumptions, in the event of multiple cycles through the interferometer, include congruence among the individual cycles (i.e., any elements or conditions directly affecting the photon's wave vector must be identical among all cycles). This chapter seeks to investigate the single event of coherent evolution across the two mode subspace from various mathematical perspectives in order to gain insight into the behavior of interaction free measurement and the assess the possibility of enhancing the established schemes.

7.2 Trigonometric Formalism: Coherent Evolution

By its very nature, polarization based interferometry inevitably utilizes a half-wave plate and/or polarizing beam splitters at some point in the schema in order to rotate the state to/from various superpositions of the orthogonal bases. In the polarization based concept of interaction free measurement discussed in the previous section, the angle of rotation due to the half-wave plate will cause the photonic state (polarization or spatially based) to coherently evolve into the orthogonal state over the duration of N cycles. The introduction of loss, however, hinders the complete evolution of the state. Figure 7.2 shows the evolution of a state vector over the duration of a single cycle through the interferometer. In each step of the process, a right triangle can by projection of the state vector onto the horizontal basis. This, in combination with the fact that each half-wave rotation is $\Delta\theta = \pi/(2N)$ leads to a physically intuitive understanding of both the states instantaneous rotation and attenuation any cycle.

Figures 7.1(a) and 7.1(b) illustrate the process of the photon passing through the half-wave plate on the kth instantaneous pass of N total passes, and represents its related rotation angle without attenuation. After being vertically attenuated as in figure 7.1(c), the resulting state is technically the $(k+1)$th state; however, from a relative perspective, the process will now repeat with this attenuated state as the "new" kth cycle. In this sense, figure 7.1(d) is equivalent to figure 7.1(a) for the next cycle. This nested process loop continues until $k = N$.

As defined, the amount of polarization rotation per cycle is inversely proportional to the total number of cycles. Also, the probability amplitude of the photon's existence in the lossy path is – at all times – nonzero and increases with each pass through the half-wave plate for a finite number of cycles. Our scheme assumes that amplitude attenuation due to loss affects only the $|V\rangle$ mode, where an object has the potential to enter the beam. For the general state $|\Psi\rangle = c_H|H\rangle + c_V|V\rangle$ the lossy mode is attenuated directly as λc_V, which effectively decreases the rotation



(a) (b) (c) (d)

FIGURE 7.1. Trigonometric analysis of state attenuation and saturation angle. (a) shows the horizontally polarized input state rotated by angle θ_k after k passes through the interferometer without loss. For the purpose of illustration, loss in introduced only on the $k+1$ cycle. (b) shows the $k+1$ pass through the half-wave plate, rotating the state by $\Delta\theta$. After this rotation, (c) illustrates a small attenuation in the vertical path only. This yields a state with less amplitude and rotation angle as compared to the lossless case, as seen in (d).

angle by $\tan^{-1}(\lambda \tan\theta)$ and the total magnitude by $c_V\lambda$, where $\lambda = \sqrt{1-L}$ is the loss factor, and L is the actual probability of loss.

This lends itself immediately to the following recursion relations for effective rotation and vector magnitude per cycle, respectively:

$$\theta_k = \tan^{-1}\left(\lambda \tan\left[\theta_{k-1} + \Delta\theta\right]\right) \tag{7.1a}$$

$$\|v_k\|^2 = \|v_{k-1}\|^2 \left(\cos^2\left[\theta_{k-1} + \Delta\theta\right] + \sin^2\left[\theta_k + \Delta\theta\right]\right) \tag{7.1b}$$

By inspection, these can be cast as composite functions representing the terminal rotation and magnitude.

$$\|v_N\| = \prod_{k=0}^{N-1} A\left(B^{(k)}\left(\Delta\theta\right)\right) \tag{7.2a}$$

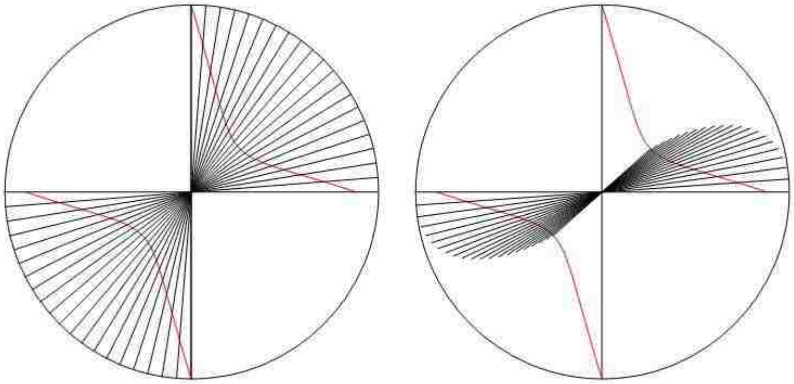
$$\theta_N = B^{(N)}\left(\Delta\theta\right) - \Delta\theta \tag{7.2b}$$

where

$$A\left(x\right) = \left(\cos^2 x + \lambda^2 \sin^2 x\right)^{1/2}$$

$$B\left(x\right) = \Delta\theta + \tan^{-1}\left(\lambda \tan x\right)$$

Given the nested nature of these composite functions, a directly analytical plot is not possible. However, simple programming loops can easily produce the numerical results for each instantaneous cycle.



(a)

(b)

FIGURE 7.2. For the lossless configuration (a), there is no deviation from $\pi/2n$ resulting in a complete orthogonal rotation and no dimished vector amplitude. For non negligible loss (b) , the vertical amplitude attentuation results in a significant reduction of rotation angle and vector amplitude. The red line illustrates the terminal state magnitude (after N cycles) for continuous values of loss from lossless (uppermost point) to total loss (intersection with horizontal axis) while maintianing a fixed number of cycles (20 shown). Notice again, the distinctly attentuated magnitude at intermediate loss values representing maximum loss probability.

Figure 7.2 is a polar representation of the state's rotational saturation and terminal amplitude due to loss, and a trace of maximum loss per cycle. Each radial on the plot represents the state vectors instantaneous square magnitude (i.e., the magnitude of the state's probability of existence) for the kth of N cycles. Also, the angle of each radial is a direct representation of the state's angle during that cycle. Here, $N = 20$, so there are 20 bisections through the first and third quadrants to represent the characteristics of the state vector at each discrete cycle. Also, the red curve represents the terminal probability of the state to exit the interferometer in either mode (i.e., the probability that the photon is not absorbed by loss). Figure 7.2(a) represents the lossless scenario and exhibits the expected coherent evolution into the orthogonal state with no probability of photon loss.

Figure 7.2(b), however, is experiencing a 20% loss probability in the vertical mode. This causes the state vector to continually decrease in magnitude and effective rotational angle. The total loss probability traced by the red curve is given by

$$P_l = 1 - \|v_N\|^2 \tag{7.3}$$

and clearly has a maximum that occurs on the interval $0 < L < 1$. The exact value, however, is non-trivial due to its dependence on the recursively composition nature of the governing equations.

7.3 Linear Algebraic Formalism: Eigensystem

The necessity for implementation of the quantum Zeno effect directly suggests a very the repetitive nature of the interferometer. Mathematically, this means that the same transformations must be reapplied N times. From the point of view of a Jones calculus formalism, this transformation can be performed by direct matrix multiplication. The electric field vector is given by

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} = \begin{bmatrix} |E_x|\, e^{i\psi_x} \\ |E_y|\, e^{i\psi_y} \end{bmatrix} \tag{7.4}$$

The optical elements involved in the process include a half-wave plate, mirrors, and the object interaction probability matrix. First, the Jones matrix for a half-wave plate is given by [24]

$$\hat{M}_{\lambda/2} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}. \tag{7.5}$$

The Jones matrix for an odd number of mirrors per path per cycle (which induces a relative phase shift of $e^{i\pi}$), is

$$\hat{M}_{mir} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{7.6}$$

Finally, the matrix representing the objects influence on the state in the vertical mode only is

$$\hat{M}_{int} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda e^{i\phi} \end{bmatrix} \tag{7.7}$$

where $\lambda = \sqrt{1-L}$ is the amplitude of the inverse loss probability and ϕ is any phase imparted on the state while interacting with the loss. Clearly, $\lambda = 0$ would be analogous to have an opaque object in the path.

The cumulative transfer matrix representing the process of one cycle through the interferometer is then given by the product of the above.

$$\begin{aligned} \hat{M} &= \hat{M}_{mir}\hat{M}_{int}\hat{M}_{\lambda/2} \\ &= \begin{bmatrix} \cos\theta & \sin\theta \\ -\lambda e^{i\phi}\sin\theta & \lambda e^{i\phi}\cos\theta \end{bmatrix} \end{aligned} \tag{7.8}$$

The implementation of repeated cycles requires that it be reapplied to the state recursively. This would be analogous to raising a scalar variable to the power of N, however special care must be taken for the case of a matrix power.

As only elements of diagonal matrices may be directly raised to a power, it becomes necessary to diagonalize the established transfer matrix. This is accomplish by a unitary transformation whose elements are composed of the eigenvectors of the transfer matrix. Solving the eigensystem and appending the eigenvectors into columns of a 2x2 matrix, this unitary operator is defined as

$$\hat{U} = \begin{bmatrix} \frac{\cos\theta(\lambda-1)+\sqrt{\cos^2\theta(\lambda-1)^2-4\sin^2\theta\lambda}}{2\sin\theta\lambda} & \frac{\cos\theta(\lambda-1)-\sqrt{\cos^2\theta(\lambda-1)^2-4\sin^2\theta\lambda}}{2\sin\theta\lambda} \\ 1 & 1 \end{bmatrix}. \tag{7.9}$$

the eigenvalues of the system are found by applying this unitary transformation to the transfer matrix as

$$\hat{D} = \hat{U}^{-1}\hat{M}\hat{U} \tag{7.10}$$

$$= \begin{bmatrix} \lambda_m & 0 \\ 0 & \lambda_p \end{bmatrix} \tag{7.11}$$

where λ_p and λ_m are the eigenvalues given by

$$\begin{aligned} \lambda_p &= \frac{1}{2}\left(\cos\theta(\lambda+1) + \sqrt{\cos^2\theta(\lambda-1)^2 - 4\sin^2\theta\lambda}\right) \\ \lambda_m &= \frac{1}{2}\left(\cos\theta(\lambda+1) - \sqrt{\cos^2\theta(\lambda-1)^2 - 4\sin^2\theta\lambda}\right). \end{aligned} \tag{7.12}$$

Now that the transfer matrix is in diagonal form, it can be recursively applied to the resultant state vector an arbitrary number of times. Once the process is complete, the inverse unitary transformation must be applied to obtain the proper output state. Specifically,

$$\hat{M}_k = \hat{U}\hat{D}^k\hat{U}^{-1}. \tag{7.13}$$

This results in a tedious equation without lending much intuition, and so it shall be omitted. Far more enlightening is the final output of the interferometer once the transfer matrix has been recursively applied N times.

$$\mathbf{v}_k = \hat{M}_k \mathbf{v}_0 = \hat{M}_k \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{7.14}$$

$$= \begin{bmatrix} \frac{\cos\theta \lambda_m \lambda_p (\lambda_m^n - \lambda_p^n) + \lambda(\lambda_m \lambda_p^n - \lambda_m^n \lambda_p)}{\lambda_m(\lambda_m - \lambda_p)\lambda_p} \\ -\frac{\sin\theta \lambda(\lambda_m^n - \lambda_p^n)}{\lambda_m - \lambda_p} \end{bmatrix}. \tag{7.15}$$

After N cycles, the half-wave plate has made a cumulative total rotation angle of $N\theta = \pi/2$. At this point, there are essentially two possibilities for obtaining the output (prior to any state measurement): after the half-wave plate, or after the second polarizing beam splitter of the Mach-Zehnder interrogator.

Figure 7.3 shows the spatial and polarization analogs for these two specific output methods. As represented in figures 7.3(a) and 7.3(b), when the output is taken after the Nth rotation, it is not permitted to interrogate the object's presence a final Nth time (N rotations and $N - 1$ measurements). Figures 7.3(c) and 7.3(d) represent the situation where the photon is allowed through the interrogating interferometer a final, Nth time (N rotations and N measurements). An interesting subtly of this second setup is that when an object is present, there will be exactly 0% chance of a detection in that mode (lower mode for 7.3(c) or $|V\rangle$ mode for 7.3(d)). Clearly this will increase the probability of interacting with the object and decrease the overall detection probability; however, as this "bright port" detection mode offers no information on the presence of an object regardless of the setup, no information is lost by having no detection at this port.

The equivalent mathematical descriptions of these scenarios differ by one power in the transfer matrix. Assuming state $\mathbf{v}$ is represented by 7.3(c) and 7.3(d), its expression is

$$\mathbf{v}_N = \hat{M}^N_{\left(\theta = \frac{\pi}{2N}\right)} \mathbf{v}_0. \tag{7.16}$$

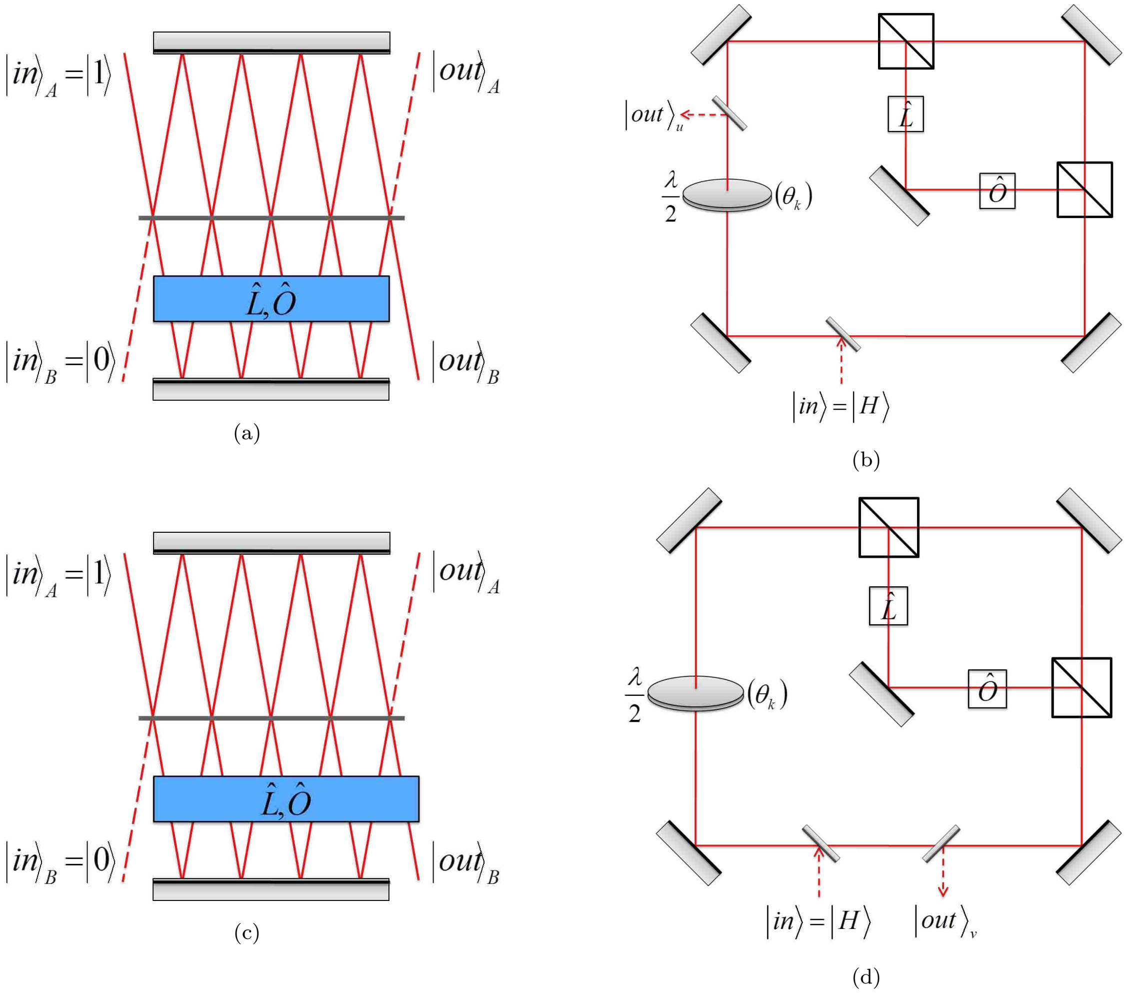


FIGURE 7.3. Illustration of whether a final object measurement is made after the state has been rotated N times such that the sum of the rotations is $N\theta = \pi/2$. Analogs for spatial and polarization bases are depicted side by side.

Defining the alternative state – depicted in 7.3(a) and 7.3(b) – as $\mathbf{u}$ and ensuring output immediately have its rotation gives

$$\mathbf{u}_N = \hat{M}_{\lambda/2}\hat{M}^{N-1}_{\left(\theta=\frac{\pi}{2N}\right)}\mathbf{v}_0. \tag{7.17}$$

The implications of these subtle difference will be discussed fully in later sections. Once the choice of output method has been made, the detection probabilities for the output state are then just the absolute square of the two components of the output vector $\mathbf{u}$ or $\mathbf{v}$.

7.4 Combinatoric Formalism: Path Counting

It's perhaps conceptually intuitive, although a bit mathematically tedious, to consider the interference effects of the single photon in the spatial arrangement as a string of possible paths that the photon could take as it reaches each beam splitter. In a classical sense, this photon could be personified as reaching a fork in the road, flipping a coin, and taking the resulting random path (a full quantum description requires a little more care to obtain proper self-superposition, but we'll get there). This results in a total of 2^N possible paths the photon could traverse.

In the simplest of examples, consider the Fabry-Perot of figure 6.2 with no losses or objects and an arbitrary number of beam splitter incident events, N. Probability dictates that each of N events have probabilities R and T of reflecting or transmitting, respectively. These are summed as independent probabilities – as the photon must take one or the other of the two paths, *classically*. Successive incident events are products of this sum, as they are not independent of the preceding events. Therefore, the resulting classically description of all possible path probabilities is given by $(R+T)^N$. Note that $R+T=1$ (and thus $(R+T)^N=1$), which is expected given no loss or objects introduced to the system. Although impressionably trivial, this binomial expansion indeed describes each and every possible *path* the photon could take. When considering the same counting argument, but associating the effects of loss to all probabilities incident on the beam splitter from the lower mode only, the complexity of the counting is immediately apparent.

The first restriction to recognize is the requirement for output at the upper port versus the lower port. Again referring to figure 6.2 whose input is in the upper mode, *odd powers of* T will be output at the *lower port* while *even powers of* T will be output to the *upper port*. Given that these are orthogonal modes, we can analyze their associated paths independently. First, let's tackle the paths that lead

TABLE 7.1. State coefficients for all possible paths through an interferometer resulting in lower mode output. A comparison is given between $N = odd$ as in (a) and $N = even$ in (b)

(a)

$N = 7$	λ^0	λ^1	λ^2	λ^3	λ^4	λ^5	λ^6	λ^7
R^6T^1	0	1	1	1	1	1	1	1
R^4T^3	0	0	5	8	9	8	5	0
R^2T^5	0	0	0	6	9	6	0	0
R^1T^7	0	0	0	0	1	0	0	0

(b)

$N = 8$	λ^0	λ^1	λ^2	λ^3	λ^4	λ^5	λ^6	λ^7	λ^8
R^6T^1	0	1	1	1	1	1	1	1	1
R^4T^3	0	0	6	10	12	12	10	6	0
R^2T^5	0	0	0	10	18	18	10	0	0
R^1T^7	0	0	0	0	4	4	0	0	0

to output into the lower port (T^{odd}). It will also be necessary to generalize for both even and odd total number of beam splitters (rotations), N. Through tedious effort, it can be shown that for $N = 7$ beam splitters and $N = 7$ loss factors, a matrix of the state coefficients can be constructed. This takes the form of table 7.1(a).

The coefficients of table 7.1 are valid for a single photon input incident on a mode orthogonal to the mode containing the loss parameter. The first column, $0\lambda^0$, explicitly shows the presence of loss after the final beam splitter (i.e., the photon cannot exit the lower mode without interacting with the loss parameter). For $N = 8$, there still remains a maximum of T^7 transmissions to ensure output at the lower port. The corresponding coefficients for output into the top mode are also given in table 7.2

By inspection, it is found that this recursion of coefficients is described by a figurate number series. Specifically, these are the anti-diagonals of a multiplication table of k dimensional triangle numbers (where T^{2k+1} and k therefore $k \in \{0, 1, ..., \lfloor N/2 \rfloor\}$). Specifically, the *nth* triangle number in k dimensions is

TABLE 7.2. State coefficients for all possible paths through an interferometer resulting in upper mode output. A comparison is given between $N = odd$ as in (a) and $N = even$ in (b)

(a)

$N = 7$	λ^0	λ^1	λ^2	λ^3	λ^4	λ^5	λ^6	λ^7
R^7T^0	1	0	0	0	0	0	0	0
R^5T^2	0	6	5	4	3	2	1	0
R^3T^4	0	0	10	12	9	4	0	0
R^1T^6	0	0	0	4	3	0	0	0

(b)

$N = 8$	λ^0	λ^1	λ^2	λ^3	λ^4	λ^5	λ^6	λ^7	λ^8
R^8T^0	1	0	0	0	0	0	0	0	0
R^6T^2	0	7	6	5	4	3	2	1	0
R^4T^4	0	0	15	20	18	12	5	0	0
R^2T^6	0	0	0	10	12	6	0	0	0
R^0T^8	0	0	0	0	1	0	0	0	0

given by

$$\mathbb{T}_n^k = \mathbb{C}_k^{n+k-1} = \frac{n^{\bar{k}}}{k!} = \frac{n^{(k)}}{k!} \tag{7.18}$$

where $\mathbb{C}$ is the binomial coefficient and $n^{\bar{k}}$ is a rising factorial equivalent to $n^{(k)}$ in Pochhammer notation. These triangle numbers also appear in Pascal's triangle along the diagonals. This is shown in figure 7.4, which is essentially a realigned Pascal triangle with shifted index.

TABLE 7.3. Triangle number series for a given k and n.

n k	0	1	2	3	4	5	$\cdots$
0	1	1	1	1	1	1	
1	0	1	2	3	4	5	
2	0	1	3	6	10	15	
3	0	1	4	10	20	35	
4	0	1	5	15	35	70	
$\vdots$							$\ddots$

Table 7.4 lists the explicit values for various dimensional triangle series, where the value of k determines the number of transmitted paths as T^{2k+1}, and n determines the number of times the photon takes the path with loss. For lower mode

output, the state coefficients are determined from anti-diagonal product of these triangle series (rows or columns, given its symmetry). That is, given a specific number of beam splitters (N) and a specific number of transmissions in the path ($2k+1$), the loss coefficients ($c_{k,n}\lambda^n$ where $n \in \{1, 2, .., N\}$) would be given by

$$c_b k, n = \mathbb{T}_n^k \mathbb{T}_n^{N-k}. \tag{7.19}$$

A similar analysis for paths resulting in an upper mode output also yield these sets of higher dimensional triangle series. The difference being that the photon *can* exit the top mode without any interaction with a loss probability. Therefore, $n = 0$ is now possible while $n = N$ is not. Another artifact of the output state is a shift in dimensionality between the triangle series product terms, which becomes

$$c_b k, n = \mathbb{T}_n^k \mathbb{T}_{n+1}^{N-k}. \tag{7.20}$$

Given that, for $n \geq 1$, the matrix for the triangle series is symmetric, the loss coefficients above are merely examples as the indices may be interchanged. The terms will begin to take a more rigorous form once further restrictions are applied. Namely, when handling the asymmetry of column $n = 0$, and the imposition of the reflection and transmission probabilities as related to upper and lower mode outputs.

Up to this point, as mentioned previously, all equations have been assuming real-valued probabilities simplicity and convenience as the counting arguments are universal. At this point it is necessary to consider the physically accurate quantum analog of the scenario. In fact, we have not been dealing with true probabilities, but complex valued probability amplitudes describing the quantum superposition of the interferometer. It is only in this manner of analysis that mathematical framework allows for self-interference effects of the single photon. Luckily, as with field quantization, the “conversion” is straightforward. One can simply make the replacements $R \rightarrow r$ and $T \rightarrow (it)$. The equation now describes the output *state* of the interferometer and not the probability of any specific occurrence.

Having found the form of the loss coefficients for a static number of transmissions, it is necessary to consider the coefficients for all paths containing all possible values of transmission. This process will yield the values of table 7.1. This is done via a nested sum over the possible transmissions and their subsequent loss coefficients. Extrapolating from the known coefficients, path restrictions, and substituting the quantum variable analogs gives

$$\langle B|out\rangle = \sum_{n=0}^{\lfloor N/2 \rfloor} r^{N-(2n+1)}(it)^{2n+1} \sum_{k=0}^{N-1} \lambda^{k+1} \mathbb{C}_n^{N-1-k} \mathbb{C}_n^{k} \tag{7.21a}$$

$$\langle T|out\rangle = r^N + \sum_{n=1}^{\lfloor N/2 \rfloor} r^{N-2n}(it)^{2n} \sum_{k=1}^{N-1} \lambda^{k} \mathbb{C}_n^{N-k} \mathbb{C}_{n-1}^{k-1} \tag{7.21b}$$

where $\langle T|out\rangle = |T\rangle$ and $\langle B|out\rangle = |B\rangle$ are the projections of the output state onto the “Top” and “Bottom” mode bases. The real-valued probability of output into these modes is now the complex square of these two quantities individually.

Note the state’s dependence on the loss transmission probability amplitude, λ. In the top state, the possibility of the photon escaping with interferometer without an loss interaction is explicitly given by the reflectance term preceding the initial summation ($\lambda^0 r^N$). Subsequent loss terms are handled by the summation for $k \in \{1, 2, .., N-1\}$, which directly implies the impossibility of have N loss interactions. Similarly, the loss parameter in the bottom state is summed N total times. From this, it is obvious that λ^{k+1} signifies loss after the final beam splitter and the corresponding impossibility of have non-interaction with loss. On the other hand, if this term were changed to λ^k, the summation would clearly indicate the absence of loss after the final beam splitter and allow for a lossless path to the bottom output state.

From this point, it is possible to recast the binomial components in terms of rising factorials in an effort to further simplification and intuitive analysis. To outline the process, consider the $|B\rangle$ state component. The summation of n truncates the sum at a lower limit of $n = 0$ and an upper limit of $n = \lfloor N/2 \rfloor$. Beyond this range, the indices may be shifted at will without consequence. To exploit this freedom, the order of summation may be interchanged; and, without loss of generality, the summation can be re-indexed to match the summation over k. This yields

$$|B\rangle = itr^{N-1} \sum_{k=0}^{N-1} \lambda^k \sum_{n=0}^{N-1} \frac{(k-N+1)^{(n)}(-k)^{(n)}}{1^{(n)}} \frac{(-T/R)^n}{n!} \tag{7.22}$$

where the Pochhammer notation for a rising factorial is used.

$$x^{(y)} = x(x+1)(x+2)\cdots(x+y-1) = \frac{(x+y-1)!}{(x-1)!} = y!\mathbb{C}_y^{x+y-1} \tag{7.23}$$

This derivation relies on the fact that the lower index of the binomial coefficient be equivalent. As you can see from equation 7.21b, this is not the case for the indices of the top state loss coefficients. Therefore, Pascal's recursion must be used to shift the index at the cost of creating an extra term of summed binomial products.

$$\mathbb{C}_k^N = \mathbb{C}_{k-1}^{N-1} + \mathbb{C}_k^{N-1} \tag{7.24}$$

Using this relationship and reordering the summations with a symmetric (arbitrary) upper limit, the top state can be written as

$$\begin{aligned} |T\rangle &= r^N + \sum_{k=1}^{N-1} \lambda^k \sum_{n=1}^{N-1} r^{N-2n}(it)^{2n}\mathbb{C}_n^{N-k}\mathbb{C}_{n-1}^{k-1} \\ |T\rangle &= r^N \left[1 + \sum_{k=1}^{N-1} \lambda^k \left(\sum_{n=0}^{N-1} \mathbb{C}_n^{N-k}\mathbb{C}_n^k(-T/R)^n - \sum_{n=0}^{N-1} \mathbb{C}_n^{N-k}\mathbb{C}_n^{k-1}(-T/R)^n \right)\right] \end{aligned} \tag{7.25}$$

which when cast into Pochhammer notation displays a similar structure to equation (7.22)

$$|T\rangle = r^N \left[1 + \sum_{k=1}^{N-1} \lambda^k \sum_{n=0}^{N-1} \frac{(-T/R)^n}{n!} \left(\frac{(k-N)^{(n)}(-k)^{(n)}}{1^{(n)}} - \frac{(k-N)^{(n)}(1-k)^{(n)}}{1^{(n)}} \right) \right] \tag{7.26}$$

Using the fact that the sum over n is truncating and therefore has arbitrary limits, this form of the state component can be immediately recognized as a hypergeometric series of the form

$${}_2F_1 \left[\begin{matrix} a, b \\ c \end{matrix} ; z \right] = \sum_{k=0}^{\infty} \frac{a^{(k)} b^{(k)}}{c^{(k)}} \frac{z^k}{k!}. \tag{7.27}$$

Using this definition and a combination of Euler's hypergeometric transformations,

$$\begin{aligned} {}_2F_1 \left[\begin{matrix} a, b \\ c \end{matrix} ; z \right] &= (1-z)^{-a} {}_2F_1 \left[\begin{matrix} a, c-b \\ c \end{matrix} ; \frac{z}{z-1} \right] \\ &= (1-z)^{-b} {}_2F_1 \left[\begin{matrix} c-a, b \\ c \end{matrix} ; \frac{z}{z-1} \right] \\ &= (1-z)^{c-a-b} {}_2F_1 \left[\begin{matrix} c-a, c-b \\ c \end{matrix} ; z \right] \end{aligned}$$

the orthogonal output state components can ultimately be expressed as

$$|B\rangle = \frac{it}{r^{N+1}} \sum_{k=0}^{N-1} \lambda^k {}_2F_1 \left[\begin{matrix} k+1, N-k \\ 1 \end{matrix} ; -T/R \right] \tag{7.28a}$$

$$\begin{aligned} |T\rangle = r^N &+ \frac{1}{r^{N+2}} \sum_{k=1}^{N-1} \lambda^k {}_2F_1 \left[\begin{matrix} k+1, N-k+1 \\ 1 \end{matrix} ; -T/R \right] + \\ &- \frac{1}{r^{-N}} \sum_{k=1}^{N-1} \lambda^k {}_2F_1 \left[\begin{matrix} k, N-k+1 \\ 1 \end{matrix} ; -T/R \right] \end{aligned} \tag{7.28b}$$

Unfortunately, due to the term $(-T/R)$, the hypergeometric series is "unbalanced" and therefore isn't a candidate for further simplifications through the use of common identities.

7.5 Performance Degradation

Now that the mathematical framework of the system has been presented in detail, some discussion of the physical effects the system arrangement and parameter influence should also be analyzed. When configured for a dark port arrangement, any detection from the dark port should ordinarily signify the presence of an object in the target mode. Any detection from this dark port in the absence of an object is therefore defined as a false positive result. An quantum interferometer is an extremely sensitive device, and as such, it is necessary to understand the effects of imperfections and establish coping mechanisms when possible.

7.5.1 Phase Sensitivity

Throughout all derivations thus far, only loss as a investigative parameter has been presented. In all likelihood, any nonuniform medium capable of inducing photon loss will typically impart some degree of relative phase shift onto the state as well. Mathematically, this doesn't alter the previous derivations beyond the inclusion of another continuous parameter, ϕ, representing a shift in phase of $e^{i\phi}$. This we have assumed all loss to be in the single, "target" arm of the sensing interferometer, this will obviously be the only mode with the phase factor. Since the calculations here involve only two modes, this is equivalent to replacing the two dimensional loss matrix from before with

$$\hat{L} = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi}\lambda \end{bmatrix} \tag{7.29}$$

where it is again assumed that any phase shift will be in the same mode as the loss and a pristine "control" mode is carefully maintained. Another generalized representation for the introduction of a phase shift would be an operator, $\hat{\phi}$, which has an arbitrary mode defined as the eigenstate and the eigenvalue $e^{i\phi}$, as before.

Due to orthogonality, the loss operator acting on any other mode produces no result.

As should be expected, this phase shift will degrade the established interference calibration and often decrease reliable performance. Also, if a phase influential abnormality is indeed present, the cyclic polarization scheme will only serve to exacerbate the effects of any phase shift.

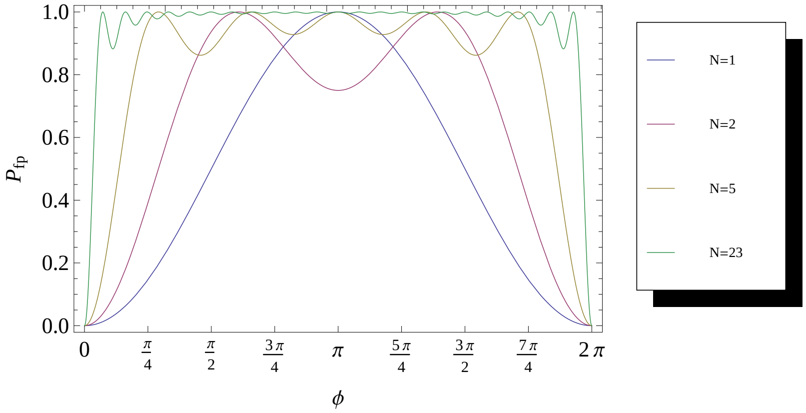


FIGURE 7.4. Probability of a detection at the dark port in the absence of an object as a function of relative phase shift. Here, the MZI is in the dark port configuration with no losses. As expected, the probability of a false positive result becomes more sensitive to phase as the number of cycles is increased. Shown above are plots for N=2,5,10,20 beam splitters (polarization rotations).

Figure 7.4 shows the extreme sensitivity of the interferometer with respect to relative phase shift. This assumes that the phase shift is due to some effect specifically independent of the photon loss probability. The plot represents the number of cycles through a polarization based interferometer, where N represents the number of object measurements made. The output is taken after a final pass through the half-wave plate. This is explicitly illustrated by figure 7.3(b). As expected, increasing the number of cycles quickly saturates the output at the dark port yielding erroneous results for the present application. Given this behavior, it is within the realm of possibility that this setup could make sensitive phase measurements in the absence of loss (with a multiple trial approach similar to [26, 27, 28]); however, the degree of specificity required for full phase analysis is nontrivial and not within the scope of this thesis.

7.5.2 Loss Sensitivity

As with the introduction of phase in the previous section, the presence of any loss in the interferometer will affect the detected output to some degree. Two main scenarios were investigated for this approach: loss solely in the target mode containing the object and equivalent loss in both modes. Again, as with phase, such a setup could be considered quite sensitive to the introduction of loss when attempting to discern the presence of an opaque object, or recommissioned as a means of imaging semi-transparent objects via a more continuous range of detection probabilities.

For an ideal interaction free measurement, any detection at the dark port guarantees the presence of an object in the interferometer. However, with the introduction of single mode loss into the target arm (in the absence of an actual object) will effectively simulate an object to a varying degree. In this case, any detection at the dark port should be considered a false positive result.

As seen in figure 7.5, an increase in target mode loss results in a monotone increase in the probability of a false positive detection. The sharp increase in the high loss regime ($L \approx 1$) is due to the fact that this region is effective an actual opaque object being detected. An increase in the number of cycles directly serves to increase the number of possible interactions the state undergoes with the environmental loss; therefore, an increase in N is directly related to an increase in the false positive detection probability.

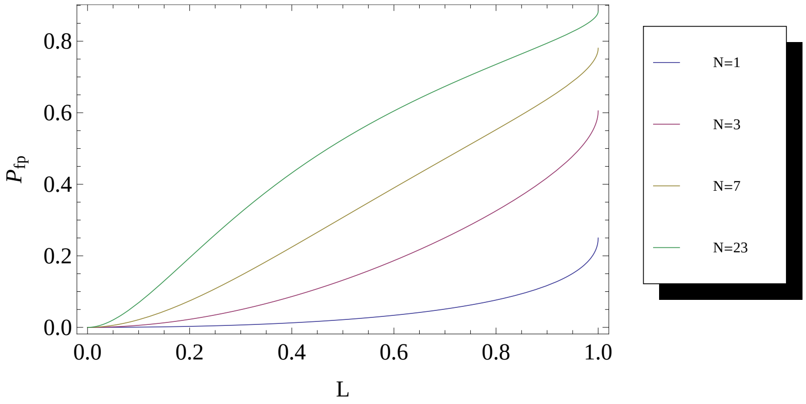


FIGURE 7.5. Probability of a detection at the dark port in the absence of an object as a function of single mode loss. Here, the MZI is in the dark port configuration with no relative phase shifts. As expected, the probability of a false positive result increases as the number of cycles is increased. Shown above are plots for N=2,5,10,20 total beam splitters (polarization rotations).

A result of particular significance to this thesis is the probability of *total* loss in the system. With respect to probability conservation, this total loss is expressed as

$$P_L = 1 - P_D - P_B \tag{7.30}$$

where P_D and P_B are the probabilities of detection at the dark port and detection at the bright port, respectively. This notation will be adopted henceforth, as it is independent of the measurement bases.

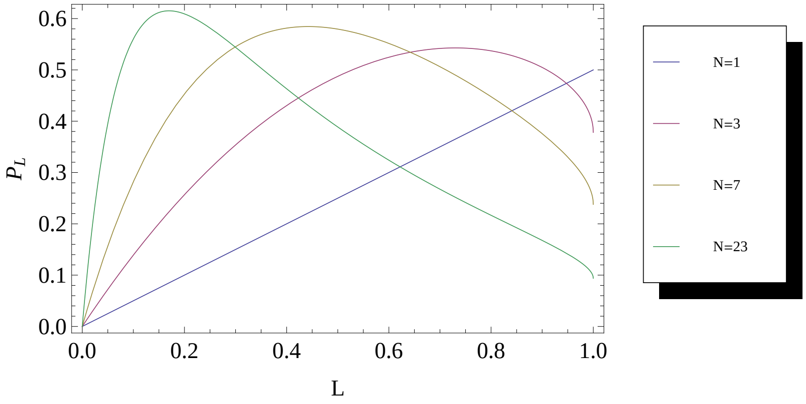


FIGURE 7.6. Nonmonotonic behavior of the photon absorption probablity as a function of single mode loss probability. N represents the number of state rotations with $N-1$ object mode interactions.

Notice from figure 7.6 that the trend of the total loss probability is finite and nonmonotonic for any number of cycles $N > 1$. That is to say, the total detection probability as determined non-discriminately from both output ports yields a minimum value at some loss $0 < L < 1$. The physical interpretation of this lies in the quantum Zeno effect. When $L = 1$, there is an effective hard measurement on the state ensuring its existence in the lossless control arm; however, $L < 1$ corresponds to a soft measurement thus allowing for an increased probability of it's entangled component to exist in the lossy target arm thereby increasing its absorption probability in subsequent cycles. Therefore, while the presence of semitransparent objects has been considered previously with respect to the effect of phase shifts on specific detection schemes [24], our approach differs in that our analysis is of this total loss probability within the interferometer due to a general loss coefficient introduced into the target mode. The possible advantages of this will be discussed in greater detail in subsequent chapters.

7.5.3 Maximizing Conditions for IFM

While it is expected that the introduction of nonideal conditions in the sensing medium will degrade performance, there are still methods and reconfigurations that enable optimized measurements. The first of such is a method of "rebalancing" the interferometer in the event of single mode loss. As seen from figure 7.5, single mode loss undermines the quantum entanglement of the modes and breaks the dark port configuration of the interferometer, which decreases result accuracy and reliability.

Figure 7.7 illustrates the scenario in which loss is independently variable in both arms of the interferometer. This is shown for the symmetric case of $N = 1$ where reflection and transmission probabilities are equivalent ($T = R = 1/2$). Due to

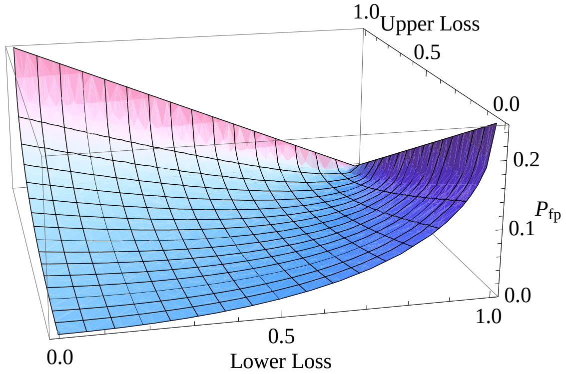


FIGURE 7.7. Reduction of false positive detection probabilities by "balancing" the modes (i.e., equivalent loss probability). The symmetry of the plot is a result of using $N = 2$ beam splitters resulting in equivalent transmission and reflection probabilities $(T = R = 1/)$.

this symmetry, the same output probability is obtained for loss introduced into the target arm or the control arm: the dark port illuminates and false positive results are obtained. However, when the losses in both modes are equivalent, the dark port arrangement is regained. Therefore, it can be concluded that the output intensity difference is a function of the difference in loss probability between the two modes. The advantage to artificially introducing excess loss is the realignment of the dark port, which ensures confidence of an object's presence if a photon is detected at the dark port. The obvious disadvantage is the increased probability that the sensor will yield a null result. Therefore, this realignment technique requires great precision and an abundance of ancillary photons to compensate for the increased null results.

Once equivalent loss is established in the modes, techniques of optimizing the number of beam splitters dependent on the amount of loss have been established [24, 34]. The analytical result for the optimized number of beam splitters can be obtained directly from a calculus based maximization of the probability of interaction free measurement. By definition – and with the assumption that the object is already present in the interferometer – interaction free measurement occurs under two conditions: the single photon did not take the path of the interferometer containing an object, and the photon exits (and is detected) out of the dark port.

In an effort to adhere to convention and for an accurate comparison, I present here a brief analysis of this maximization for a N-cyclic interferometer with $N-1$ cavity/object interactions. Through the use of statistical counting methods, the mathematical representation of the collection of independent possible paths for the photon becomes

$$P_{ifm} = R^N (1-L)^{N-1} \tag{7.31}$$

where

$$R = \cos^2\left(\frac{\pi}{2N}\right) \tag{7.32}$$

Clearly, there is no maximum in the loss plane, but there is a maximum in the trend line for N. In order to maximize this "success" probability with respect the the number of cycles, we take $\partial P_{ifm}/\partial N = 0$ and solve for L. This yields the dual mode loss probability as a function of total number of cycles (with the addition of an extra pass through the half-wave plate at the output) for optimum interaction free measurement probability and is given by

$$L(N) = \exp\left[-\frac{\pi \tan\left(\frac{\pi}{2N}\right)}{n}\right] \sec^2\left(\frac{\pi}{2N}\right). \tag{7.33}$$

Unfortunately, however, this does not account for negative outcome probabilities such as false positive, null results (including bright port output and absorption due to loss), and direct object interactions. According to the conventional definition, the efficiency of interaction free measurement is

$$\eta = \frac{P_{suc}}{P_{suc} + P_{fail}} = \frac{R_{S/F}}{R_{S/F} + 1} \tag{7.34}$$

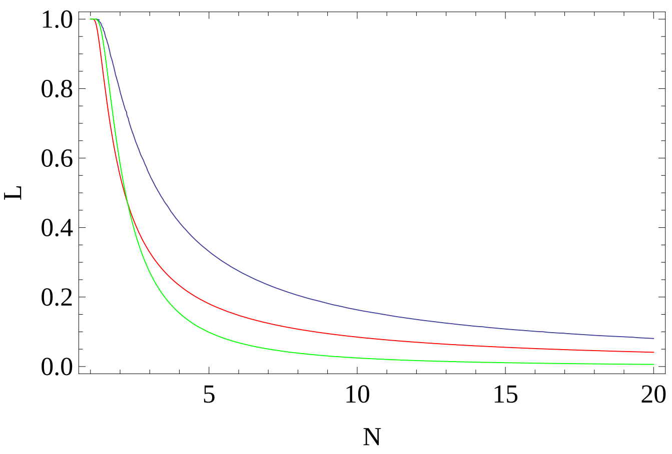


FIGURE 7.8. Plot of the relationship between dual mode loss and number of beam splitters that produce the maximum ratio of successful interaction free measurement to failed measurement due to interaction with the object. The blue curve represents the numerical solution for the maximum, the red curve is the analytical solution by way of small angle approximation, and the green curve maximizes the probability of interaction free measurement without recourse to failed measurements.

where P_{suc} is the probability of a successful interaction free measurement, P_{fail} is the failed probability due to direct object interaction, and $R_{S/F} = P_{suc}/P_{fail}$ is the ratio of successful to failed detection probabilities. Therefore, maximizing this ratio will directly maximize the efficiency of the sensor. While this does not have a direct analytical solution, small angle approximations can be employed that asymptotically converge on the solution in the limit of large N [24, 34]. This optimum number of beam splitters is given by

$$N_{opt} = \frac{\kappa}{\ln(1/\alpha)} \tag{7.35}$$

where κ satisfies the transcendental equation $\kappa = 1 - e^{-2\kappa} \approx .796812...$

Figure 7.8 shows the trend of these curves juxtaposed to the numerical result for the optimum number of beam splitters. The green curve is the analytical maximum with respect to the probability of interaction free measurement alone, the red curve is the approximate relationship to maximize efficient in the high N regime, and the blue curve is the numerical solution to the maximization. Although the conventional definition of efficiency is effective, it does disregard the subtle complexities of the situation (i.e., the fact that there are more than simply two outcomes to the procedure). It is not always immediately clear exactly how to include these other possible probabilities, but alternate methods will be proposed later.

Chapter 8
Symmetric Hypothesis Testing

8.1 Statistical Resolution

Since the interferometric process of interaction free measurement is probabilistic and involves only a single photon, it becomes necessary to repeat the experiment numerous times in order to gain statistics about the state of the device (i.e., whether an object is present in the interferometer or not). As discussed previously, photonic "interaction" with a lossy mode does not necessarily guarantee photon loss, but does affect the wave function of the entangled state thereby altering detection probabilities to various extents. By accumulating statistics about the state of the device's output, one can become increasingly certain about the state of the device itself.

At the root of statistical decision theory lies the necessity to accumulate statistical probability distributions of all possible outcomes to a particular problem at hand.

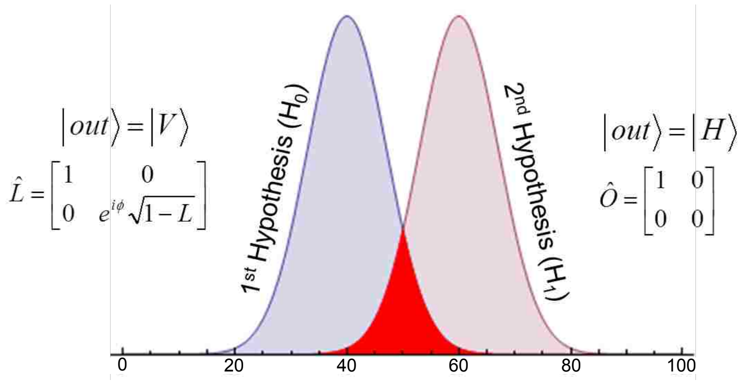


FIGURE 8.1. Statistical likelihood of discerning between two independently probable events. The dark red area in the middle is located at the probabilistic overlap of the two events, and is therefore the minimum uncertainty in distinguishing between the two.

Figure 8.1 is a rendition of a hypothetical probability distribution for the outcome of a system involving two possible hypotheses. The single axis represents the number of times a single outcome has occurred. If the two hypothetical scenarios are sufficiently distinct, there will be a certain resolution between the statistical single outcome. Suggestively added to the diagram are the descriptors for the horizontal and vertical mode cavities of the interferometer. With the accumulation of an adequate number of independent trials, the presence of an object may be distinguished from environmental loss up to an arbitrary confidence threshold.

The actual scenario of lossy interaction free measurement is in slight in contrast to the oversimplification of figure 8.1 in that there may exist more than one pertinent outcome. While there are assuredly only two hypotheses – the presence or absence of an object – there are multiple possible outcomes from which to draw a conclusion. Namely, there is output at the bright port, output at the dark port, absorption by loss, and interaction with the object.

While the existence of these four possible outcomes is an accurate theoretical description, practical application is limited since the only actual detections occur at the output. Therefore, it poses great difficulty to determine whether the photon was absorbed by the loss factor or encountered the object. This is especially true if the photon were in the mode which contains both loss and an object. Even in the event an object is present, there is still a probability that the photon will be absorbed prior to reaching the object. This is insignificant from an experimental perspective where the photon yields a null result in both cases. However, if this involved the hypothetical Vaidman bomb [30], the two scenarios are as contradictory as life and death!

8.2 Classical Chernoff Bound

As alluded to in figure 8.1, the two hypotheses H_0 and H_1 correspond to the two scenarios of interaction free measurement: the object is absent (H_0) or present (H_1). With no prior knowledge of the system, one can safely assume that these hypotheses are weighted with equal probability (symmetric testing) and are described by possible outcomes $b = \{A, L, B, D\}$ with probabilities $P_0(b)$ and $P_1(b)$ corresponding to hypotheses H_0 and H_1, respectively. By means of statistical hypothesis testing, the classical Chernoff bound defines an upper limit on the error of accepting an incorrect hypothesis describing an observed probability distribution in terms of the probability of possible outcomes and the number of trials attempted. In this sense, the two hypotheses are discrete parameters based on discrete quantifiers of their outcomes. Therefore, the classical error probability is composed of the minimization of a discrete set of probabilities given by

$$P_e = \frac{1}{2} \sum_b \min \left\{ P_0(b), P_1(b) \right\}. \tag{8.1}$$

where the minimization is not continuous, but involves choosing the single hypothesis (P_0 or P_1) that minimizes the error probability. The parameter, b, represents all possible measurement outcomes of the statistical system. The factor of 1/2 results from assuming equally weighted *a priori* hypothesis probabilities. Using the fact that $p, q \leq p^s q^{1-s}$, the discrete set minimization can be expressed as a continuous function minimization. Note that simultaneous inclusion of both hypotheses in a continuous fashion doesn't directly reflect the physical error probability of the system. Rather, via the above inequality, is places a theoretical upper bound on the error probability and defines the classical Chernoff bound [29].

$$P_e \leq \frac{1}{2} \min_{s \in [0,1]} \sum_b P_0^s(0) P_1^{1-s}(b) \equiv P_{cc} \tag{8.2}$$

The classical Chernoff bound is a function of the two distinct probability distribution of a single independent trial's outcome. It establishes the theoretical metric by which to gauge the experimentally observed outcomes of the system. Therefore, while a single trial may fall well within the range of an acceptable conclusion, the statistical nature inherent in the process typically demands multiple trials in order to obtain a reasonable confidence in the conclusion.

Once multiple measurements of the system are performed (by conducting M multiple trials), the theoretical probability distribution becomes the product of the individual outcome probabilities. Therefore, the Chernoff bound for multiple trials, M, is expressed as

$$P_e \leq \frac{1}{2} \min_{s \in [0,1]} \left(\sum_b P_0^s(0) P_1^{1-s}(b) \right)^M . \tag{8.3}$$

As the number of trials is increased arbitrarily, it becomes useful to employ an analysis of large numbers on a logarithmic scale. Exploiting the fact that the natural logarithm is monotone increasing, it is clear that the process of minimization and logarithmic operation commute. Ergo, the logarithm of the classical Chernoff bound defines the Chernoff distance as

$$C(P_0, P_1) \equiv - \min_{s \in [0,1]} \log \left[\sum_b P_0^s(b) P_1^{1-s}(b) \right] \tag{8.4}$$

This form of the expression for hypothesis resolution is convenient when applied to the exponential rate equation

$$P_e \; e^{-MC(P_0,P_1)}. \tag{8.5}$$

This expresses the general trend of the error probability in the limit of many trials. By application of a factor of 1/2, the exact Chernoff bound of equation (8.2) can be recovered.

$$P_{cc} = \frac{1}{2} e^{-MC(P_0,P_1)}. \tag{8.6}$$

For the special case that the two hypotheses only have two possible outcomes, the probability distribution of the Chernoff bound is described by the binomial distribution

$$\begin{aligned} P_0(N_{0,1}) &= \mathbb{C}^{N}_{N_{0,1}} p^{N_{0,1}} \bar{p}^{N-N_{0,1}} \\ P_1(N_{0,1}) &= \mathbb{C}^{N}_{N_{0,1}} q^{N_{0,1}} \bar{q}^{N-N_{0,1}} \end{aligned} \tag{8.7}$$

An example of this could be flipping two coins, one of which is biased. After enough trials, and in comparison to an unweighted binomial distribution (expected of an unbiased coin), it will become clear which of the coins is biased. Referring to figure 8.1, the two distinct distributions would represent the two coins, while the axis represents the number of times the coin was observed to fall on heads. The region of overlap between the distributions is minimum probability of error. The the binary Chernoff distance, in the case of a hypothesis testing apparatus with only two outcomes, registered with probabilities $p_1(1) = p$ and $p_1(2) = \bar{p}$ or $p_0(1) = q$ and $p_0(2) = \bar{q}$, is

$$C_2(p,q) = \xi \ln \frac{\xi}{p} + \bar{\xi} \ln \frac{\bar{\xi}}{\bar{p}}, \tag{8.8}$$

where

$$\xi = \ln\left(\bar{q}/\bar{p}\right) / \left(\ln\left(p/\bar{p}\right) + \ln\left(\bar{q}/q\right)\right)$$

and $\bar{x} = 1 - x$. Therefore, knowing p and q is sufficient for error estimation.

8.3 Application to IFM and Practicality

While the classical Chernoff bound is a very robust and encompassing means of analyzing an interaction free measurement device, the implementation of such an analysis protocol can be cumbersome and often requires numerical analysis techniques.

In order to investigate the error limit on MZI based interaction free measurement, the probability of both hypotheses (presence and absence of object) must be considered simultaneously in the calculation. Obtaining the expressions for outcome probabilities from derivations detailed in previous chapters, and assuming the simplification condition of uniformity among optical components and phaseless loss conditions, leads to the construction of table 8.3. Table 8.3 has adopted

TABLE 8.1. Outcome probabilities for a cyclic interferometer with $N-1$ cavity interactions. The null hypothesis, h_0, represents the absence of an object from the sensor.

	H_0	H_1
P_d	$P_0(D)$	$\cos^{2N}\theta$
P_b	$P_0(B)$	$\cos^{2(N-1)}\theta\sin^2\theta$
P_a	0	$(1-L)(1-\cos^{2(N-1)}\theta)\sin^2\theta$
P_l	$L(1-\cos^{2(N-1)}\theta)\sin^2\theta$	$L(1-\cos^{2(N-1)}\theta)\sin^2\theta$

the notation of the Chernoff equations where the subscript represents the hypothesis in question and "is a function of" the specific outcome under observation. Hence, the probabilities of the null hypothesis (object absent) as determined from measurement at the dark and bright port ($P_0(D)$ and $P_0(B)$, respectively) are the complex squares of the hypergeometric equations (7.28) given by

$$P_0(B) = \left| \frac{it}{r^{N+1}} \sum_{k=0}^{N-1} \lambda^k {}_2F_1 \left[\begin{matrix} k+1, N-k \\ 1 \end{matrix} ; -T/R \right] \right|^2$$

$$\begin{aligned} P_0(D) = \Bigg| r^N &+ \frac{1}{r^{N+2}} \sum_{k=1}^{N-1} \lambda^k {}_2F_1 \left[\begin{matrix} k+1, N-k+1 \\ 1 \end{matrix} ; -T/R \right] + \\ &- \frac{1}{r^{-N}} \sum_{k=1}^{N-1} \lambda^k {}_2F_1 \left[\begin{matrix} k, N-k+1 \\ 1 \end{matrix} ; -T/R \right] \Bigg|^2 . \end{aligned}$$

Clearly, even with the mentioned simplifying assumptions, a calculation of the Chernoff bound involving these two formulae alone becomes quite difficult, even by numerical methods.

As an illustrative example, consider further simplification to an idealized MZI apparatus containing only two beam splitters and no loss that performs interaction free hypotheses testing based on the four established theoretically possible outcomes: ($b = A$) the probability of absorption because of photon loss, ($b = L$) the probability of object interaction, ($b = D$) the probability of an IFM, and ($b = B$) the probability of learning nothing where the photon exits through the bright port of the interferometer. The idealized probabilities of table 8.3 are represented as functions of the two beam splitter "rotations," where the two parameters θ_1 and θ_2 denoting the initial and final beam splitters can be altered independently. The importance of no photon loss without an object, $P_0(A) = P_0(L) = 0$, and the dark-port condition, $P_0(D) = 0$, becomes now obvious in the light of (8.1). These assumptions ensure that the error of false acceptance comes from the probability of a photon to exit through the bright port in the presence of an object ($P_1(B) = \cos^4\theta_1$). Since this bright port probability is the only source of erroneous results, the error probability is given by

TABLE 8.2. Outcome probabilities for lossless interaction free measurement using an MZI with two 50:50 beam splitters.

	H_0	H_1
P_b	1	$\cos^4\theta_1$
P_d	0	$\cos^2\theta_1 \sin^2\theta_2$
P_a	0	$\sin^2\theta_1$
P_l	0	0

$$P_e = \frac{1}{2}P_1(B) = \frac{1}{2}\cos^4\theta_1.$$

For our idealized example, let's temporarily assume the use of 50:50 beam splitters ($R = T$), which implies $\theta_1 = \theta_2 = \pi/4$. In this case, given that there is no minimization to perform, the Chernoff bound is exactly equal to the error probability

$$P_{cc}(N = 2, L = 0) = P_e = 1/8 = 12.5\%$$

Given that this is the most theoretically idealized scenario, one could conclude that 12.5% is the lowest possible error probability for any $N = 2$ IFM measurement. While this claim will be supported later, a conceptual analysis is quite sufficient. For this ideal system, the absence of an object reliably produces a dark port output. However, when a a bright port output is detected, it cannot be definitively concluded that there is no object. Likewise, measuring the dark port and receiving no output does not ensure the absence of an object. Furthermore, it must be assumed that there is only a 25% chance of any photon reaching one of the output ports (this is the default assumption so as to not exclude the possible presence of the object). Now, suppose a photon was successfully detected at the bright port. This means that there is only 25% certainty that there is no object. As far as this single measurement is concerned, there also exists a 25% chance that the object *is* in the interferometer.

Clearly, the worst possible error probability is 50%. This suggests that there is no statistical information and any conclusion is simple a random guess. Any skew from this 50% error probability (toward 1 or 0) directly implies that some statistical bias has been observed and the error probability decreases to min $P_e, 1 - P_e$, which begins to resemble elements of the Chernoff equations.

To conclude the example, a bright port detection has been made which, in itself, yields a 1 in 4 chance of erroneous conclusion. However, before any measurement was made, there already existed an intrinsic 1 in 2 chance of choosing incorrectly. Therefore, the combined probability of an incorrect choice of hypothesis is $1/4 \cdot 1/2 = 12.5\%$, or 1 in 8 chance of choosing incorrectly.

It is important to note the ideal simplifications of this example; namely, losslessness. The bright port was chosen as an example here for the specific reason that it alone doesn't divulge any information, and therefore allows for erroneous conclusions. In actuality, the purely quantum mechanical nature of the single photon interference effect collapses the probability distributions necessary for error calculations in the ideal scenario. Specifically, if even a single photon is observed at the dark port, there is 0% error of choosing the incorrect hypothesis. While this was indeed the exact alluring theory behind the inception of interaction free measurement, any practical imperfections broaden the hypothesis probability distributions and sharply degrade the certainty of conclusion.

8.3.1 Efficiency vs Confidence

From table 8.2, it is clear that only P_b contributes a nonzero term in the summation. After minimization of s, the Chernoff bound on the maximum error probability reduces to $\cos^4(\theta_1)/2$. As illustrated in figure 8.2, for $\theta_1 \rightarrow 0$ (minimum probability of photon-object interaction) the error probability saturates to 50%. Intuitively, this represents a complete lack of photonic interrogation forcing a hypothesis con-

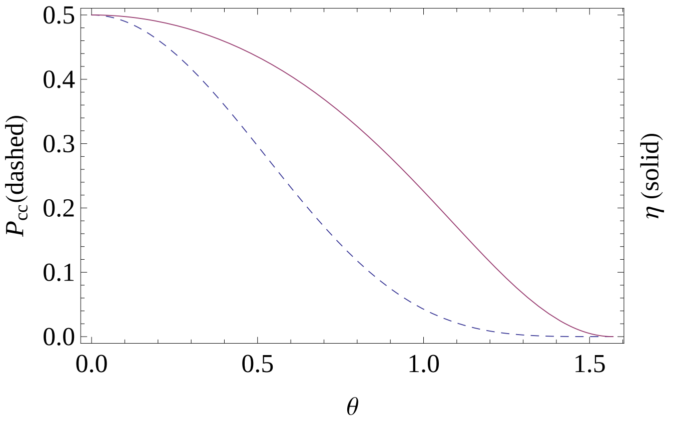


FIGURE 8.2. Comparison of the conventionally established efficiency metric and maximum error probability given by the classical Chernoff bound as a function of initial beam splitter angle (θ_1) with the second beam splitter couples by $\theta_2 = \pi/2 - \theta_1$.

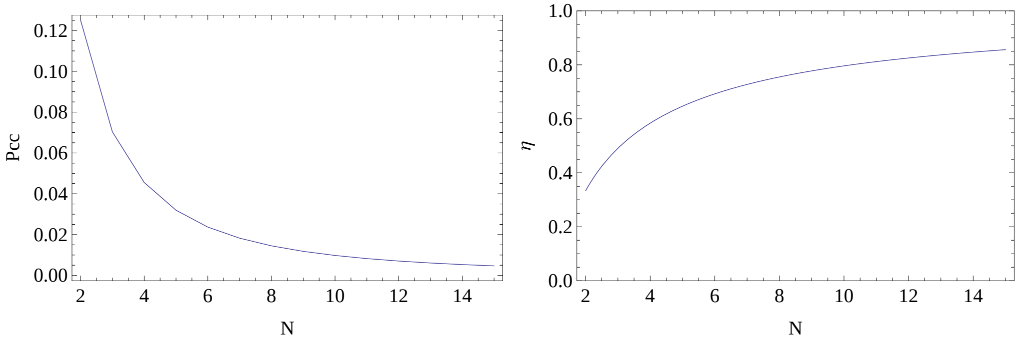
clusion with 50% success/failure probability. Conversely, for $\theta_1 \to \pi/2$ (maximum probability of photon-object interaction), there is virtually zero conclusion error. In this configuration, the detection port is irrelevant since the presence of an object will guarantee absorption of the photon. Figure 8.2 also displays the efficiency of the detector as discussed in the previous section.

A direct comparison of the efficiency and Chernoff error should not be made directly without consideration for their representative quantities. Both curves have a similar trend, and both strive to represent the effectiveness of the device. Of utmost importance, however, is the fact that the efficiency is desired at its maximum, while the Chernoff bound is desired at a minimum. Also, given that the definition of efficiency excludes all outcomes that aren't "conclusive" of IFM or object interaction. This, by direct implication, requires only the scenario of an object present in the interferometer, and offers no contrasting comparison which would establish a base for inherent environmental imperfections. Furthermore, there is no generally accepted form of the efficiency incorporating the possibility of multiple trials.

The Chernoff bound, on the other hand, doesn't measure how effectively the presence of an object is measured with respect to its probability of interaction. Instead, it focuses on the inclusion of any experimentally practical factors and gives a rigorous theoretical limit on the error of *choosing* the incorrect hypothesis. By definition, this bound is a comparative measure of the systems behavior with and object and without an object simultaneously through the minimization of the product of the two. Its concern for all possible outcomes ensures increased accuracy of such a comparison; and as such, places more rigorous concern on being confident about a conclusion drawn from the data.

By definition, conventional IFM schemes involve solely detection at the output ports, without specific regard for the means of photon loss, since such an event is not experimental detected at the output. Therefore, the Chernoff bound remains largely a theoretical limit with respect to inclusion of all possible outcomes; however, nothing in its definition prevents mathematical redefinition of the outcomes to encompass multiple events (e.g., grouping photon loss and object interaction to reduce the expression to three experimentally verifiable outcomes).

Although the Chernoff error bound and the current definition of efficiency have many difference – and contradictions evens – their usefulness as metrics for effective system analysis is not undermined.



(a)

(b)

FIGURE 8.3. A comparison of the trends of the classical Chernoff bound on error and the conventional definition of efficiency with respect to an increasing number of cycles through the interferometer. Ideal conditions are assumed (i.e., $L = 0, \phi = 0$).

Figure 8.3 show the trend of the error bound and the detection efficiency as a function of number of cycles through the interferometer, N. As should be expected, both exhibit improvements for large N due to the quantum zeno effect.

To date, the conventional analysis of effectiveness for MZI based IFM device has centered around an early definition of efficiency. While previous analyses have relied primarily on this measure of efficiency, it is clear here that maximum efficiency is inevitably accompanied by maximum error probability, and visa versa. This relationship suggests the necessity for a compromise of reliability and efficient functionality of the detector in question.

8.3.2 Performance Sensitivity

Ideal operation of the interferometer requires lossless conditions to ensure the environment has no photonic interactions. Realistically, however, it is necessary to analyze the effects of photon loss and phase fluctuations on the reliability of detection. As stated previously, the balancing condition for a MZI requires the total beam splitter rotation to be $\pi/2$, but it is not immediately clear how any deviation from this dark port arrangement will affect the error probability.

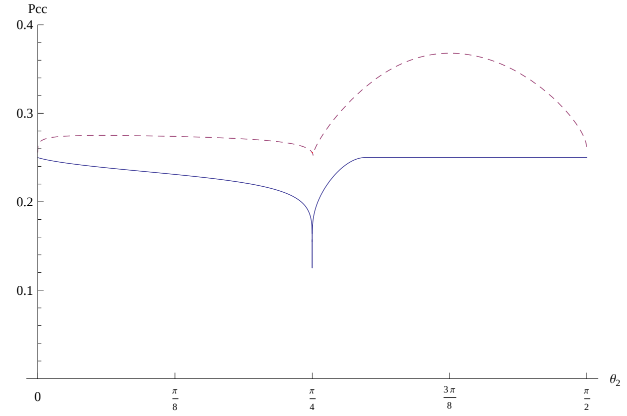


FIGURE 8.4. Chernoff bound on error probability for $\theta_1 = \pi/4$ as a function of θ_2 with no loss (solid) and 1% loss prior to object (dashed).

Figure 8.4 illustrates a straight forward investigation of the classical Chernoff bound for a single trial ($M = 1$). Without loss of generality, the first beam splitter is maintained as 50:50, while the second is varied from transparent to opaque. Clearly, the minimum error occurs at the dark port arrangement where both beam splitters are 50% reflective. For a lossless scenario (solid line in figure 8.4), the minimum possible error is 12.5%. The extreme sensitivity to this configuration causes even the smallest deviation from the dark port arrangement to increase the error dramatically. The device's sensitivity to loss should also be noted. The dashed line in figure 8.4, representing only 1% probability of photon loss ahead of the object increases the minimum error probability to 25.5%.

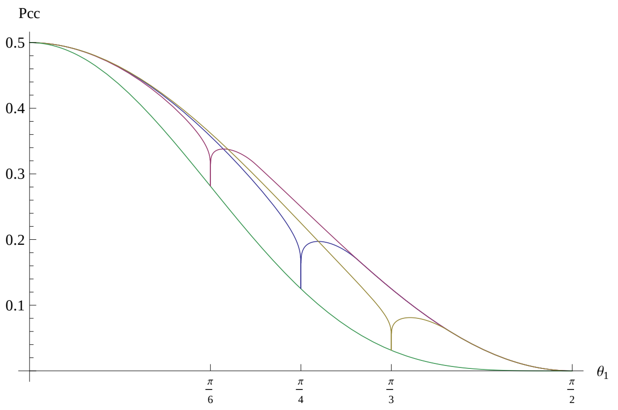
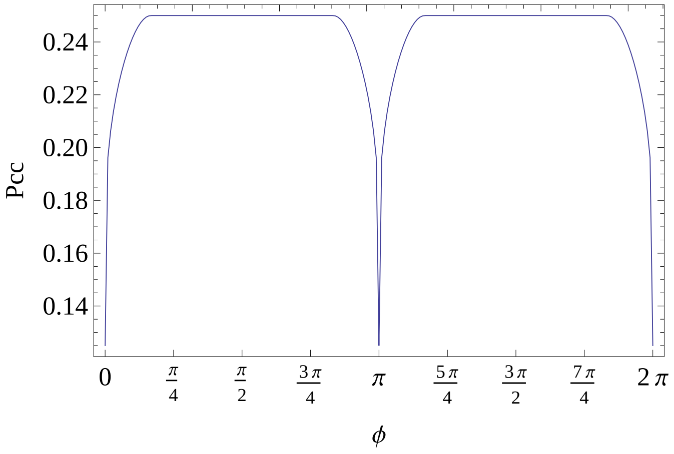


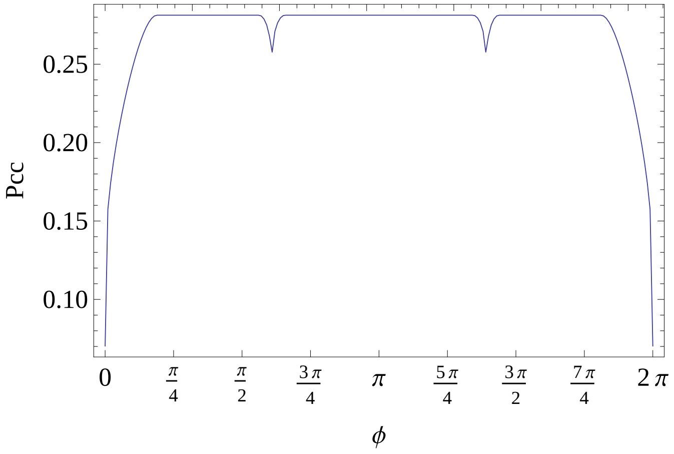
FIGURE 8.5. Chernoff bound on error probability for $\theta_2 = \pi/3, \pi/4, \pi/6$ as a function of θ_1. The local minima in error probability occur for each curve when the sum of the two beam splitter parameters equals $\pi/2$, and the lower bounding curve represents a continuously variable second beam splitter yielding the theoretical minimum for all angles.

For completeness, the effect of varying the initial beam splitter reflectivity and the resulting error characteristics should also be considered. As seen in figure 8.5, the absolute maximum probability of 50% occurs when photon is certain to route through the arm containing no object and no loss. This is a sensible result since, in principle, one could never gain greater than 50% certainty of the objects presence without allowing the photon any probability of "probing" the target mode. Similarly, when the photon is guaranteed to enter the target mode, one can know the object's presence with absolute certainty by simply detecting whether or not the photon exists the interferometer. Neglecting these trivial cases, a locally minimum cusp yields the same error probability as figure 8.4 existing at the dark port configuration.

A final concern with respect to performance degradation is the introduction of a relative phase between the orthogonal modes. As shown in previous chapters, such a phase (even in the absence of loss) will ensure altered state propagation and erroneous results in the form of false positives. In direct relation to these skewed outcome probabilities is a very steep decrease in hypothesis confidence.



(a)



(b)

FIGURE 8.6. Chernoff bound on error rapidly increases with any nonzero single mode phase shift, and saturates at a value dependent on the number of cycles. Figure (a) is the plot for N=2 beam splitters, and figure (a) is for N=3 beam splitters.

As seen in figure 8.6, while increasing the number of cycles reduces the minimum value of the error probability (about $\phi = 0$), it tends to increase the saturated maximum error value for "off resonance" phase regions. This meaning that there exists N-1 regions of phase alignment resulting in the minimum theoretical error probability found at $\phi = 0, 2\pi$. For figure (a), this is not immediately obvious simply because the region is so confined that it couldn't be resolved to sufficient precision for plotting by the current numerical approach employed.

8.3.3 Effects of Multiple Trials

Reference [24] shows that in the quantum Zeno implementation of the multi-pass polarization based Mach-Zehnder interferometer, the probability of a successful interaction-free measurement (with negligible system loss) can become arbitrarily close to unity for any semitransparent object to be measured within a certain tolerance of single mode phase shift and low system loss. Specifically, this loss was introduced equivalently in both modes. However, when system loss is not negligible, for any quantum efficiency at the detectors, the probability of a successful interaction-free measurement decreases more rapidly than the probability of an incorrect conclusion for large N.

To compensate for such false conclusions, it is proposed that the experiment be repeated numerous times, and defines a new efficiency proportional to the quantum efficiency of the detectors based on the probabilities of "successful" or "failed" measurement outcomes. Here, it is important to note that all previous analyses of interaction-free measurement and imaging are based on probabilistic detections of a single photon. Such probabilities becomes known only in the limit of an infinite number of photons (equivalent to repeating the experiment an infinite number of times), and as such inevitably contain some degree of error/uncertainty.

The problem with infinitely repeating experimental runs is the guarantee to eventually create an undesired interaction with the object to be measured/imaged, therefore defeating the objective of obtaining information by interaction-free means. The Chernoff is based on probability distributions, and is general enough to allow for any reasonable probability distribution. As briefly discussed in previous sections, the bound can also be recast into a binary form as in equation (8.8). However, it holds that any reasonable probability distribution must integrate to unity. Figure 8.7 is a plot of a generic binomial distribution (red curve) with its

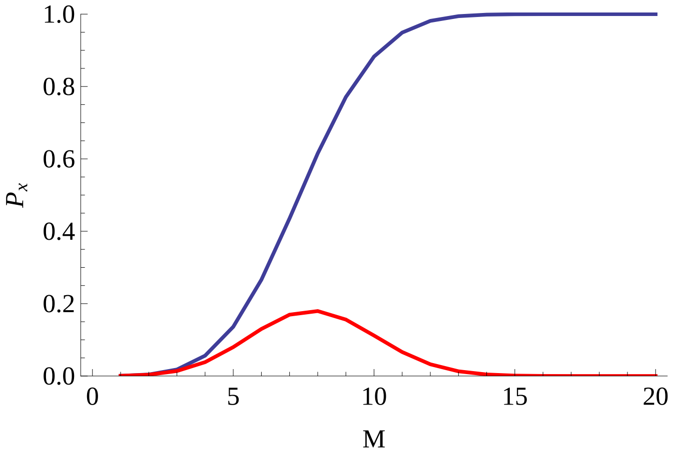


FIGURE 8.7. Probability of photonic interaction with the measured object P_x as a function of the number of repeated independent trials M. The red curve shows the binomial distribution between interaction and non-interaction with the object, and the blue curve is the accumulated probability of interaction. Data shown here is for N=5, M=20, L=0.

associated cumulative distribution (blue curve). Obviously this trend would hold (albeit with slightly different properties) for any choice of M, N, L. The probability of the photon being absorbed is essentially reduced to a binomial distribution given by

$$\sum_{m=0}^{M} \mathbb{C}_m^M P_s^m (1 - P_s)^{M-m}$$

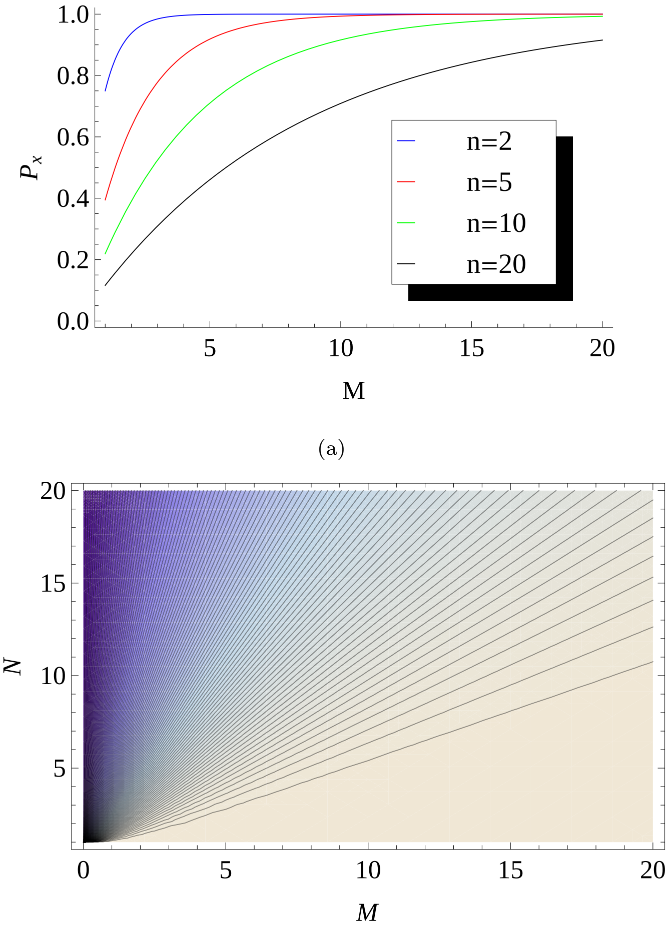
where P_{abs} is the absorption probability for N cycles through the interferometer given in (6.14) as

$$1 - (\cos^2\theta)^{(}N - 1).$$

Therefore, the probability of photon-object interaction after M trials is found to be

$$\begin{aligned} P_x(N, M) &= 1 - [1 - P_x(N)]^M \\ &= 1 - \cos^{2M(N-1)}\theta. \end{aligned}$$

The prominent term in this absorption expression is governed by $0 \leq \cos^{2kN}\theta < 1$ for any finite number of cycles $1 \leq k \leq N$ through the interferometer. As seen



(a)

(b)

FIGURE 8.8. Probability of photon-object interaction after 20 cycles (N=19 total beam splitters) and 40 independent trials (M) is shown to be approximately 99.3%.

in figure 8.8(a), the absorption probability is monotone increasing with negative concavity such that it asymptotically approaches unity. Figure 8.8(b) strengthens this assertion by confirming the trend for all values of M, N.

While increasing the number of cycles will lessen the probability of absorption per M cycles, any practical application of such an MZI based IFM device must take into consideration this inevitable measurement failure when demanding reliability factors of efficiency and conclusion confidence.

Chapter 9
An Invisible Quantum Tripwire

9.1 Introduction

This chapter will focus on the use of beam splitters arranged as a Mach-Zehnder interferometer (MZI) for various sensing applications. As shown in figure 9.1, propagation through the initial beam splitter of the MZI allows a single photon to become path entangled with the vacuum which permits highly sensitive detection of relative factors between the two modes. Another related use of this interferometric sensor is multi-photon entangled two mode states with application to lossy environment metrology. This proposal provides an analysis of the quantized loss mechanism and the resulting density matrix formulation and its effectiveness for quantum metrology.

In our work, we discovered a curious nonlinear behaviour of photon's transmission in a Zeno enhanced but lossy IFM apparatus. This discovery leads us to an IFM protocol robust against photon loss and dephasing. In addition, we recast the entire protocol in terms of statistical hypothesis testing, allowing us to quantify the operation of the device as a reliable yet undetectable intruder alert system — the invisible quantum tripwire.

9.2 IFM Revisited

The Elitzur-Vaidman "Bomb" gedanken experiment posits that there exists a bomb with a single-photon sensitive detonator and the goal is to optically detect the presence of such a bomb without detonation. In contrast to the expectations of the classical approach, where such a goal could not be reached, quantum optics allows for a solution — measurement without interaction.

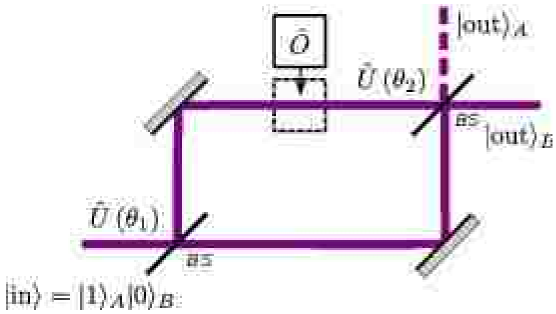


FIGURE 9.1. A lossless Mach-Zehnder interferometer in a dark port arrangement, $\theta_1 + \theta_2 = \pi/2$, and a zero phase difference between its arms, constitutes a simple IFM setup with efficiency $\eta \leq 1/2$. This scheme allows for interaction-free hypotheses testing of a path being blocked (h_1) or it being clear (h_0).

This measurement is based on a fascinating property of a single photon to interfere with itself while being indivisible. Imagine a lossless Mach-Zehnder interferometer (MZI) with beam splitters described by a two mode coupling matrix

$$\hat{U}(\theta_i) = \begin{bmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{bmatrix} \tag{9.1}$$

and the possibility of a photon-sensitive object to be placed in the detection arm (see figure 9.1). This detection arm stays invisible to the object for as long as a photon has not been absorbed by the object. There are two possible scenarios: the path is blocked or it is clear. If the path is clear, a single photon, after the first beam splitter $\hat{U}(\theta_1)$, can travel both arms of an interferometer and interfere with itself at the second beam splitter $\hat{U}(\theta_2)$. Under a proper choice of beam splitters, $\theta_1 + \theta_2 = \pi/2$, and a zero phase difference, such an interference will result in zero probability of the photon to leave the MZI in mode A (dark port), that is $P_0(D) = 0$. If the path is blocked by an object $\hat{O}$, there is a definite destruction of the interference as well as the probability for an object to absorb a photon, $P_1(A) = \sin^2\theta_1$. Loss of the photon tells us that an object is there, but this is a measurement with an interaction. Without interference there exists a non-zero probability for a photon exiting the MZI through the dark port, $P_1(D) = \cos^2\theta_1 \cos^2\theta_2$. Detection of a photon in a dark port constitutes a measurement without an interaction. The efficiency of a given measurement is $\eta = P_1(D)/[P_1(D) + P_1(A)]$, since an object is detected with probability $P_1(D) + P_1(A)$, while detection without interaction is carried out with probability $P_1(D)$. In the presented setup, there is a limit on the highest efficiency $\eta = \cos^2\theta_1/(1 + \cos^2\theta_1) \leq 1/2$, which is achieved at the limit where $P_1(D) \to 0$, $P_1(B) \to P_0(B) = 1$ and single trial detection becomes improbable.

9.3 Performance and Invisibility

The error of false acceptance in a lossless MZI with a dark port is minimized by an increase of the first beam splitter's reflectance ($\theta_1 \to \pi/2$). It means that all the photons are routed into the detection arm. Hence, interaction with an object becomes unavoidable and the photon path becomes visible. In the opposite case, $\theta_1 \to 0$, the probability of an interaction with the object is significantly reduced, at the expense of high statistical error. In order to compensate for the increased statistical error, multiple trials are required. For the photon path to stay invisible to the object, every photon must be received at the output, which happens with the probability $\bar{P}_{\text{vis}} = \exp(-MC_{\text{vis}})$, where the visibility distance, $C_{\text{vis}} = -\ln\cos^2\theta_1$, was introduced for an easy comparison with the Chernoff distance, $C(P_0, P_1) = -2\ln\cos^2\theta_1$. Judging by these distances, it is possible for the detection to be hidden from the object, $\bar{P}_{\text{vis}} \gg 0$, while revealing the presence of the object with a high level of certainty $P_e \to 0$. Sadly, any deviation from the ideal setup — such as loss, phase shifts, or non-perfect dark port arrangement makes the Chernoff and visibility distances comparable; thus effectively preventing the invisibility of a tripwire based on IFM in a setup presented in figure 9.1.

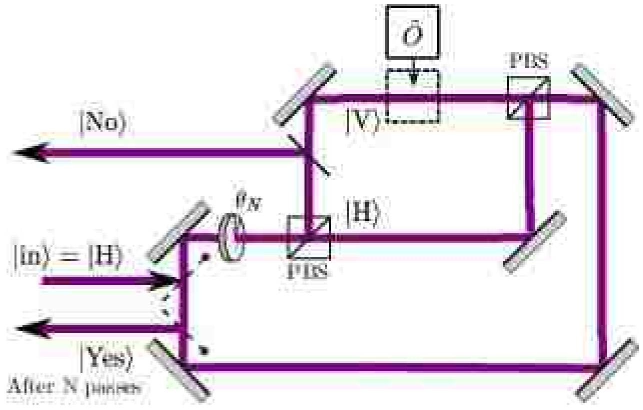


FIGURE 9.2. IQT apparatus based on a N-pass IFM in the polarization interferometer. With each pass, a photon's polarization is rotated by an angle θ_N. The presence of an object prevents accumulation of polarization rotation and is similar to the quantum Zeno effect. An additional beam splitter inside the polarization interferometer models unavoidable loss in the arm accessible by the object as well as controlled loss that is adjusted to provide best performance of the IQT apparatus.

9.4 Invisible Tripwire

Nevertheless, an invisible quantum tripwire (IQT) is possible. We realize it through a combination of an efficient IFM apparatus and a proper interrogation technique. A possible IQT apparatus is presented in figure 9.2 and is based on a N-pass IFM apparatus, which offers improved efficiency η due to the quantum Zeno effect [22]. A crucial part of IQT apparatus is, however, a quantum interrogation technique that deals much better with high sensitivity of the N-pass IFM to photon loss [34], as well as eliminates the dark-port condition. This technique is based on the partial Zeno effect and actually adds a controllable amount of loss to the detection arm by means of a beam splitter with tunable reflectivity. Any attempt to register a photon (that constitutes a tripwire) as well as crossing the path of a photon, would immediately engage the quantum Zeno effect resulting in drastic reduction of the photon loss. This effect will increase the rate at which photons exit the system and trigger the alarm, with a confidence level given by the Chernoff bound.

9.5 Theoretical Description

From this analysis, it is clear that the maximum distinguishability is obtained when the difference in detection probability at the output ports is maximum. This approach relies on the quantum Zeno effect ($N > 2$) to produce a low probability of interaction with the opaque object (maximizing detection at the output). Furthermore, the Chernoff distance is maximized when the detection probability at the output is minimum in the absence of an object in the interferometer. As discussed in section 7.5.2, this can be achieved by introducing an artificial loss ($0 < L < 1$) into the target arm (tripwire). Intuitively, this effectively creates an instantaneous "switch" in detection probability when an opaque object crosses the tripwire. Furthermore, since the artificial loss is introduced prior to the target, the probability of photons entering the target area is minimum and thus the tripwire is "invisible".

The N-pass IFM apparatus itself is based on a polarization interferometer that operates in the basis of linear polarizations $|\mathrm{H}\rangle$ and $|\mathrm{V}\rangle$. The path of vertical polarization constitutes a tripwire. The evolution of a photon's polarization state is described by successive multiplication of matrices $\hat{U}\left(\theta_N\right)$, $\hat{L}(\lambda)$, $\hat{O}(h)$ corresponding to polarization rotation by θ_N and loss, λ, of a photon in the detection arm:

$$\hat{U}\left(\theta_N\right) = \begin{bmatrix} \cos\theta_N & -\sin\theta_N \\ \sin\theta_N & \cos\theta_N \end{bmatrix} \qquad \text{and} \qquad \hat{\mathrm{L}}\left(\lambda\right) = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{bmatrix} \tag{9.2}$$

as well as the presence $\hat{O}(h_1) = \hat{L}(1)$ or absence $\hat{O}(h_0) = \hat{L}(0)$ of an object. If the input state of a photon is $|\psi_0\rangle$ then after a single pass it will be $|\psi_1\rangle = \hat{O}(h)\hat{L}\hat{U}\left(\theta_N\right)|\mathrm{H}\rangle$. The probability to detect a photon with polarization X after N passes is $P_{\mathrm{X}} = \langle\psi_N|\mathrm{X}\rangle\langle\mathrm{X}|\psi_{\mathrm{N}}\rangle$, while the probability of total transmission is

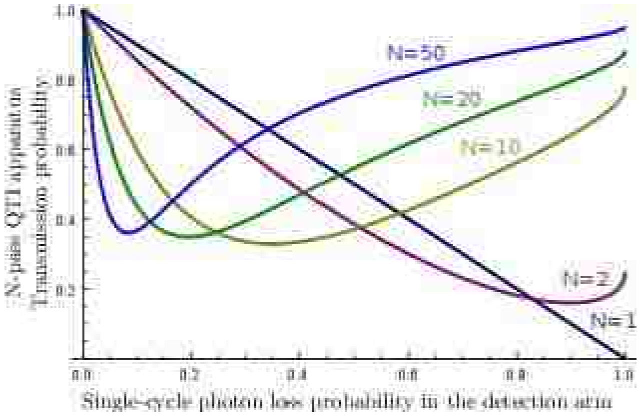


FIGURE 9.3. The single-photon transmission probability in a N-pass IQT apparatus P_{tr} for $N\theta_N = \pi/2$ as a function of single-cycle probability of photon loss in the detection arm. Loss in a N-pass IQT is optimized for this partial Zeno effect to take place. The detection of an object is based on increase of transmission.

$P_{\text{tr}} = \langle\psi_N|\psi_N\rangle$, where $|\psi_N\rangle$ is obtained by repeating a single-pass evolution N times.

In the IFM apparatus, a photon is initially horizontally polarized, $|\text{H}\rangle$. With each pass, polarization is rotated by an angle θ_N, which increases a photon's probability to be in the detection arm, where the photon interacts with a beam splitter before being sent along the tripwire. We present the transmission probability P_{tr} as a function of a single-cycle probability of photon loss in the detection arm, λ, in the absence of an object (see figure 9.3). P_{tr} is given for a different number of passes but with the same angle of evolution $N\theta_N = \pi/2$. A 100% photon loss corresponds to the presence of an object in the detection arm. One can see, transmission in this case improves with the number of passes due to the quantum Zeno effect. The region of small λ demonstrates how an artificial lossless case behaves since even a small amount leads to a significant drop in the transmission probability. Interestingly, the smallest transmission probability is for relatively high loss, but it is not high enough for the quantum Zeno effect to become apparent. This partial

Zeno effect corresponds to a special type of quantum state evolution in the presence of a soft measurement.

Our quantum interrogation technique is based on this special evolution. A controllable amount of loss λ is introduced in the detection arm by means of a beam splitter with tunable reflectivity. This additional loss in the presence of an object reduces the probability of a photon striking the object during a trial, $P_{\text{str}} = (1-\lambda)(1-\cos^{2N}\theta_N)$. Furthermore, we assume that reflectivity of a beam splitter and inherent phase shifts are constantly adjusted such that detection of a photon at the output is minimal. Such an adjustment is made in order to counteract changes in the environment as well as for the partial Zeno to be maintained. The latter would obviously not be possible in the presence of an object. Thus hypotheses testing is based on two outcomes: an almost 0% probability to detect a photon at the output in the absence of an object and 100% in its presence.

Recall from equation (8.8) that he Chernoff distance, in the case of a hypotheses testing apparatus with only two outcomes, registered with probabilities $p_1(1) = p$ and $p_1(2) = \bar{p}$ or $p_0(1) = q$ and $p_0(2) = \bar{q}$, is

$$C_2(p,q) = \xi \ln \frac{\xi}{p} + \bar{\xi} \ln \frac{\bar{\xi}}{\bar{p}}, \tag{9.3}$$

where $\xi = \ln(\bar{q}/\bar{p}) \,/\, (\ln(p/\bar{p}) + \ln(\bar{q}/q))$ and $\bar{x} = 1-x$. Therefore, knowing p and q is sufficient for error estimation. The transmission probability could be calculated analytically only in the presence of the object, $p = \cos^{2N}\theta_N$. However, in the absence of an object, the transmission probability, q, is experimentally available information, which is constantly provided by the IQT apparatus.

There are two primary goals of the IQT apparatus: detection of an object with high certainty, $P_e \to 0$, while staying invisible, $\bar{P}_{\text{vis}}(M) \approx 1$. Currently, it is not possible to satisfy both goals, thus the following compromise between confidence level and invisibility is assumed. We would like $\bar{P}_{\text{vis}}(M) > P_e$, which means a lower likelihood of hitting our object with a photon than accepting the wrong hypothesis, while maintaining a confidence level above a blind guess: $1 - P_e > 0.5$.

9.6 Efficiency Analysis

The efficiency of an IFM was originally defined as a ratio of successful IFM measurement probability to the sum of the probability for a success and failed measurement. Note that a failed measurement is on in which a photon interacted with the object – indeterminate measurements (detection at the bright port) are discarded. According to this definition, the basic spatial-mode interferometer with two beam splitters only achieves 50% efficiency [25].

$$\eta = \frac{P_s}{P_s + P_f} = \frac{R_1 R_2}{R_1 R_2 + T_1} = \frac{1 - T_1}{2 - T_1} \tag{9.4}$$

Numerous variations on the original setup have proposed to increases this efficiency. The most notable of these is the implementation of the quantum Zeno effect. As discussed in section 6.4, repeated passes through the MZI can be implemented in the polarization basis as depicted in 6.3. This requires a tunable HWP rotation angle that is inversely proportional to the number of cycles. This scheme reaches unit efficiency in the limit of infinite cycles [23]; however, this suggests that the rotation of the state polarization is infinitesimal which reduces its practicality analogously to the limit of $T_1 \to 0$ in the spatial mode scheme above. To further complicate the traditional approach, for multiple cycles and nonzero loss, the output state is nonorthogonal. This considerably complicates the detection of the output and compromises the reliability of the result.

In light of our discovery of the nonmonotonic loss tendency determined by nondiscriminative output measurements, however, we proposed an alternative to the predefined methods and an alternate analysis of the sensors reliability: symmetric hypothesis testing. This is accomplished via the classical Chernoff bound for binary probability outcomes [29]: the photon is absorbed by loss (or an object) denoted by probability $p(0)$, or the photon exits the sensor (and is detected at either port) described by probability $p(1)$. Performing multiple single photon trials accumulates statistics eventually satisfying a predetermined confidence requirement between the two possible hypotheses: an object has been detected (with probability p_1), or there is no object present (with probability p_0). Altogether, these probabilities are $p_1(0) = p$ and $p_1(1) = 1 - p = \bar{p}$, or $p_0(0) = q$ and $p_0(1) = 1 - q = \bar{q}$. The distinguishability between the two possible hypotheses $p_{0,1}$ is represented by the Chernoff distance, and is defined as

$$C_{bin}(p, q) = \xi \ln \frac{\xi}{p} + \bar{\xi} \ln \frac{\bar{\xi}}{\bar{p}} \tag{9.5}$$

where

$$\xi = \frac{\ln(\bar{q}/\bar{p})}{\ln(p/\bar{p}) + \ln(\bar{q}/q)}.$$

9.7 Improving Output Detection

Throughout this thesis, there has been careful and deliberate labeling of schematics and calculations to distinguish between having N or $N - 1$ loss/object cavity interactions (see figures 7.3). Depending on the application, either arrangement could prove more advantageous. While spatial and polar interferometric setups are analogous, discussions here will address the polarization basis specifically to mitigate confusion. In this terminology, the will make N total cycles through the interferometer and N passes through the half-wave plate. This ensures the cumulative rotation angle $N\theta = \pi/2$ is identical for both. The difference then becomes whether the photon is allowed to through the MZI cavity one final time after the Nth rotation (state $\vec{u}$), or if it is extracted immediately after the rotation (state $\vec{v}$).

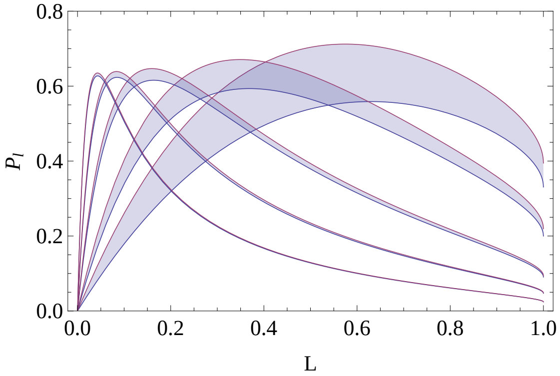


FIGURE 9.4. Plot of the total photon loss probability versus loss coefficient. The blue (lower) curve represents state $\vec{u}$ detection, with N object measurements for $\theta = \pi/(2N)$. The red (upper) curve represents state $\vec{v}$ detection, with $N-1$ object measurements and $\theta = \pi/(2N)$. From right to left, the curve pairs are calculated for $N = 5, 10, 25, 50, 100$ cycles.

For conventional IFM, detection of state $\vec{v}$ is likely to be undesirable for a couple of reasons. Most importantly perhaps is the rather significant increase in the probability of interacting with the object. In figure 9.4, this can be seen as the difference between the curves at $L = 1$. For $N = 2$ cycles (not pictured) with $\theta = \pi/4$, detection of $\vec{v}$ is 1.5 times more likely to strike the object than detection of $\vec{u}$. Also, in the presence of appreciable loss, visual analysis of figure 7.3 indicates significant increases in total photon loss probability. Both factors contribute to increases in failed trials and null outputs.

Of final concern on the application of this scheme to conventional IFM is the fact that the presence of an object will guarantee *no* output at the bright port. Conversely, if *any* output is detected at the bright port, there is 100% certainty that the object is not present. That is to say there is 0% chance of a false negative. Probability of detection out of the dark port (IFM or false positive) remains identical regardless of detection of $\vec{u}$ or $\vec{v}$. In the presence of arbitrary loss, however, the classical Chernoff bound is highly dynamic. Specifically, the bound is a function of the environmental loss which, under circumstances of equivalent detection probability as loss probability, could result in the maximum limit of 50% error. However, distinguishing the hypothesis from dark port detection is the only method of getting a probability distribution for both hypotheses. Dark port detection is monotone increasing and loss probability (the only other statistical possibility) is monotone decreasing. This implies that increased beam splitters will reduce total loss while increasing horizontal port detection for all values of loss. This is counterproductive to efforts aimed at establishing a meaningful distribution for the hypotheses.

The trip wire proposed here, however, negotiates the possibility of a hard loss at the final bright port by non-discriminative output detection. The problem of arbitrary environmental losses is handled through the intentional introduction of artificial loss into the target arm as intrinsic to the conceptual schematic. By carefully maintaining a maximum loss probability, the trip wire is both invisible and predictably reliable with respect to hypothesis confidence upon measurement. The final concern remains whether the inclusion of the Nth cavity coupling an improvement or detractor for this scheme.

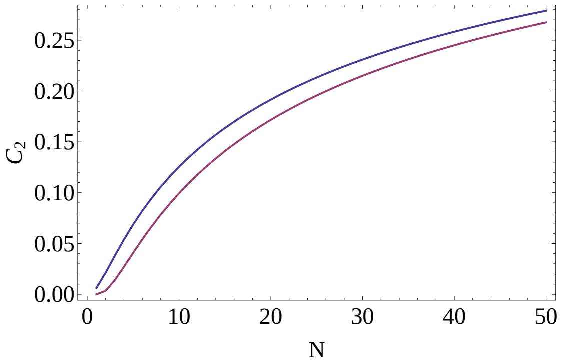


FIGURE 9.5. Binary Chernoff distance calculated from comparison of maximum total loss probability to 100% (object simulation) as a function of the number of interferometric cycles. The blue (lower) curve represent detection of state $\vec{u}$, while the red (upper) curve is for detection of $\vec{v}$.

Figure 9.5 shows a comparison of the binary Chernoff information (measure of hypothesis distinguishability) as calculated from the two output methods in question (state $\vec{u}$ prior to final MZI coupling, and state $\vec{v}$ including the final MZI coupling before detection). Per design, the difference between the maximum possible loss probability (in the absence of an object) and the probability of loss (due to loss or interaction) when an object is present is defined as the comparative conditions. As seen from figure 9.5, the detection of state $\vec{v}$, which *includes* the final MZI coupling, yields more information and increases conclusion confidence.

9.8 Results

In our apparatus, it is assumed that a tripwire becomes visible after a single event of a photon striking an object. Therefore, the probability of a tripwire to stay invisible after M trials is $\bar{P}_{\text{vis}}(M) = \exp(-MC_{\text{vis}})$ as before where the visibility distance, $C_{\text{vis}} = -\ln(1 - P_{\text{str}})$, is defined in terms of the probability to strike an object, as described earlier.

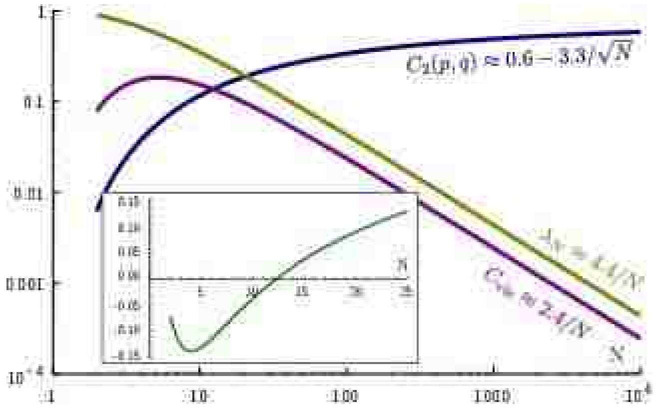


FIGURE 9.6. Chernoff $C_2(p,q)$ and visibility C_{vis} distances as a function of number of passes N as well as the amount of loss, λ_N in the detection arm required for partial Zeno to take place. Inset is a difference between those distances. Invisible detection becomes possible when this difference becomes positive.

We numerically simulated the performance of the IQT apparatus based on the state evolution described above. For a given number of passes N and θ_N, we numerically minimized the single-trial transmission probability (in the absence of an object P_{tr}) by adjusting controllable loss λ. Then we used this parameter to calculate the Chernoff and visibility distances $C_2(p,q)$ and C_{vis}. Figure 9.6 summarizes these results for a total angle of evolution $N\theta_N = \pi/2$ as a function of number of passes. This reveals that at least 13 passes are necessary for visibility distance to become smaller than Chernoff distance thus allowing for $\bar{P}_{\mathrm{vis}}(M) > 2P_e$. Table 9.1 presents numerical values of the visibility distance, the ratio of the distances, as well as the operational amount of loss in the detection arm, λ_N. It again shows that at least 13 passes are required before the statistical error starts going to zero faster than the probability of staying invisible. It also shows that a requirement of the total angle of rotation to be $N\theta_N = \pi/2$, which is a requirement for the standard N-pass IFM apparatus, could be dropped. One can actually use θ_N as an additional parameter for the optimization of IQT apparatus. In the case of $\pi/4$, the visibility distance is shortened by a factor of four. The shorter the distance the more trials are necessary, thus allowing for longer acquisition times and better averaging out of any additional errors. In addition, one can see that the Chernoff distance actually becomes greater relative to the visibility distance, which signifies that for the same probability of invisibility, statistical error could be made smaller for the $\pi/4$ case than it was possible with a greater total angle of rotation. Fi-

TABLE 9.1. Ratio of the distances, visibility distance with a corresponding controllable loss for two cases of $N\theta_N$.

	$N\theta_N = \pi/2$			$N\theta_N = \pi/4$		
N	$\frac{C_2(p,q)}{C_{\mathrm{vis}}(N)}$	$C_{\mathrm{vis}}(N)$	λ	$\frac{C_2(p,q)}{C_{\mathrm{vis}}(N)}$	$C_{\mathrm{vis}}(N)$	λ
5	0.29	0.184	0.575	0.28	0.057	0.523
10	0.75	0.154	0.349	0.79	0.042	0.314
11	0.85	0.147	0.324	0.92	0.039	0.291
12	0.96	0.140	0.302	1.00	0.038	0.271
13	1.07	0.133	0.282	1.14	0.035	0.253
20	1.91	0.098	0.195	2.08	0.025	0.174
50	6.16	0.045	0.084	6.73	0.011	0.075

nally, the amount of controlled loss in the detection arm is relatively high, which is comforting for practical realizations.

Chapter 10
Conclusion

In conclusion of the discussed numerical metrological study, it was found that a variety coherent-like states are optimum in the finite photon set of quantum states. The N00N state was reproduced in the lossless regime. If maintained in a lossy regime, the density-matrix description of photon loss may be avoided and the state-vector approach can be adopted [15, 14, 42]. However, we strove to implement the more general form of loss through a density matrix formalism. Previously, we have used the density-matrix approach to lossy interferometers for particular input states of light (namely, the $M\&M'$ states [17]). Our method here, however, applies to any input state with a fixed number of photons. Thus, it allows optimization of the input state in the presence of an arbitrary amount of propagation loss in the two arms of an interferometer.

Other analytical states have since been proposed which exhibit better resilience in the presence of loss, but provide no claim of being optimum within a specific loss regime [14, 15, 16, 17]. We have generalized our approach by a systemized numerical search for the optimum quantum state in a two-mode interferometer in the presence of loss. We calculate the classical Fisher information which, via the Cramer-Rao bound, which is inversely proportional to the variance in single mode phase of the interferometer.

In summary, we have performed unconstrained optimization of a lossy two-mode interferometer. We conclude that input N00N states are optimal for nearly zero loss [13], and that finite-photon number two mode coherent states are optimal — with shot-noise sensitivity — for large loss. Our results suggest that, if sensitivity is the only metric of success, ordinary coherent input state interferometry is best for high loss. This leaves open super-sensitive schemes employing squeezed light at the detector [43] or super-resolving schemes employing photon number resolving detectors [44].

Another means of sensing discussed herein focused specifically on the single photon state. This study was built from preexisting discovers of the possibility of determining an object's presence (of various opacities and phase shifts) through interaction free measurement. This work investigated the practical possibility of applying such a scheme to discerning the presence of opaque objects in the presence of loss through the discovery of a unique feature of total photon loss probability when loss is introduced to a single mode of the interferometer.

From this discovery, we have presented an IQT apparatus that is robust against both loss of photons and random phase accumulations in the detection arm. Interaction-free hypotheses testing in an IQT apparatus allows for stealth operation: detection of an intrusion while being virtually undetectable by an intruder. In addition, our apparatus does not require analysis of a photon's polarization state and does not rely on an exact $\pi/2$ rotation. Rather, the dependence is on the tuning of a practically controllable parameter incorporated into such a device which this allows for more robust detection methods. Therefore, such an IQT apparatus holds great promise for practical applications related to security.

It should be noted that accurate probabilities require multiple trials M, and that the number of cycles must be increased to achieve the required confidence level C_{bin}. However, these conditions risk increasing photon-target interaction thereby defeating the invisibility of the tripwire. The maximum probability of choosing an erroneous hypothesis in terms of the binary Chernoff distance is $P_e(M) = \frac{1}{2}\exp[-MC_{bin}]$. In order to make a fair comparison, we introduce an invisibility distance $C_{vis} = -\ln[1 - P_{str}]$ and invisibility probability $\bar{P}_{vis}(M) = \exp[-MC_{vis}]$. The visibility distance is a function of the probability of a photon to strike the target object $P_{str} = (1 - \lambda)(1 - \cos^{2N}\theta)$.

The method of operation begins with determining the amount of loss for a given number of cycles that will maximize the total absorption probability. This loss value is then used to calculate C_{bin} and C_{vis}. In order to be affective, we wish to choose N large enough for $\bar{P}_{vis}(M) > 2P_e(M)$. This is the limit at which the invisibility becomes greater than the maximum error of accepting the incorrect hypothesis. Because of the analogousness of these functions, this condition will hold for all M conditioned on $C_{vis} < C_{bin}$. As exemplified by table 9.1, this condition is satisfied for $N > 13$ cycles (assuming an orthogonal total rotation).

References

[1] Robert H. Good, *Classical Electromagnetism*. Hayward: Saunders College Publishing, 1999.

[2] Raymond A. Serway, Clement J. Moses, and Curt A. Moyer. *Modern Physics*. 2nd ed. New York: Saunders College Publishing, 1997.

[3] Eugene Hecht, *Optics*. 3rd ed. New York: Saunders College Publishing, 1998.

[4] Stephen Gasiorowicz, *Quantum Physics*. 3rd ed. Hoboken: John Wiley & Sons, 2003.

[5] Marlan O. Scully and M. Suhail Zubairy, *Quantum Optics*. Cambridge: Cambridge University Press, 1997.

[6] Christopher Gerry and Peter Knight, *Introductory Quantum Optics*. Cambridge: Cambridge University Press, 2005.

[7] D. F. Walls and G. J. Milburn, *Quantum Optics*. New York: Springer-Verlag, 1995.

[8] Z.Y. Ou and L. Mandel, Am. J. Phys. **57**, 66 (1989).

[9] R. A. Campos, B. E. A. Saleh, M. C. Teich, Phys. Rev. A **40**, 1371 (1989).

[10] R. P. Feynman, R. B. Leighton and M. Sands, *The Feynman Lectures on Physics*: Reading, MA (1965), Vol. 3, Chs. 3-4.

[11] K. T. McDonald and L. J. Wang, e-print: arXiv:quant-ph/0312032

[12] A. N. Boto *et al.*, Phys. Rev. Lett. **85**, 2733 (2000).

[13] H. Lee, P. Kok, and J. P. Dowling, J. Mod. Opt. **49**, 2325 (2002); J. P. Dowling, Contemp. Phys. **49**, 125 (2008).

[14] G. Gilbert, M. Hamrick, and Y. S. Weinstein, J. Opt. Soc. Am. B **25** 1336 (2008).

[15] M. A. Rubin, and S. Kaushik, Phys. Rev. A **75**, 053805 (2007).

[16] A. D. Parks *et al.*. Rev. Sci. Instr. **76**, 043103 (2005).

[17] S. D. Huver, C. F. Wildfeuer, and J. P. Dowling, Phys. Rev. A **78**, 063828 (2008).

[18] H. Uys and P. Meystre, Phys. Rev. A **76**, 013804 (2007).

[19] R. Loudon, *The Quantum Theory of Light* 3rd ed. (Oxford University Press, Oxford, UK 2000).

[20] R. J. Glauber, Phys. Rev. **131**, 2766 (1963).

[21] A. Perelomov, *Generalized Coherent States and Their Applications, Texts and Monographs in Physics*: Springer, Berlin 1985.

[22] P. G. Kwiat, H. Weinfurter, T Herzog, and A. Zeiling, Phys. Rev. Lett. **74**, 4763 (1995).

[23] P. G. Kwiat *et al.*, Phys. Rev. Lett. **83**, 4725 (1999).

[24] Ju-Seog Jang Phys. Rev. A **59**, 2322 (1999).

[25] A. G. White, J. R. Mitchell, O. Nairz, and P. G. Kwiat, Phys. Rev. A **58**, 605 (1998).

[26] L. Pezze and A. Smerzi, Phys. Rev. A 73, 011801(R) (2006).

[27] T. Kim and H. Kim, JOSA B, Vol. 26, Issue 4, pp. 671-675 (2009).

[28] A. Luks and V. Perinova, J. Phys.: Conf. Ser. 36 103 (2006).

[29] H. Chernoff, Annals of Mathematical Statistics **23** 493 (1952).

[30] A. C. Elitzur and L. Vaidman, Found. Phys. **23**, 987 (1993).

[31] P. G. Kwiat, H. Weinfurter, and A. Zeilinger, Sci. Am. (Int. Ed.) **275**, 52 (1996).

[32] Alan J. DeWeerd, Am. J. Phys. **70**(3), 272 (2002).

[33] R. M. Angelo, Found. Phys. **39**, 109 (2009).

[34] T. Rudolph, Phys. Rev. Lett. **85**(14), 2925 (2000).

[35] J. Calsamiglia and R. Munoz-Tapia, Phys. Rev. A **77**, 032311 (2008).

[36] P. M. Anisimov1, D. J. Lum, S. B. McCracken, H. Lee and J. P. Dowling, New J. of Phys. **12**, 083012 (2010).

[37] T. W. Lee *et al*, Phys. Rev. A **80**, 063803 (2009).

[38] G. A. Durkin and J. P. Dowling, Phys. Rev. Lett. **99,** 070801 (2007).

[39] U. Dorner *et al.*, Phys. Rev. Lett. **102,** 040403 (2009).

[40] P. Kok *et al.*, Rev. Mod. Phys. **79,** 135 (2007).

[41] D. B. Uskov *et al.*, Phys. Rev. A **79,** 042326 (2009).

[42] Y. Gao and H. Lee, J. Mod. Opt. **55,** 3319 (2008).

[43] C. M. Caves, Phys. Rev. D **23,** 1693 (1981).

[44] Y. Gao, C. F. Wildfeuer, P. M. Anisimov, H. Lee, and J. P. Dowling, e-print: arXiv:quant-ph/0907.2382v2 (2009).

Vita

Blane McCracken was born on November 5 1981, in Florence, South Carolina. He finished his undergraduate studies at Francis Marion University in May 2004 with a Bachelor of Science degree in computational physics. While attending Francis Marion, he was accepted into an REU internship at Montana State University, where he performed computational research on the magnetic properties of solar flares. After graduating, Blane devoted two years to musical performance, computer programming and network administration, and the successful completion of his private pilot's license. In August 2006 he came to Louisiana State University to pursue graduate studies in theoretical quantum physics. During his time at LSU, Blane presented his research in many national and international conferences, and interned at Northrop Grumman on experimental quantum optics. He is currently a candidate for the degree of Doctor of Philosophy in theoretical quantum optics, which will be awarded in May 2012.