Coherent Multi-Photon Interference and Compensation of Polarization Dispersion

by

John Michael Fini

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2001

Author . . . .
Department of Electrical Engineering and Computer Science
May 4, 2001

Certified by . . . .
Hermann A. Haus
Institute Professor
Thesis Supervisor

Certified by . . . .
Peter L. Hagelstein
Associate Professor
Thesis Supervisor

Accepted by . . . .
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Coherent Multi-Photon Interference and Compensation of Polarization Dispersion

by

John Michael Fini

Submitted to the Department of Electrical Engineering and Computer Science
on May 4, 2001, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Science and Engineering

Abstract

This thesis will explore strategies for coherent manipulation of multi-photon packets. Correlated multi-photon states can arise in nonlinear optical devices. A nonlinear quantum interferometer which includes these states can have interesting and strikingly different behavior from a conventional interferometer. The special quantum states set up in these devices are degraded by loss. In this thesis, we will set up theory for describing the interesting quantum behavior of these devices and the limitations imposed by loss decoherence.

The underlying structure of interferometer states is shared by all two-mode systems—including classical polarization states of a single-mode fiber. We are exploring birefringent systems as a possible implementation of nonlinear quantum coupler ideas, but also because of its practical importance in optical communications. Polarization mode dispersion an important source of signal distortion in high-bitrate communications arising from unwanted birefringence in the fiber. We will describe theoretical analysis of compensated PMD systems, as well as our theoretical and experimental effeots to develop novel PMD compensators.

Thesis Supervisor: Hermann A. Haus
Title: Institute Professor

Thesis Supervisor: Peter L. Hagelstein
Title: Associate Professor

Acknowledgments

I have gotten to the point where I can write this thesis only because of the enormous good luck I have had over the years with supportive friends and insightful teachers. First I would like to thank my oldest and dearest friend, Kavitha, for helping me become what I am. Through many changes and accomplishments, I have relied on her to nurture parts of me that others could not see.

I cannot thank my family, and especially my mother, enough. I thank them, most of all for always wanting more for me. For encouraging me whenever I have wanted to push myself, or just encouraging me to take life less seriously. The path towards this degree probably started with Ms. Straus and the rest of an inspiring team of techers at Hunter College High School. I hope they continue opening up opportunities as they have for me. I have also been lucky to have a wonderful experience learning from Hermann Haus and Peter Hagelstein. I benefited not only in the obvious ways, from their enthusiasm and powerful minds. In them, I found rare examples of caring, respect, and dignity. In a very imperfect world, I somehow stumbled upon two men that I can call role models.

I also owe a lot to the communities that have embraced me here at MIT. The members of the optics group, the students of Senior House, and members of the MIT Tae Kwon Do Club all gave me the feeling of a fun and interesting home within this otherwise drab institute. This list of people and thoughts is still horribly incomplete. I hope that all those who have touched my life realize that this page does not begin to capture my feelings or memories.

This work was supported by the National Science Foundation, the Air Force Office of Scientific Research, the Office of Naval Research, and the 3M Corporation.

Contents

List of Figures

Chapter 1

Introduction

In this thesis, we will explore some of the ways people think about light. Naturally, physicists have thought a lot about this topic, and the current understanding is very sophisticated: there are theories with every level of complexity, dealing both with the complicated interactions with matter and of the fundamental issues behind a correct field theory. We do not aim to cover new ground in this direction. Rather, we explore some interesting perspectives that have recently arisen for describing quantum optics problems: photon configuration-space theory and photon DeBroglie waves.

These perspectives give new insight and interpretations to specific quantum optics models. In addition, Hagelstein's recent work on photon configuration space provides more complete answers to some of the most fundamental questions in quantum optics. Here we are more interested in possible applications; a new class of devices that manipulate light in a qualitatively non-classical way is suggested by these theoretical developments. In this thesis, we explore the possibility of implementing strange quantum measurements in a variety of systems.

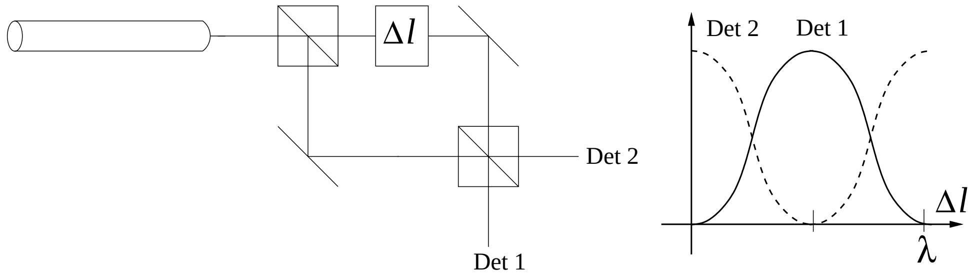


Figure 1-1: Interference of light is most easily understood using classical wave optics. In a standard Mach-Zehnder interferometer, we see interference fringes with the photon wavelength λ.

1.1 Traditional quantum optics

Generally speaking, light can be understood as a classical wave. The simple and elegant formulation given by Maxwell's equations (along with various models of matter) generally describes the world very well, and is our classical understanding of light. It says that the light intensity seen by our eyes or by a detector is the power delivered by the electromagnetic fields, which vary in time and space according to known rules.

Quantum rules are occasionally needed to complete our understanding. A simple, classical interferometer is shown in Figure 1-1. The classical description of this interferometer found in an optics textbook accurately describes most of the behavior of a real device. However, a key concept of quantum uncertainties is part of the real physics missing from that description. To understand the measurements in detail, one might instead draw a picture like Fig. 1-2, and say that the uncertainties in the field obey the rules of quantum mechanics. Just as a quantum mechanical particle cannot have a well-defined trajectory—its position and momentum cannot be simultaneously and precisely known—the quantum field does not have a well-defined amplitude.

But the picture, Fig. 1-2, does not represent the most general type of quantum optical system. It shows the case of a quantum system which behaves qualitatively like its classical counterpart. In general, the dynamics of a nonlinear optical system can be richer and more complex than this semiclassical view suggests. To see this, consider that a single-mode classical field has only one degree of freedom, the com-

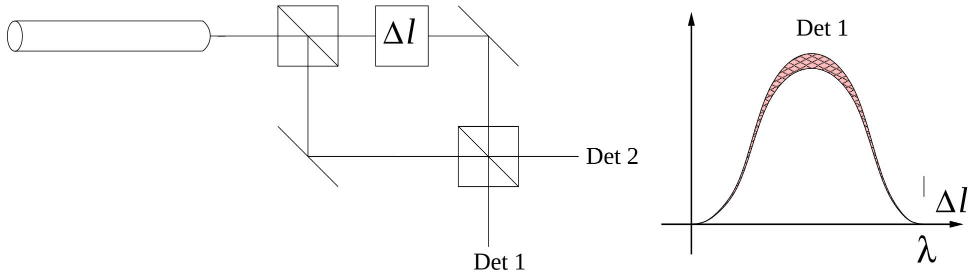


Figure 1-2: In a quantum interferometer, the fields are subject to fundamental quantum uncertainty. The input quantum state determines the counting statistics, and thus the uncertainty, but the qualitative behavior of the device is the same as the classical case.

plex field amplitude $a(t)$. A single-mode quantum field can exist in a superposition of number states $|n\rangle$, so that the state is described by an infinite-dimensional complex amplitude vector. Since the state space is much larger, one can in theory construct quantum systems that have complicated Hamiltonians and highly nontrivial dynamics. For example, one could construct a self-consistent mathematical model for an interferometer which has wildly different behavior depending on the input photon number. Although the traditional formalism for describing quantum optical states permits such complex, non-classical dynamics, physicists almost universally study semiclassical systems.

The reason is simple: in the real experiments that they have analyzed, a semiclassical picture works very well; as in Fig. 1-2, quantum properties may have an important effect on the measurement, but they do not qualitatively change the dynamics of the device. Ultimately, this is because optical nonlinearities are weak. In the next chapter we will see that the dynamics of a linear optical system are essentially classical. In real-world systems with small nonlinearities, more non-classical states of a system are not accessible. The dynamics can generally be handled with some semiclassical approximation, and deviations from the classical field remain small as the system evolves, as with the case of squeezing. Squeezing experiments demonstrate an important and potentially useful effect—they break the limits of shot noise in an interferometer—but work within the semiclassical regime. Such a system necessarily involves a quantum analysis, but has qualitative behavior which is still essentially classical.

1.1.1 A new look: photon configuration space

Light is made up of particles called photons. These particles are immediately apparent in the traditional quantum optics literature; in particular, everywhere in these theories, the particles are being created and destroyed by photon operators. Well-known manipulations of photon operators—the methods of second quantization—allow us to analyze and solve optical models. But while we see the photons everywhere, they do not seem to act like the familiar particles of Sophomore physics. The photons that we meet in the literature sometimes have momenta, but almost never have position variables, or anything remotely resembling a trajectory (however uncertain) that would let us visualize it as a normal particle. When we ask why this is, the literature tells us that photons are fundamentally different from particles with mass, and that we must avoid thinking about certain photon qualities, like position, for reasons which are complicated and theoretical.

This thesis is in part motivated by the notion that photons can and should be thought of more like "regular" particles. To prove that they can, Hagelstein recently answered the question: "Why don't people ever talk about photon wavefunctions?" The answer is a bit involved [31]. The basic idea is as follows:

To form a correct description of an optical system using a photon wavefunction, you would want to prove that the description is equivalent to the more fundamental quantum electrodynamics description. In the middle of this proof, one discovers a number of problems, and one might understandably declare that photon wavefunctions are too problematic, and thus meaningless.

What Hagelstein pointed out is that these daunting technical issues bear a very close resemblance to the issues that come up when one derives an electron-wavefunction model from QED! That is, there are issues, but they are not nearly as different for photons and electrons as people generally believe. Hagelstein has carried out the solution of the technical problems, and outlined various properties of a photon wavefunction. The notion of a wavefunction can thus be carefully defined and reconciled with known properties of photons—for example, one should not infer that since a photon has a wavefunction the problems with defining a photon position operator go away. A photon still does not have a well-defined position operator (as is well known), but a wavefunction can be defined to include this property.

The second assertion is that people should use more particle-like descriptions of optical systems. Several reasons come to mind: One is the great success of particle and wavefunction-based methods in atomic and nuclear physics. If these methods were so useful in other areas, why arbitrarily exclude them from quantum optics? At least in some cases, second-quantized descriptions using creation and annihilation operators are much more complicated than the equivalent "first-quantized" (that is, wavefunction or coordinate-based). Much of our basic intuition comes from the simple first-quantized problems we studied before we knew about creation and annihilation. Could we not discover a much more basic intuitive understanding of some systems, then, by thinking more in terms of particles?

The name for this particle-oriented type of theory is "photon configuration space." While the term itself may be off-putting, I emphasize that it refers to the kind of simple theory generally used as an introduction to quantum mechanics, where particles have position and momentum coordinates, etc. Finally, we stress that a configuration space model is generally derived from, and exactly equivalent to some second-quantized counterpart. Configuration-space theory does not describe different physics, but it describes a system in a different way. Naturally, the hope is to uncover aspects of the physics which were obscured in the second-quantized description.

This brings us to the obvious question: Does the configuration space perspective offer anything new and useful? Much of the research done towards this thesis has been motivated by this question. In the next chapter, Chapter 2, we present background for the problems of optical propagation and interference. Chapter 3 gives a detailed example of configuration space applied to soliton propagation in fiber. It presents simple thought experiments for manipulating quantum solitons and obtaining new and potentially useful effects. In Chapter 4, we identify severe difficulties of such an approach due to loss decoherence. In the process, we identify an interesting new interpretation of soliton noise in the presence of loss and gain. This is an example of how a configuration-space theory can gives us useful intuition to apply towards other problems. We then move on to the topic of DeBroglie-wave interference. We describe this briefly below.

1.1.2 Photonic DeBroglie waves

The first published photon configuration space theory for quantum-optics applications was probably [51]. In that article, Lai and Haus describe a soliton state in terms of a wavefunction $\Psi(x_1, \ldots, x_n)$ for the positions x_j of n photons along the length of the fiber. In this formulation, the Kerr nonlinearity appears as an attraction between the photons: a pairwise, delta-function potential, $V = V_0 \sum_{j \neq k} \delta(x_j - x_k)$. Under this attractive force, the photons can be trapped into a bound state not unlike that of an atom or nucleus.

A bound, many-photon object has the potential to be very interesting. Naturally, in a many-photon soliton, one can observe relatively uninteresting, semiclassical behavior. If each photon approximately sees a classical "mean field" generated by the others, then little correlation will develop between photons, and the resulting state will be unremarkable. On the other hand, we can imagine manipulating this quantum pulse as a whole, placing it in a superposition of states, and even observing interference not of the individual photons, but of the n-photon object. In the next chapter, we will elaborate on this "photonic DeBroglie wave" concept. We will see that it is both strikingly different than the usual quantum optical manipulations of light, and potentially quite useful.

Chapters 5 and 6 take a detailed look at a particular photonic DeBroglie wave interferometer model. We have numerically and theoretically confirmed exciting new physics present in some simplified interferometer models, and tried to assess the plausibility of a real-world implementation. Chapter 7 briefly outlines the connection between the two-mode problems of Chapters 5 and 6, and the problem of polarization dynamics addressed in the remainder of this thesis. To some extent, this connection is obvious: the two polarizations of a “single-mode” waveguide constitute a two mode system with exactly the same mathematical framework as an interferometer. Similar issues of nonlinear dynamics and squeezing have been studied in both kinds of systems.

1.2 Polarization and optical fiber communications

If polarization dynamics are a potentially interesting implementation of the above quantum ideas. However, classical polarization dynamics in an optical communications line is an issue of intense practical interest. Unwanted varying birefringence in optical fiber is a key source of signal distortion limiting communications bandwidths today. This distortion is called polarization mode dispersion. In the final chapters, we summarize our theoretical and experimental contributions in the area of polarization mode dispersion. Chapter 8 provides some quick background of how PMD is described theoretically, and how people try to mitigate the distortion. In Chapter 9, we present a theoretical analysis of the accumulation of PMD in a cascade of compensated channels. Finally in Chapter 10, we discuss estimation of PMD parameters using polarimeter measurements. This is part of an effort towards feed-forward PMD compensation conducted with Patrick Chou and Poh-Boon Phua.

Chapter 2

Quantum propagation and interference of light

The bulk of this thesis examines specific implementations of untraditional quantum optics ideas: configuration space and photonic DeBroglie methods. Here we give some background on the quantum theory of pulse propagation and optical interference. We briefly discuss the basic approach traditionally used for quantum optics problems: second-quantization with some semiclassical approximation. We also note that the traditional quantum optics literature is consistent with the possibility of qualitatively new effects. The literature tends to emphasize semiclassical dynamics because realistic systems generally fall safely within this regime. The type of effects that would be encountered in another regime, that of Photonic DeBroglie wave devices are fascinating, and outlined briefly below.

2.1 Standard method: second-quantization

The theory of quantum optics has almost all been in terms of a second-quantized description. This refers simply to a description written in terms of photon annihilation and creation operators. As a simple example, one might look at a set of electromagnetic modes with photon energies $\hbar\omega_k$, given by the Hamiltonian,

$$\hat{H} = \sum_k \hbar\omega_k \hat{a}_k^\dagger \hat{a}_k. \tag{2.1}$$

The Hamiltonian and commutation relations,

$$[\hat{a}_k, \hat{a}_j^\dagger] = \delta_{kj} \tag{2.2}$$

are the starting point for the well-known methods of performing quantum optics calculations. By including various coupling and nonlinear terms in the Hamiltonian, one can model and understand a rich variety of optical systems.

2.1.1 Linear optical systems

Quantum dynamics do not play a particularly interesting role in the dynamics of linear optical systems. A system is linear if it has a photon-conserving quadratic Hamiltonian,

$$\hat{H} = \sum_{j,k=1}^{M} E_{jk} \hat{a}_j^\dagger \hat{a}_k \tag{2.3}$$

The dynamics of all such M-mode systems are essentially the same. To see this, note that the Hermitian matrix with entries E_{jk} can be diagonalized by a unitary transformation $\boldsymbol{U}$,

$$\boldsymbol{E}' = \boldsymbol{U}\boldsymbol{E}\boldsymbol{U}^\dagger \quad \text{(diagonal)}. \tag{2.4}$$

Now define the new set of photon operators for the new eigenmodes of the field,

$$\hat{b}_j = \sum_k U_{jk} \hat{a}_k \tag{2.5}$$

and obtain the diagonalized form of the Hamiltonian,

$$\hat{H} = \sum_{k=1}^{M} E'_{kk} \hat{a}_k^\dagger \hat{a}_k. \tag{2.6}$$

From here the eigenvalue decomposition of any quantum state is trivial, and so the general quantum dynamics follows the standard methods for exactly diagonalized systems.

That the exact solution of the eigenstates amounts to the diagonalization of an $M \times M$ matrix is a very nontrivial statement. The classical system has M modes, but the quantum problem has many more dimensions. That is, an M-mode field has basis states $|n_1, n_2, \ldots, n_M\rangle$, where any of the n_j's can in principle extend from 0 to ∞. Even for a finite number of photons, the state space is large. The matrix $\boldsymbol{E}$ that we diagonalize is clearly not the Hamiltonian of the quantum problem. Linear systems are a special case where the potentially complex quantum dynamics reduce to the simpler classical mode decomposition. The only difference between a quantum and a classical field is that the quantum field is still fundamentally made up of discrete energy quanta. Uncertainties arise from the counting statistics of the photons.

There is another reason why this type of system is not as "interesting" as a nonlinear optical system. Any linear system with an input coherent state has a coherent state for all time. Since coherent states essentially act like classical fields, one goal of quantum optics is to explore systems which can transform ordinary coherent states into more interesting "quantum" states. Only when a system has a nonlinearity can it develop interesting quantum dynamics given a "classical" (that is coherent-state) input field. The proof is simple; given the above Hamiltonian, the coherent state $|\alpha_k(t)\rangle$ satisfies the Schrödinger equation as long as α satisfies the classical coupled-mode evolution,

$$i\frac{d}{dt}\alpha_j(t) = \sum_k E_{jk}\alpha_k(t) \tag{2.7}$$

This can be confirmed by direct substitution.

2.1.2 Nonlinear systems and squeezing

Consider the degenerate parametric amplifier analyzed, for example, in [34]. This system has a non-quadratic Hamiltonian, which couples pump photons (with operator

$\hat{b}$) to pairs of signal photons (with operator $\hat{a}$),

$$\hat{H} = V\hat{b}^\dagger\hat{a}\hat{a} + V^*\hat{a}^\dagger\hat{a}^\dagger\hat{b}. \tag{2.8}$$

By writing the Heisenberg equation of motion, one quickly finds that the coupling of the pump and signal is not trivial. One cannot easily transform the Hamiltonian into any exact, diagonalized form as we could in the linear case. While the evolution equations have a compact, intuitive form, they are not at present exactly solvable.

Yet the quantum properties of parametric oscillators are well understood. As with other real-world optical systems, the optical nonlinearities are small and allow a semiclassical analysis. For the same basic reason, physicists have generally been able to get an excellent understanding of squeezing and other quantum optical phenomena from semiclassical or linearized theories. In this case, the small optical nonlinearity requires an intense pump field. Quantum uncertainty in this many-photon field can then be considered a higher-order correction, and an approximate solution is obtained by setting $\hat{b}$ to a complex constant [34]. with this approximation, the Heisenberg evolution takes the simple form,

$$\frac{d}{dt}\begin{bmatrix} \hat{a} \\ \hat{a}^\dagger \end{bmatrix} = \begin{bmatrix} 0 & C \\ C & 0 \end{bmatrix}\begin{bmatrix} \hat{a} \\ \hat{a}^\dagger \end{bmatrix} \tag{2.9}$$

This equation *is* exactly solvable, and in fact gives the well-known Bogolyubov transformation,

$$\begin{bmatrix} \hat{a}(t) \\ \hat{a}^\dagger(t) \end{bmatrix} = \begin{bmatrix} \mu & \nu \\ \nu^* & \mu \end{bmatrix}\begin{bmatrix} \hat{a}(0) \\ \hat{a}^\dagger(0) \end{bmatrix}. \tag{2.10}$$

This is the essence of squeezing in a parametric amplifier; once the evolution is shown to give a Bogolyubov transformation, one can immediately connect with the standard literature on squeezed states. Naturally, there is a literature filling in various details and higher-order corrections of this result.

Another type of linearization is seen in soliton theory. again, the idea is that soliton formation in materials with small nonlinearities requires many photons. The quantum nature of the field then leads to only tiny deviations from the classical evolution. This is formalized in [35] by writing the quantum field explicitly as a classical part and a small quantum correction, $\hat{\phi}(x) = \phi_0(x) + \hat{v}(x)$. In these semiclassical analyses, one is assuming from the outset that the qualitative behavior must be classical (generally a quite reasonable assumption). In this thesis, we would like to explore the possibility of another, truly nonclassical, regime. A regime where the quantum deviations $\hat{v}$ slowly grow until they are no longer a small perturbation. Such a regime requires a different perspective, which we discuss in the next section.

2.1.3 Two-mode interferometer

We are particularly interested in two-mode quantum problem, since it captures in a simple way the essential aspects of interference. Once we have two modes, light can evolve along two "paths" in a system and show interference between these paths.

A basis of two-mode states is constructed from the photon operators and the vacuum state $|0,0\rangle$,

$$|\psi\rangle = \sum_{n_1,n_2} c_{n_1,n_2} |n_1, n_2\rangle \tag{2.11}$$

$$|n_1, n_2\rangle = \frac{1}{\sqrt{n_1! n_2!}} (\hat{a}_1^\dagger)^{n_1} (\hat{a}_2^\dagger)^{n_2} |0,0\rangle \tag{2.12}$$

Linear coupling between two modes can now be described, for example, by

$$\hat{H} = \hbar\omega_0(\hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2) + V(\hat{a}_1^\dagger \hat{a}_2 + \hat{a}_2^\dagger \hat{a}_1) \tag{2.13}$$

and allows us to understand the quantum behavior of a beamsplitter, for example. Along with the phase shift operator,

$$\hat{P} = e^{i\phi(\hat{n}_1 - \hat{n}_2)/2}. \tag{2.14}$$

this allows us to understand the basic limitations on interferometry imposed by quantum mechanics.

Historically, the problem of defining the phase shift $\Delta\phi$ resolvable in an interferometer with a given optical power has been very important. The model outlined above reveals that that "shot noise" in an interferometric measurement is a result of the statistics of photons choosing one path or another as they pass through the interferometer [10, 70]. That is, the statistics of photons independently choosing a path at the beamsplitters ultimately gives rise to a limiting uncertainty in output intensity. Careful analysis of this problem demonstrated that appropriate combinations of special quantum states at both ports of an interferometer could lead to sub-shot-noise measurement precision [10, 6, 82, 87]. This is one of the main successes of quantum optics.

The current theoretical understanding of the quantum interferometer is quite sophisticated. It includes methods of representing states and visualizing the transformations such as given in [87]. It is clear that the authors of these papers understood that quantum states can in principle display complex, qualitatively non-classical dynamics. However, examples in the real world are hard to come by, due in part to decoherence effects that we discuss in Chapters 4 and 6. The community has come to think of things semiclassically, so that now the suggestion of any effect that is not semiclassical is seen as being very counter-intuitive. The untraditional perspectives that we discuss in this thesis try to present an intuitive way of understanding problems outside of the semiclassical regime. It remains to be seen whether these perspectives will fulfill their potential for uncovering new possibilities in real systems.

The brief background that we have given above give some of the basic context of our problem. In Appendix A we outline some interesting related problems that may help motivate and clarify the scope of this research.

2.2 Photonic DeBroglie Waves

A standard Mach-Zehnder interferometer was shown in Fig. 1-1. The two modes are coupled linearly by a beamsplitter, and a phase shift $\Delta\phi$ is applied, for example by introducing a path length difference, Δl, such that $\Delta\phi = 2\pi\Delta l/\lambda$. The interference observed at the detectors of course shows characteristic wavelength λ. The quantum statistics of such a linear device are well understood [87, 9]. They can be summarized surprisingly well by a simple linearity argument: when many photons enter a linear interferometer, each chooses its output port independently of the others, with probabilities proportional to the classical intensity. Since each photon follows the classical intensity, the dynamics are essentially classical, with quantum "uncertainties" arising through the counting statistics.

Very different physics arise if, somehow, each photon's path through the interferometer depends on the other photons. For example, imagine that we have designed a special nonlinear beamsplitter which effectively binds the photons together, but allows them to couple between waveguides as a single quantum unit. The behavior that would result, shown in Fig. 2-1, may seem counterintuitive. The detector currents show interference with characteristic wavelength λ/N, where N is the number of photons. One can gain some physical intuition by thinking in terms of a bound N-photon particle with momentum $N2\pi\hbar/\lambda$. Like a bound group of nucleons, the wavelength varies as one over the momentum, and is thus inversely proportional to the number of constituent particles, N. Mathematically, this "counterintuitive" behavior actually comes right out of standard theory. It is well known that the phase shift operator is given by

$$\hat{P} = e^{i\phi(\hat{n}_1 - \hat{n}_2)/2}. \tag{2.15}$$

Suppose one could generate a correlated superposition state, where all photons are placed in one waveguide $|N,0\rangle$ or the other $|0,N\rangle$, but are never divided between the two:

$$|\Psi\rangle = \alpha|N,0\rangle + \beta|0,N\rangle \tag{2.16}$$

This state is inherently very sensitive to phase shifts between the modes [5]. It experiences N times the usual, classical phase-shift. For example, with $\alpha = \beta = 1/\sqrt{2}$,

$$\hat{P}|\Psi\rangle = e^{iN\phi/2}\left(|N,0\rangle + e^{iN\phi}|0,N\rangle\right)/\sqrt{2}. \tag{2.17}$$

so that the final state oscillates from $|N,0\rangle + |0,N\rangle$ to $|N,0\rangle - |0,N\rangle$ and back again as ϕ varies from 0 to $2\pi/N$. This result seems counterintuitive only because

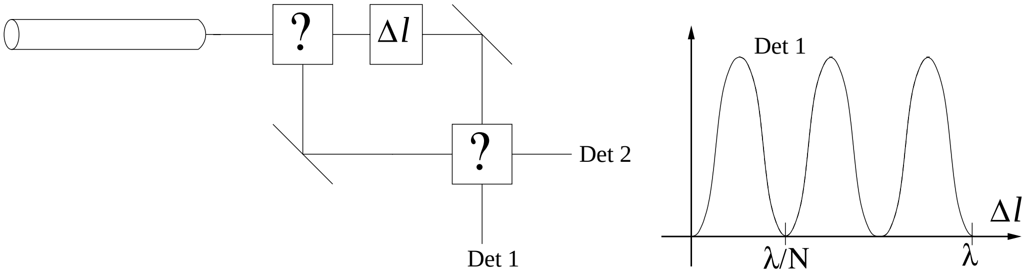


Figure 2-1: The nonlinear Mach-Zehnder interferometer pictured here has special beamsplitters in which N photons are effectively bound together. One can show that such a bound N-photon particle exhibits a striking λ/N interference effect.

optical nonlinearities are typically weak, so that strongly correlated states such as $|\Psi\rangle$ typically do not arise. Below, we discuss a device model for generating the superposition state and resolving the phase-shifted states.

The possibility of designing such devices with qualitatively non-classical behavior has generated much enthusiasm. Recent papers [77, 25, 7] have built on earlier results using entangled photon pairs; papers such as [50, 65] report early photonic DeBroglie wave measurements, although they precede the use of that name. While photon-pair results are interesting, their generalization to the more interesting multiple-photon case ($N > 2$) is nontrivial. Existing proposals [4, 62] may be difficult to implement.

In Hagelstein's proposed nonlinear coupler, one achieves N-photon correlated states in a natural way: the nonlinear "attraction" between photons essentially leads to bound states in the coupler analogous to the soliton bound states of the nonlinear Schrödinger equation. A number of effects may make the concept difficult to implement. For example, to maintain coherent superpositions, one must essentially avoid losing even a single photon. This translates to a very strict requirement on the ratio of loss and nonlinearity coefficients.

Similarly, in pulse propagation, quantum uncertainties are typically thought of a small perturbation of qualitatively classical dynamics. However, at least within a simple mathematical model, one can clearly construct quantum superposition states: the soliton can be in a coherent superposition of position states, for example. If the positions were well-separated, we would again be able to observe qualitatively new, non-classical effects analogous to the interferometer effects described above. In the next four chapters, we will explore these theoretical possibilities.

Chapter 3

Lossless pulse propagation

Optical solitons have been a successful tool for exploring quantum mechanics because of their elegant basic physics. The simplest model has a balance of fiber dispersion and Kerr nonlinearity, given by the Hamiltonian [51],

$$\hat{H} = \hbar \left[\frac{-\omega''}{2} \int dx \hat{\phi}^\dagger(x) \frac{\partial^2}{\partial x^2} \hat{\phi}(x) + \kappa \int dx \hat{\phi}^\dagger(x) \hat{\phi}^\dagger(x) \hat{\phi}(x) \hat{\phi}(x) \right] \tag{3.1}$$

We can relate this Hamiltonian to our classical intuition by writing the Heisenberg equation of motion: the field operators evolve according to the quantum version of the nonlinear Schrödinger equation,

$$i \frac{d}{dt} \hat{\phi}(x,t) = \frac{-\omega''}{2} \frac{\partial^2}{\partial x^2} \hat{\phi}(x,t) + 2\kappa \hat{\phi}^\dagger(x,t) \hat{\phi}(x,t) \hat{\phi}(x,t) \tag{3.2}$$

Quantum solitons resemble their classical counterparts—the above equation governs the quantum (Heisenberg) field operator, but if we remove the hats we get a classcial nonlinear wave equation, which has soliton solutions:

$$\phi_0(x,t) = A\mathrm{sech}((x - x_0(t))/x_s) e^{ip_0 x + i\theta(t)}. \tag{3.3}$$

These are subject to the soliton condition, $2|A|^2\kappa = |\omega''|/x_s^2$ and well-known evolution equations $\dot{x}_0 = \omega'' p_0$ and $\dot{\theta} = (\omega''/2)(1/x_s^2 - p_0^2)$. We use $\omega'' < 0$ and $\kappa > 0$.

The derivation of the above "moving frame" equation is discussed in [51]. When using this equation, we must keep in mind that pulse timing and temporal shape have been converted into position x and propagation along the fiber length z has been converted to an evolution in "time" $t = z/v_g$. What appears mathematically as momentum, $-i\hbar d/dx$, corresponds better physically to frequency. Thus we refer to the "spectrum" of the pulse at many points in the discussion of the momentum uncertainty.

3.1 Configuration space and soliton states

A key concept in understanding quantum states in optical fiber is the separation of the collective coordinates from other, internal degrees of freedom. The collective variables of the field are quantities such as the total photon number and average position. Other degrees of freedom can be thought of as the relative positions of photons along the length of the fiber, as we discuss below. The remarkable thing is that the evolution of collective and relative variables separate completely, dramatically simplifying the dynamics. This is true in the classical dynamics, and plays an important role in the quantum system as well. In particular, soliton dynamics are especially simple, since the relative coordinates are in a bound eigenstate. In this section, we give a brief review of a number of key concepts related to quantum states and the separation of collective and internal variables.

3.1.1 n-photon wavefunction

Consider a state of n photons. These photons can in general be in any mode of the one-dimensional system, and may not form a soliton at all. The separation we are looking for is most clearly formulated using the photon wavefunctions of [51, 32]. The most general n-photon state is described by a symmetric wavefunction $\Psi(x_1, \ldots, x_n)$ of the individual photon positions, x_1, etc,

$$|\Psi\rangle = \frac{1}{\sqrt{n!}} \int dx_1 \ldots dx_n \Psi(x_1, \ldots, x_n) \hat{\phi}^\dagger(x_1) \ldots \hat{\phi}^\dagger(x_n)|0\rangle \tag{3.4}$$

Applying the Hamiltonian (3.1) to the above state, we find the equivalent "configuration-space" Hamiltonian [51, 32],

$$\hat{H}_n(t) = \hbar \left[\sum_k \frac{-\omega''}{2} \frac{d^2}{dx_k^2} + \sum_{j<k} 2\kappa\delta(x_k - x_j) \right]. \tag{3.5}$$

That is, this Hamiltonian acts directly on $\Psi(x_1, \ldots, x_n)$ and is mathematically equivalent to the second-quantized Hamiltonian on the n-photon sector,

$$\hat{H}|\Psi\rangle = |\hat{H}_n \Psi\rangle, \tag{3.6}$$

as can be confirmed by direct substitution.

3.1.2 Collective coordinates

The basic idea of separating variables is that the collective position,

$$X = \frac{1}{n} \sum_{k=1}^{n} x_k \tag{3.7}$$

is special: as with many massive particle systems, forces between the photons move their individual positions, but do not move their center-of-mass. More precisely, X is the conjugate variable to the total field momentum, P, which is conserved by the Hamiltonian [51]. X then evolves exactly like a free particle.

We would like a simplified evolution, where X evolves separately,

$$\Psi(x_1, \ldots, x_n; t) = \psi_n(X, t)\Phi_n(x_1 - X, \ldots; t). \tag{3.8}$$

Mathematically, this results from separability of the Hamiltonian. It can be written exactly as a sum of collective and relative parts with no interaction of the two [32]:

$$\hat{H}_n = \hat{H}_c + \hat{H}_r \tag{3.9}$$

The free particle evolution of X is then independent of the state of the relative coordinates—independent of the shape of the pulse. The evolution takes an arbitrary wavefunction $\psi(X,0)$ as an initial condition to the free-particle Schrödinger equation [32, 43],

$$i\hbar\frac{d}{dt}\psi_n(X,t) = \hat{H}_c\psi_n(X,t) = \frac{-\hbar\omega''}{2n}\frac{d^2}{dX^2}\psi_n(X,t) \tag{3.10}$$

Naturally, the relative coordinates evolve according to their (more complicated) Hamiltonian, $\hat{H}_r$

$$i\hbar\frac{d}{dt}\Phi_n = \hat{H}_r\Phi_n. \tag{3.11}$$

A soliton state refers to a special state of the *relative* coordinates, a bound eigenstate of Φ_{sn},

$$E_{sn}\Phi_{sn} = \hat{H}_r\Phi_{sn}. \tag{3.12}$$

For a soliton state, the center position X and photon number are the only remaining degrees of freedom (along with their conjugate variables). The relative coordinates $x_1 - X$, etc., are constrained to their specific bound state. The bound state Φ_{sn} has been discussed at length in [51, 52].

3.1.3 Example: two photons

The two-photon case is worth mentioning because it demonstrates the general properties of the separation and bound-state, but can be exactly solved and easily visualized. The two-photon Hamiltonian,

$$\hat{H}_n = \hbar\left[\frac{-\omega''}{2}\left(\frac{d^2}{dx_1^2}+\frac{d^2}{dx_2^2}\right) + 2\kappa\delta(x_1-x_2)\right]. \tag{3.13}$$

can be readily written in terms of the collective and relative coordinates, $X = (x_1 + x_2)/2$ and $\xi = x_1 - x_2$. To relate the momenta, we use

$$\frac{\partial}{\partial x_1} = \frac{1}{2}\frac{\partial}{\partial X} + \frac{\partial}{\partial \xi} \tag{3.14}$$

$$\frac{\partial}{\partial x_2} = \frac{1}{2}\frac{\partial}{\partial X} - \frac{\partial}{\partial \xi} \tag{3.15}$$

$$\frac{d^2}{dx_1^2}+\frac{d^2}{dx_2^2} = \frac{1}{2}\frac{d^2}{dX^2} + 2\frac{d^2}{d\xi^2} \tag{3.16}$$

$$\tag{3.17}$$

so that

$$\hat{H}_c = \frac{-\hbar\omega''}{4}\frac{d^2}{dX^2} \tag{3.18}$$

$$\hat{H}_r = \hbar\left[-\omega''\frac{d^2}{d\xi^2} + 2\kappa\delta(\xi)\right]. \tag{3.19}$$

The Schrödinger equation is clearly satisfied by $\Psi(x_1, x_2) = \psi(X,t)\Phi(\xi)e^{-i\Omega t}$ if these functions obey,

$$i\hbar\frac{\partial}{dt}\psi(X,t) = \frac{-\hbar\omega''}{4}\frac{d^2}{dX^2}\psi(X,t) \tag{3.20}$$

and

$$\Omega\Phi(\xi) = \left[-\omega''\frac{d^2}{d\xi^2} + 2\kappa\delta(\xi)\right]\Phi(\xi). \tag{3.21}$$

The two-photon soliton can thus be reduced to the solution of two textbook 1-D quantum problems. The bound state $\Phi(\xi)$ is the exponential "tent" function, $\Phi(\xi) = \sqrt{A}\exp(-|A\xi|)$, $A = \kappa/|\omega''|$.

3.1.4 Large n: Hartree approximation

The Hartree approximation allows us to apply our understanding of the classical dynamics of a system to the quantum problem. Within the approximation, for large n, the individual photons in a soliton take the classical field (3.3) as their wavefunction. As shown in [83, 51], Φ_{sn} is approximately the product wavefunction

$$\Phi_{sn}(x_1 - X, x_2 - X, \ldots) \approx \phi_0(x_1 - X)\phi_0(x_2 - X)\ldots \tag{3.22}$$

A product wavefunction does not capture all aspects of the bound state (such as photon correlations); it should be used to describe the distribution of a single photon whenever one is singled out. This confirms our classical intuition: for example, when viewed on a standard detector, the pulse will have the expected classical shape, since each photon is randomly distributed according to the classical intensity. In this sense, the relative coordinates retain the classical properties of the soliton, while the

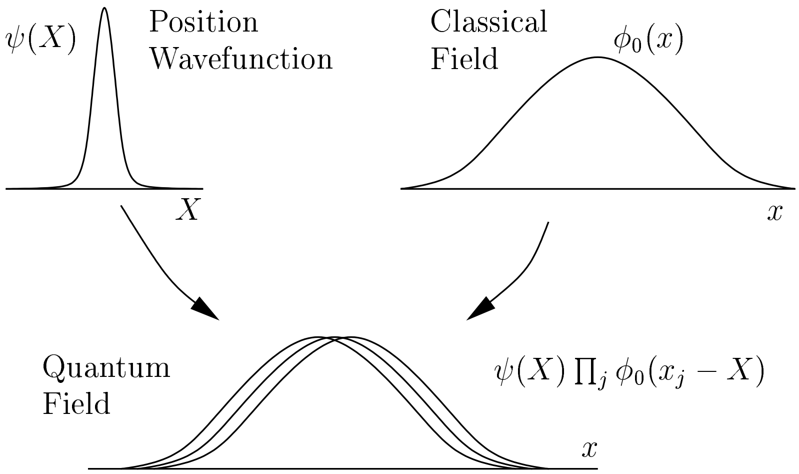


Figure 3-1: The many-photon wavefunction describing a quantum optical pulse is composed of collective position $\psi(X)$ and relative position parts. In the Hartree approximation, the relative position of each photon about X follows the classical field $\phi_0(x_j - X)$.

collective coordinates (momentum, position, photon number, and phase) carry the quantum uncertainties, as depicted in Figure 3-1.

3.1.5 Classical states and strange quantum states

The separation discussed above is important in two respects: First, it is the underlying reason why simple, closed-form evolutions have been found for the collective coordinates. For example, in deriving the Gordon-Haus effect [29] and certain soliton squeezed-state effects, the analysis is manageable because the collective variables do not couple to the continuum. Second, the separation implies the possibility of generating intriguing quantum soliton states. These states defy our usual (semiclassical) intuition.

Consider the standard coherent-state soliton. This is the state that most strongly resembles a classical soliton. The quantum uncertainty of the collective position and momentum can be calculated from the linearized theory [35]. More intuitively (but in exact agreement with the linearized calculation), the coherent-state, or "shot-noise," uncertainty is that of n uncorrelated photons distributed in the classical field (3.22):

$$\langle \Delta X^2 \rangle = \left\langle \left(\frac{1}{n} \sum_j \Delta x_j \right)^2 \right\rangle = \frac{1}{n^2} \sum_j \langle \phi_0 | \Delta x_j^2 | \phi_0 \rangle = \frac{\langle \Delta x^2 \rangle_{\text{classical}}}{n} \tag{3.23}$$

Similarly for the momentum,

$$\begin{aligned} \hat{P} &\equiv \sum_k \left[-i\hbar \frac{\partial}{\partial x_k} \right] = \sum_k \hat{p}_k && (3.24) \\ \langle \Delta P^2 \rangle &= \sum \langle \phi_0 | \Delta p_j^2 | \phi_0 \rangle = n \langle \Delta p^2 \rangle_{\text{classical}} && (3.25) \end{aligned}$$

On the other hand if the photon positions or momenta are *correlated*, we can achieve non-coherent-state uncertainties. Below we explore the possibility of generating momentum squeezed states, which we define as having sub-shot-noise uncertainty in the momentum $\hat{P}$.

The quantum position uncertainty of a large-n soliton is typically orders of magnitude smaller than the classical pulsewidth. A moderate amount of momentum squeezing may broaden the position wavefunction $\psi(X)$, but not nearly enough to approach the classical pulsewidth. The quantum uncertainty is thus a tiny perturbation of the classical field, even for fairly large squeezing. On the other hand, noise-driven position uncertainties can be large, but are accompanied by a loss of coherence. This type of uncertainty is best thought of as a classical noise (though it may ultimately arise from quantum mechanisms).

A more remarkable possibility is illustrated in Figure 3-2. Here the collective position wavefunction $\psi(X)$ has two peaks separated by a distance greater than the classical pulse width. Physically, this means that all n photons are either localized at one position or the other, but are not divided between the two. When observed on a photodetector, the pulse will randomly “choose” one position or the other, but will not be broken up. Furthermore, these two states are in a coherent superposition: They can, in theory, be manipulated to give n-photon interference effects analogous to those in [65, 22]. Such a state is highly non-classical and unusual, since it has macroscopically observable states in quantum superposition. For precisely the same

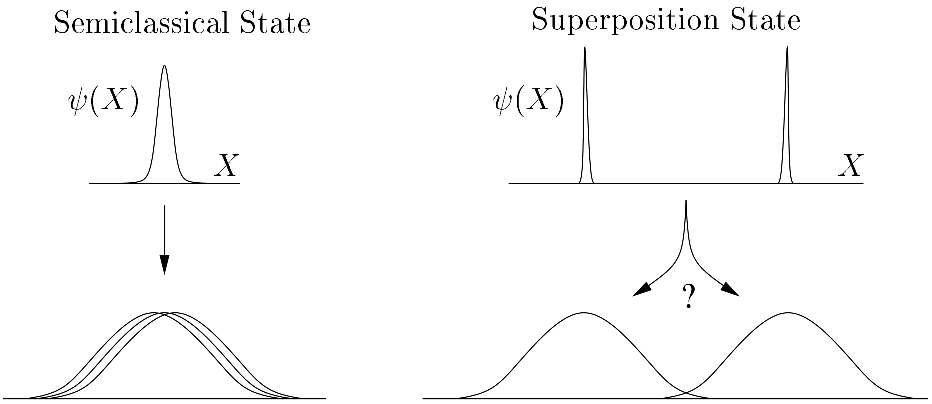


Figure 3-2: The position wavefunction is far narrower than the pulsewidth for a coherent state, or any state arising from semi-classical dynamics. Other, highly non-classical states are possible theoretically: The "superposition state" depicted here has position uncertainty greater than the classical pulse width. Such a state represents a quantum superposition of macroscopically distinguishable states.

reason, it is extremely fragile.

The possibility of such remarkable and counter-intuitive states within our mathematical model follows directly from the above discussion. We have seen that $\psi(X, t = 0)$ is in principle completely unconstrained. An intriguing question, then, is how one might arrange for a classical state to transform into one of these remarkable states. Below we discuss a modified momentum squeezer that can make this transformation in an ideal system. We also discuss how fiber losses generally prevent such extreme non-classical states.

3.2 Unperturbed soliton evolution

The unperturbed soliton evolution can be understood without the explicit use of a photon-configuration-space theory [35, 44, 19]. The conservation of momentum and number, and the spreading of their conjugate variables, position and phase, can be calculated in a number of ways. For example, the linearized evolution of the collective variable projections gives [35],

$$\begin{aligned}
\langle \Delta \hat{p}^2(t) \rangle &= \langle \Delta \hat{p}^2(0) \rangle & (3.26)\\
\langle \Delta \hat{x}^2(t) \rangle &= \langle \Delta \hat{x}^2(0) \rangle + 4\langle \Delta \hat{p}^2(0) \rangle t^2 & (3.27)\\
\langle \Delta \hat{n}^2(t) \rangle &= \langle \Delta \hat{n}^2(0) \rangle & (3.28)\\
\langle \Delta \hat{\theta}^2(t) \rangle &= \langle \Delta \hat{\theta}^2(0) \rangle + \frac{n_0^2|c|^4}{4}\langle \Delta \hat{n}^2(0) \rangle t^2. & (3.29)
\end{aligned}$$

The projections give us equations with the form of free-particle dynamics for small quantum deviations. However, we have seen above that the position evolution can be separated *exactly*, not a result of the linearization. In [32], Hagelstein derived free-particle dynamics for position and phase. The configuration-space calculation assumed only the Hartree approximation and large-n. Thus an ideal, unperturbed soliton would show position and phase spreading even in the long-time limit, when the uncertainties become fairly large.

We would like to emphasize that photon configuration-space theory is not a different set of models for quantum optics problems. Instead it is a different way of handling the same models. A CS Hamiltonian is generally derived from, and mathematically equivalent to, some second-quantized model. In this section, we discuss the results of [32] in this light. We show that once intuition is taken from the "particle picture" of solitons, the mathematical results can be obtained using strictly second-quantized methods.

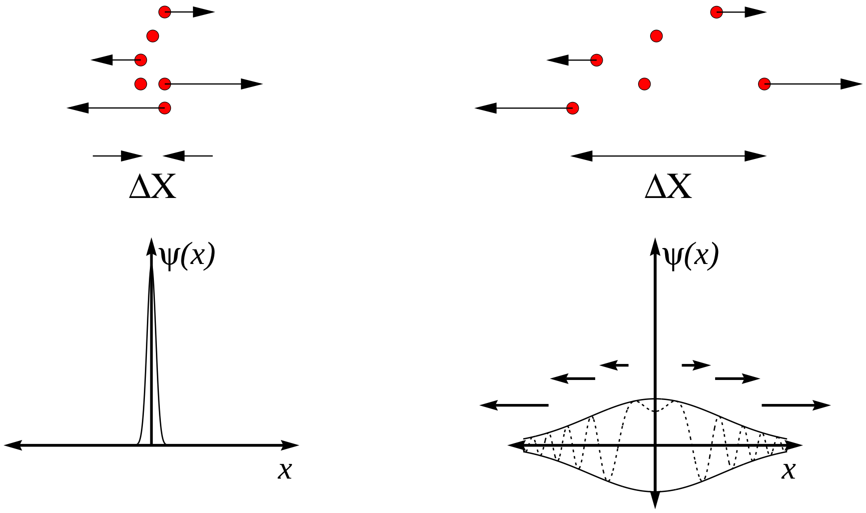


Figure 3-3: Free particle dynamics are depicted here for the classical case of a ballistically expanding gas and for the quantum case of a Gaussian wave packet. The real part (dotted) and magnitude (solid) of the wavefunction are shown. The evolution of soliton position in a nonlinear, dispersive fiber can be understood in terms of these simple pictures, despite their complicated appearance in second-quantized language.

3.2.1 Position Dynamics

Despite all of the mathematics, the basic soliton evolution is extraordinarily simple. The basic intuition is that the soliton moves as a free-particle, a particle with no forces acting on it. If the momentum of the particle is initially uncertain, its position will spread in a ballistic fashion, like initially confined particles in a gas, or a group of joggers running at constant but different speeds. This simple concept is depicted in Figure 3-3.

To derive ballistic dynamics, we need nothing more than the Heisenberg equation of motion:

$$i\hbar\frac{d}{dt}\hat{q} = [\hat{q}, H], \tag{3.30}$$

applied to position and momentum. The second quantized definitions are equivalent to Eqs. (3.7) and (3.24):

$$\hat{N} = \int dx\hat{\phi}^\dagger(x)\hat{\phi}(x) \tag{3.31}$$

$$
\begin{aligned}
\hat{X} &= \hat{N}^{-1}\int dx\hat{\phi}^{\dagger}(x)x\hat{\phi}(x) && (3.32)\\
\hat{P} &= -i\hbar\int dx\hat{\phi}^{\dagger}(x)\frac{\partial}{\partial x}\hat{\phi}(x) && (3.33)
\end{aligned}
$$

Thus X and P are canonically conjugate $[\hat{X},\hat{P}]=i\hbar$, and have dynamics,

$$
\begin{aligned}
\frac{d}{dt}\hat{X} &= \frac{\omega''}{\hbar}\hat{N}^{-1}\hat{P} && (3.34)\\
\frac{d}{dt}\hat{P} &= 0 && (3.35)
\end{aligned}
$$

Spreading is obtained by solving the above system, in terms of the initial values (subscripted with 0):

$$\hat{X}(t)=\hat{X}_0+\frac{\omega''}{\hbar}\hat{N}_0^{-1}\hat{P}_0t \tag{3.36}$$

$$\langle\hat{X}^2(t)\rangle=\langle\hat{X}_0^2\rangle+\frac{\omega''}{\hbar}\hat{N}_0^{-1}\langle\hat{X}_0\hat{P}_0+\hat{P}_0\hat{X}_0\rangle t+(\omega''/\hbar)^2\langle(\hat{N}_0^{-1}\hat{P}_0)^2\rangle t^2 \tag{3.37}$$

This agrees with the linearized result (3.27) for the appropriate normalization and initial conditions.

3.2.2 Phase dynamics

Number and phase behave more-or-less like momentum and position. However, the analysis of phase is complicated by a number of issues. Phase is not a well-defined quantum variable. Even for a one-mode system, there are problems in defining the quantum variable canonically conjugate to photon number. This has been recognized for a long time. Further, our problem is multi-mode: we must then identify what mode of the field the phase applies to.

These difficulties require that we look at the large-n limit for a meaningful phase operator. This is what is typically done for single-mode problems. Large-n solitons are also subject to the Hartree approximation, which is very useful in defining and analyzing phase.

The basic argument has three parts. The first is that, within the Hartree approximation, there is a single, well-defined, soliton mode. The quantum phase of the soliton can thus be identified with this mode, the normalized classical soliton $\phi_0(x)$.

The second part is that when the photon number is large, and the phase uncertainty is not too large, one can use the conjugate relation

$$[\Delta\hat{\theta}, \hat{N}] = i \tag{3.38}$$

as an (approximate) defining relation for $\Delta\hat{\theta}$. We will use this commutation relation rather than an explicit expression for θ to derive the dynamics.

Finally, the Hartree approximation [83] gives us the energy of a soliton as a simple function of n, $E_n = \hbar n(n-1)^2\kappa^2/6\omega''$. Within this approximation, we can identify a number-phase part of the Hamiltonian:

$$\hat{H}_{n,\theta}(\hat{N}) = \hbar\hat{N}(\hat{N}-1)^2\kappa^2/6\omega'' \tag{3.39}$$

From Eqs. (3.38) and (3.39), we can immediately write down

$$\frac{d}{dt}\Delta\hat{\theta} = [\Delta\hat{\theta}, \hat{H}_{n,\theta}]/i\hbar = \frac{d}{d\hat{N}}\hat{H}_{n,\theta} \equiv \hat{v}_\theta \tag{3.40}$$

$$\frac{d}{dt}\hat{N} = [\Delta\hat{N}, \hat{H}_{n,\theta}]/i\hbar = 0 \tag{3.41}$$

The right hand side of Eq. (3.40) can now be identified as the velocity of the phase. The velocity $\hat{v}_\theta$ is an algebraic function of the photon number, and is thus conserved. The spreading in phase is then related to the uncertainty in this variable.

$$\langle\Delta\hat{\theta}^2\rangle = \langle(\Delta\hat{\theta}_0 + t\Delta\hat{v}_\theta)^2\rangle \tag{3.42}$$

This derivation did not require the equivalent of a linearization approximation. However, it is not as fundamental as the position dynamics, which follows directly from the Hamiltonian and the definitions of the operators. The identification of the number-phase Hamiltonian (3.39) is specific to soliton states, not a fundamental separation of variables as in Eq. (3.9). Similarly, the resulting ballistic phase spreading is valid apart from a linearized regime, but not for all time. At some point, the phase

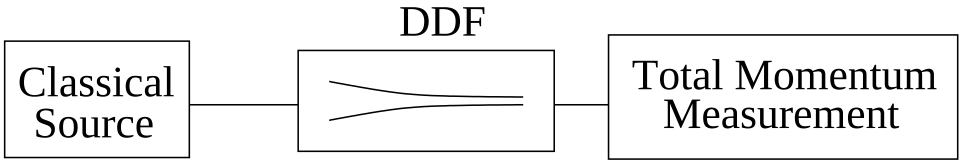


Figure 3-4: Coherent-state pulses are launched into a dispersion-decreasing fiber. The balance of dispersion and nonlinearity is shifted gradually, causing the pulse to narrow. The pulse thus experiences spectral broadening, but experiences negligible change in total momentum (which is conserved). The spread in classical momentum (spectral width) and quantum total momentum are thus incommensurate, and can be detected as an effective squeezing of the total momentum.

uncertainty will approach 2π and the conjugate relationship (3.38) will fail.

3.3 Momentum squeezer

We now consider the schematic measurement of Figure 3-4. The setup for generating squeezing is extremely simple: an input stream of pulses is put through a dispersion-decreasing fiber (DDF). The total momentum of output pulses can be observed using appropriate heterodyne detection [35]. The input pulses are assumed to be in a coherent state, and thus have "normal," shot-noise statistics. The mechanism of squeezing in the DDF is that the spectrum of each pulse is significantly broadened (via pulse compression), while the total momentum of the pulse as a whole is rigorously conserved. The collective momentum uncertainty thus remains at the coherent-state level of the narrow input spectrum, much lower than the coherent-state level of the output spectrum. We now discuss this problem in detail.

3.3.1 Momentum dynamics

As with the standard fiber problem, we solve the evolution by separating the collective and relative degrees of freedom. The evolution of the collective coordinates reveal the conservation of momentum, and the evolution of the relative coordinates will determine the broadening of the classical, or single-photon, pulse spectrum.

We model a dispersion-decreasing fiber using the usual nonlinear fiber model but with a time-varying factor adjusting the dispersion along the length z of the fiber: $\omega''(t = z/v_g) = \gamma(t)\omega_0''$ and $\gamma(0) = 1$. The Hamiltonian is then

$$\hat{H}(t) = \hbar\left[-\gamma(t)\frac{\omega_0''}{2}\int dx\hat{\phi}^\dagger(x)\frac{\partial^2}{\partial x^2}\hat{\phi}(x) + \kappa\int dx\hat{\phi}^\dagger(x)\hat{\phi}^\dagger(x)\hat{\phi}(x)\hat{\phi}(x)\right] \tag{3.43}$$

Some subtleties in relating this model to physical fiber parameters can be found in [49]. The separation of collective coordinates can be performed just as before. The Schrödinger equation for the soliton position is

$$i\hbar\frac{\partial}{\partial t}\psi_n(X) = -\gamma(t)\frac{\hbar\omega_0''}{2n}\frac{d^2}{dX^2}\psi_n(X) \tag{3.44}$$

but in fact, we do not need to solve this equation to find the evolution of total momentum. Defining the momentum by Eq. (3.24) and again using the configuration-space form of the Hamiltonian,

$$\hat{H}_n(t) = \hbar\left[\sum_k \frac{-\gamma(t)\omega_0''}{2}\frac{\hat{p}_k^2}{\hbar^2} + \sum_{j<k} 2\kappa\delta(x_k - x_j)\right], \tag{3.45}$$

we argue that the momentum evolution is trivial. This is because the potential only involves *differences* between the photon positions x_k and x_j. The changes in their momenta are equal and opposite. Mathematically, we have

$$\begin{aligned}\left[\sum_l \hat{p}_l, \delta(x_k - x_j)\right] &= [\hat{p}_k, \delta(x_k - x_j)] + [\hat{p}_j, \delta(x_k - x_j)] && (3.46)\\ &= -i\hbar[\dot{\delta}(x_k - x_j) - \dot{\delta}(x_k - x_j)] = 0. && (3.47)\end{aligned}$$

Thus the total momentum is conserved,

$$[\hat{P}, \hat{H}(t)] = 0 \tag{3.48}$$

and its evolution is trivial, as expected:

$$\langle \Delta \hat{P}^2 \rangle = \langle \Delta \hat{P}^2 \rangle_{t=0}. \tag{3.49}$$

As mentioned above, the shape of a pulse is determined by the relative-coordinate wavefunction, $\Phi_n(x_1 - X, \ldots)$, and essentially mimics the classical solution. Each individual photon lives in a position wavefunction $\phi_0(x - X, t)$. Each also has a momentum distributed according to the Fourier transform of ϕ_0. In a dispersion-decreasing fiber, the single-photon wavefunction ϕ_0 evolves according to the classical equation. The interaction with the other $n - 1$ photons is essentially included as a mean-field [51],

$$\frac{\partial}{\partial t}\phi_0(x,t) = -\gamma(t)\frac{\omega_0''}{2}\frac{\partial^2}{\partial x^2}\phi_0(x,t) + 2\kappa(n-1)|\phi_0(x,t)|^2\phi_0(x,t) \tag{3.50}$$

Just as a standard detector would see the classical soliton pulse shape, a standard spectrometer would see the spectrum arising from this classical evolution.

Pulse propagation in a classical DDF has been studied previously [76, 49, 12]. For example, if a soliton is launched into a fiber with gradual dispersion variations, the pulse remains a soliton of width $x_s(t) = |\omega''(t)|/\kappa n$ consistent with the local fiber parameters, $\omega''(t)$ and $\kappa(t)$, at each point. It is adiabatically compressed with minimal radiation losses. The spectral broadening factor in this case is

$$\langle \Delta \hat{p}^2(t) \rangle / \langle \Delta \hat{p}^2(0) \rangle = (x_s(0)/x_s(t))^2 = \gamma^{-2}(t). \tag{3.51}$$

For non-soliton pulses, spectral broadening can be obtained numerically or experimentally, if necessary. We define the spectral broadening ratio in the general case by

$$g^{-2}(t) \equiv \langle \Delta \hat{p}^2(t) \rangle / \langle \Delta \hat{p}^2(0) \rangle \tag{3.52}$$

3.3.2 Squeezing: sub-shot noise uncertainty

The mechanism of spectral broadening—momentum exchange between the photons of the pulse—conserves the total momentum of the pulse. Each individual photon ends up with a more uncertain momentum than it started, but for every exchange event that red-shifts one photon, the other scattered photon is blue-shifted. The argument is essentially unchanged in the case of non-adiabatic or non-soliton pulse compression. The photon momenta become correlated, as they must if the total momentum uncertainty is to remain fixed. Since the photons are not exchanging momentum with an external reservoir, but rather with each other, coherence as well as conservation is preserved. Naturally, other processes such as loss can lead to decoherence, and will be discussed in Section 4.5.

To quantify the effective "squeezing," we define the squeezing ratio to be the ratio of the momentum uncertainty to the coherent-state or shot-noise level (denoted "coh"),

$$\langle \Delta P^2(t) \rangle_{\rm coh} \equiv n \langle \Delta p^2(t) \rangle. \tag{3.53}$$

Since $\hat{P}$ is unchanged, but classical spectrum is broadened, this ratio drops below unity,

$$S(t) \equiv \frac{\langle \Delta P^2(t) \rangle}{\langle \Delta P^2(t) \rangle_{\rm coh}} = \frac{n \langle \Delta p^2(0) \rangle}{n \langle \Delta p^2(t) \rangle_{\rm coh}} = g^2(t). \tag{3.54}$$

Adiabatic soliton compression gives, for example,

$$S(t) = \gamma^2(t). \tag{3.55}$$

Squeezing of position uncertainty (timing jitter) is clearly also of interest. One might wonder if a dispersion-increasing fiber could lead to pulse expansion and position squeezing, just the reverse of the process described above. But pulse position X is not a conserved quantity, and in fact spreads progressively with time, as discussed above. Position uncertainty grows on a time scale on the order of the soliton period. For adiabatic soliton-compression, this is faster than the broadening of the classical pulse, and so the compression will have little effect. It is unclear whether any simple modification of this strategy could give timing-jitter squeezing.

3.3.3 Momentum measurement

Heterodyne detection of soliton variables, including momentum, has been outlined in [35]. This detection method requires a strong local oscillator with stable timing and frequency compared to the uncertainties of the signal. Setting up this measurement will likely be more difficult than setting up the generation of momentum-squeezed light itself. A number of other strategies could also be pursued to experimentally confirm the correlated-photon states. This is a subject of future study.

In addition to the usual fiber losses, we may want to include detector loss and light coupled to the continuum: that is, the variation of the dispersion is never perfectly adiabatic, and so continuum modes will be excited. In theory, continuum photons can be included in the normal momentum expression (3.24). However, the observed momentum may be different from the total field momentum, depending on the measurement apparatus. For example, if the soliton-quadrature is used [35], the shed continuum should be considered lost from the system. We leave a detailed analysis of the detection system for future work.

3.4 Extreme squeezing and many-photon interference

The degree of momentum squeezing attainable using the above method is ultimately limited by the propagation loss and the specifics of the detector. In the next section, we explore the fundamental limitation imposed by fiber loss. Here we discuss some fascinating properties of the field that would arise in a purely lossless system. These are the qualitatively nonclassical properties of the type of field depicted in Figure 3-2, where the soliton is in a coherent superposition of macroscopically distinguishable states.

Consider a system of n photons squeezed by a ratio greater than $\sqrt{n}$. The first

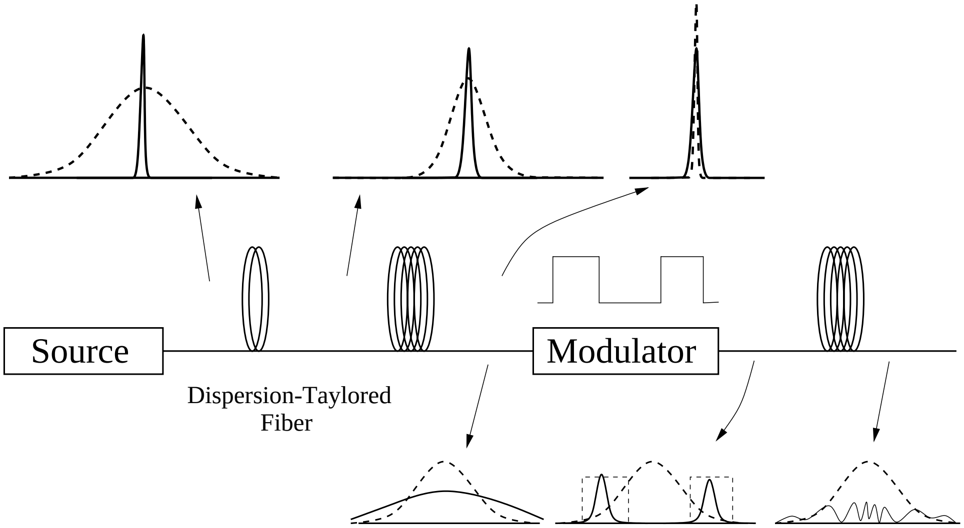


Figure 3-5: A fiber version of two-slit interference is sketched. Quantum wavefunctions $\psi(X)$ (solid lines) and classical pulse shapes (dashed lines) are depicted at several points in a dispersion-tailored fiber system. The transition from moderate to extreme squeezing is shown in the top three axes. For extreme squeezing, possible in an ideal, lossless system, the extent of $\psi(X)$ has actually exceeded the classical pulse width. Pulses are in a quantum superposition of macroscopically distinguishable position states. In the lower three axes (note the change of scale), we imagine that an ideal modulator can project out a two-peaked component of the wavefunction $\psi(X)$, which subsequently displays non-classical “DeBroglie-wave,” interference.

qualitative difference we should recognize is that the quantum position uncertainty is no longer a small perturbation of the field; for this reason, the semiclassical approach of [35], starting with $\hat{\phi} = \phi_0 + \delta\hat{\phi}$, does not apply to this case. The general soliton states discussed above,

$$|\Psi\rangle = \sum_n \int dX \psi_n(X) |\Phi_{n,X}\rangle \tag{3.56}$$

are applicable. This regime might seem counterintuitive since it goes against the semiclassical picture. The individual photons are living in a wavefunction that is centered on an uncertain position. It is helpful to think of an atom analogy, where this kind of picture is more intuitive and familiar: there is nothing strange about imagining an atom with position uncertainty greater than an atomic radius. In this case, the electrons are in a definite bound state about an uncertain center position. This is exactly analogous to the soliton of Fig. 3-2.

3.4.1 Multi-photon two-slit interferometer

Figure 3-5 depicts a thought experiment which takes this reasoning a step further. In the first part, a DDF is used to broaden the soliton by a factor greater than $\sqrt{n}$, so that the pulse is in a quantum superposition of macroscopically distinct position states. In the second part, the pulse passes through a time-gated analog of a two-slit interferometer.

The "double slit" can be thought of as an ideal absorption modulator. This modulator presents a time-varying absorption with two completely lossless "pass" windows. Like a conventional double slit, the modulator blocks the pulse most of the time—whenever it does not line up with a pass window. With some moderately small probability, the pulse is positioned correctly to pass through the two time windows. The output wavefunction $\psi_{n,X}$ is thus projected onto a two-peaked state representing the component that is not blocked. The two-peaked distribution can then evolve in another (variable dispersion) fiber, giving rise to interference of the collective wavefunction $\psi(X)$.

If such a system were possible, it would have remarkable properties. By manipulating the many-photon pulse as a whole, and not the individual photons, we would obtain a qualitatively different kind of interference from linear interferometers. The characteristic wavelength would be the photonic De Broglie wavelength λ/n, corresponding to the *total* momentum of the soliton, [65, 32, 22].

Clearly there are practical difficulties with this setup; for example the proposed "ideal modulator" is impossible to fabricate. A realistic design would likely look quite different from the above sketch, and will need to overcome formidable barriers such as loss decoherence; the reason why these thought experiments run against our intuition is because we live in a world where optical nonlinearities are generally weak compared to various decoherence mechanisms. We do not attempt such a realistic design here. A successful design will need to incorporate more sophisticated modulators as well as detectors.

Chapter 4

Loss and Soliton propagation

Loss involves exchange of energy and momentum with an external reservoir, so that these quantities are no longer conserved. This leads to decoherence, and has important implications on interference, as we will discuss.

On a more basic level, the change in photon number seems difficult to reconcile with our general approach, of describing the field using a multi-photon wavefunction. In our configuration-space point of view, loss of even a single photon means that we must somehow connect an initial n-photon wavefunction with a final $(n-1)$-photon wavefunction, which exists in a different Hilbert space altogether. Since loss is significant in any practical device of interest, the usefulness of the configuration-space approach requires some resolution of this issue.

The resolution can be found in [21, 32], and has a simple, high-level message: take a photon out of a large-n soliton, and it is still a soliton with only a small change in its collective coordinates. This is justified by calculating overlap integrals in the large-n, Hartree approximation. It means that collective wavefunctions are meaningful even though the total evolution deviates from Eq. 3.1, and the relative wavefunction transitions across photon number sectors.

4.1 Pictures of Loss: second- and first-quantized descriptions

Once considering this problem for fundamental reasons, we discovered an alternative and quite different physical picture corresponding to the important Gordon-Haus effect. This is an example of how the configuration-space theory, though mathematically equivalent to the second-quantized version, results in different explanations and pictures, and provides new insight. For the bulk of this chapter, we discuss the calculation of Gordon-Haus jitter using the configuration-space loss formalism. This tests the new formalism on a previously-solved problem, and allows us to develop our intuition.

4.1.1 Standard Models

The standard quantum model for loss couples a system to a reservoir with a continuum of oscillators [16]. If the system of interest is called A and the reservoir B, then the Hamiltonian is written

$$\hat{H} = \hat{H}_{\mathrm{A}} + \hat{H}_{\mathrm{B}} + \hat{H}_{\mathrm{I}}, \tag{4.1}$$

where I denotes the interaction, typically a simple one-photon exchange, and the reservoir has a large number of closely-spaced energies

$$\hat{H}_{\mathrm{B}} = \sum_k \hbar\omega_k \hat{d}_k^\dagger \hat{d}_k. \tag{4.2}$$

The spacing $\omega_{k+1} - \omega_k$ goes to zero in the continuum limit. This kind of model allows us to derive the evolution of the system's state from basic quantum equations.

A simpler but quite useful model is the "beamsplitter model" of loss is depicted in Figure 4-1. This model allows us to visualize important features of general loss interactions using a simple, intuitive special case. For example, when light is "lost" from system A, it invariably goes somewhere. There is some mode or quantum system into which the photons are coupled. When this mode is depicted as the output of a

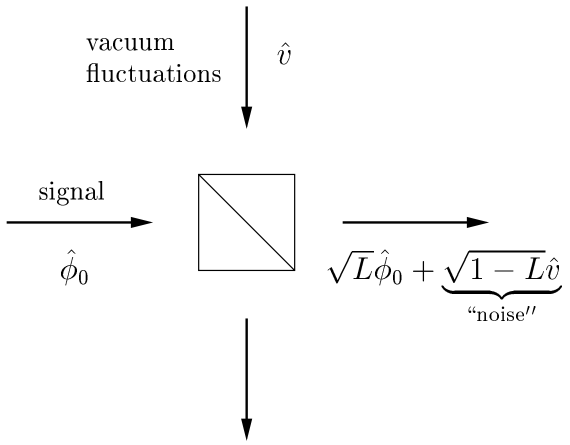


Figure 4-1: Quantum mechanically loss represents coupling of energy into an external reservoir. In the simple "beamsplitter" representation, the mode of the reservoir receiving the energy is depicted as the port of a beamsplitter. The quantum fluctuations (for example, vacuum fluctuations) initially present in the reservoir couple into the system, acting as a source of noise. This simple model can be used to obtain quantitative limits on loss-induced noise.

beamsplitter, it is immediately clear that this other mode also couples into system A. This is one explanation why loss is inevitably accompanied by noise: whenever A is coupled to a reservoir mode B, the quantum noise in B is also coupled into A [33, 36].

From this simple model, one can derive quantitative noise limits for a single mode system . If the power loss factor is $L \equiv \cos^2(\theta)$, then the transformation of photon operators $\hat{a}$ (system A) and $\hat{b}$ (reservoir B) is

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{bmatrix} \cos(\theta) & i\sin(\theta) \\ i\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{pmatrix} \hat{a}_0 \\ \hat{b}_0 \end{pmatrix} \tag{4.3}$$

Uncertainties are obtained simply by plugging in the above [34]. For example, the statistics of the output power $\propto \hat{a}^\dagger\hat{a}$ can be derived assuming the reservoir is in the ground state:

$$\langle \hat{a}^\dagger \hat{a} \rangle = L\langle \hat{a}_0^\dagger \hat{a}_0 \rangle \tag{4.4}$$

$$\langle (\hat{a}^\dagger \hat{a})^2 \rangle = L^2 \langle (\hat{a}_0^\dagger \hat{a}_0)^2 \rangle + (1-L)\langle \hat{a}^\dagger \hat{a} \rangle. \tag{4.5}$$

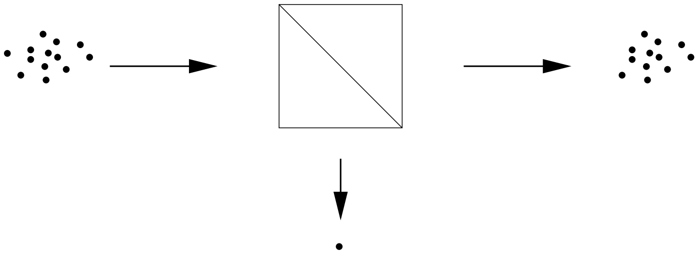


Figure 4-2: In the particle picture of loss, vacuum fluctuations do not appear explicitly. Instead the quantum nature of the field appears in the discrete energy quanta. Uncertainties in the outputs of a beamsplitter are understood in terms of photon-counting statistics.

The uncertainty in number is then

$$\langle \Delta^2 \hat{a}^\dagger \hat{a} \rangle = L^2 \langle \Delta^2 \hat{a}_0^\dagger \hat{a}_0 \rangle + (1 - L) \langle \hat{a}^\dagger \hat{a} \rangle \tag{4.6}$$

The $1 - L$ term is the addition of vacuum fluctuations coupled in from the reservoir.

The second-quantized calculation of uncertainties for more complicated systems is actually quite similar. The coherent-state soliton uncertainties [35] and Gordon-Haus noise [29] are traditionally calculated by projecting vacuum fluctuations onto the modes of interest.

4.1.2 A new picture

The above description of loss-induced quantum noise is intuitive in certain respects: it allows us to treat vacuum fluctuations, a familiar if not transparent term, just like any classical noise source. But where do vacuum fluctuations come from, ultimately? They unavoidably arise when we try to formalize a field that has discrete energy quanta. Interestingly, we have found that loss-induced noise can be derived directly from the discreteness of the energy exchange [21]. The basic intuition follows from Figure 4-2. In this particle picture of loss, vacuum fluctuations do not appear explicitly, but their effect is reproduced exactly. This is another example of a wave-particle duality in quantum mechanics. Figures 4-1 and 4-2 are both correct, and describe the same thing, but in a very different way. For example, the quantum uncertainty in output power in the particle picture is a matter of photon-counting statistics. The mechanism for loss-induced amplitude noise (calculated above) is clear: individual photons randomly choose between the outputs of the beamsplitter. We now go through the details of the Gordon-Haus calculation.

4.2 Gordon-Haus noise calculation

We now present the analysis that shows that the particle picture, discussed above, describes the familiar loss-induced quantum noise. Loss-induced noise results from a series of single-photon loss events. We begin by analyzing a single event perturbatively, calculating the disturbance from the interaction with matter given by the Hamiltonian (4.9). We then combine the effects of many events to derive the total momentum noise.

The uncertainty of a soliton's center frequency is the noise most relevant to the Gordon-Haus effect. Since the center frequency and the soliton's total energy are intimately related, we present our analysis in terms of energy variables and energy exchange. In this way, we can utilize fundamental conservation properties. Once we have calculated the energy quanta exchanged with the loss reservoir, we can easily relate energy to whatever related variables we want: center frequency, soliton momentum, and group velocity.

One of the larger goals of this thesis is to develop useful applications of a photon configuration-space, or particle, point of view. In this section, our analysis clearly draws on particle intuition. However, we try to use primarily the more familiar second-quantized notation for the calculation, so that our results are clear to those otherwise unfamiliar with photon wavefunctions.

4.2.1 Outline of the calculation

The simple and familiar technique of first-order perturbation theory captures single-photon processes without adding unnecessary complexity. Perturbation theory is well suited to the analysis of low-loss fibers: The soliton propagates undisturbed except for an occasional single-photon exchange with an interaction site.

We calculate the first-order state evolution given by the Hamiltonian (4.9). The field is described by the soliton Hamiltonian $\hat{H}_A$, essentially the same as Eq. (3.1).

Here we include the missing terms in the dispersion relation

$$\omega(k) = \omega(k_0) + \omega'(k_0)(k - k_0) + \omega''(k_0)(k - k_0)^2/2. \tag{4.7}$$

to emphasize the group velocity of the pulse $v_{g0} \approx \omega'(k_0)$. Roughly speaking, the group velocity is a conversion factor between energy and momentum perturbations, $\hbar\Delta\omega \approx v_{g0}\hbar\Delta k$. The field Hamiltonian is then,

$$\begin{aligned} \hat{H}_A &= \hbar \int dx \hat{\phi}^\dagger(x) \left[\omega_0 - i\omega' \frac{\partial}{\partial x} - \frac{\omega''}{2} \frac{\partial^2}{\partial x^2} \right] \hat{\phi}(x) \\ &\quad + c \int dx \hat{\phi}^\dagger(x) \hat{\phi}^\dagger(x) \hat{\phi}(x) \hat{\phi}(x). \end{aligned} \tag{4.8}$$

The interaction is an electronic transition described by [21],

$$\hat{H}_I = -\int \hat{\jmath} \cdot \hat{A} \ d^3\vec{r} \ \ \approx \ \ V(\hat{b}^\dagger + \hat{b})[\hat{\phi}^\dagger(x) + \hat{\phi}(x)], \tag{4.9}$$

where x is the position of the loss site. The reservoir could be described, for example, by Eq. (4.2) in a limit of many closely-spaced energy levels. The details of the reservoir determine the relation between the operator $\hat{b}$ involved in the interaction and operators $\hat{d}_k$ of the natural reservoir excitations. This is discussed further below and in [16].

Our initial state is a direct product of soliton and matter states

$$|\Psi_i\rangle = |\Phi_n\rangle_{\text{soliton}} \otimes |\beta\rangle_{\text{matter}}. \tag{4.10}$$

We proceed via the standard perturbative state expansion. This is written in the interaction picture as

$$|\Psi_f\rangle = |\Psi_i\rangle + \frac{1}{i\hbar} \int_{t_i}^{t_f} d\tau \tilde{H}_I |\Psi_i\rangle + \ldots. \tag{4.11}$$

We can rewrite this state in terms of the Schrödinger-picture interaction operator,

$$|\Psi_f\rangle = |\Psi_i\rangle + \frac{1}{i\hbar}\int_{t_i}^{t_f} d\tau e^{-i\hat{H}_0(t_i-\tau)/\hbar}\hat{H}_I e^{-i\hat{H}_0(\tau-t_i)/\hbar}|\Psi_i\rangle. \tag{4.12}$$

The basic effects of photon exchange on the soliton can be studied in a systematic way from this starting point.

Consistent with a first-order, perturbative approach, the interaction probability $\sin^2(\theta)$ will be small for any particular interaction sight. Loss of many photons occurs through the combined effects of many interaction sites.

Uncertain exchanged energy The basic idea is to look at the perturbed component $|\Psi_1\rangle$ of the state,

$$|\Psi_f\rangle \approx |\Psi_i\rangle + \theta|\Psi_1\rangle \tag{4.13}$$

and determine the properties of the exchanged energy ϵ. Naturally this exchanged energy is uncertain. It is exactly this uncertainty that constitutes the noise added to the soliton by the interaction. In our calculation, below, we define a "spectrum" $T(\epsilon)$ for ϵ. This spectrum gives us the uncertainty of each exchanged photon, and is calculated directly from $|\Psi_1\rangle$.

In order to connect with the soliton center frequency (which relates to Gordon-Haus timing jitter) we define the operator,

$$\hat{h} \equiv \hat{H}_{\mathrm{A}}\hat{N}^{-1}. \tag{4.14}$$

This is field energy per photon. For a given photon number, this includes the constant binding energy of the soliton and the dispersive part related to soliton frequency. Naturally, a change in this variable $\Delta\hat{h}$ can be understood intuitively as a shift in momentum or carrier frequency (essentially, $\Delta\omega = \Delta h/\hbar$). We discuss these related variables in Appendix B.

The intuition behind the calculation is fairly simple: If the total energy E_A is perturbed by an exchange ϵ,

$$E_{Af} = E_{Ai} - \epsilon, \tag{4.15}$$

then the perturbation to $\hat{h}$ can be calculated simply by accounting for the factor $1/n$. We might intuitively expect the final uncertainty added to $\hat{h}$, that is the noise-per-photon, to look like

$$\delta\langle\Delta\hat{h}^2\rangle \sim \frac{\langle\Delta\epsilon^2\rangle}{n^2}. \tag{4.16}$$

This is almost exactly what we get from the full analysis.

4.2.2 Spectrum of lost photons

All solitons can be constructed from the fundamental soliton states of Lai and Haus [51, 43]. These are states of definite number n, energy E_A, and momentum $P(E_A, n)$, written $|n, E_A\rangle$. At the heart of loss and gain effects, there is the problem of scattering between initial and final fundamental soliton states. Once this scattering of fundamental states is understood, the evolution of general soliton states follows by superposition.

We now derive the evolution of the state $|n, E_{Ai}\rangle$ into an $(n-1)$-photon state. In particular, we find the scattering amplitude between $|n, E_{Ai}\rangle$ and $|n-1, E_{Af}\rangle$ applicable to a loss site. In the language of [57, Chapter XVII], this is nothing more than $\langle b|U^{(1)}|a\rangle$. Throughout this chapter, we restrict ourselves to initial states of definite photon number for simplicity.

We start with an energy eigenstate of both the field and reservoir,

$$|\Psi_i\rangle = |n, E_{Ai}\rangle \otimes |E_{Bi}\rangle \tag{4.17}$$

It is convenient to use a normalized form of the first-order perturbed state (4.12). We can accomplish this simply by using

$$|\Psi_f\rangle = \cos(\theta)|\Psi_i\rangle + \sin(\theta)|\Psi_1\rangle \tag{4.18}$$

where the perturbed state component is

$$\begin{aligned}\sin(\theta)|\Psi_1\rangle &= \frac{1}{i\hbar}\int_{t_i}^{t_f} d\tau e^{-i\hat{H}_0(t_i-\tau)/\hbar}\hat{H}_I e^{-i(E_{Ai}+E_{Bi})(\tau-t_i)}|\Psi_i\rangle && (4.19)\\ &= \frac{1}{i\hbar}\int_{t_i}^{t_f} d\tau e^{-i(\hat{H}_0-E_i)(t_i-\tau)/\hbar}\hat{H}_I|\Psi_i\rangle && (4.20)\end{aligned}$$

To proceed we must evaluate $\hat{H}_I|\Psi_i\rangle$. The integration in time will then select out final states which conserve total energy. Specifically, for an absorption site $(\hat{b}|E_{Bi}\rangle = 0)$,

$$\hat{H}_I|\Psi_i\rangle = V\left(\hat{\phi}(x)|n, E_{Ai}\rangle\right) \otimes \left(\hat{b}^\dagger|E_{Bi}\rangle\right) \tag{4.21}$$

These expressions can be evaluated using standard second-quantized methods, as follows.

Reservoir excitation We consider a loss reservoir with a characteristic resonance ω_l and decay rate Γ. This can arise, for example, if the atomic transition with lowering operator $\hat{b}$ coupled to a continuum with operators $\hat{c}_k$. The model

$$\hat{H}_B = \hbar\left[\omega_l\hat{b}^\dagger\hat{b} + \sum_k \omega'_k\hat{c}_k^\dagger\hat{c}_k + \sum_k \Omega_k(\hat{b}^\dagger\hat{c}_k + \hat{c}_k^\dagger\hat{b})\right], \tag{4.22}$$

has been solved in [16]. The diagonalization of $\hat{H}_B$ gives

$$\hat{b} = \lim_{\Delta\omega\to 0}\sum_k\left[\frac{\Gamma\Delta\omega/2\pi}{(\Gamma/2)^2 + (\omega'_k)^2}\right]^{1/2}\hat{d}_k \tag{4.23}$$

where $\Delta\omega$ is the spacing of the ω'_k's. This is one way of deriving the lifetime of an atomic excitation,

$$\langle\hat{b}^\dagger\hat{b}\rangle = e^{-\Gamma t}. \tag{4.24}$$

Whether we use this or another reservoir model, we can expand the perturbed state in energy eigenstates of the bath; that is, since the energy states are complete, there is some function β such that

$$\hat{b}^\dagger|E_{Bi}\rangle \equiv \int dE_{Bf}\beta(E_{Bf} - E_{Bi})|E_{Bf}\rangle \tag{4.25}$$

For the reservoir model described above, $|\beta(\Delta E)|^2$ is the often-calculated Lorentzian spectrum of the exponential atomic decay. In any case, β contains all information about the reservoir relevant to our problem.

Gain will be treated with essentially the same model. Gain is achieved when the atom is held in an excited state by a fully excited bath, $\langle \hat{c}_k^\dagger \hat{c}_k\rangle = 1$. The essential effects which we discuss in this chapter are the same for loss and gain. Rather than derive our results twice, we provide a full analysis of loss and argue in Section 4.2.5 that gain calculations are similar.

Photon annihilation The calculation of $\hat{\phi}(x)|n, E_{Ai}\rangle$ is more complicated. It represents the state of the field after a photon is removed at a particular position. Naturally, the field has many degrees of freedom, and so an exact calculation may be complicated.

Intuitively, we expect that removing only one photon out of many leaves us with a slightly perturbed soliton. If this is correct, then we can expand the perturbed state in fundamental soliton states. In fact, it is clear that the photon number is $n-1$. The final state can then be constructed as a superposition of states $|n-1, E_{Af}\rangle$; that is, we can find some function ϕ such that

$$\hat{\phi}(x)|n, E_{Ai}\rangle \approx \int dE_{Af}\phi(E_{Ai} - E_{Af})|n-1, E_{Af}\rangle. \tag{4.26}$$

The orthogonality of fundamental soliton states along with this expansion implies

$$\phi(E_{Ai} - E_{Af}) = \langle n-1, E_{Af}|\hat{\phi}(x)|n, E_{Ai}\rangle. \tag{4.27}$$

This intuitive leap actually represents a very important argument. It is not obvious *a priori* that the above expansion will work. It includes all soliton states, since the fundamental solitons $|n', E'\rangle$ span this set. But it does not include states with an excited continuum, since these are not pure soliton states. For the moment, we proceed with the calculation assuming that the above expression is correct. In section 4.2.4, we discuss the important approximation more carefully.

In that section, we also calculate the relevant matrix element,

$$\phi(\epsilon) = \phi_0(\epsilon) \tag{4.28}$$

where ϕ_0 is the classical soliton spectral amplitude. Here again, details of the calculation are somewhat tedious, but the end result is very intuitive: immediately after losing a photon, the field energy must be shifted by an amount $\epsilon = \hbar\omega$ corresponding to the lost photon. Since the individual photons of the soliton live within the classical spectrum of the soliton, the amplitude $\phi(\epsilon)$ simply mimics this spectrum. The classical spectrum is well known, and leads to easy calculation of relevant uncertainties.

Energy conservation

Energy exchanged in the interaction must be both lost from system A and absorbed by system B. We have just characterized the energy spectra of these two operations by $\phi(\epsilon)$ and $\beta(\epsilon)$. We now combine these in the expression (4.20) to obtain the spectrum of exchanged photons.

$$\begin{aligned} &\sin(\theta)|\Psi_1\rangle = \\ &\quad \frac{V}{i\hbar}\int dE_{Af}\int dE_{Bf}\beta(E_{Bf}-E_{Bi})\phi(E_{Ai}-E_{Af}) \\ &\qquad \int_{t_i}^{t_f} d\tau e^{-i(E_{Af}+E_{Bf}-E_{Ai}-E_{Bi})(t_i-\tau)/\hbar}|n-1, E_{Af}\rangle \otimes |E_{Bf}\rangle \end{aligned} \tag{4.29}$$

The integration in time produces a $\delta[E_{Af}+E_{Bf}-(E_{Ai}+E_{Bi})]$, enforcing energy conservation and leaving us eventually with a single integral,

$$\sin(\theta)|\Psi_1\rangle = \int d\epsilon T(\epsilon)|n-1, E_{Ai}-\epsilon\rangle \otimes |E_{Bi}+\epsilon\rangle. \tag{4.30}$$

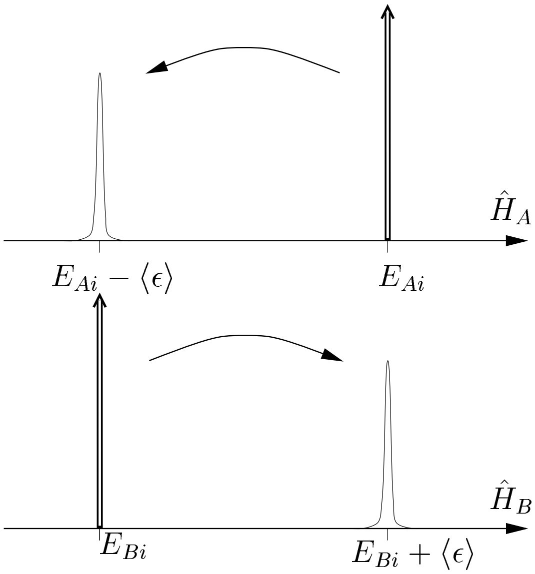


Figure 4-3: Systems A and B, initially in energy eigenstates, exchange an uncertain amount of energy. The individual energies, $\hat{H}_A$ and $\hat{H}_B$, have equal spreads in the final state. The total energy $\hat{H}$ is conserved and therefore still definite even after the interaction.

Here the spectrum of *exchanged* photons T is defined as

$$T(\epsilon) = -2\pi i V \phi(\epsilon)\beta(\epsilon). \tag{4.31}$$

This form reflects our common-sense statement above: since the photon must be both lost from A and absorbed by B, the rate of exchange is related to the product of the availability of photons at a given energy, $\phi(\epsilon)$, times the absorption rate of the reservoir at that energy β. We develop this notion further below.

The probability of exchange at this loss site is

$$\text{Exchange Probability} = \sin^2(\theta) = \int d\epsilon |T(\epsilon)|^2. \tag{4.32}$$

We can also think of $T(E_{Ai}, E_{Ai} - \epsilon)$ as a scattering amplitude between fundamental solitons. The transition from $|n, E_{Ai}, E_{Bi}\rangle$ to the perturbed state is depicted in Figure 4-3.

Once we have obtained T, we can compute whatever we want involving $\hat{H}_A$. In the broadband case of practical interest, $T(\epsilon)$ mimics the sech shape of the classical soliton spectrum, as discussed in the following sections.

4.2.3 Energy “Kicks”

Exchange of energy between the field and matter leads to increasing uncertainty in the soliton energy, even if loss and gain are in balance. Having obtained a convenient expression for the perturbed state, we can now quantify the “kick” of energy uncertainty given by each exchanged photon. In particular, we are interested in the disturbance to the energy-per-photon $\hat{h}$, as the energy is perturbed,

$$E_A \longrightarrow E_A - \epsilon. \tag{4.33}$$

The result is surprisingly simple and intuitive, taking the form,

$$\delta \left\langle \Delta \hat{h}^2 \right\rangle = \frac{\langle \Psi_1 | \Delta \epsilon^2 | \Psi_1 \rangle}{(n-1)^2} \tag{4.34}$$

where ϵ is the energy of the exchanged photon, as discussed below. The expectation on the state $|\Psi_1\rangle$ is evaluated as an average over the exchange spectrum $T(\epsilon)$ calculated in the previous section. The deviation $\Delta\epsilon$ is defined with respect to the initial average energy-per-photon, $\Delta\epsilon = \epsilon - E_A/\hbar$.

Example: average total energy

We start with a simple example of calculating expectations using $|\Psi_1\rangle$ and $T(\epsilon)$. We will use the results of the previous sections, which apply to initial energy eigenstates. The extension to the general case is discussed in Section 4.4.1, but the results relevant to energy kicks are identical.

Suppose that we would like the mean shift in field energy,

$$\langle \hat{H}_A \rangle_f - \langle \hat{H}_A \rangle_i.$$

For the initial state, we have $\langle \hat{H}_A \rangle_i = E_{Ai}$. The final state has contributions from the unperturbed and perturbed state components,

$$\langle \hat{H}_A \rangle_f = \cos^2(\theta) \langle \Psi_i | \hat{H}_A | \Psi_i \rangle + \sin^2(\theta) \langle \Psi_1 | \hat{H}_A | \Psi_1 \rangle \tag{4.35}$$

Cross terms between the two vanish since $|\Psi_i\rangle$ and $|\Psi_1\rangle$ have different photon numbers. Using Eq. (4.30), we find

$$\sin^2(\theta) \langle \Psi_1 | \hat{H}_A | \Psi_1 \rangle = \int d\epsilon |T(\epsilon)|^2 (E_{Ai} - \epsilon) \tag{4.36}$$

If we recall the exchange probability relation (4.32), we obtain

$$\begin{aligned} \langle \hat{H}_A \rangle_f - \langle \hat{H}_A \rangle_i &= \sin^2(\theta) \left(\langle \Psi_1 | \hat{H}_A | \Psi_1 \rangle - E_{Ai} \right) && (4.37) \\ &= -\int \epsilon |T(\epsilon)|^2 d\epsilon. && (4.38) \end{aligned}$$

Built into this expression are the probability of a photon exchange, Eq. (4.32), and the average energy perturbation per photon exchanged:

$$\langle \Psi_1 | \epsilon | \Psi_1 \rangle \equiv -\frac{\int \epsilon |T(\epsilon)|^2 d\epsilon}{\int |T(\epsilon)|^2 d\epsilon} = \frac{\langle \hat{H}_A \rangle_f - \langle \hat{H}_A \rangle_i}{\sin^2(\theta)}. \tag{4.39}$$

The noise added to the total soliton energy is similarly obtained from T,

$$\langle \Delta \hat{H}_A^2 \rangle_f - \langle \Delta \hat{H}_A^2 \rangle_i = \int \epsilon^2 |T(\epsilon)|^2 d\epsilon, \tag{4.40}$$

However, the total soliton energy is less directly related to the soliton velocity than the energy-per-photon $\hat{h}$. We now move on to the perturbation of this variable.

The Kick

We write the initial energy eigenstate as $E_{Ai} = n h_i$, so that

$$\langle \Delta \hat{h}^2 \rangle_i = \langle (\hat{h} - h_i)^2 \rangle_i = 0 \tag{4.41}$$

Similarly, the final state contributes uncertainty only from its perturbed component $|\Psi_1\rangle$,

$$\langle\Delta\hat{h}^2\rangle_f - \langle\Delta\hat{h}^2\rangle_i = \sin^2(\theta)\langle\Psi_1|(\hat{h} - h_i)^2|\Psi_1\rangle. \tag{4.42}$$

Plugging in the perturbed state (4.30), we have

$$\begin{aligned}
\langle\Delta\hat{h}^2\rangle_f - \langle\Delta\hat{h}^2\rangle_i &= \int d\epsilon |T(\epsilon)|^2 \left(\frac{nh_i - \epsilon}{n-1} - h_i\right)^2 && (4.43)\\
&= \int d\epsilon |T(\epsilon)|^2 \left(\frac{h_i - \epsilon}{n-1}\right)^2 && (4.44)
\end{aligned}$$

If we define $\Delta\epsilon = \epsilon - h_i$, we can write the relevant noise formally as

$$\langle\Delta\hat{h}^2\rangle_f - \langle\Delta\hat{h}^2\rangle_i = \sin^2(\theta)\frac{\langle\Psi_1|\Delta\epsilon^2|\Psi_1\rangle}{(n-1)^2}. \tag{4.45}$$

Significance of perturbed energy

In the following sections, we will see that the disturbance (4.45) is all we need to calculate the Gordon-Haus effect. It gives us the disturbance of each photon exchange,

$$[\delta\langle\Delta\hat{h}^2\rangle]_{\text{single-photon}} = \frac{\langle\Psi_1|\Delta\epsilon^2|\Psi_1\rangle}{(n-1)^2} \tag{4.46}$$

which can be combined to give the net effect. In deriving this expression, we have formalized something which is entirely obvious from the particle picture, Figure 4-4. The disturbance is just the effect of algebraically removing one contribution to the average energy. The underlying cause of noise is then the *graininess* of the loss of energy. This is also depicted in the simple picture of Figure 4-2 discussed at the beginning of this chapter.

The soliton is disturbed by the random energy offsets of individual exchanged photons. Ultimately, the "spectrum," $|T|^2$, of the exchanged energy is a simple product of the spectra of the matter system and the soliton, depicted in Figure 4-5. This has the simple interpretation that the probability of ϵ being exchanged is just the probability that such a quantum is available times the loss rate at that energy.

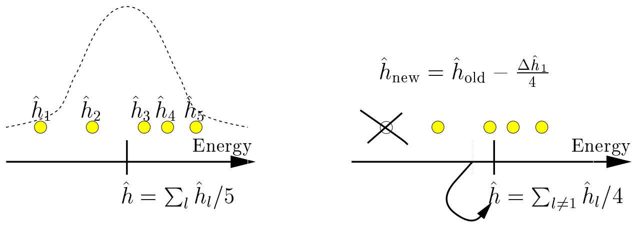


Figure 4-4: The configuration space approach describes a soliton as a collection of local photon particles. In the above particle picture of loss, individual photon energies are initially in a Hartree (energy-domain) wavefunction. The effect of a photon annihilation on the average energy $\hat{h}$ is a simple matter of algebra.

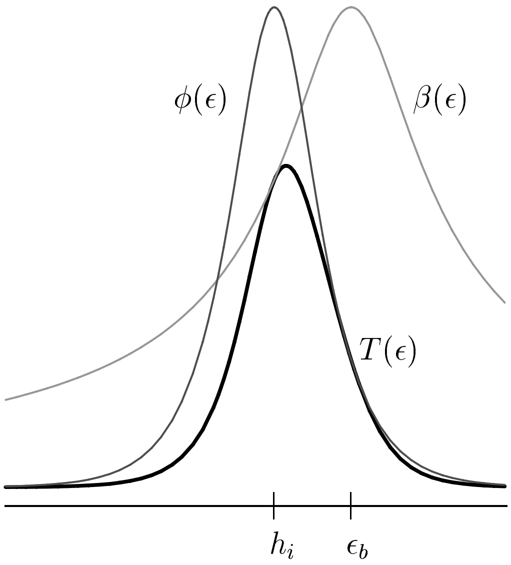


Figure 4-5: The transition spectrum is simply a product of β and ϕ. The number of photons transferred is thus the available density times the absorption.

The mechanism of noise described above might initially seem different from that of the beamsplitter model, Figure 4-1. There, loss (and gain) inevitably caused coupling of quantum fluctuations of external modes to the field. While the mechanisms depicted in Figures 4-1 and 4-4 may *seem* different, we recall that quantum fluctuations of a field are a manifestation of the discreteness of its excitations. A second-quantization approach naturally emphasizes field fluctuations that are only implicitly present in a first-quantized model. In first-quantization, vacuum fluctuations and ASE are implicitly included once we write down a wavefunction for integer numbers of particles. Ultimately, first and second quantization present different tools for dealing with exactly the same physics.

4.2.4 A closer look at the loss spectra

So far we have analyzed the high level argument taking us from a basic model of interaction to the per-photon soliton frequency noise. A number of subtle but important issues have been glossed over. In this section we return and take a closer look at these details. We discuss mathematical approximations and their impact on the physics.

State decomposition

The perturbed field state,

$$\hat{\phi}(x)|n, E_{Ai}\rangle \tag{4.47}$$

plays a central role in calculations involving loss and gain. The operator $\hat{\phi}(x)$ annihilates a photon locally and instantaneously from the soliton, disturbing the center-of-mass position of the soliton as well as the relative positions of the remaining individual photons.

So far we have simply assumed that the approximation (4.26) is valid. As mentioned above, the physical intuition behind this approximation is that one lost photon constitutes only a small disturbance to the soliton, and so the states resulting from the operation by $\hat{\phi}(x)$ are still soliton states. This is not exact. We have no reason to think that $\hat{\phi}(x)|n, E_{Ai}\rangle$ has a precisely zero overlap with other states. One can easily imagine significant overlap with states composed of a perturbed soliton and excited continuum, for example. We address this and other issues below.

A first approximation If the perturbed state is composed primarily of a soliton state component, then Eq. (4.26) will apply. The energy states $|n-1, E_f\rangle$ can be considered a complete basis, and

$$\phi(\epsilon) \approx \langle n-1, E_{Ai} - \epsilon|\hat{\phi}(\epsilon)|n, E_{Ai}\rangle. \tag{4.48}$$

The states involved here are not localized, but the operator acts locally at the coordinate x of the loss site. We have the additional difficulty that the matrix element involves a correlation of the energies of photons in an eigenstate. The Hartree approximation generally ignores such correlations. However, we can obtain the result indirectly by calculating an overlap of localized soliton states and then doing a Fourier transform.

For a loss site at position x, we want

$$\phi(\epsilon; x) = \langle n-1, E_{Ai} - \epsilon|\hat{\phi}(x)|n, E_{Ai}\rangle. \tag{4.49}$$

The energy states can be expanded in terms of localized soliton position states $|n, X\rangle$,

$$|n, E_{Ai}\rangle = v_{g0}^{-1/2}|n, P\rangle = (2\pi\hbar v_{g0})^{-1/2} \int dX e^{iPX/\hbar}|n, X\rangle, \tag{4.50}$$

where the group velocity v_{g0} comes into the normalization as a conversion from energy to momentum. The matrix element is now,

$$\phi(\epsilon; x) = (2\pi\hbar v_{g0})^{-1} \int dX dX' e^{i(PX-P'X')/\hbar}\langle n-1, X'|\hat{\phi}(x)|n, X\rangle. \tag{4.51}$$

This can be evaluated using a Hartree product state,

$$|n,X\rangle \approx \frac{1}{\sqrt{n!}}\left[\int dx'\phi_0(x'-X)\hat{\phi}^\dagger(x')\right]^n|0\rangle. \tag{4.52}$$

Here $\phi_0(x'-X)$ is the "wavefunction" of each photon about the center position, as discussed in [51, 32]. By commuting the annihilation operator through the n creation operators, we obtain

$$\hat{\phi}(x)|n,X\rangle \approx \sqrt{n}\phi_0(x-X)|n-1,X\rangle. \tag{4.53}$$

As we substitute this into Eq. (4.51), we must make sure that the wavefunction $\phi_0(x)$ is sensible; in particular, we must choose the nominal momentum p_0 of the wavefunction. Clearly, since the Hartree state is approximating an energy eigenstate, we should choose p_0 consistent with the initial energy E_{Ai}, that is $p_0 = h_i/v_{g0}$.

Plugging in, we have

$$\phi(\epsilon,x) = (2\pi\hbar v_{g0})^{-1}\sqrt{n}\int dX e^{i(P-P')X/\hbar}\phi_0(x-X;p_0) \tag{4.54}$$

And taking the Fourier transform,

$$\phi(\epsilon,x) = (2\pi\hbar v_{g0})^{-1}e^{i(P'-P)x/\hbar}\phi_P(P'-P-p_0). \tag{4.55}$$

Here the classical momentum distribution ϕ_P has been introduced,

$$\phi_P(P) \equiv \int dX e^{-iP(x-X)/\hbar}\phi_{\text{class}}(x-X;p_0) \tag{4.56}$$

Finally, we would like to convert back to units of energy. We substitute $P-P'$ for ϵ/v_{g0}, and plug in the sech function,

$$\phi(\epsilon,x) = Ae^{-i\epsilon x/\hbar v_{g0}}\text{sech}\left(\frac{\epsilon-h_i}{\hbar\Delta\omega_{clas}}\right). \tag{4.57}$$

where $h_i = E_{Ai}/n$ is the energy-per-photon corresponding to p_0, A is a constant, and

$\Delta\omega_{\mathrm{clas}}$ is the classical soliton spectral width.

The above derivation is somewhat lengthy, since we must tease out photon correlations from the Hartree approximation. The high level result, however, is quite understandable. Individual photons live in energy wavefunctions centered about a collective soliton energy-per-photon $\hat{h} = \hat{H}_A/n$. When a photon is removed, its energy is distributed according to this energy wavefunction, with uncertainty $\Delta\omega_{\mathrm{clas}}$. This picture is analogous to a model of electrons in wavefunctions about an atomic center. In both cases, the Hartree approximation can make a particle formalism quite easy to work with and powerful even when one particle is being singled out, as in Eq. (4.47). Our use of configuration-space methods in optics is motivated by their success in atomic and nuclear problems.

Configuration spaces for different Fock sectors A soliton described by an n-photon wavefunction can evolve into a state described by an $n-1$ photon wavefunction. Any configuration-space calculation of expectations such as

$$\langle\Psi_f|\Delta\hat{H}_A^2|\Psi_f\rangle - \langle\Psi_i|\Delta\hat{H}_A^2|\Psi_i\rangle \tag{4.58}$$

necessarily means connecting wavefunctions from two different configuration spaces.

For lossless systems, we considered the evolution of a general n-photon wavefunction $\Psi(x_1, \ldots, x_n)$ on the set of n photon position coordinates. Once loss is introduced, the coordinate system itself can change. That is, when a photon is lost, the final wavefunction exists in a Hilbert space with fewer particles:

$$\Psi_i(x_1, \ldots, x_n) \longrightarrow \Psi_f(x_1, \ldots, x_{n-1}). \tag{4.59}$$

To analyze the center-of-mass position operator, for example, we must first recognize that it has a *different* configuration-space representation before and after the absorption event (4.59):

$$X_i = \sum_{k=1}^{n} x_k/n \quad \longrightarrow \quad X_f = \sum_{k=1}^{n-1} x_k/(n-1). \tag{4.60}$$

Having said this, we understand that writing Eqs. (4.26) and (4.28) was actually quite significant, although not mathematically laborious. We have found a transformation taking a wavefunction for X_i into a wavefunction for a different operator, X_f, in a different configuration space. Specifically, an initial complex exponential in X_i (momentum eigenstate) is mapped to a sech in X_f. Finally, focusing on a particular lost coordinate, x_n, is not only important formally, but is a key piece of intuition behind loss-induced noise.

Excited internal states The state (4.47) has an altered photon number and perturbed relative coordinates. It *is not* purely a superposition of soliton components with relative coordinates in the ground state. Loss of a photon disturbs the "potential" seen by the other photons, which can then become unbound. The complete decomposition would look much more complicated; we might begin by writing,

$$\begin{aligned}\hat{\phi}(x)|n, E_{Ai}\rangle &= \int dE_{Af}\phi(E_{Ai} - E_{Af})|n-1, E_{Af}\rangle \qquad (4.61)\\ &\quad + \sum_j \int dE_{Af}\phi_j(E_{Ai} - E_{Af})|E_{Af}, n-j-1\rangle \otimes |\eta_j\rangle. \qquad (4.62)\end{aligned}$$

Here $n-1$ is the total number of photons, and j is the number of photons which have been shed into continuum, or unbound modes. The state of the continuum is written simply as $|\eta_j\rangle$ here, but would in fact have a non-trivial description of its own. As depicted in Figure 4-6, energy in unbound states tends to disperse from the soliton [35]. If continuum photons were important, and the above description necessary, the analysis of loss would be very much more complicated than what we presented earlier. We then have a strong motivation to show that the simpler expansion is correct.

In Appendix D, we argue that non-soliton contributions to the perturbed state (4.47) are of order $1/n$ or smaller. This mathematical result is of great importance. It means that Eq. (4.26) is approximately a straightforward superposition of solitons, as we assumed in deriving $\phi(\epsilon)$ above. In this limit, the relative part of the wavefunction

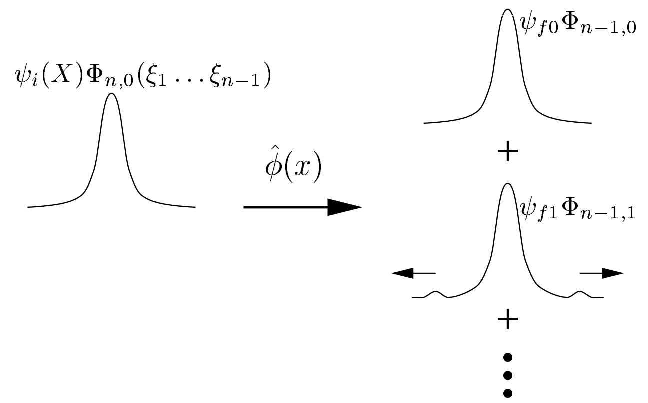


Figure 4-6: The local operator $\hat{\phi}(x)$ acting on an initial soliton state generates an $n-1$ soliton component [with ground internal state $\Phi_{n-1,0}$], as well as components with excited internal energy. Excited state components have j unbound photons, which disperse away from the soliton. We have shown that the $j=0$ soliton state dominates the sum.

is trivial; in the language of Chapter 3,

$$\begin{aligned}\Psi_i(x_1,\ldots,x_n) &= \psi_i(X_i)\Phi_{sn}(x_1-X,\ldots) && (4.63)\\ \Psi_f(x_1,\ldots,x_{n-1}) &\approx \psi_f(X_f)\Phi_{s,n-1}(x_1-X,\ldots), && (4.64)\end{aligned}$$

and so the field evolution is just a transformation of the center-of-mass wavefunction

$$\psi_i(X) \longrightarrow \psi_f(X). \tag{4.65}$$

Analytical estimate

A simple analytical estimate of $T(\epsilon)$ can obtained [21] using Gaussian shapes for β and ϕ. Physically, this means we assume a Gaussian lineshape of the loss or gain and approximate the sech shape of the soliton spectrum by a Gaussian. Mathematically,

$$|T(\epsilon;h_i)|^2 \approx T_0 e^{-(\epsilon-\epsilon_b)^2/2\sigma_b^2} e^{-(\epsilon-h_i)^2/2\sigma_c^2} \tag{4.66}$$

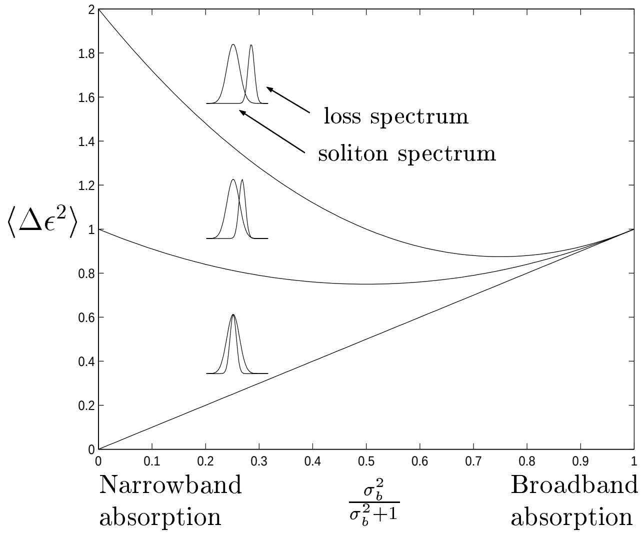


Figure 4-7: Graph of energy perturbation as a function of absorption bandwidth. The perturbation $\langle \Delta\hat{\epsilon}^2 \rangle$ is normalized to units of the classical soliton bandwidth ($\sigma_c^2 \equiv 1$) and plotted as a function of the parameter $\sigma_b^2/(\sigma_b^2 + \sigma_c^2)$. This parameter goes to unity if the absorption bandwidth is much broader than the soliton spectral width, in which case $\langle \Delta\hat{\epsilon}^2 \rangle \to \sigma_c^2$. Three curves are given, corresponding to the offsets of the absorption peak from the center of the soliton spectrum: $\epsilon_b - h_i = 0$, 1, and 2.

leads to

$$\langle (\epsilon - h_i)^2 \rangle \approx \frac{\sigma_c^2 \sigma_b^2}{\sigma_c^2 + \sigma_b^2} + \frac{\sigma_c^4 (\epsilon_b - \langle \hat{h} \rangle_i)^2}{(\sigma_c^2 + \sigma_b^2)^2} \tag{4.67}$$

Of course, more accurate estimates can be obtained numerically using the correct shapes for ϕ and β. Due to pathological behavior of the Lorentzian, a modified spectrum must be used, corresponding to more physical behavior far off-resonance. The Gaussian approximation, plotted in Figure 4-7, is sufficient for our purposes, and essentially exact for broadband absorption and gain spectra.

The exchange spectrum generally mimics the narrower of the contributing spectra. This is certainly sensible in the limiting cases: a narrow-band absorber will exchange only definite quanta of energy, while a broadband absorber will accept whatever energies are present in the soliton. As a result, the noise added to $\hat{h}$ will not generally be larger than σ_c^2. In principle, a narrow absorption line far off of the soliton frequency [that is, with large $(\epsilon_b - h_i)^2$] can add a large disturbance per photon. However, such absorption events occur with low probability, and will tend to be outweighed by better matched interactions.

4.2.5 Gain

We have derived the needed expectations for loss interactions. Photon emission does have some basic differences from absorption, but the effects that we are interested in are the same, as we argue below.

The crucial step is to show that $\hat{\phi}^\dagger(x)|E_{Ai}\rangle$ is dominated by a soliton component with internal ground state, as discussed in Section 4.2.4. If this is true, then just as we had Eqs. (4.26) and (4.48), we now have

$$\hat{\phi}^\dagger(x)|E_{Ai}\rangle = \int dE_{Af}\phi^+(E_{Af} - E_{Ai})|E_{Af}\rangle \tag{4.68}$$

and

$$\phi^+(\epsilon; E_{Ai}) = \langle E_{Ai} + \epsilon|\hat{\phi}^\dagger(\epsilon)|E_{Ai}\rangle = [\phi(\epsilon; E_{Ai} + \epsilon)]^*. \tag{4.69}$$

For large n, the spectra of states perturbed by loss and gain are approximately the same,

$$\phi^+(\epsilon) \approx \phi(\epsilon). \tag{4.70}$$

An exchanged photon, whether annihilated by $\hat{\phi}$ or created by $\hat{\phi}^\dagger$, simply falls within the single-photon spectrum $\phi(\epsilon; E_{Ai})$ about the average energy E_{Ai}/n. Expectations such as Eq. (4.46) follow in exactly the same way for loss and gain.

The initial argument, that the soliton contribution dominates, is somewhat different for gain. One way of seeing this is that $\hat{\phi}^\dagger$ has both spontaneous and stimulated emission built into it. The action of $\hat{\phi}$ is always to take a photon from the soliton mode, in which all available photons reside. In contrast, $\hat{\phi}^\dagger$ can either add a photon to the soliton mode, or induce spontaneous emission.

For large photon number, however, the dominant effect will come from photons emitted into the highly-occupied soliton "mode," or single-photon wavefunction. Another way of looking at this is that an emitted photon will only be felt by the soliton if it overlaps with other photon wavefunctions. Restricting ourselves to this dominant effect, gain becomes exactly the reverse of loss: a photon is either added to or removed from the soliton mode by the action of the photon operators. The condition for the dominance of the soliton term applies for gain as well.

It is worth noting that conflicting definitions of spontaneous emission have been used here. We need to be careful with the language that has evolved in this field. In quantum optics, "amplified spontaneous emission" has been used as a broad term describing the effects of fluctuations coupled in by loss and gain interactions (the intuition behind Figure 4-1). With this as our definition, the configuration space calculation presented here exactly confirms the current understanding of ASE-induced momentum uncertainty. As we have mentioned, our noise mechanism, exchange of discrete, uncertain quanta, is in fact a manifestation of the presence of field fluctuations, or ASE.

However, in the calculation of $\hat{\phi}^\dagger(x)|E_{Ai}\rangle$, a different use of "spontaneous" has been used. The action of $\hat{\phi}^\dagger$ is to generate spontaneous and stimulated photons, as described above. What we have argued above is that only the stimulated component matters when the photon number is large. From this point of view, spontaneous emission has been neglected entirely, and is not a significant source of noise. The discrepancy is only a semantic one.

4.3 Timing Jitter

We have developed the machinery to understand the effects of loss on solitons at the microscopic level. We now demonstrate the applicability of this problem to practical propagation by reproducing the well-known Gordon-Haus effect.

According to the Gordon-Haus analysis, timing jitter is the ultimate result of momentum noise,

$$\frac{\partial \langle \Delta \hat{P}^2 \rangle}{\partial t} = S_P. \tag{4.71}$$

Below, we show that quantitatively the same noise can be derived from our results of Section 4.2.3

$$\frac{\partial \langle \Delta \hat{P}^2 \rangle}{\partial t} = [\delta \langle \Delta \hat{h}^2 \rangle]_{\text{single-photon}} \times \text{Rate of Exchanges} \tag{4.72}$$

in the limit of broadband loss and gain spectra. We conclude that the simple picture of Figure 4-4 is the basic mechanism underlying timing jitter induced by loss and gain.

Gordon-Haus basics

The basic Gordon-Haus effect can be understood as follows: momentum is perturbed by white noise injected into the soliton as in Eq. (4.72). The rate can be calculated from perturbative field equations [29, 35]. If we think of the (Heisenberg) position operator as

$$\hat{X} = \hat{X}_0 + \omega' t - \frac{\omega''}{\hbar n} \int dt \hat{P}$$

then the dominant uncertainty at long propagation times is

$$\langle \Delta \hat{X}^2 \rangle \propto \left\langle \left(\int dt \Delta \hat{P} \right)^2 \right\rangle. \tag{4.73}$$

This system results in a t^3 dependence of the position uncertainty,

$$\langle \Delta \hat{X}^2 \rangle \propto S_P t^3. \tag{4.74}$$

Combining the kicks

The formalism of this chapter leads to the Gordon-Haus analysis above once we combine the individual perturbations of interaction events. We show here that the Gordon-Haus result is reproduced exactly in the limit that the loss and gain spectrum are broad with respect to the soliton spectrum.

According to the momentum "kick" analysis, each loss or gain event contributes the conditional perturbation

$$\frac{\langle \Delta \hat{h}^2 \rangle_f - \langle \Delta \hat{h}^2 \rangle_i}{\sin^2(\theta)} \approx \frac{\langle \Delta\epsilon^2 \rangle}{n^2}. \tag{4.75}$$

The noise resulting from many events will just be the sum of these independent contributions. After propagating distance δz in a loss-compensated fiber, a pulse with n photons experiences $n\alpha\delta z$ loss events and an equal number of gain events, where α is the loss parameter of the material. The total perturbation to $\hat{h}$ can then be written as

$$\delta \langle \Delta \hat{h}^2 \rangle = \frac{\langle \Delta\epsilon^2 \rangle}{n^2} \times 2\alpha n \delta z. \tag{4.76}$$

The expectation is a shorthand for the uncertainty of each exchanged photon, as discussed above

$$\langle \Delta\epsilon^2 \rangle \equiv \langle \Psi_1 | \Delta\epsilon^2 | \Psi_1 \rangle \equiv \frac{\int \Delta\epsilon^2 |T(\epsilon)|^2 d\epsilon}{\int |T(\epsilon)|^2 d\epsilon} \tag{4.77}$$

In the case of broadband gain and loss, the exchange spectrum simply follows the classical soliton spectrum,

$$T(\epsilon) \propto \phi_0(\epsilon). \tag{4.78}$$

In this case, the uncertainty in each exchanged quanta is simply $\hbar$ times the uncertainty in frequency of the individual loss photon. But the "frequency uncertainty" of a photon is just a complicated name for the classical spectral width, $\Delta\omega_{\text{clas}}^2$

$$\langle \Delta\epsilon^2 \rangle = \hbar^2 \Delta\omega_{\text{clas}}^2. \tag{4.79}$$

Finally, we have

$$\delta\langle\Delta\hat{h}^2\rangle = \frac{2\hbar^2\Delta\omega_{\text{clas}}^2}{n}\alpha\delta z. \tag{4.80}$$

Clearly, the present result Eq. (4.80) is of the form (4.72). To compare the results quantitatively, we express both in the same units. The frequency noise implied by Eq. (4.80) is

$$\delta\langle\Delta\hat{h}^2\rangle/\hbar^2 = \frac{2\Delta\omega_{\text{clas}}^2}{n}\alpha\delta z = \frac{\alpha\delta z}{3n\tau_0^2}. \tag{4.81}$$

This is exactly what we get when we rescale the Gordon-Haus dimensionless frequency noise

$$\delta\langle\Delta\hat{\Omega}^2\rangle = \frac{\alpha\delta z}{3n}, \tag{4.82}$$

for the sech(t) soliton.

Can we avoid the noise? In the limit of broadband loss and gain, the above derivation assures us that the previous Gordon-Haus analysis is inescapable, regardless of any special initial states we might devise. We can adjust the parameters (τ_o, α, and ω'') or expand our model, but have no new loopholes. For practical fibers, the broadband assumption is quite safe. It is intriguing that a regime of noise reduction related to soliton propagation with spectral filters [56] seems possible in principle.

Reduced Gordon-Haus noise would result, for example, for a narrow gain spectrum (unlikely in real fibers). The mechanism is straightforward: if photons are added (or lost) at precisely the average frequency of the soliton, they do not perturb this frequency. In a sense this is because the spectrally narrow gain essentially implements *filtering* of the soliton pulse, discussed previously [56]. However, rather than restoring soliton components which have wandered from the desired frequency, the rate of wandering is reduced.

The noise S_p is proportional to $\langle\Delta\epsilon^2\rangle$ by Equation (4.76). In general, the uncertainty "kicks" are related to the width of the soliton spectrum σ_c and the width of the loss (gain) spectrum σ_b. When the loss and gain are broad, the noise is determined by the soliton spectrum only. We call this value S_{p0}, the level of momentum noise given in the literature and in Eq. (4.82). If the loss and gain bandwidth are

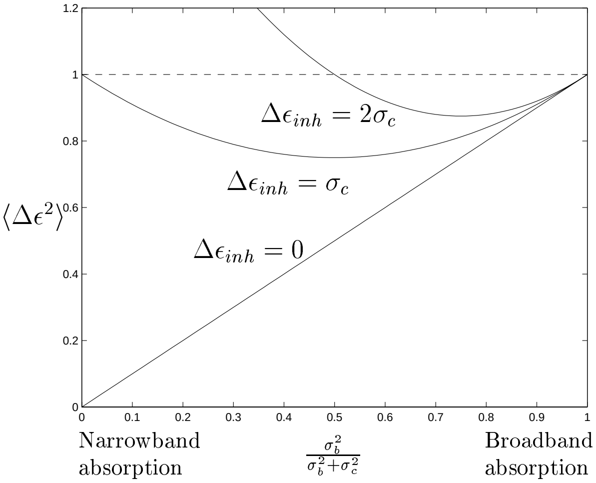


Figure 4-8: Momentum noise is normalized to the Gordon-Haus level, and plotted as a function of the gain/loss bandwidth, parameterized by $f_b = \sigma_b^2/(\sigma_b^2+\sigma_c^2)$. The three curves correspond to three different levels of *inhomogeneous* broadening: $\Delta\epsilon_{inh}^2 = 0$, σ_c^2, and $2\sigma_c^2$.

significant, they limit the spectrum of the exchanged photons. We think of the noise as being a function of the dimensionless parameter, $f_b = \sigma_b^2/(\sigma_b^2 + \sigma_c^2)$. The smaller this parameter is, the smaller the loss bandwidth relative to the soliton spectrum.

Using the Gaussian approximation to $\langle\Delta\epsilon^2\rangle$, Equation (4.67), we can write

$$S_p = S_{p0}\frac{\langle\Delta\epsilon^2\rangle}{\sigma_c^2} \tag{4.83}$$

$$\approx S_{p0}[f_b + (1-f_b)^2\Delta\epsilon_b^2]. \tag{4.84}$$

The offset of the loss resonance is $\Delta\epsilon_b = \hbar\omega_l - h_i$. As expected, when the loss is broad, $f_b \to 1$, we get $S_p \to S_{p0}$. However, when f_b is less than one, the momentum noise can drop below the usual level.

Figure 4-8 plots the normalized momentum noise. If all loss and gain sites have resonances centered exactly on h_i ($\Delta\epsilon_b = 0$), we can read off, for example, a 50% noise reduction at $f_b = 1/2$ ($\sigma_b = \sigma_c$). If the loss is only homogeniously broadened, $\Delta\epsilon_b$ is a constant offset. *Inhomogeneous broadening* can be included by using an average spread, $\Delta\epsilon_{inh}^2$ in place of $\Delta\epsilon_b$ in Eq. (4.84). As a result, even if lifetime broadening is negligible $\Gamma \sim \sigma_b \approx 0$, noise reduction requires that inhomogeneous broadening be small: $\Delta\epsilon_b < \sigma_c$.

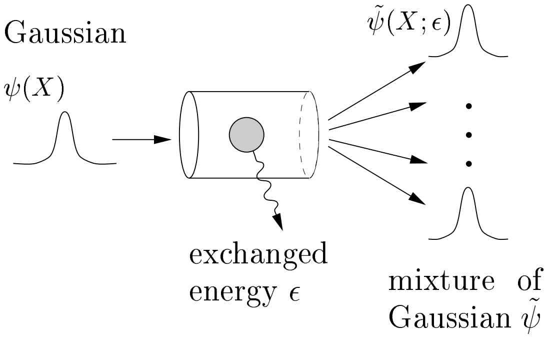


Figure 4-9: We depict one possible input-output picture of a loss event. A state with an initially Gaussian X-wavefunction is mapped by loss onto a mixture of perturbed Gaussian states. The result is easily extended to multiple loss events. This picture is relevant to the problem of soliton wavefunction manipulation and observation.

4.4 Decoherence, statistical mixtures, and measurements

So far, we have focussed primarily on loss-induced energy perturbations. We have also restricted ourselves to initial field states of definite energy. The underlying formalism which we have introduced is quite general; it can be applied to solitons with general center-of-mass wavefunctions and captures all effects of photon exchange on the soliton state.

In this section, we will outline the more general use of the formalism. Perhaps most importantly, our description of the perturbed soliton state should plug into calculations directly, avoiding intermediate steps. With this in mind, Figure 4-9 shows one possible mapping of input to output field states, applicable to localized solitons. This particular mapping is based on Gaussian $\psi(X)$ wavefunctions. Entanglement with the matter state is handled in a convenient and natural way by regarding the output as a statistical mixture of perturbed Gaussians. This is described mathematically below.

4.4.1 Input-output picture of photon exchange

Having obtained the results of previous sections, the mapping of Figure 4-9 comes out easily. It allows a simple input-output picture to be applied quantitatively to propagation problems.

We begin by writing the initial state of a soliton with general center-of-mass wavefunction [43, 32]. We recall that h_i and momentum P each represent the soliton center frequency in different units. Then *any* soliton state expanded in terms of its momentum wavefunction ($\psi_P(P)$, the Fourier transform of $\psi(X)$)

$$|\Psi_i\rangle = \int dP \psi_P(P) |P, E_{Bi}\rangle, \tag{4.85}$$

can also be written as

$$|\Psi_i\rangle = \int dh_i a(h_i) |E_{Ai} = nh_i, E_{Bi}\rangle, \tag{4.86}$$

since for fixed n, h_i is an algebraic function of P ($\hat{P} \approx \hat{H}_X/\omega'$) Since the initial state is a sum of energy eigenstates, the final state is just the sum of each evolved in time.

As before, interaction leaves us with either zero or one photon exchanged,

$$|\Psi_f\rangle = \cos(\theta)|\Psi_0(t_f)\rangle + \sin(\theta)|\Psi_1(t_f)\rangle. \tag{4.87}$$

The perturbed state component is a straightforward superposition,

$$\sin(\theta)|\Psi_1(t_f)\rangle = \int dh_i a(h_i) \int d\epsilon T(\epsilon; h_i) |nh_i - \epsilon, E_{Bi} + \epsilon\rangle \tag{4.88}$$

When we write the "unperturbed" state component we are careful about how we normalize the states. We recall that for an input eigenstate, the normalization gave

$$\cos(\theta)|\Psi_0\rangle \to \left(1 - \int d\epsilon |T(\epsilon)|^2\right)^{1/2} |\Psi_i\rangle. \tag{4.89}$$

It is helpful to maintain conservation of total energy conservation without including a full second-order perturbative correction. This is accomplished simply by superposing the normalized unperturbed states of each energy,

$$\cos(\theta)|\Psi_0(t_f)\rangle = \int dh_i a(h_i) \left(1 - \int d\epsilon |T(\epsilon; h_i)|^2\right)^{1/2} |h_i, E_{Bi}\rangle \tag{4.90}$$

so that the total probability corresponding to each initial energy is time-invariant.

Or analysis of energy kicks presented earlier now extends almost unchanged to the case of these general soliton states. It had better: in realistic situations, a soliton is often localized and will therefore certainly not be in an energy eigenstate. The momentum noise analysis of the preceding sections is actually only meaningful if it agrees with the case of small but finite initial energy uncertainty, as we now show. The kick

$$\langle \Delta\hat{h}^2\rangle_f - \langle \Delta\hat{h}^2\rangle_i = \int dh_i \int d\epsilon |a(h_i)|^2 |T(\epsilon; nh_i)|^2 \left(\frac{\epsilon - h_i}{n-1}\right)^2 \tag{4.91}$$

is formally identical to Eq. (4.45). The above general energy perturbation reduces to Eq. (4.44) in all cases currently of interest: as long as the uncertainty in h_i is small compared to the width of $T(\epsilon)$.

What is perhaps more interesting is interpreting the perturbed state in line with Figure 4-9. In Appendix C we obtain an interesting approximate form for the perturbed state (4.88),

$$\sin(\theta)|\Psi_1(t_f)\rangle = \underbrace{\int d\epsilon T(\epsilon; \bar{h}) |E_{Bi} + \epsilon\rangle}_{\text{statistical spread}} \otimes \underbrace{\int dh_i \tilde{a}_\epsilon(h_i) |nh_i - \epsilon\rangle}_{\text{perturbed wavefunction}}. \tag{4.92}$$

The approximation applies to the usual physical case where the quantum spread in the center frequency is much smaller than the classical spectrum. By separating the reservoir and field states, we bring out some important physics of the interaction: The above state is an entanglement of an absorber with the field. Since the entanglement is with a reservoir, it is equivalent to generating a statistical mixture of soliton states (as discussed in Appendix C). Each component of the mixture has an approximate

Gaussian wavefunction $\tilde{a}$ in the energy basis. It is interesting to note that the expansion in coherent components is not unique: a different expansion is obtained by using a different basis for the reservoir state.

Since this theory transforms Gaussians to Gaussians, it is easily cascaded. In the limit that $\tilde{a}$ is still approximately equal to a, multiple interactions will give a perturbed state. One could begin to write a theory for such states after k lost photons

$$\int d\epsilon_1 \ldots d\epsilon_k |\text{Reservoir State}(\epsilon_1, \ldots, \epsilon_k)\rangle \otimes \int dh_i \tilde{a}_k(h_i; \epsilon_1, \ldots, \epsilon_k) \left| nh_i - \sum_j \epsilon_j \right\rangle \tag{4.93}$$

The rules for deriving the coherent wavefunctions $\tilde{a}(h_i, \ldots)$ would follow from the basic principles we have outlined in this chapter and in Appendix C. Naturally, the rules for transforming $a \rightarrow \tilde{a}_k$ equivalently let us transform the soliton position wavefunction $\psi_i(X) \rightarrow \psi_{n-k}(X)$.

4.4.2 Measurement and uncertainty

So far we have not explicitly looked at interaction events as detection of photons. Remarkably, such considerations alone allow us to arrive at Eq. (4.80), and thus provide an additional interpretation of the above results. Specifically, the perturbation in soliton energy, and thus momentum, can be estimated as nothing more than the Heisenberg disturbance in the soliton momentum upon measurement of its position.

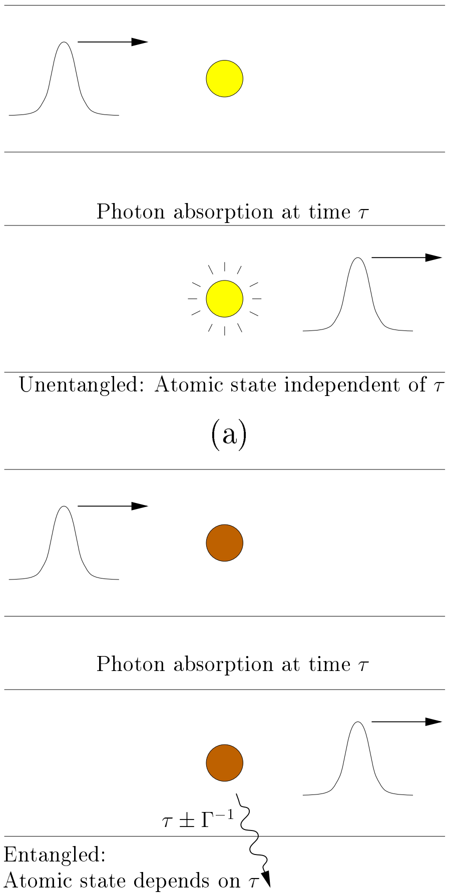
We have mentioned that our absorber is exactly the same model used for a detector of bandwidth Γ. Although the absorbers are not detectors which are read out experimentally, they set up an entanglement between soliton states and states of the macroscopic bath. The interaction site has then served as a detector with regard to Heisenberg uncertainties. We only outline this interpretation here, but a more careful analysis of the states supports our conclusions.

Figure 4-10 depicts the system before and after an absorption event in the case of broad and narrow-band interactions. The sharp resonance (slow detector) has the absorber state essentially independent of the exact time of the interaction, and so no entanglement of states results. For the broadband absorption, the quanta is quickly passed to the macroscopic bath, so that an independent state of system B exists essentially for each time slot of width Γ^{-1}. This is precisely the kind of entanglement which represents measurement. We can say that the arrival time of the photon has been measured to precision Γ^{-1}.

Since the absorbed photon is left behind, we are not interested in the measurement disturbance in its momentum, rather the momentum of the soliton. By measuring the arrival time of a photon at a local absorber, we have measured the soliton position to within $\Delta x_{\text{classical}}$, the classical width of the pulse. This position measurement is associated with a Heisenberg uncertainty kick in momentum. We calculate this kick to be

$$\Delta P_{\text{kick}} = \frac{\Delta p_{\text{classical}}}{n} \tag{4.94}$$

where $\Delta x_{\text{classical}}, \Delta p_{\text{classical}}$ is the minimum uncertainty pair of the single photon wavefunction . The factor of n corresponds to the soliton mass being n times the mass of



(a)

(b)

Figure 4-10: The effect of a photon absorption on the soliton state depends on the degree of entanglement of the absorption time with the bath state. For weak coupling of atom to bath (a), the final bath state is insensitive to the absorption time, and no entanglement results. For strong coupling, photons absorbed more than a few decay times Γ^{-1} apart result in essentially orthogonal atom-bath states. For small Γ^{-1}, this entanglement is the very essence of a precise measurement of the absorption time τ.

a photon. A single absorption then gives a perturbation

$$\langle \Delta \hat{P}_s^2 \rangle |_i^f \approx \frac{\langle \Delta \hat{p}^2_{\mathrm{classical}} \rangle}{n^2}, \tag{4.95}$$

which is exactly equivalent to Eq. (4.80).

We have then interpreted Gordon-Haus noise in terms of a basic quantum principle. Noise injected in soliton momentum is just the Heisenberg disturbance which must accompany measurement of the of photon (and thus soliton) position.

4.5 Momentum squeezer revisited

We now return to momentum squeezer described in Section 3.3. A central argument presented there was the conservation of total momentum of the optical pulse. Fiber loss violates this conservation by allowing a net exchange of energy and momentum with external reservoirs. This does not prevent momentum squeezing, but does place fundamental limits on the amount of squeezing we can achieve for given fiber parameters.

4.5.1 Loss and momentum noise

Precisely the same mechanism of momentum noise that we have just been discussing applies to the momentum squeezing problem. We can thus apply the results of the preceding sections, Eq. (4.72), directly to this problem. It is convenient to express the noise $\langle \Delta P^2 \rangle$ as a fraction of the classical soliton momentum spread at a given time $\langle \Delta p^2(t) \rangle$, so that we can connect with Eqs. (3.53) and (3.54). We recall that this is proportional to the coherent-state momentum uncertainty, denoted by "coh." The momentum disturbance in time dt is the uncertainty of $n\alpha v_g dt$ lost photons:

$$\delta \langle \Delta \hat{P}^2 \rangle = \frac{\delta z}{v_g} \frac{d}{dt} \langle \Delta \hat{P}^2 \rangle = \alpha n \delta z \langle \Delta p^2(t) \rangle \tag{4.96}$$

where α is the fiber loss in units km^{-1}. Combining Eqs. (3.54) and (4.96), and integrating the noise in $\delta z = v_g dt$ we get the squeezing ratio including loss-induced momentum noise,

$$\begin{aligned} S(t) &\equiv \frac{\langle \Delta P^2(t) \rangle}{\langle \Delta P^2(t) \rangle_{\text{coh}}} & (4.97) \\ &= \frac{n\langle \Delta p^2(0) \rangle + \alpha v_g n \int d\tau \langle \Delta p^2(\tau) \rangle}{n \langle \Delta p^2(t) \rangle} & (4.98) \\ &= \frac{n\langle \Delta p^2(0) \rangle + \alpha v_g n \langle \Delta p^2(0) \rangle \int d\tau g^{-2}(\tau)}{n\langle \Delta p^2(0) \rangle g^{-2}(t)} & (4.99) \\ &= g^2(t) + \alpha v_g t_{\text{eff}}. & (4.100) \end{aligned}$$

The effective loss time is defined as

$$t_{\text{eff}} \equiv \int_0^t d\tau (g^2(t)/g^2(\tau)). \tag{4.101}$$

This quantity takes into account the momentum perturbations caused as the pulse spectrum varies. That is, a photon lost at time $\tau = 0$ causes a small disturbance due to the narrow spectrum of the pulse at that time, and thus contributes little to t_{eff}. A photon lost at time $\tau = t$ causes a larger disturbance, since the spectrum is fully broadened [21].

4.5.2 Dispersion profile design

Using the above squeezing expression, one can begin to design a dispersion profile for achieving large momentum squeezing. The total fiber length $z = v_g t$ and profile $g(t)$ are chosen to maximize spectral broadening and to avoid excessive loss-induced momentum noise. The literature on soliton compression using DDF [76, 49, 12] gives us an excellent starting point in finding an optimal design.

The total length of the device will be dictated by the nonlinear coefficient. That is, for the device to generate a significant amount of spectrum, the interaction time must be long enough for nonlinear phase to accumulate,

$$t_{\text{eff}} > \Delta x_{\text{pulse}}/n\kappa. \tag{4.102}$$

Here $n/\Delta x$ is the "intensity" corresponding to our normalized nonlinearity κ. There is thus a limit on squeezing associated with given loss and nonlinearity coefficients,

$$S(t) > \alpha v_g \Delta x_{\text{pulse}}/n\kappa. \tag{4.103}$$

As we see below, a specific case of soliton pulse-compression bears out this general rule.

4.5.3 Exponential soliton compression

Adiabatic compression of solitons requires that $g(t)$ vary slowly compared to the soliton period $\propto \Delta x_s/n\kappa$. For example, an exponential-compression scheme

$$g^2(t) = g^2(0)e^{-2pt}. \tag{4.104}$$

is described in [49]. We can also take into account the broadening of soliton pulses due to loss. In place of the lossless relation $g = \gamma$, we then have

$$\gamma(t) = g(t)e^{-2\alpha t} = e^{-(p+2\alpha)t}. \tag{4.105}$$

The adiabatic condition is determined by the nominal classical soliton,

$$\phi_0(x,t) = \sqrt{n/2}\operatorname{sech}(x/x_s)e^{-\Omega t}, \tag{4.106}$$

which accumulates phase at the rate

$$\Omega = |\omega''|/2{x_s}^2 = n\kappa/2x_s \tag{4.107}$$

Adiabatic compression then satisfies $p \ll n\kappa/2x_s$ at every point along the fiber. Naturally, x_s is maximum at $t = 0$, so that this point will give the adiabatic condition.

For exponential compression, t_{eff} approaches the asymptotic value $1/2p$ as t increases,

$$t_{\text{eff}} = (1 - e^{-2pt})/2p, \tag{4.108}$$

since early loss events contribute little momentum noise. Figure 4-11 shows the squeezing ratio plotted as a function of distance along the fiber for several loss values. After traveling for sufficiently long in such a fiber, the squeezing ratio (4.100) becomes

$$S(t) \longrightarrow \alpha v_g/2p. \tag{4.109}$$

By design, p will be as large as possible, but less than the nonlinear rate, $n\kappa/2x_s(0)$.

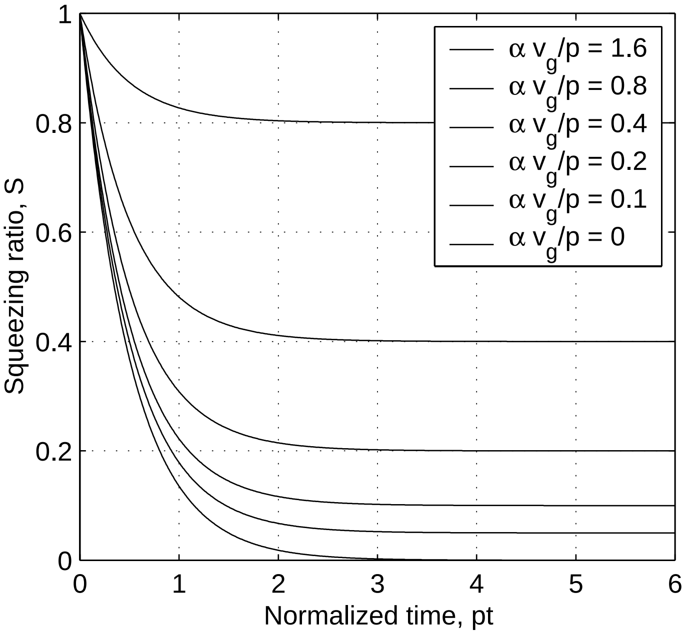


Figure 4-11: The squeezing ratio is plotted as a function of normalized time for an exponential pulse compression: $g(t) = e^{-pt}$. The squeezing ratio reaches an asymptotic value as the pulse propagates down the fiber. This squeezing limit is imposed by loss-induced momentum noise, and is simply the ratio of the loss rate and the compression rate, $S_{\min} = \alpha v_g/2p$.

When the loss to nonlinearity ratio is small, a large degree of squeezing can thus be achieved. This limit clearly demonstrates the limitation imposed by that ratio,

$$S(t) \gg \alpha v_g x_s(0)/n\kappa. \tag{4.110}$$

The squeezing ratio can be rewritten in an elegant form in terms of dimensionless parameters. We first satisfy the adiabatic condition by setting

$$p = \mu n\kappa/2x_s(0), \tag{4.111}$$

where $\mu \ll 1$. We then define the normalized loss coefficient:

$$\theta = \alpha v_g x_{\min}/n\kappa, \tag{4.112}$$

and the compression ratio

$$z = x_{\min}/x_s(0) = e^{-pt_{\text{final}}} < 1. \tag{4.113}$$

The squeezing ratio (4.100) at $t = t_{\text{final}}$ is then simply,

$$S = z^2 + \theta\mu^{-1}(z^{-1} - z). \tag{4.114}$$

To achieve large squeezing, we raise μ and reduce α and $x_{\min}$ as much as possible: μ is constrained by the breakdown of adiabaticity, α by the materials and fabrication capabilities, and the pulse width by the breakdown of our fiber model for femtosecond pulses. Once we find a satisfactory ratio θ/μ, z can be varied by choosing the initial pulse width, compression rate, and total length. In Figure 4-12, we sketch the optimum squeezing attainable as a function of θ/μ. The analytical solution can be obtained from Mathematica, but is too cumbersome to include here. (If a crude analytical estimate is needed, $S^2 + (\theta - 1)^2 = 1$ fits fairly well.) For example, using some reasonable numbers $\alpha \sim 1/20\text{km}$, $\mu \sim 1/5$, and taking $v_g x_{\min}/n\kappa \sim 0.1\text{km}$ from [84], one should be able to achieve $\theta/\mu \sim 1/40$, and thus 7 or 8 dB of squeezing. To exceed 15 dB of squeezing would be far more difficult, requiring $\theta/\mu \approx .002$.

4.5.4 Optimizing squeezing

Exponential compression is by no means the only scheme one should consider. Adiabatic compression with a variety of dispersion profiles has been described in the literature [58, 55]. This makes better use of the adiabatic condition by increasing the compression rate p as the soliton period shortens. In addition, non-soliton methods may prove more efficient, as long as an appropriate momentum-measurement device is available.

Naturally, optimizing a specific design will require a more careful analysis. For example, continuum generation should be included, along with various effects applicable to femtosecond-pulses [12]. Though the detailed results will be different, the

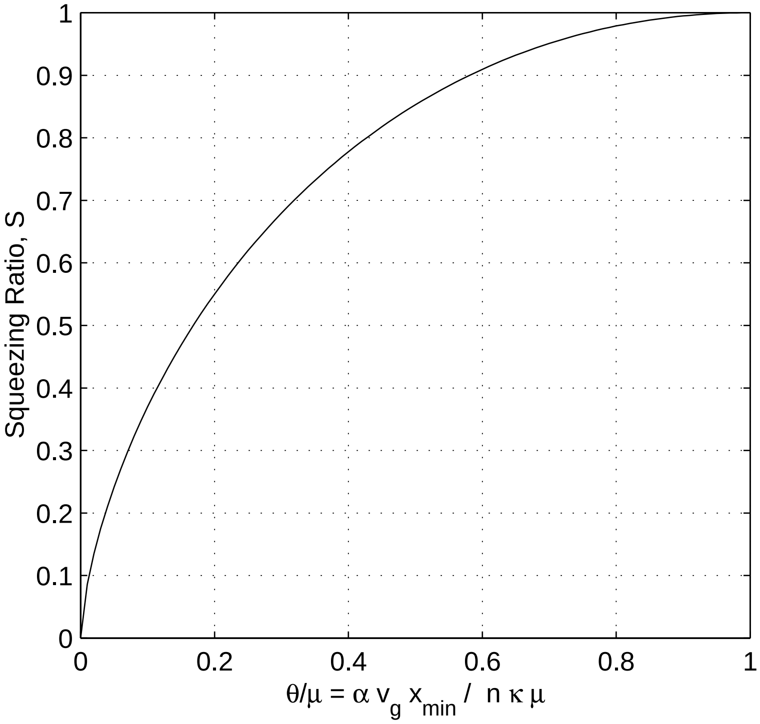


Figure 4-12: The squeezing ratio is plotted for adiabatic soliton compression with an exponentially decreasing dispersion. The horizontal axis is θ/μ, essentially a normalized loss ratio, where $\mu < 1$ is the adiabatic constant and $\theta = \alpha v_g x_{\min}/n\kappa$ is a figure of merit combining the loss, nonlinearity and pulse width.

general relations and trends that we have outlined above will apply.

4.5.5 Extreme squeezing

For extreme squeezing, a squeezing ratio less than $1/n$ is required, and so

$$\alpha v_g < \kappa/\Delta x_{\text{pulse}}. \tag{4.115}$$

The basic idea is that even a single lost photon constitutes a measurement of the position of the pulse (to within a pulsewidth), and thus breaks the superposition. The loss length must then be longer than the nonlinear interaction length of *single photons*. This extreme loss sensitivity is characteristic of many-photon DeBroglie-wave effects. Generally speaking, realistic losses are far too large for the effects to be observable. Improved materials, detectors, and experimental designs may allow future progress towards this goal.

Chapter 5

Multi-photon tunneling interferometer

Hagelstein has proposed a nonlinear optical coupler that would demonstrate photonic DeBroglie-wave effects. This coupler constitutes the special beamsplitter that one would need to build an interferometer for DeBroglie waves as depicted in Figure 2-1. We now focus on analysis and simulation of a specific implementation of this nonlinear interferometer concept.

In this chapter, we start with a very simple model of a nonlinear waveguide coupler, and show that interesting DeBroglie-wave effects emerge without special approximations. As in a soliton system, the Kerr interaction between photons acts as a binding force. If it is strong enough compared to other interactions between the field modes, all photons will switch from one waveguide to the another coherently as a single unit.

In particular, we use second-quantized model in order to connect the configuration-space concepts with a more familiar starting point. This allows us to utilize the related literature on tunneling in spin systems. One can also perform a purely configuration-space analysis; in fact, Hagelstein has obtained preliminary results using this alternative approach.

5.1 Basic Physics

Consider a nonlinear directional coupler (NLDC) composed of two Kerr waveguides with overlapping evanescent fields. A general model would be quite complicated, including transverse bound and continuum modes, temporal mode shaping, loss, and fabrication imperfections, all using quantized field models. The basic underlying physics can be understood much more simply: like others who have studied the quantum nonlinear coupler [39, 11], we begin with a two-mode Hamiltonian $\hat{H}$ governing the propagation of a quantum state along the length of the coupler. Thus we include only the linear coupling μ and Kerr nonlinearity κ essential to the function of the NLDC,

$$\hat{H}_{\text{field}}(t) \;=\; -\mu(t)(\hat{a}_1^\dagger \hat{a}_2 + \hat{a}_2^\dagger \hat{a}_1) - \kappa(\hat{a}_1^\dagger \hat{a}_1^\dagger \hat{a}_1 \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2^\dagger \hat{a}_2 \hat{a}_2). \tag{5.1}$$

In using this model, modes 1 and 2 are assumed at the outset to have well-defined transverse modes and temporal pulse shape. The operators $\hat{a}_1^\dagger$ and $\hat{a}_2^\dagger$ create photons in these pre-defined left and right waveguide modes, respectively. The linear coupling part of the Hamiltonian $\mu(\hat{a}_1^\dagger \hat{a}_2 + \hat{a}_2^\dagger \hat{a}_1)$ is the standard interaction between modes that leads to normal beamsplitter or coupling behavior. The nonlinear part adds photon correlations and, as we will see, makes the dynamics much more interesting.

Classical versions of this model have been used to describe dual-core fibers and semiconductor waveguides, for example. The distance $d(z)$ between the fiber cores or waveguides is varied as a function of length z along the propagation. The degree of overlap of the evanescent fields then gives rise to a linear coupling $\mu(t = z/v_g)$ which seems to vary with time for a pulse in the moving frame. The particular material implementation will need to satisfy strict requirements on loss, index uniformity, etc. We currently envision the device as a special, two-core fiber with high nonlinearity, although the requirements of our quantum system are beyond the forseeable performance of this technology.

By using this simplified two-mode description, we have brushed some important effects under the rug. While this is a useful simplification for gaining understanding, we do not imagine that loss or temporal pulse shape can realistically be ignored. These issues will all have to be addressed before our scheme can be considered a realistic design.

5.1.1 Trapped states

A good deal of attention has been paid to the above Hamiltonian. There have been some results on the optical coupler [39, 11, 48], for example squeezing, but much more on other physical systems with equivalent mathematical models such as molecular spins, for example [71, 69, 78]. Since the models are equivalent, the extensive understanding of quantum spin dynamics applies directly to our problem.

We are interested in the case where the nonlinearity is large, so that the Kerr effect provides a binding energy. This is analogous to the case of soliton propagation—when a soliton propagates, photons dig a "potential well" for each other via the nonlinear index change. All photons remain "trapped" in this well even in the presence of chromatic dispersion. Similarly, when a pulse enters on one side of a Kerr coupler, a nonlinear index change is induced in that waveguide. There will then be an index difference between the two physically identical waveguides. If this index difference is large enough, it will cause a phase mismatch of the small linear coupling, and light will be classically trapped in the initial waveguide. This effect has been observed experimentally—coupling between waveguides in the NLDC is suppressed when the input intensity is large enough [79]. For discussion of the trapping condition using a spin representation, see [78].

We can think of the trapping effect as the result of an energy of localization included in the Hamiltonian (5.1). That is, the nonlinear term assigns a "potential energy" to the degree of localization of photons. To see this, we rewrite the nonlinear operator for fixed total photon number N:

$$\begin{aligned}
\hat{a}_1^\dagger\hat{a}_1^\dagger\hat{a}_1\hat{a}_1 + \hat{a}_2^\dagger\hat{a}_2^\dagger\hat{a}_2\hat{a}_2 &= \hat{n}_1^2 + \hat{n}_2^2 + \hat{n}_1[\hat{a}_1^\dagger, \hat{a}_1] + \hat{n}_2[\hat{a}_2^\dagger, \hat{a}_2] && (5.2)\\
&= \frac{1}{2}[(\hat{n}_1 + \hat{n}_2)^2 + (\hat{n}_1 - \hat{n}_2)^2] + \hat{n}_1 + \hat{n}_2 && (5.3)
\end{aligned}$$

$$= \frac{1}{2}(\hat{n}_1 - \hat{n}_2)^2 + f(N). \quad (5.4)$$

The "potential energy" $-(\kappa/2)(\hat{n}_1 - \hat{n}_2)^2$ is minimized when all photons are in one waveguide or the other. When the nonlinearity is large, this potential tends to maintain the localization of light in one waveguide. This leads to classical trapping, but also to quantum correlated-photon states.

5.1.2 Multiphoton tunneling

A highly-nonlinear two-mode coupler is analogous to a particle in a double well, shown in Fig. 5-1. If light is initially localized in one waveguide, there is an energy barrier which classically traps it there. However, just as a quantum mechanical particle can tunnel across the barrier, a quantum coupler can evolve from a "trapped" state localized in one waveguide to a state localized in the other by tunneling across the barrier. In fact, the spin tunneling problem (exactly equivalent to our coupler model) has been rigorously mapped [69] onto a 1D double-well problem. Essentially all of the intuition from particle tunneling applies to multi-photon tunneling. The fact that multi-photon tunneling should occur seems natural from the configuration-space point of view; one can think of a bound collection of photons as a quantum particle whose "center of mass" evolves along the coupler. The collective dynamics of these interacting particles can be approached much like similar problems for massive particles [32, 31].

In the typical symmetric double-well problem, tunneling arises because the localized states are not true eigenstates of the system. A localized state is composed of symmetric and antisymmetric eigenstates, which are nearly degenerate. The small energy splitting leads to slow tunneling from one localized state to the other. In the NLDC as well, eigenstates obey the natural symmetry of the system. The two waveguides of the coupler are identical, and so all eigenstates must have a symmetry upon exchange of the waveguides: $\langle N-k, k|\Phi\rangle = \pm\langle k, N-k|\Phi\rangle$. States localized on one waveguide are not eigenstates, but superpositions of the symmetric and antisymmetric states. While localized states can be "trapped" on the time scale of linear

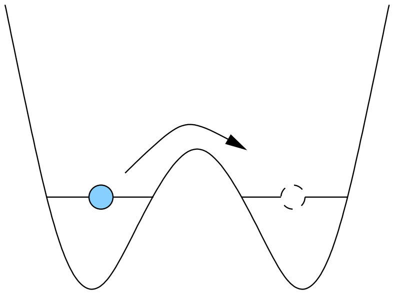


Figure 5-1: The state of a two-mode oscillator can be rigorously mapped onto the standard 1D tunneling problem. In the 1D problem, the particle has insufficient energy to pass the barrier classically, and is thus in a trapped state. A quantum particle can escape from this trapped state through tunneling. Photons in a two mode coupler can similarly evolve from a state localized on the left to a state localized on the right even in the trapping regime, where the nonlinear barrier is high.

oscillations, they eventually tunnel.

The basic mathematics is exactly the same for the coupler as for the double-well problem. An initially (left) localized state can be decomposed into the eigenstates,

$$|\Psi(t=0)\rangle = |\ell\rangle = (|s\rangle + |a\rangle)/\sqrt{2}. \tag{5.5}$$

These have a small energy splitting $\hbar\Omega$,

$$\begin{aligned} \hat{H}|s\rangle &= (E_0 - \hbar\Omega/2)|s\rangle && (5.6) \\ \hat{H}|a\rangle &= (E_0 + \hbar\Omega/2)|a\rangle. && (5.7) \end{aligned}$$

The time-evolution of this state is easily written in terms of eigenstates:

$$|\Psi(t)\rangle = e^{-iE_0 t/\hbar}(e^{i\Omega t/2}|s\rangle + e^{-i\Omega t/2}|a\rangle)/\sqrt{2}. \tag{5.8}$$

The energy difference $\hbar\Omega$ leads to dephasing of the antisymmetric and symmetric states. Since the energy splitting is small, the state will appear to be trapped on time scales on the order of μ^{-1}. However, on a longer time scale, the dephasing causes an oscillation between the left and right localized state, $|\ell\rangle$ and $|r\rangle = (|s\rangle - |a\rangle)/\sqrt{2}$, at the tunneling frequency Ω:

$$|\Psi(t)\rangle = e^{-iE_0 t/\hbar}[\cos(\Omega t/2)|\ell\rangle + i\sin(\Omega t/2)|r\rangle]. \tag{5.9}$$

The end result is exactly the kind of coherent superposition (2.16) which motivated us to analyze this device. Tunneling of multi-photon bound particles arises naturally from the physics of a strong Kerr nonlinearity, and provides the kind of states that are needed for observing photonic DeBroglie wave effects. As with other tunneling systems, the rate is very sensitive to all system parameters. This sensitivity creates a number of difficulties for implementing such a device.

5.1.3 Adiabatic variation:

A conflict arises in designing a NLDC: in order to maximize tunneling, we want to keep the linear coupling fairly large, thus making the overlap between $|\ell\rangle$ and $|r\rangle$ large. However, we may want the states at the input, output, and phase shifter to be as localized as possible, that is $|\ell\rangle \sim |N, 0\rangle$ and $|r\rangle \sim |0, N\rangle$. This condition requires small linear coupling. We have explored Hagelstein's proposal to adiabatically vary the linear coupling to satisfy both requirements.

The adiabatic theorem says that if the Hamiltonian of a system is changed slowly enough, an initial eigenstate of the Hamiltonian $\hat{H}(t = 0)$ evolves to remain an eigenstate of $\hat{H}(t)$ at each moment. Figure 5-2 shows the adiabatic interferometer configuration. The idea is to adjust the symmetric and antisymmetric states, $|s\rangle$ and $|a\rangle$, without disturbing coherent superpositions of the two or scattering into other states.

More specifically, the adiabatic principle says that there will be little scattering between the states a and b as long as the variations in the potential energy V satisfy $\hbar dV/dt \ll \Delta E_{ab}^2$. For small μ and n of interest in the following simulations, we can use the simpler condition $\hbar/\Delta t \ll \Delta E_{ab}$. The energy difference between the ground state, with $(n_1 - n_2)^2 = n^2$, and next excited state, with $(n_1 - n_2)^2 = (n-2)^2$ is

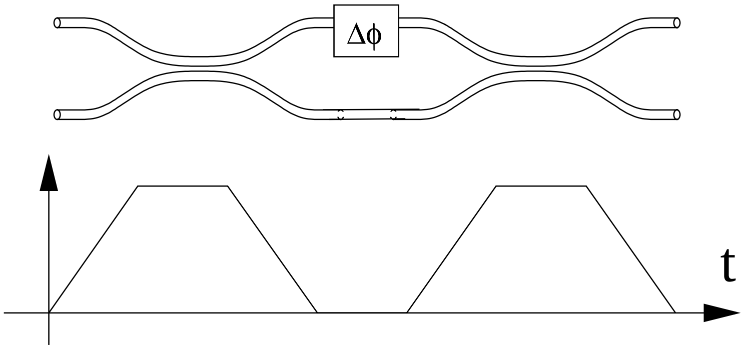


Figure 5-2: The adiabatic nonlinear Mach-Zehnder interferometer is composed of two matched NLDC's and a standard, linear phase-shift. The linear coupling μ in the NLDC's is turned on slowly by bringing the waveguides in closer proximity, allowing adiabatic evolution of the quantum field state.

approximately $\Delta E \approx 2\kappa(n-1)$. The variation of μ should occur on a time scale

$$\Delta t > \frac{\hbar}{2\kappa n} \tag{5.10}$$

The intuitive result is that the coupler must be longer than the nonlinear phase period. For example, the coupler should be many soliton periods in a soliton system. In the following numerical simulations, we confirm the principle that scattering is negligible for these slow variations. We return to the more general adiabatic condition in the next chapter.

5.2 Numerical simulation

The following calculation provides a test of the theoretical description we have outlined, confirming that a Kerr coupler includes the basic physics required for a photonic DeBroglie-wave device. We can also begin to explore quantitatively the limitations present in this simple model, such as the coupler length required for acceptable "adiabatic" coupling.

5.2.1 Schrödinger Equation with varying μ

We have simulated the interferometer of Fig. 5-2, which includes two NLDC's and a standard phase shifter. We start with initial states with photons in one waveguide only, and perform the three state transformations sequentially. The phase shift is trivially calculated using the relation (2.15). The NLDC has a more complicated, but straightforward, evolution according to the time-varying Hamiltonian. The nonlinear parameter κ was set to 1, and the linear coupling was slowly varied as seen in Fig. 5-2. For simplicity, we used a piecewise-linear variation: the first NLDC was simulated as

$$\mu(t) = \begin{cases} \mu_{\max} t/T & ; \quad 0 < t < T \\ \mu_{\max} & ; \quad T < t < 2T \\ \mu_{\max}(3T - t)/T & ; \quad 2T < t < 3T \end{cases} \tag{5.11}$$

The second NLDC is an identical copy of the first. The coupler length was chosen such that $T \gg \hbar/\kappa n$ to satisfy the adiabatic condition. In Figure 5-4, below, we show simulation results using very large values of T, to bring out features of the small-μ regime. Other simulations, discussed later, confirm that more moderate device lengths are sufficient.

We solved the Schrödinger equation by direct time-domain integration using leapfrog differencing. The field evolution separates naturally into photon number sectors. The N-photon, two mode quantum state can be represented by a normalized $N+1$-element vector $\boldsymbol{c}$, which evolves according to

$$i\frac{d}{dt}\boldsymbol{c} = \boldsymbol{A}\boldsymbol{c}. \tag{5.12}$$

One can easily identify the matrix elements of $\boldsymbol{A}$, $A_{k,j} = \langle N-k, k|\hat{H}/\hbar|N-j, j\rangle$, [71]. The results were insensitive to step size, indicating that our differencing accurately approximates the differential equations.

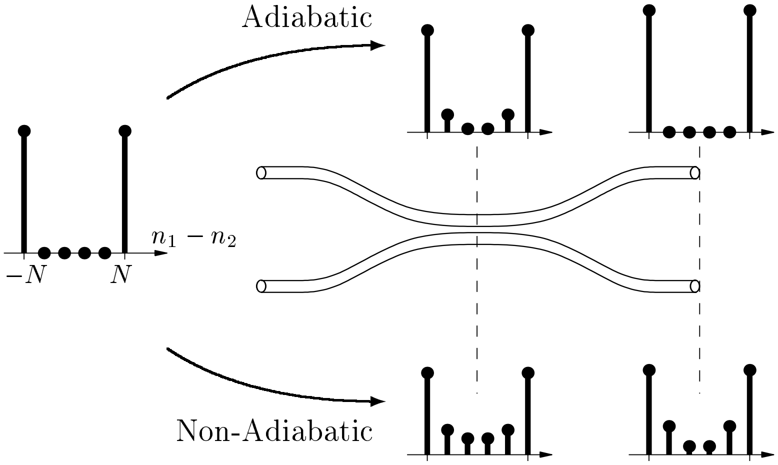


Figure 5-3: Simulations demonstrate the adiabatic principle. An initial, symmetric ground state remains in the instantaneous ground state as long as the Hamiltonian of the NLDC is slowly (adiabatically) varied. If the variation is too rapid, untrapped states are excited, and the photon correlations are largely lost. We depict states with $N = 8$ total photons by plotting the probabilities $|\langle n_1, n_2|\Psi\rangle|^2$ *vs.* $n_1 - n_2$. The input, center, and output of a single NLDC are shown left to right for the adiabatic and nonadiabatic cases.

5.2.2 Basic results

We first show adiabatic manipulation of eigenstates. The lowest symmetric and antisymmetric states were manipulated without giving rise to unwanted states by slowly turning on the linear coupling, as shown in Fig. 5-3. This is numerical evidence that we can manipulate multi-photon correlated states.

More interesting is the direct confirmation of multi-photon tunneling interferometry. In Fig. 5-4, we show the probabilities $|\langle\Psi|N-k, k\rangle|^2$ for a 5-photon state evolving through an interferometer. Several features are worthy of note. Superposition states of the type discussed above were indeed generated at the output of the first NLDC. Although components with $|n_1 - n_2| \neq N$ are present within each NLDC (while $\mu > 0$), the contribution of unwanted states drops essentially to zero once the linear coupling is turned off. Finally, we see the desired phase sensitivity—complete switching from $|N, 0\rangle$ to $|0, N\rangle$ with a phase shift of only π/N. This indicates that the remarkable

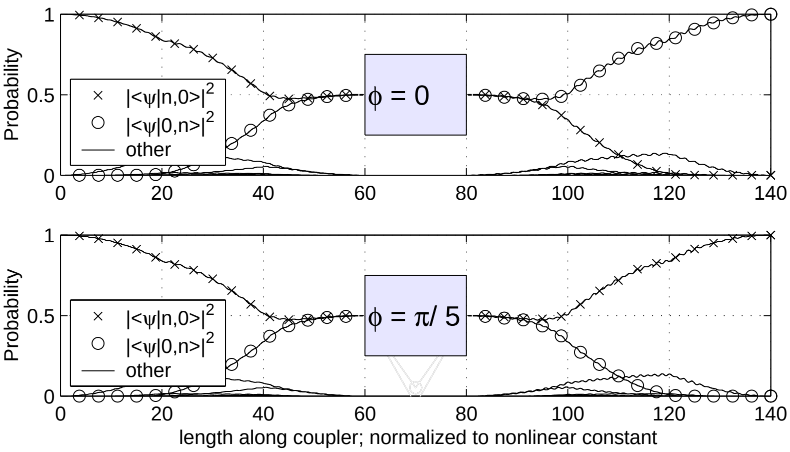


Figure 5-4: The quantum state of a two-mode field $|\Psi\rangle$ evolves according to the first NLDC, phase shifter, and the second NLDC in sequence. Probabilities for the basis states $|N-k, k\rangle$ are plotted versus length along the device, and show striking nonlinear quantum effects. Adiabatic design of the couplers leaves us with superpositions of the $|N, 0\rangle$ and $|0, N\rangle$ only (other states are only temporarily excited within the couplers). Thus we have achieved manipulation of N-photon particles, bound together by the Kerr effect. As predicted, this gives us full switching for a phase shift of only π/N. $\mu_{\max} = .703\hbar$; $T = 20 = 10\hbar/\kappa$; $\kappa = \hbar/2$; $N = 5$.

λ/N interference has been achieved.

5.3 Prospects for measurement

The simulations discussed in the previous section confirm the basic physical principles behind multi-photon tunneling. An actual measurement would differ from the above model for a variety of practical and fundamental reasons. Here, we address a few of the most important differences.

5.3.1 Number uncertainty

Previous simulation of the NLDC showed tunneling for fixed-N, but suggested that uncertainties in N essentially eliminated the effect [11]. The basic issue is that tunneling rates are extremely sensitive to system parameters, including the photon number. A design optimized for $N = N_0$ may show essentially no tunneling for $N = N_0 + 1$.

To understand the impact of number uncertainty on tunneling, we first consider some general properties of number-conserving systems. Conservation of photon number means that the Fock sectors (state components with different photon number) evolve separately. In our lossless model, that is, $|\Psi(t)\rangle = \sum_N |\Psi_N(t)\rangle$ and each Fock component evolves independently,

$$i\hbar \frac{\partial}{\partial t} |\Psi_N(t)\rangle = \hat{H} |\Psi_N(t)\rangle \tag{5.13}$$

since the Hamiltonian does not change photon number. Similarly, a photodetector measurement is performed by an operator $\hat{D}$ that does not mix photon number components, $[\hat{D}, \hat{n}_1 + \hat{n}_2] = 0$. An interesting consequence is the expectation of such a measurement:

$$\langle \Psi(t)|\hat{D}|\Psi(t)\rangle = \sum_N \langle \Psi_N(t)|\hat{D}|\Psi_N(t)\rangle. \tag{5.14}$$

This equation expresses the independence of the Fock sectors: Since state components with different photon numbers do not mix, mean photocurrents are simply a statistical average of the means for each Fock sector. Thus any physical effect present for a fixed photon number is also present for the uncertain-number case, but may be obscured by the averaging.

The simulation results [11] correctly pointed out this washing out of tunneling when several Fock sectors are averaged. One should not misinterpret the result to mean that tunneling is not present in the mixed-N case. In fact, for $\mu \neq 0$, all localized states in *all* Fock sectors either tunnel or undergo untrapped oscillations. The expectation $\langle \hat{n}_1 - \hat{n}_2 \rangle$ is a sum of oscillating, zero-mean contributions (5.14) from the Fock sectors. It will never be *permanently* localized, no matter what parameters are chosen. However, since the tunneling rate has a rapid exponential fall-off with photon number, tunneling may be effectively zero for any reasonable time scale for photon number above some threshold. The localization ($\langle \hat{n}_1 - \hat{n}_2 \rangle > 0$), is then effectively permanent in such cases.

Our simulations confirm this basic intuition. The tunneling rate is indeed seen to be very sensitive to photon number. The variation in tunneling rate leads to loss

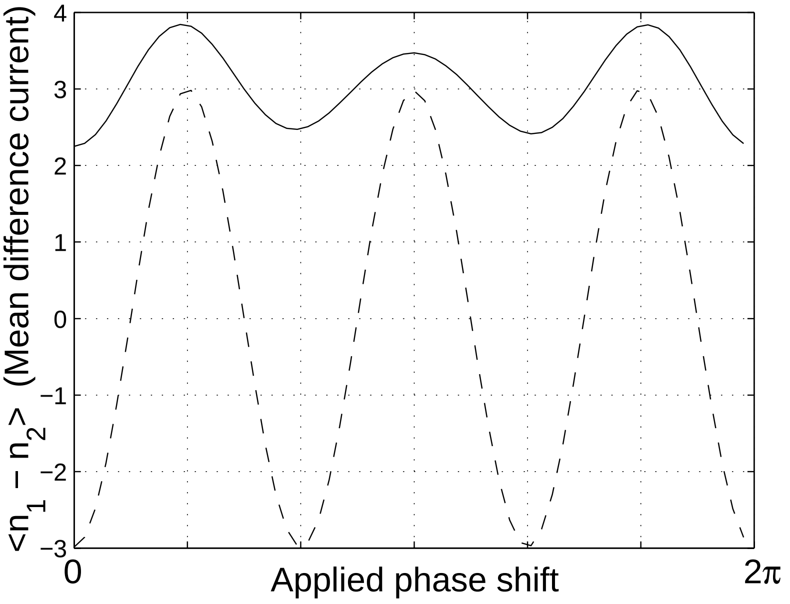


Figure 5-5: The mean difference current, proportional to $\langle \hat{n}_1 - \hat{n}_2 \rangle$, is plotted here as a function of the phase shift of the interferometer. The dashed curve results from an input field with exactly 3 photons, and has the ideal λ/N-fringe shape. The solid curve results from an input coherent state with $\langle \hat{n}_1 + \hat{n}_2 \rangle = 3$. The fringes are apparent despite significant loss of contrast, due primarily to components with higher photon number that tunnel with very low probability. $\mu_{\max} = .526\hbar$; $T = 3 = 1.5\hbar/\kappa$; $\kappa = \hbar/2$.

of contrast in an interference measurement. For $\langle \hat{n} \rangle = 3$, Poisson number statistics leave the desired interference effects intact, albeit with reduced contrast. A plot of $\langle \hat{n}_1 - \hat{n}_2 \rangle$ *vs.* $\Delta\phi$ is shown in Fig. 5-5 and clearly shows the effect we described in Fig. 2-1. For larger n, the loss of contrast due to uncertain number may wash out fringes almost entirely. This is an important issue to be addressed in future work.

5.3.2 Loss and detection

It is well known that correlated superposition states are extremely sensitive to loss [86, 60, 18]. In particular, the completely correlated state (2.16) is destroyed by loss of even a single photon. Other correlated states may have somewhat less strict requirements, but the relevant length for decoherence still seems to be of order $1/\alpha n$, where α is the material loss rate. Unless some clever design is found, the per-photon nonlinearity must be larger than the loss. This may ultimately prevent any observation of photonic DeBroglie waves using the multiphoton tunneling concept. We discuss this issue in the following chapter.

5.3.3 Other practical issues

To actually observe these effects, they must appear in a real physical system, not a two-mode model. Many details of the design must be fleshed out, perhaps most importantly the pulse-shape dynamics. We have considered soliton and zero-dispersion implementations, both of which have unresolved issues. In the soliton case, one must make the two-mode dynamics self-consistent with the pulse shape: that is, the soliton shape in each waveguide depends on the splitting of energy between the two. For the zero-dispersion case, one must carefully identify the relevant time scale over which photons interact, which may differ from the pulse width [40].

5.4 Spin representations and tunneling

In some respects, a spin representation is much more natural description of the nonlinear coupler. The extensive literature on the mathematical mode we have been discussing is primarily focussed on spin variables, and connecting with these results is important. The spin representation uses variables quadratic in the photon operators to describe the two mode state. The link is intimately connected with the Jones-space and Stokes-space connection described in the later chapters of this thesis.

The three-dimensional, Hermitian spin vector has coordinates

$$\hat{J}_x = \frac{1}{2}(\hat{a}_1^\dagger \hat{a}_2 + \hat{a}_2^\dagger \hat{a}_1) \tag{5.15}$$

$$\hat{J}_y = \frac{1}{2i}(\hat{a}_1^\dagger \hat{a}_2 - \hat{a}_2^\dagger \hat{a}_1) \tag{5.16}$$

$$\hat{J}_z = \frac{1}{2}(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2). \tag{5.17}$$

We can easily prove the spin commutation relations,

$$[\hat{J}_x, \hat{J}_y] = i\hat{J}_z \quad \text{(and circular permutations)} \tag{5.18}$$

$$\hat{n} = \hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2 \tag{5.19}$$

$$[\hat{n}, \hat{J}_k] = 0 \tag{5.20}$$

directly from $[\hat{a}_j, \hat{a}_j^\dagger] = 1$. The Hamiltonian (5.1) can now be written in the elegant form

$$\hat{a}_1^\dagger \hat{a}_2 + \hat{a}_2^\dagger \hat{a}_1 = 2\hat{J}_x \tag{5.21}$$

$$\hat{a}_1^\dagger \hat{a}_1^\dagger \hat{a}_1 \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2^\dagger \hat{a}_2 \hat{a}_2 = 2\hat{J}_z^2 + \hat{N} + \hat{N}^2/2 \tag{5.22}$$

$$\hat{H}_{field} = -2\mu(t)\hat{J}_x - 2\kappa \hat{J}_z^2 - \hat{H}_0. \tag{5.23}$$

Here $\hat{H}_0 = -\kappa(\hat{N}+\hat{N}^2/2)$ depends only on the total photon number, and therefor does not relate to photon switching between the waveguides. The energy of localization, or binding energy, is now clearly just the term $-2\kappa\hat{J}_z^2$.

It is useful to think of the spin trajectory $\vec{J}(t)$ that one could define for a classical state, where the photon operators take on definite values. The magnitude of the spin vector is $n/2$, and so the trajectory remains on the surface of a sphere. The trapping regime can be understood by picturing the intersection of this sphere (which represents number conservation) with the surface $E_0 = -\mu n^2/2 - 2\mu J_x - 2\kappa J_z^2$ (which represents conservation of energy). If κJ_z^2 is initially large enough, then the intersection of these surfaces (and thus the state trajectory) is localized around the initial J_z [78]. The corresponding quantum system has been mapped onto a standard 1D tunneling problem, as depicted in Figure 5-6.

Tunneling rates: We have reproduced the calculations outlined in the literature of tunneling rates. We have found that they are not only generally quite small, but become a smaller and smaller fraction of the nonlinear phase rate as we move to higher photon number. In Figure 5-7, we show some calculated tunneling rates normalized

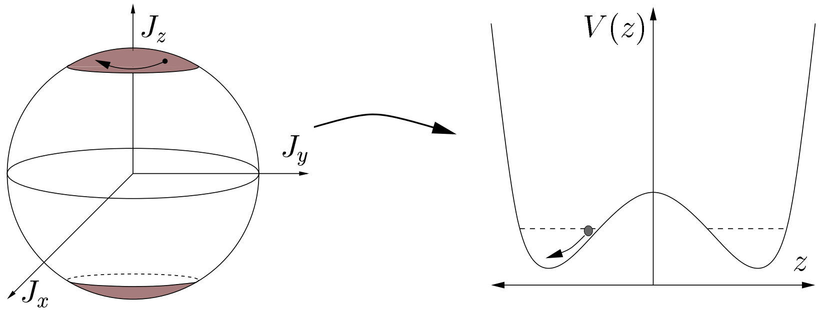


Figure 5-6: The quantum spin system describing multi-photon tunneling can be mapped onto a 1D particle problem using characteristic functions. Here we depict the three-dimensional spin representation and the double well of the equivalent single-particle Schrödinger equation derived by Scharf [69].

to the nonlinear phase rate, that is with $\kappa n = \hbar$. Thus, in a large-n design, we may be pushed to device lengths even longer than one would expect from the simple adiabatic argument presented above.

The range of moderate tunneling rates is shown in Fig. 5-8, and consists of very low photon number and μ near the critical value for trapping. Both of these conditions seems necessary to avoid prohibitively small tunneling rates. Unfortunately, in this regime, most other states of the system are untrapped. Any scattering into these states will cause a background of non-tunneling oscillation between the waveguides, and may obscure any interesting effects.

5.5 Adiabatic condition for many photons

Above we argued that for small n and μ, the adiabatic condition can be expressed as

$$\kappa n \Delta t / \hbar \gg 1. \tag{5.24}$$

This has intuitive appeal: it says simply that the nonlinearity must be significant on the length scale of the coupler, or else the switching behavior will be quasi-linear, and thus classical. However, this simplistic expression can be misleading as we move to more relevant n and μ. In this section, we discuss the adiabatic condition more

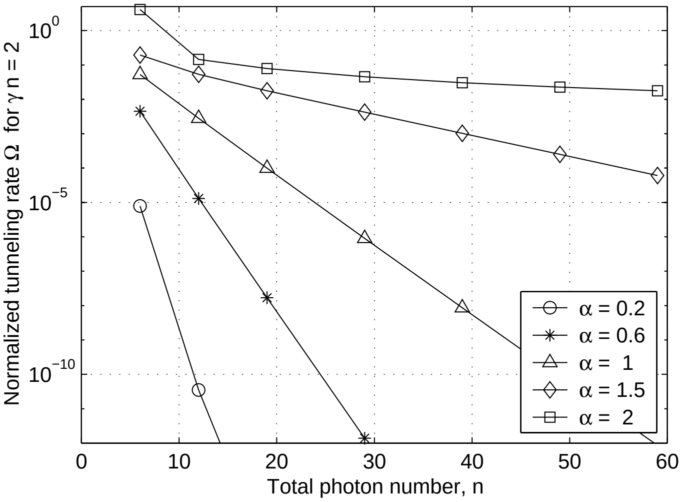


Figure 5-7: Tunneling rates fall off exponentially with photon number n in the trapped regime. Several values of $v = 2\mu/\hbar$ are shown, and $\kappa n = \hbar$. Only weakly trapped systems, $v \approx v_{\text{crit}} = 2$, have a polynomial Ω *vs.* n relation.

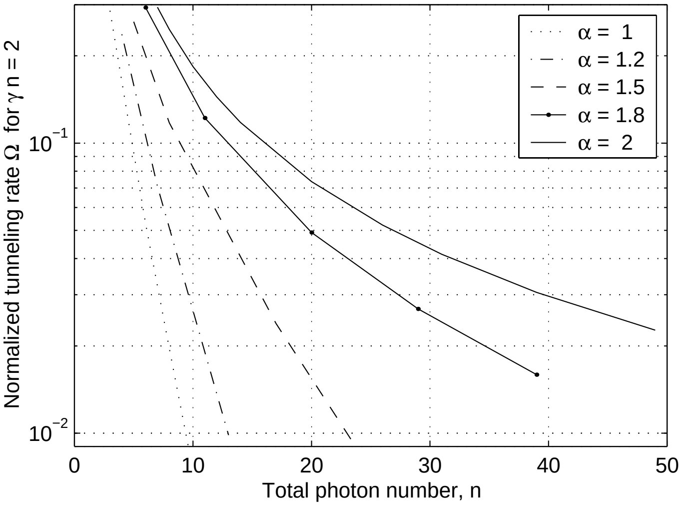


Figure 5-8: Tunneling rates in the potentially useful weak-trapping regime. Again, $\kappa n = \hbar$ and the symbol α represents $2\mu/\hbar$.

carefully. We argue that, regardless of the photon number, the total interaction time required to observe adiabatic tunneling is on the order of $\kappa/\hbar$. That is, the coupler length must be on the order of the nonlinear phase rate *per photon*. Typically this length is extremely large due to weak nonlinearities of optical materials.

Balancing the tunneling length and the adiabatic length

The total length of an adiabatic tunneling coupler must be understood in terms of the balance of two requirements: switching cannot occur faster than the tunneling rate, or faster than the adiabatic limit. In the calculation of Figure 5-4, we took μ to be a fixed, small constant, and then adjusted T until it matched one quarter of the effective tunneling period. That is, we adjusted T until the state had high overlap with the desired superposition state,

$$|\psi\rangle \approx |s\rangle \equiv (|n,0\rangle - i|0,n\rangle)/\sqrt{2}. \tag{5.25}$$

But of course μ need not be some arbitrary fixed constant. In fact, we can achieve larger and larger tunneling rates by raising μ. Thus we can reach $|s\rangle$ using smaller and smaller values of T by increasing μ.

However, the adiabatic condition says that the coupling, proportional to μ, must not vary too quickly. As μ is increased, the adiabatic length becomes longer. At some point, μ will become so large that the adiabatic length is longer than the tunneling period. Since T must satisfy both, raising μ does not result in a shorter device beyond this point. The optimum choices of μ and T will occur when the adiabatic length and the tunneling length are of the same order. We have gone through the optimization process numerically for several values of n and will discuss the results below.

The weakly trapped regime

We have discussed in the preceding section that the tunneling rate is vanishingly small outside of a narrow range of parameters. For more than a few photons, the rate is essentially zero for small μ and then suddenly turns on around the critical trapping condition

$$\mu \approx \mu_{\text{crit}} = \frac{\kappa n}{2}. \tag{5.26}$$

Thus the "balance" of tunneling and adiabatic lengths occurs at around this value. By design, then, we will always want to operate in the weakly trapped regime (or untrapped regime).

Adiabatic condition for weakly trapped states

The rule of thumb (5.10) that we used for the adiabatic condition for small n and μ does not extend beyond this regime. The correct condition is

$$\Delta t \gg \frac{\hbar \mu_{\max} \langle b | \hat{J}_x | a \rangle}{\Delta E_{ab}^2}, \tag{5.27}$$

and is expressed in terms of scattering from state a to state b. We can then obtain a crude but useful estimate for the adiabatic length in the weakly trapped regime. We assume $\Delta E_{ab} \approx 2\kappa n$ as before, but now argue that

$$|\langle b | \hat{J}_x | a \rangle| \sim n/2. \tag{5.28}$$

To understand this, consider the spin-representation of the quantum state. One might naively assume $|a\rangle$ is equal to the initial state $|n, 0\rangle$ and obtain $|\langle b | \hat{J}_x | a \rangle| \sim n/2 \approx \sqrt{n}/2$, as is true for small μ. This is the n-photon quantum state that best approximates $\vec{J} = (n/2)\hat{i}_z$, and therefore gives small values of $\langle a | \hat{J}_x^2 | a \rangle$, and thus of $|\langle b | \hat{J}_x | a \rangle|$ among possible quantum states. But in the weakly-trapped regime, the initial state is transformed into a ground state $|a\rangle$ which has very little resemblance to $|n, 0\rangle$. The direction of $\vec{J} = (n/2)\hat{i}$ now deviates significantly from $\hat{i}_z$, and a significant component lies in the x-direction. The component is then $\sim n$, as indicated by Eq. (5.28).

The adiabatic condition is

$$\Delta t \gg \frac{\hbar \mu_{\max}}{\kappa^2 n}. \tag{5.29}$$

Combining this with the trapping condition, Eq. (5.26), we have

$$\Delta t \gg \frac{\hbar}{2\kappa}. \tag{5.30}$$

This result is quite striking. If it is correct and applies generally to all relevant regimes, it means that the device length for adiabatic coupling must be greater than the single-photon nonlinear length. This a very strict condition, raising serious practical difficulties in addition to loss decoherence. For example, even if it could be made lossless, a standard fiber would need to be millions of kilometers long to obey the above constraint.

Numerical Results

A series of NLDC simulations were performed to confirm the above over-simplistic arguments. For photon numbers between $n = 3$ and $n = 12$, we basically saw exactly the expected trends.

In Figure 5-9, we see the coupler length as a function of μ. A value of the total coupler length $3T$ was obtained for each μ by minimizing the deviation from the desired state $|s\rangle$,

$$R = 1 - |\langle\psi|s\rangle|^2. \tag{5.31}$$

$$T(\mu_{\max}) = \arg\min_{T} R(\mu_{\max}, T). \tag{5.32}$$

This optimum length essentially reflects the effective tunneling rate of the piecewise-linear $\mu(t)$. For all simulated cases, it drops monotonically with μ. However, when μ becomes too large, deviations from adiabaticity become larger. Figure 5-10 shows the deviation R as a function of $\mu_{\max}$. Generally the overlap gets progressively worse for higher μ.

We should mention that all of the optima (5.32) were forced to correspond to the first tunneling period $T(\mu) \approx \pi/2\Omega$. For completeness, one could also consider devices where T is greater than a tunneling period, such as $T(\mu) \approx 9\pi/2\Omega$ of Figure 5-11. In these cases, the adiabatic condition is being met by making T larger than

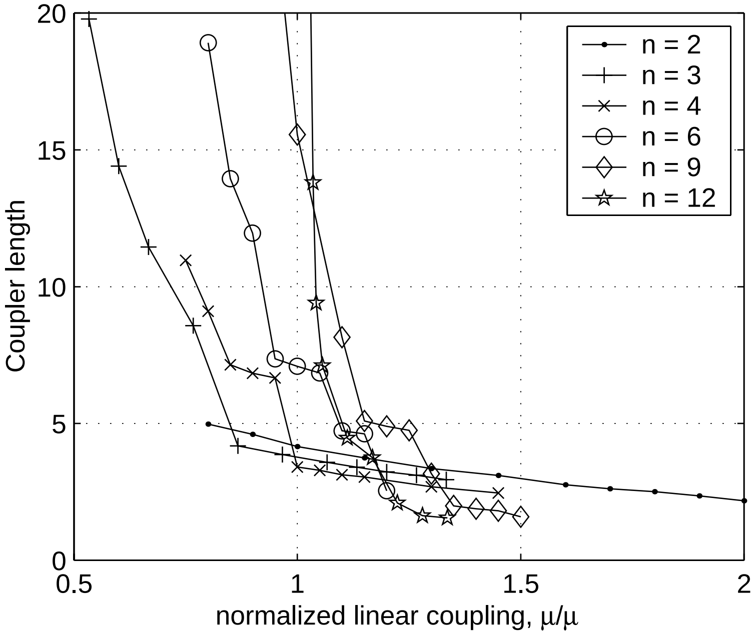


Figure 5-9:

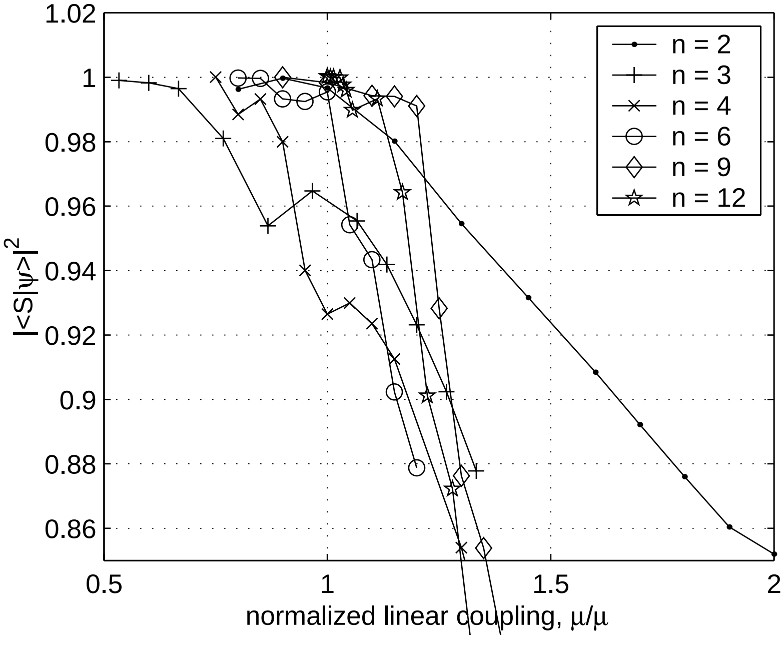


Figure 5-10:

the tunneling period. The fact that this is necessary means that μ is too large: rather than being in "balance," the adiabatic length is too large for the tunneling length.

There are two key features of Figures 5-9 and 5-10 that confirm our theoretical discussion. First, we see that the interesting regime occurs when $\mu \sim \mu_{\text{critical}}$. That is, we will always raise μ by design until we reach some acceptable tolerance for the deviation R, and this seems to always occur at μ_{critical}, as expected. (In fact, as n get larger, we seem to be pushed progressively higher, into the untrapped regime). Second, we see that for a fairly wide range of n's, the ultimate device lengths are always of order $\hbar/\kappa = 2$. There are variations, but no consistent trend with increasing n.

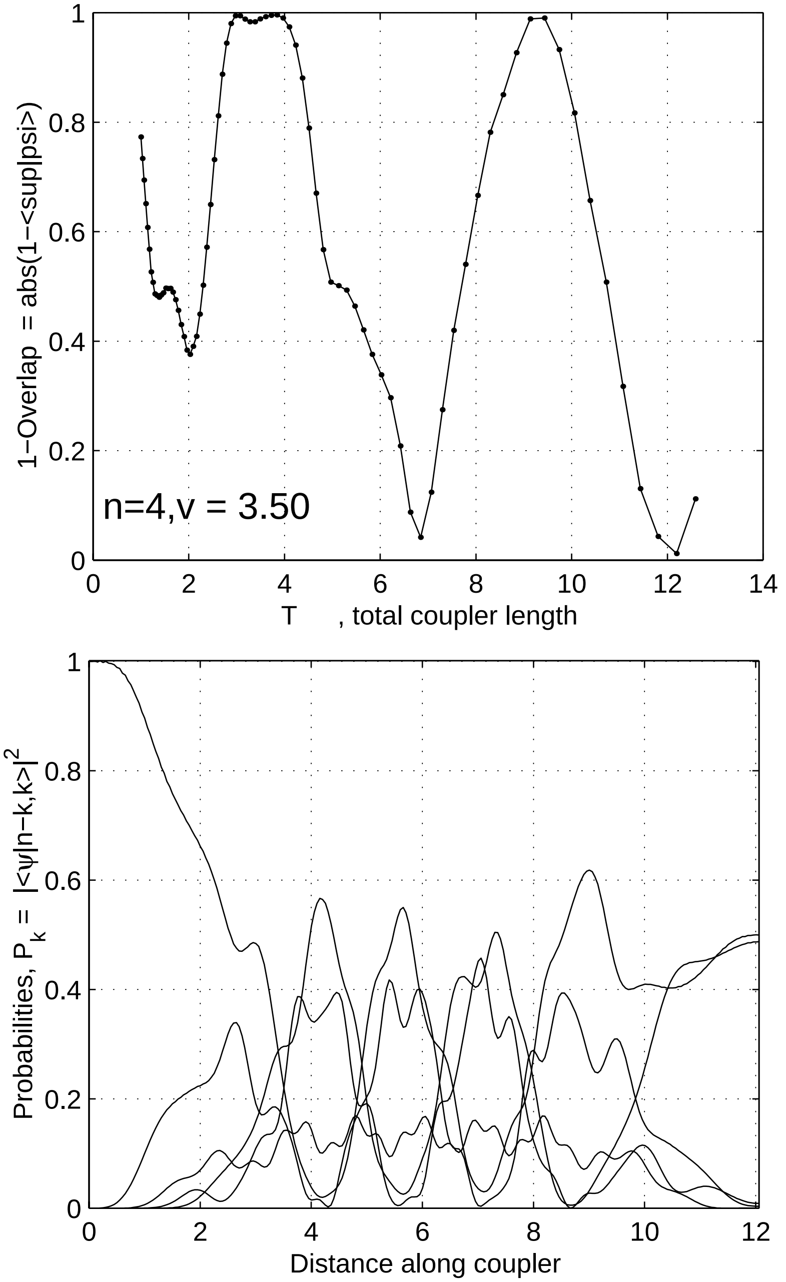


Figure 5-11: For large μ, the tunneling length is shorter than the adiabatic length. To reach the desired superposition state, the coupling length must allow one or more full tunneling oscillations. Successive optima are shown in the plot of overlap vs. T. The third minima corresponds to a coupler with $T \approx 12$ ($\kappa/\hbar \equiv 1/2$), whose state evolution is shown.

Chapter 6

Loss in a quantum interferometer

Classical interference effects are quite robust to loss. If the loss in the arms of a standard interferometer is balanced, the interference pattern is simply attenuated, but qualitatively the same. Even an imbalance simply leads to a proportional reduction of contrast, but is essentially a practical detail as long as some light is getting through each arm.

This situation can be fleshed out in detail with a quantum analysis of a linear interferometer with coherent-state inputs. The detector sees a mode in a coherent state, whose interference properties are exactly those predicted by the classical theory. The quantum impact of loss is simply that there are fewer photons—which matters only when the shot noise becomes important.

Another way of looking at the classical evolution of the coherent state uses the particle intuition of the previous chapters. A coherent state is one where the photons are uncorrelated, each existing in its own mode (or wavefunction) without any interaction with other photons. Thus, when a photon is lost, it simply never reaches the detector. Since it is not correlated or interacting with the other photons, its fate has no impact on them whatsoever. They pass through the (linear) interferometer as if the lost photon had never existed.

The situation is completely different for the states,

$$|\Psi\rangle = \alpha|N,0\rangle + \beta|0,N\rangle \tag{6.1}$$

described in the previous chapter. Here the photons are completely correlated: if one is in the left (or right) arm of the interferometer, then all photons are in that same arm. Any such fully correlated superposition is destroyed by a single loss event, because loss of a photon constitutes a measurement of which arm it was in. In this sense, many-photon superposition states are analogous to states with extremely large squeezing.

Single-photon loss condition

By the above argument, the loss sensitivity for photonic DeBroglie wave effects involving fully correlated states such as Eq. (6.1) is strict and easy to calculate: the loss seen by the fully correlated state must be less than one photon with some reasonable probability. In terms of the propagation length L, we have $\alpha L < 1/n$. As with the momentum squeezer, the propagation length cannot be chosen arbitrarily. It must be long enough to satisfy the adiabatic condition. The loss condition then means that the loss rate must be lower than the per-photon nonlinear phase rate, by at least the factor $1/n$:

$$\alpha v_g < \frac{1}{nT} \sim \frac{\kappa}{n\hbar}. \tag{6.2}$$

We can get a crude estimate of the available physical constants by looking at a simple soliton propagation example; more accurate numbers require an understanding of pulse-shaping dynamics in a nonlinear directional coupler, which is beyond the scope of this paper. The loss length in a high-quality optical fiber is defined as the inverse of the power loss rate

$$L_{\text{loss}} \equiv 1/\alpha \approx 20\text{km}. \tag{6.3}$$

We define the per-photon nonlinear length as characteristic length of accumulation of nonlinear phase *per photon*,

$$L_{\text{NL}} \equiv v_g \hbar / 2\kappa. \tag{6.4}$$

κ is connected to the nonlinear constant of the waveguide γ via the pulse width τ_s, by reasoning that $\hbar\omega n/2\tau_s$ is the peak power. For mode area A_{eff}, material Kerr constant n_2, and frequency typical for fiber solitons [1], one has

$$\gamma = \frac{n_2\omega_0}{cA_{\text{eff}}} \approx 3\text{W}^{-1}\text{km}^{-1}. \tag{6.5}$$

The per-photon nonlinear length for a 1 ps soliton is then

$$L_{\text{NL}} \equiv \frac{8\tau_s}{\hbar\omega\gamma} \sim 2 \times 10^7\text{km}. \tag{6.6}$$

The ratio $L_{\text{NL}}/L_{\text{loss}}$, which must be less than $1/n$ to satisfy the single-photon loss condition, is then

$$L_{\text{NL}}/L_{\text{loss}} \sim 10^6. \tag{6.7}$$

We can easily imagine orders of magnitude improvement by reducing the mode area A_{eff}, using femtosecond pulses, etc. Regardless, adiabatic, loss-free propagation will be difficult or impossible in fiber. Other materials should certainly be considered as well. For example, chalcogenide and GaAs each have nonlinearities orders of magnitude stronger than conventional fiber. The losses in these materials have so far been far greater than for fiber, however, so that there is no net reduction of $L_{\text{NL}}/L_{\text{loss}}$.

Towards robust DeBroglie wave effects

The design as described so far faces a seemingly insurmountable obstacle because of loss. We have begun a process of exploring the basic assumptions and limitations that prevent a realistic design. While we have identified some interesting physics and potential directions for future investigations, we have found no evidence that the obstacle can be overcome. The observation of a photonic DeBroglie effect, if possible at all, is likely to require special sources and detectors, and consists of carefully selected detection events. To some extent, such an experiment would be more akin to photon correlation measurements than to the elegant thought experiments of Figures 2-1 and 3-5.

Below, we discuss the most important generalization of our method that may help mitigate loss effects: the use of more loss-robust states that may still capture the desired strange quantum properties. Another major strategy is the use of special detectors to filter out events that would otherwise wash out the desired interference. Selective detection can help deal with loss, as it has for existing two-photon interferometers. By detecting coincidence rates for two photons, one essentially removes the background generated by cases where one photon is lost. One could thus observe interference of two-photon pulses with high contrast even if a very small fraction of the pulses experience no loss. This concept could be generalized in a number of ways to the multi-photon case, and is a subject of current research. The problem remains that there must be some appreciable probability of the desired events.

6.1 Quantum states less sensitive to loss

Both the quantum effects that we seek to observe and the undesirable loss-sensitivity are properties of the special correlated-photon states generated in the interferometer. We would like to understand the fundamental trade-off between loss-sensitivity and quantum strangeness, apart from the specifics of any particular implementation. In this section, we choose a particular metric for quantum strangeness, so that we can explore the trade-off quantitatively. To simplify our analysis, we do not consider the evolution of quantum states in the interferometer at all. We merely assume that the interferometer generates a quantum two-mode state $|\psi\rangle$ just before the phase shifter in Figure 5-2. We then crudely apply a finite lumped loss to the state, and ask whether the special properties of the state are intact.

Inherent phase resolution of a state

A key motivation for studying quantum interferometer states is their potential for precision phase measurement. Along these lines, the inherent phase resolving power of a quantum state is a useful measure of its desirable quantum properties. We have used resolving power as our quantitative measure of the usefulness of a state.

An exact metric for a state's phase resolving power is available from theory of quantum hypothesis testing [38, 5]. In that framework, one asks: With what certainty can one detect a small phase shift ϕ operating on a quantum state described by density operator ρ? That is, with what probability can one distinguish ρ from the phase-shifted $e^{i\phi\Delta\hat{n}}\rho e^{-i\phi\Delta\hat{n}}$? Leaving practical issues aside, the mathematical answer relates simply to the orthogonality of these states to be distinguished, as measured by the distinguishability, $0 \leq D \leq 1$,

$$D = \frac{1}{2}Tr\{|\rho - e^{i\phi\Delta\hat{n}}\rho e^{-i\phi\Delta\hat{n}}|\}. \tag{6.8}$$

Here $\Delta\hat{n} \equiv \hat{n}_2 - \hat{n}_1$ is the difference of photon numbers for the two paths, and the notation $Tr\{|\ldots|\}$ is a terse shorthand for summing the absolute values of eigenvalues [38]. We apply this known formalism to the problem at hand, generalizing the results of [5].

Comparison

We need to test the robustness of superposition states that have significant but not complete photon correlations, so that we can establish a resolution-robustness tradeoff. We define the set of initial states with definite $|\Delta\hat{n}| = K$ and photon number n:

$$|n, K\rangle = (|(n+K)/2, (n-K)/2\rangle + |(n-K)/2, (n+K)/2\rangle)/\sqrt{2} \tag{6.9}$$

These are the symmetric eigenstates of the nonlinear coupler in the limit of zero linear coupling (in the adiabatic coupler, theses are eigenstates after the coupling has been switched off), and are a natural starting point for our investigation. We expect that small K states have a large background in both modes and are thus more robust, while large K states have higher sensitivity to phase in the absence of loss.

The basic intuition is that, as long as each component of the superposition has some background photons in both paths, then all components will survive moderate losses. In the limit that photon number of the arms is approximately equal, loss no longer constitutes a measurement of the state, and no longer significantly disrupts the superposition. However, in this same limit, the phase resolving ability vanishes completely, since the quantum phase transformation is proportional to $\phi(\hat{n_1} - \hat{n_2})$. Using the formalism below, we confirm and quantify the trade-off between raw resolving power and robustness.

In our results section below, we present the calculation for this particular family of quantum states. This gives us a sense for what might be possible in a nonlinear interferometer without actually calculating the state evolution in the presence of loss. Estimates more applicable to a specific interferometer implementation can be calculated using the same method but starting with a more appropriate family of states. Naturally, this approach can give us an upper bound on performance, but does not address issues such as the adiabatic transition that would be included in a full simulation.

Beamsplitter loss model

Loss is included as if there were a beamsplitter in each arm of the interferometer. The formulation in terms of four-mode states is straightforward. If the reservoir is initially in the ground state,

$$|\psi_i\rangle = \sum_{n_1,n_2} c_{n_1,n_2} |n_1, n_2\rangle \otimes |0, 0\rangle_{\text{reservoir}}. \tag{6.10}$$

then the final state can be written in terms of the well-known binomial coefficients of a beamsplitter discussed, for example, in [34, Sec. 9.2]

$$|n\rangle \otimes |0\rangle \to \sum_j L_{n,j} |n-j\rangle \otimes |j\rangle, \tag{6.11}$$

$$|L_{n,j}|^2 = \left(\begin{array}{c} n \\ j \end{array}\right) \mathcal{L}^{n-j} (1-\mathcal{L})^j. \tag{6.12}$$

The final state after loss is

$$|\psi_f\rangle = \sum_{n_1,n_2} c_{n_1,n_2} \sum_{j_1,j_2} L_{n_1,j_1} L_{n_2,j_2} |n_1 - j_1, n_2 - j_2\rangle \otimes |j_1, j_2\rangle_{\text{reservoir}}. \tag{6.13}$$

This has the form of a statistical mixture of states $|\bar{\psi}_{j_1,j_2}\rangle$:

$$|\psi_f\rangle = \sum_{j_1,j_2} |j_1, j_2\rangle_{\text{reservoir}} \otimes |\bar{\psi}_{j_1,j_2}\rangle \tag{6.14}$$

$$|\bar{\psi}_{j_1,j_2}\rangle \equiv \sum_{n_1,n_2} c_{n_1,n_2} L_{n_1,j_1} L_{n_2,j_2} |n_1 - j_1, n_2 - j_2\rangle. \tag{6.15}$$

That is, the density operator of the interferometer state (the partial trace over the reservoir) has pure-state components $|\bar{\psi}_{j_1,j_2}\rangle$,

$$\rho \equiv \mathrm{Tr}_{\mathrm{R}}\{|\psi_f\rangle\langle\psi_f|\} = \sum_{j_1,j_2} |\bar{\psi}_{j_1,j_2}\rangle\langle\bar{\psi}_{j_1,j_2}|. \tag{6.16}$$

6.1.1 Results

Figure 6-1, plots distinguishability versus ϕ for a few example states, clearly demonstrating the expected trends: For $n = 63$ and no loss, the optimal phase resolution is achieved by $|n = 63, K = 63\rangle$ [which is a fully correlated state (2.16)]. However, after 5% power loss in each path of the interferometer, the performance of this state is badly degraded. The $K = 15$ state at the same level of loss still reaches near-unity distinguishability ($D = .91$, corresponding to a theoretical detection error of 5%). In fact, because of robustness to loss, the $K = 15$ state gives dramatically better probability of detection for *all* phase shifts. The degradation of correlated states is more dramatic as we move to higher n and larger losses. Also shown is the performance of a "beamsplitter state," the state emerging from a 50/50 beamsplitter with 63 photons input in one port. This curve gives the performance of uncorrelated photon states relevant to a linear interferometer. The loss-degraded $K = 15$ state outperforms the beamsplitter state, demonstrating that the shot-noise limit can be surpassed even in the presence of moderate losses.

Analytical limit

Ultimately, we would like an estimate of the minimum resolvable phase $\Delta\phi$ as a function of the power loss level $\mathcal{L}$. Numerical calculations involving density matrices

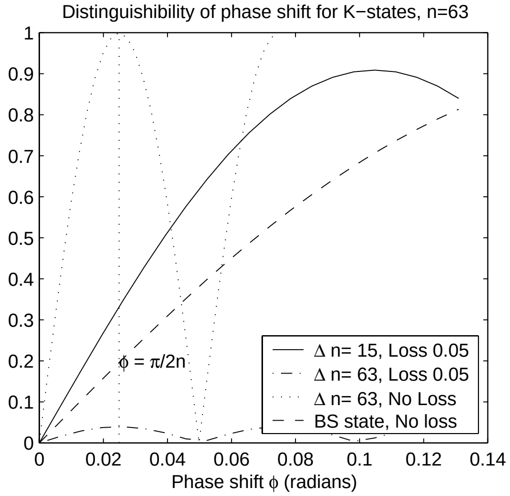


Figure 6-1: Distinguishability of phase shift ϕ for several quantum states. Total photon number before loss, $n = 63$. The "K-states" have characteristic $|\Delta n| = |\hat{n}_1 - \hat{n}_2| = K$ and represent either fully correlated ($K = 63$) or partially correlated ($K = 15$) quantum states. Shown for comparison is the "beamsplitter state," the output of a conventional beamsplitter with n input photons.

are impractical for large n. We can use calculations for moderate n to test analytical approximations. Following the spirit of [21], we have estimated the effect of finite loss $\mathcal{L}$ as a simple combination of $n\mathcal{L}$ individual photon loss events in the limit of small K/n. This simple intuitive estimate agrees surprisingly well with the exact calculation.

Consider the orthogonality properties of a phase shift applied to $|n, K\rangle$,

$$\langle n, K|e^{i\phi\Delta\hat{n}}|n, K\rangle = \frac{1}{2}(e^{i\phi K} + e^{-i\phi K}). \tag{6.17}$$

If the phase shift $\phi = \pi/2K$ is chosen, then the shifted and unshifted states are orthogonal,

$$\langle n, K|e^{i\phi\Delta\hat{n}}|n, K\rangle = \cos(\pi/2) = 0 \tag{6.18}$$

and the distinguishability is unity,

$$D = (1 - |\langle n, K|e^{i\phi\Delta\hat{n}}|n, K\rangle|^2)^{1/2} = 1. \tag{6.19}$$

That is, the phase shift can in principle be resolved with probability one using such an initial state, since the shifted and unshifted states are orthogonal.

Now consider the same state subjected to single photon loss:

$$|\bar{\psi}\rangle = \hat{a}_1|\psi\rangle = \sqrt{\frac{n+K}{2}}\left|\frac{n+K-2}{2}, \frac{n-K}{2}\right\rangle + \sqrt{\frac{n-K}{2}}\left|\frac{n-K-2}{2}, \frac{n+K}{2}\right\rangle \tag{6.20}$$

The phase-shift no longer produces perfect orthogonality:

$$\frac{\langle\bar{\psi}|e^{i\phi\Delta\hat{n}}|\bar{\psi}\rangle}{\langle\bar{\psi}|\bar{\psi}\rangle} = \frac{1}{n}\left(\frac{n+K}{2}e^{i\phi(K-1)} + \frac{n-K}{2}e^{-i\phi(K+1)}\right) \tag{6.21}$$

$$\left|\frac{\langle\bar{\psi}|e^{i\phi\Delta\hat{n}}|\bar{\psi}\rangle}{\langle\bar{\psi}|\bar{\psi}\rangle}\right|^2 = |\cos(\phi K) + (iK/n)\sin(\phi K)|^2 \tag{6.22}$$

Evaluated again at $\phi = \pi/2K$, the overlap is simply K^2/n^2, and the distinguishability is

$$D = \sqrt{1 - K^2/n^2}. \tag{6.23}$$

We extend this result to obtain our simple analytical estimate. We assume $1 \ll K \ll n$, since in this regime one can achieve loss robustness. To first order in K/n, assume that each lost photon reduces the distinguishability by the factor calculated above

$$\frac{D_{\text{final}}}{D_{\text{initial}}} = \sqrt{1 - K^2/n^2}. \tag{6.24}$$

Now extrapolate the effect of $\mathcal{L}n$ loss events as a simple product of the factors for each photon

$$D \approx \left(1 - \frac{K^2}{n^2}\right)^{\mathcal{L}n/2} \approx e^{-K^2\mathcal{L}/2n}. \tag{6.25}$$

For moderate losses, the resolvable phase is again connected to K by $\Delta\phi = \pi/2K$. If we assume distinguishability must remain above some threshold D_{thresh}, we can then relate the loss $\mathcal{L}$ to a maximum value of K, and thus a minimum resolvable phase

$$\Delta\phi^2 = \frac{\mathcal{L}\pi^2}{8n\ln(1/D_{\text{thresh}})}. \tag{6.26}$$

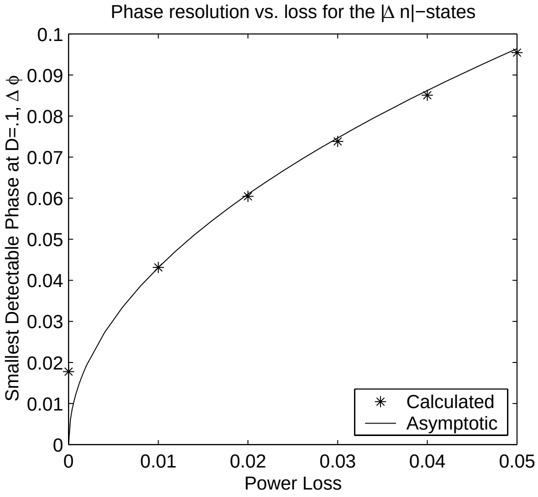


Figure 6-2: Phase resolution achievable as a function of power loss, using k-states. We see good agreement between the calculated values and the asymptotic curve. We have used $n = 63$, $D_{\text{tol}} = 0.9$

The theory is validated by the comparison with numerics shown in Figure 6-2. We expect this trend to be useful in designing interference devices, even when more general states are considered.

Detection methods

While the inherent phase resolution of a state is interesting, we must keep in mind that building an appropriate detection system may be difficult. Distinguishability guarantees that a measurement is mathematically possible, not that a realistic measurement device can be found. Multi-photon tunneling devices present one solution to this problem, by providing a phase-dependent switching characteristic. For a very small number of photons, it is also possible to more exhaustively project out components of the state space—for $n = 2$, one can use coincidence detectors or entanglements with matter; the two-mode, two-photon Fock space is three dimensional and therefore somewhat manageable. The task of constructing general unitary transformations or n-photon correlated measurements for large n is daunting for large n.

6.2 Prospects for a measurement, revisited

The trends identified in the preceding sections can help us understand variations of the NLDC-based interferometer. For example, we consider an adiabatic tunneling coupler that produces a $\Delta n = K$ state at the phase modulator $|n, K\rangle$ with $\sqrt{n} < K \ll n$. We might again use an adiabatically-varying Kerr-coupler to achieve this state, where a (weakly) left-localized state will evolve into either a left or right-localized state, depending on the phase shift. The eigenstructure of such a system is exactly that discussed in the previous chapter, except that now, many states of the system must remain trapped as the linear coupling is turned on, not just the ground state. One way to think about this is that, as the linear coupling is turned on and off, the nonlinearity must be strong enough to have a significant effect not only on the completely localized states, but on $|n, K\rangle$ as well. Otherwise, the coupler is essentially linear, and cannot have any photonic DeBroglie effects. The Kerr index shift due to K photons must thus be significant on the distance scale of the device. We will again assume that μ is on the order of the trapping condition, but now it is the trapping condition for the K-states:

$$2n\mu \approx \kappa K^2. \tag{6.27}$$

Let us now compare two potential designs: $K \ll n$ and $K = n$. Assuming the same nonlinearity κ and photon number n, the more weakly correlated states will need a much longer propagation to accumulate the same nonlinear phase. However the losses that one can tolerate are larger, since the background photons in each arm prevent decoherence when only a few photons are lost. According to our simple approximation (6.25), the acceptable loss goes from $\mathcal{L} \sim 1/n$ in the $K = n$ case to $\mathcal{L} \sim n/K^2$—up by the factor $(n/K)^2$.

Similarly, we can extend the adiabatic analysis of Section 5.5 to the case $K \neq n$,

$$\Delta t \gg \frac{\hbar \mu_{\max} \langle b|\hat{J}_x|a\rangle}{\Delta E_{ab}^2} \tag{6.28}$$

We again use $\Delta E_{ab} \approx 2\kappa K$ and the approximation (now even more easily justified)

$$\langle b|\hat{J}_x|a\rangle \sim n/2. \tag{6.29}$$

Then

$$\Delta t \gg \frac{\hbar \mu_{\max} n}{\kappa^2 K^2}. \tag{6.30}$$

Along with the trapping condition (6.27) the adiabatic condition becomes

$$\Delta t \gg \frac{\hbar}{2\kappa}. \tag{6.31}$$

Thus, the use of K-states does not change the adiabatic condition. This presents serious practical difficulties, as mentioned in the previous chapter.

We conclude that the severe loss requirements cannot be avoided using K-states. The coupler length in this case is still of order $\hbar/\kappa$, and so the loss goes as $\mathcal{L} \sim \alpha/\kappa$. To make this factor less than one in a real system is a seemingly impossible task. But our ability to see DeBroglie effects falls off exponentially with $\mathcal{L}K^2/n$. Thus either $\mathcal{L}$ or K^2/n must be small. One might imagine that very small values of K can be used, so that $K^2/n < 1$. But a state with $K = \Delta n \sim \sqrt{n}$ is not an exceptional quantum state at all. The "superposition" in this state consists of Δn values within the normal coherent-state range, and various problems arise with measurement. In this limit, it makes more sense to set up a squeezed-state measurement than try to draw some loose resemblance to tunneling.

There remain several avenues for investigation. One direction is suggested by Hagelstein's earlier analysis of these device. This analysis was done using configuration-space methods: Rather than start with a second-quantized model, and proceeding with an analysis of spin variables, the configuration space starts with photon wavefunctions and obtains tunneling behavior using different (but presumably related) approximations. Similar effects were observed for the NLDC as well as a Y-coupler, although no careful comparison has yet been done. The configuration space results suggest that greatly improved performance might be possible if the linearly coupling

$\mu(t)$ is not ramped up as a piecewise-linear function of time. A smooth time variation may show significant improvement.

In addition, many of the trends discussed above should still be confirmed by testing our approximations and comparing with numerics. The feasibility analysis we have tried to summarize here involves more than simply one number $L_{\mathrm{NL}}/L_{\mathrm{loss}}$, and various parts of the argument should be checked for weaknesses. The loss tolerance is also based on a crude theory, extrapolated well beyond the region where it was tested numerically. This theory should be tested in a much more systematic way to explore the full space of n, K, and $\mathcal{L}$ that give numerically tractable problems.

Chapter 7

Quantum polarization systems

So far, we have focussed on a particular model of a photonic DeBroglie-wave interferometer, assuming two waveguide modes with small evanescent field overlap. The basic formalism that we are using applies more broadly to any two-mode bosonic field. We have already mentioned this with reference to the literature on molecular spin tunneling and other mathematically related problems. Another implementation raises some interesting issues: that of two orthogonal polarizations of light in a single-mode waveguide. From the point of view of second-quantization, this case is no different. Having identified two modes, we perform the quantization according to the usual recipe and obtain equivalent bosonic operators $\hat{a}_1$ and $\hat{a}_2$. A model with the same basic form as Eq. (5.1) comes out, and is analyzed using the same kinds of methods.

7.1 Spin variables and polarization fields

We introduced the spin operators (5.15-5.17) as a natural set of variables for describing interferometer states. We also saw that a geometrical representation in three-dimensional spin-vector space was a useful way of visualizing certain aspects of the physics. In fact, this spin vector space should seem familiar to those interested in classical polarization-sensitive systems. In the classical limit, the spin vector simply becomes the Stokes vector, a common representation of polarization,

$$\vec{J} = \frac{1}{2} \begin{bmatrix} \hat{a}_1^\dagger & \hat{a}_2^\dagger \end{bmatrix} \vec{\boldsymbol{\sigma}} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix} \longrightarrow \vec{s} = \boldsymbol{a}^\dagger \vec{\boldsymbol{\sigma}} \boldsymbol{a}. \tag{7.1}$$

Here, a compact vector notation is used, where $\vec{\boldsymbol{\sigma}}$ is a vector of Pauli spin matrices and $\vec{J} \equiv [\hat{J}_x \ \hat{J}_y \ \hat{J}_z]^T$ is a vector of quantum operators (despite not having a hat). The spin description is described further in Appendix A. The classical Stokes-vector representation will be used extensively when we discuss PMD in the remaining chapters.

One aspect of this spin representation of polarization may seem strange: in the single-photon case, the $\hat{J}_k$ operators reduce to two-by-two matrices, characteristic of spin-1/2 systems. It is well known that photons are spin one particles, but that a photon with a given propagation vector $\vec{k}$ has only two possible orthogonal polarizations. This apparent inconsistency has been discussed much better elsewhere. We provide some comments on the subject in Appendix A.

Polarimeter measurements

In a classical system, it is possible to measure the Stokes vector using a combination of beam splitters, waveplates, and photodetectors. For example, one might split off three different copies of the field with beamsplitters, and measure the power along various polarizations. Naturally, this type of measurement breaks down for a quantum field with few photons or correlated photons (we cannot split off "copies" of a photon pair with an arrangement of beamsplitters).

The problem is not in the specific setup; the spin operators do not commute and therefore fundamentally cannot be measured simultaneously. The commutation rule was given in Eq. (5.18). This implies a Heisenberg uncertainty relation [2],

$$\langle \Delta J_k^2 \rangle \langle \Delta J_l^2 \rangle \geq |\langle \hat{J}_m \rangle|^2 \tag{7.2}$$

where k, l, m represent any orthogonal triplet (such as x, y, z). The uncertainties should then generally scale as $\sqrt{|\vec{J}|}$. We now look at the spin uncertainties for specific quantum states.

7.2 Spin vector and polarization squeezing

To better understand the uncertainty in the polarization vector, we look at the special case of a coherent state. As usual, this is the natural reference for other quantum states, and is the basis for defining a squeezing ratio. The two-mode coherent-state is given by

$$\hat{a}_j|\alpha_1, \alpha_2\rangle = \alpha_j|\alpha_1, \alpha_2\rangle. \tag{7.3}$$

The expectations are calculated with the usual manipulations. One can easily derive,

$$\langle \vec{J} \rangle = \frac{1}{2} \begin{bmatrix} \alpha_1^* & \alpha_2^* \end{bmatrix} \vec{\boldsymbol{\sigma}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}. \tag{7.4}$$

The uncertainty of a component of the spin vector is the same for all directions in the coherent state; that is, for a unit vector $\hat{v}$ in any direction,

$$\langle \Delta^2(\hat{v} \cdot \vec{J}) \rangle = \frac{1}{4}(|\alpha_1|^2 + |\alpha_2|^2) = \frac{n}{4}. \tag{7.5}$$

Thus the average spin vector has length $n/2$, with an uncertainty in all directions of $\sqrt{n}/2$.

The uncertain Stokes vector is depicted in Figure 7-1 for the coherent state as well as a polarization squeezed state, as described in more detail in [13]. Such depictions of course call to mind the techniques of [87] for visualizing interferometer dynamics and phase-resolution.

7.3 Nonlinear birefringent coupler

A number of authors [39, 13, 2] have shown that Kerr-type nonlinearities can lead to polarization-squeezed states. Linear coupling between polarization modes is simply birefringence, and is easily introduced in a Kerr waveguide. The polarization-

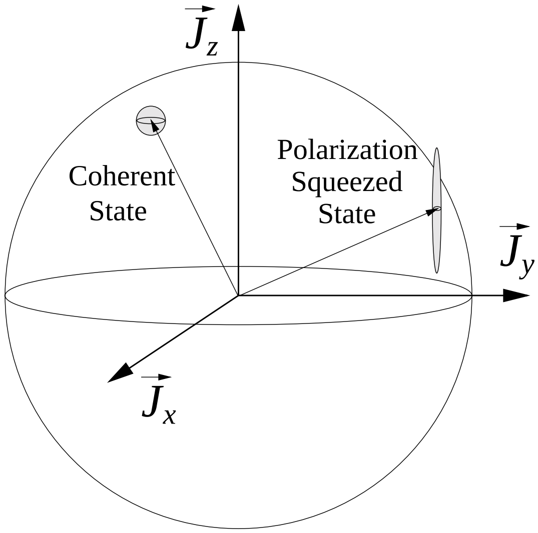


Figure 7-1: For a quantum field, the vector of spin variables plays the role of the classical Stokes vector. Uncertainties in this polarization vector are depicted here for the coherent state , and for a polarization-squeezed state. The Heisenberg relation governs the tradeoff between uncertainties in orthogonal components of the vector.

squeezing systems examined in those papers then have the same basic ingredients that we needed in a model to get interesting tunneling physics.

On the other hand, these models do not exactly give the Hamiltonian (5.1). For one thing, the polarization modes are not spatially separated. The Kerr effect can then give rise to cross-phase modulation as well as self-phase modulation. In [39], this is accounted for by additional terms such as $\hat{a}_1^\dagger \hat{a}_1 \hat{a}_2^\dagger \hat{a}_2$ in

$$\hat{H} = \hbar\omega(\hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2) + \frac{1}{2}\hbar g \left[(\hat{a}_1^\dagger)^2 \hat{a}_1^2 (\hat{a}_2^\dagger)^2 \hat{a}_2^2 + \frac{4}{3}\hat{a}_1^\dagger \hat{a}_1 \hat{a}_2^\dagger \hat{a}_2 + \frac{1}{3}(\hat{a}_1^\dagger)^2 \hat{a}_2^2 + \frac{1}{3}(\hat{a}_2^\dagger)^2 \hat{a}_1^2 \right] \tag{7.6}$$

This can be expressed in a more general and more compact form using spin operators,

$$\hat{H} = -A\hat{J}_z^2 + B\hat{J}_y^2 - h\hat{J}_x. \tag{7.7}$$

The nonlinear part of the spin Hamiltonian is biaxial. From Horak's model we can derive the ratio $A/B = 2$. In general, we might expect the ratio to depend on waveguide parameters, but be of order 1.

The introduction of a nonzero B has been considered in the spin tunneling literature. It leads to an enhancement of tunneling rates [20] that could potentially improve the effectiveness of a tunneling coupler. On the other hand, this term may alter the eigenstate structure at zero linear coupling. This could make it more difficult to match the input field to an eigenstate of the coupler. Despite these differences, we expect the basic features discussed in Chapters 5 and 6 to remain the same: The tunneling rates calculated in [20] for $A/B = 2$ are still vanishingly small. Presumably, then, one still must raise the linear coupling to the critical level, so that the weakly trapped system can see appreciable tunneling. It is possible that some clever combination of linear and nonlinear coupling can qualitatively change the scaling of the adiabatic condition derived in 6.31. For the moment this seems like a remote possibility.

The loss properties of such a device are also ultimately similar, although there are some interesting differences. For our well-separated waveguides, we assumed that loss came about from the coupling of each localized mode $\hat{a}_1$ and $\hat{a}_2$ to a *separate* reservoir. Decoherence could then be thought of as a measurement. The presence of energy in reservoir 1 implied that a photon had been lost from mode 1, yielding information about the quantum state. The case of completely overlapping polarization modes might initially seem to be different. The two localized modes could experience loss through identical coupling to the same reservoirs. One could then hope that loss decoherence would not occur, since a lost photon does not imply anything about which waveguide the photon was in. After more careful analysis, this hope was unfulfilled. On the one hand, it is clear that some loss mechanisms, such as scattering, "remember" the polarization of the light, and essentially couple different polarizations to different reservoirs, even though the modes overlap spatially. On the other hand, coupling to the same mode, though different, still has degrading effects on the coherence.

We expect the adiabatic and loss conditions are qualitatively the same for polarization-mode coupling. We look at this as one of several possible implementations for the multi-photon tunneling concept, all of which are plagued by essentially the same problems.

Chapter 8

Polarization mode dispersion in optical fiber communications

Polarization mode dispersion is the distortion resulting from unwanted birefringence in optical fibers. Typical communications fiber is single-mode fiber (SMF), designed to have a perfectly symmetric core and thus the same group velocity for all polarizations. Temperature variations and vibrations can lead to stresses which break the symmetry and lead to small local birefringence. Thus as a signal passes through a fiber link, it sees different group velocities not only for the frequency components of the signal (chromatic dispersion), but for the polarization components as well.

In the first-order approximation of PMD, the fiber is modeled as a single birefringent element: it has orthogonal "fast" and "slow" axes and a differential group delay τ between the two. This is done by identifying a two-mode group velocity matrix at a single optical frequency—typically the center frequency of the signal. As with chromatic dispersion, PMD varies with frequency. As bandwidth is increased, different frequency components of the signal will see significantly different PMD. As a system is pushed to higher bitrate, not only does the first-order PMD become more severe (since the differential delay τ becomes a larger fraction of the bitperiod), but the first-order approximation starts to break down. Higher-order PMD must then be considered. In the limit that τ exceeds the bitrate, the first-order approximation is invalid and the distortion no longer resembles a simple differential delay (see Figure

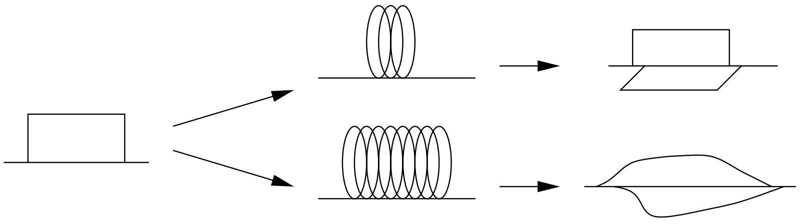


Figure 8-1: The intuitive picture of first-order PMD is that the two principal polarizations have different velocities due to the slight fiber birefringence. The distorted pulse is then simply two polarization components with a relative delay. In a longer fiber, this first-order approximation will break down. Any single delay will fail to capture the signal distortion, and higher-order PMD must be considered.

8-1).

In this chapter, we review the basic formalisms used for linear PMD, and briefly discuss compensation. In fact, the linear PMD analysis is illustrative but not complete. Nonlinearities are generally important in any practical communications system. The importance of analyzing PMD together with nonlinearities and with polarization dependent loss (PDL) is becoming clearer [80, 46], but presents a much more complicated picture.

8.1 Fundamentals of PMD

The theory of fiber PMD is well developed and steadily progressing. A number of excellent papers describe the basic phenomenon in the Jones space or Stokes space representation. However, the connection between the two representations is often left out, and some of the subtle, basic properties of PMD still cause some confusion. In this section, we review some of the background required to better understand the Stokes-space representation and to see our results in context.

To help clarify the two representations, we use a boldface font for vectors $\boldsymbol{a}$ and matrices $\boldsymbol{M}$ in two-dimensional complex Jones space. Three-dimensional, real Stokes vectors are written in plain font with vector symbols $\vec{\Omega}$, or hats $\hat{s}$ for unit vectors.

Jones space

We begin with an optical communications signal. In Jones space, a complex 2-vector represents the amplitudes of two polarization modes. The input $\boldsymbol{a}$, is taken to be a pure polarization signal, with frequency-independent, normalized direction $\boldsymbol{a}_0$,

$$\boldsymbol{a}(t) = \int \frac{d\omega}{2\pi}\boldsymbol{a}(\omega)e^{-i\omega t} = \int \frac{d\omega}{2\pi} f(\omega)e^{-i\omega t}\boldsymbol{a}_0 \tag{8.1}$$

The function $f(\omega)$ gives the spectrum of the signal. The signal is in general subjected to a number of effects as it passes through the system. The PMD literature primarily focuses on the case of a *linear* and *lossless* system, characterized by the unitary matrix $\boldsymbol{T}(\omega)$,

$$\boldsymbol{b}(t) = \int \frac{d\omega}{2\pi}\boldsymbol{T}(\omega)\boldsymbol{a}(\omega)e^{-i\omega t} = \int \frac{d\omega}{2\pi} f(\omega)e^{-i\omega t}\boldsymbol{T}(\omega)\boldsymbol{a}_0. \tag{8.2}$$

For a narrow enough signal bandwidth, PMD can be characterized by the Taylor expansion

$$\boldsymbol{T} = \boldsymbol{T}(\omega_0) + (\omega - \omega_0)\boldsymbol{T}_\omega(\omega_0) + \ldots \tag{8.3}$$

The first-order PMD approximation is

$$\boldsymbol{T}(\omega) \ \approx \ [1 + (\omega - \omega_0)\boldsymbol{T}_\omega(\omega_0)\boldsymbol{T}^\dagger(\omega_0)]\boldsymbol{T}(\omega_0) \tag{8.4}$$

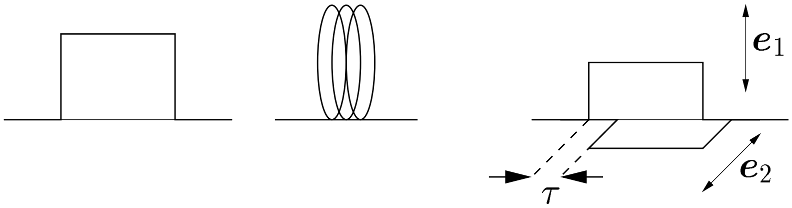


Figure 8-2: The first-order effects of polarization mode dispersion can be interpreted using this simple physical picture. The group delay of a two-mode wave is mathematically expressed by a 2×2 Hermitian matrix $\boldsymbol{M}$. The eigenvectors of this matrix generally experience different group delays. The difference between the eigenvectors of $\boldsymbol{M}$ is called the differential group delay, and represents a time delay induced between the polarization components.

$$\approx \quad [1 + i(\omega - \omega_0)\boldsymbol{M}(\omega_0)]\boldsymbol{T}(\omega_0) \tag{8.5}$$

where the PMD matrix is defined as $\boldsymbol{M}(\omega_0) = -i\boldsymbol{T}_\omega(\omega_0)\boldsymbol{T}^\dagger(\omega_0)$. When $\boldsymbol{T}$ is unitary (lossless), $\boldsymbol{M}$ is Hermitian. Such perturbative treatments have been very useful because they suggest an intuitive picture. For example, the first order expression above has a natural physical interpretation, shown in Figure 8-2: an input signal goes through a nominal, frequency-independent transformation $\boldsymbol{T}(\omega_0)$ followed by a polarization-dependent time-delay. To see this, we think of the Hermitian matrix $\boldsymbol{M}$ as having two real eigenvalues with units of time. These eigenvalues are the delays τ_k $(k = 1, 2)$ seen by the principal states of polarization (the eigenstates $\boldsymbol{e}_k$),

$$[1 + i(\omega - \omega_0)\boldsymbol{M}(\omega_0)]\boldsymbol{e}_k \approx e^{i\Delta\omega\tau_k}\boldsymbol{e}_k. \tag{8.6}$$

The difference of the delays is τ, the differential group delay (DGD).

The perturbative approach has been useful also because PMD is necessarily a small effect in a functioning communications system. That is, for the usual fiber statistics [26], whenever a signal suffers only moderate distortions, the narrow-band Taylor expansion is valid. As signal bandwidths are pushed higher, and more sophisticated compensators are designed, non-perturbative formulations of PMD will become more useful. The PMD matrix for a broader-band system is defined as before,

$$i\boldsymbol{M}(\omega) = \boldsymbol{T}_{\omega}(\omega)\boldsymbol{T}^{\dagger}(\omega), \tag{8.7}$$

but now $\boldsymbol{M}(\omega)$ can have a complex frequency variation over the signal bandwidths.

Stokes space

A Stokes vector $\vec{s}$ is a three-dimensional real vector. It can be thought of as a polarimeter measurement—representing the power along projected components of the optical signal. Alternatively, we can think of it as a convenient mathematical mapping of Jones-vector polarizations, with convenient symmetry properties. For a pure-frequency signal, the magnitude of $\vec{s}$ is simply the power, and the direction is the mapping of the polarization onto the Poincare sphere [47].

In Stokes space, the basic PMD representation does not start with Figure 8-2, but with Figure 8-3. A continuous-wave, tunable-frequency input field is launched, and the normalized output Stokes vector $\hat{s}$ is measured on a polarimeter. The PMD vector $\vec{\Omega}(\omega)$ describes the motion of the output signal polarization as frequency is tuned, assuming the input polarization remains fixed,

$$\frac{d}{d\omega}\hat{s} = \vec{\Omega} \times \hat{s}. \tag{8.8}$$

This is the fundamental PMD relation in Stokes space[64]. The PMD vector can of course be related to the Jones-space PMD matrix. The connection between Jones and Stokes space is

$$\hat{s}(\omega) \equiv \boldsymbol{a}_0^{\dagger}\boldsymbol{T}(\omega)^{\dagger}\vec{\boldsymbol{\sigma}}\boldsymbol{T}(\omega)\boldsymbol{a}_0 \tag{8.9}$$

where $\vec{\boldsymbol{\sigma}}$ is a vector of Pauli matrices. We can now relate $\boldsymbol{M}$ to the frequency variation of $\hat{s}$

$$\frac{d}{d\omega}\hat{s} = i\boldsymbol{a}_0^{\dagger}\boldsymbol{T}(\omega)^{\dagger}[\vec{\boldsymbol{\sigma}}\boldsymbol{M} - \boldsymbol{M}\vec{\boldsymbol{\sigma}}]\boldsymbol{T}(\omega)\boldsymbol{a}_0 \tag{8.10}$$

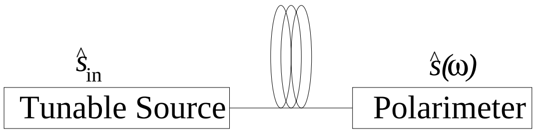


Figure 8-3: The basic PMD formulation in Stokes space is the frequency variation of the Stokes vector, $d\hat{s}/d\omega = \vec{\Omega} \times \hat{s}$. This corresponds to the measurement depicted above, where the frequency is varied by tuning a narrowband source. The input polarization $\hat{s}_{\text{in}}$ is frequency independent, and $\hat{s}$ is the output polarization measured on a polarimeter.

With a few more steps, one can confirm that this is equivalent to Eq. (8.8) with

$$\boldsymbol{M}(\omega) = \frac{-1}{2}\vec{\Omega}(\omega) \cdot \vec{\boldsymbol{\sigma}} - \beta(\omega)\mathbf{1}, \tag{8.11}$$

where β is the isotropic contribution—the group delay due to chromatic (polarization-independent) dispersion. More detail on the connection between Stokes space and Jones space can be found in [30].

We will see that the theory of polarimeter measurements for a finite-bandwidth signal is a natural extension of the Stokes-space fundamental equations. These include the effects of depolarization when different frequency components add out-of-phase.

Cascaded systems

As a simple example of how a system can be described in these two representations, consider a link made up of two segments in series. The basic Jones-space description of a cascade is simple: the transformation of the second segment $\boldsymbol{T}_2$ simply follows the first $\boldsymbol{T}_1$,

$$\boldsymbol{T} = \boldsymbol{T}_2\boldsymbol{T}_1. \tag{8.12}$$

The Jones-space PMD matrix for a cascade follows immediately from Eq. (8.7),

$$i\boldsymbol{M}(\omega) = \boldsymbol{T}_2\frac{d\boldsymbol{T}_1}{d\omega}\boldsymbol{T}_1^\dagger\boldsymbol{T}_2^\dagger + \frac{d\boldsymbol{T}_2}{d\omega}\boldsymbol{T}_1\boldsymbol{T}_1^\dagger\boldsymbol{T}_2^\dagger = i\boldsymbol{T}_2(\omega)\boldsymbol{M}_1(\omega)\boldsymbol{T}_2^\dagger(\omega) + i\boldsymbol{M}_2(\omega), \tag{8.13}$$

where we have used the unitarity of the $\boldsymbol{T}$ matrices: $\boldsymbol{T}_k\boldsymbol{T}_k^\dagger = 1$. The total PMD is then something like the sum of the PMD matrices of the parts. The only complication is that the PMD of the first segment is seen through the second segment, and is thus transformed by $\boldsymbol{T}_2$. Since this transformation is itself frequency dependent, the cascading rules for higher-order PMD matrices become successively more complicated. For example,

$$\boldsymbol{M}_\omega = \boldsymbol{T}_2\boldsymbol{M}_{1\omega}\boldsymbol{T}_2^\dagger + \boldsymbol{M}_{2\omega} + \boldsymbol{T}_{2\omega}\boldsymbol{M}_1\boldsymbol{T}_2^\dagger + \boldsymbol{T}_2\boldsymbol{M}_1\boldsymbol{T}_{2\omega}^\dagger. \tag{8.14}$$

Making appropriate substitutions, we can rewrite this in the more compact form,

$$\boldsymbol{M}_\omega = \boldsymbol{T}_2\boldsymbol{M}_{1\omega}\boldsymbol{T}_2^\dagger + \boldsymbol{M}_{2\omega} + i[\boldsymbol{M}_2, \boldsymbol{T}_2\boldsymbol{M}_1\boldsymbol{T}_2^\dagger] \tag{8.15}$$

where the bracketed expression is a commutator, $[A, B] \equiv AB - BA$.

The Stokes-domain PMD vector is basically just a rewriting of above PMD matrix. Thus we can translate the cascading rule (8.13) as,

$$\vec{\Omega}(\omega) = \mathcal{R}_2(\omega)\vec{\Omega}_1(\omega) + \vec{\Omega}_2(\omega) \tag{8.16}$$

where we have represented the polarization rotation of segment 2 by the 3×3 matrix $\mathcal{R}_2(\omega)$. The second-order cascading rule can be derived from the commutation properties of the spin matrices $\vec{\boldsymbol{\sigma}}$,

$$\vec{\Omega}_\omega = \mathcal{R}_2\vec{\Omega}_{1\omega} + \vec{\Omega}_{2\omega} + \mathcal{R}_2\vec{\Omega}_1 \times \vec{\Omega}_2. \tag{8.17}$$

8.1.1 Signal distortion and PMD

When a pulse propagates in a linear communications channel, it picks up frequency-dependent phases represented by $\vec{\Omega}(\omega)$ and $\beta(\omega)$. If PMD is small, the first-order approximation gives a simple way to think about pulse distortion. The simple time-domain description of two delayed polarization components can be plugged directly into calculations of eye opening, etc. For fibers with more complicated PMD, we would like some simple metrics for quantifying distortion.

Temporal moments $\overline{t^n}$ of the distorted optical signal, depicted in Figure 1, are an elegant metric. These are defined as moments of the energy distribution,

$$\overline{t^n} \equiv \int t^n |\boldsymbol{b}(t)|^2 dt. \tag{8.18}$$

Using Fourier transform properties, Karlsson [41] has derived expressions for the moments which include all orders of PMD and chromatic dispersion. For example, we obtain a convenient expression for the pulse delay,

$$\bar{t} = \int t|\boldsymbol{b}(t)|^2 dt, \tag{8.19}$$

by plugging in Eq. (8.2):

$$\begin{aligned}
\bar{t} &= \int dt \;\; t \int \frac{d\omega'}{2\pi} \int \frac{d\omega}{2\pi} f^*(\omega') f(\omega) e^{i(\omega-\omega')t} \boldsymbol{a}_0^\dagger \boldsymbol{T}^\dagger(\omega') \boldsymbol{T}(\omega) \boldsymbol{a}_0 && (8.20)\\
&= \int \frac{d\omega'}{2\pi} \int \frac{d\omega}{2\pi} \boldsymbol{a}_0^\dagger \boldsymbol{T}^\dagger(\omega') f^*(\omega') f(\omega) \boldsymbol{T}(\omega) \boldsymbol{a}_0 \int dt \;\; t e^{i(\omega-\omega')t} && (8.21)\\
&= -i \int \frac{d\omega}{2\pi} \boldsymbol{a}_0^\dagger \boldsymbol{T}^\dagger(\omega) f^*(\omega) \frac{d}{d\omega}[f(\omega)\boldsymbol{T}(\omega)] && (8.22)
\end{aligned}$$

This naturally breaks up into contributions due to the input field amplitude $f(\omega)$, chromatic dispersion β, and the PMD $\vec{\Omega}$,

$$\begin{aligned}
\bar{t} &= -i \int \frac{d\omega}{2\pi} f^*(\omega) \frac{d}{d\omega} f(\omega) - i \int \frac{d\omega}{2\pi} |f|^2 \boldsymbol{a}_0^\dagger \boldsymbol{T}^\dagger \boldsymbol{T}_\omega \boldsymbol{a}_0 && (8.23)\\
&= -i \int \frac{d\omega}{2\pi} f^*(\omega) \frac{d}{d\omega} f(\omega) - i \int \frac{d\omega}{2\pi} |f|^2 (-\beta(\omega) - \frac{1}{2} \hat{s}(\omega) \cdot \vec{\Omega}). && (8.24)
\end{aligned}$$

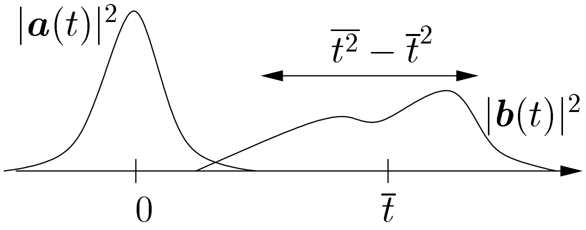


Figure 8-4: Chromatic and polarization dispersion contribute phase which leads to pulse distortion. Distortion can be described by temporal moments, such as the offset $\overline{t}$ and the pulse width $\overline{t^2} - \overline{t}^2$.

The second moment $\overline{t^2}$ similarly includes the transform-limited width given by the spectrum $|f|$, as well as various phase contributions: there is a differential phase due to PMD $\vec{\Omega}$, and a phase delay $\tilde{\beta}$ due to initial pulse chirp and chromatic dispersion, $\tilde{\beta} \equiv \beta + \frac{d}{d\omega}\arg(f)$. Assuming an input pulse with pure polarization, $\boldsymbol{a}(\omega) = f(\omega)\boldsymbol{a}_0$, the result is [41]

$$\overline{t^2} = \int \left\{ \left(\frac{d|f|}{d\omega} \right)^2 + |f|^2 \left| \tilde{\beta}\hat{s} + \frac{1}{2}\vec{\Omega} \right|^2 \right\} \frac{d\omega}{2\pi}. \tag{8.25}$$

We will use these expressions to analyze accumulated distortion in cascaded, compensated lines.

8.2 Strategies for compensation

A number of strategies have been successfully pursued for dealing with PMD. When installation of new fiber is an option, high-quality fiber has become available which dramatically reduces the effect. For improving performance over an existing fiber, compensators fall into several categories:

- Principal-state transmission: When a signal is launched along one of the principal states (the fast or slow axis of the total birefringence), the lowest-order PMD distortion is eliminated. Unfortunately, the principal states cannot be measured locally at the transmitter. In this method, the principal states are measured at the receiver (or through some reflectometry method) and the transmitted polarization is aligned using a polarization controller. A limitation is that the principal state estimate is fundamentally delayed by the speed of light, and so rapid PMD fluctuations on very long links cannot be compensated.

- Electronic compensators: This family of receiver-end methods operates on the electrical signal after photo-detection. Methods vary in sophistication. The basic idea is that given a model of the the input signal and the effects of birefringence, some combination of linear filters and time-dependent thresholding circuitry can adjust for the effects of PMD. For example, the threshold for discerning a “1” and a “0” on a given bit can be raised if the previous bit was a zero, since PMD causes spillover between the bits. A shortcoming is that as bitrates are increased, faster electronics are needed.

- Modulation and compression schemes: Pulses can be reshaped using a bit-synchronous phase or amplitude modulator [66]. This requires clock recovery and can have the side-effect of generating spectrum, which can lead to additional signal distortion if the compensator is used in a cascaded system. This technique must also be tailored to the data format.

- Optical compensators: Optical compensators cascade a controllable PMD with the uncontrolled PMD of the fiber. They can be used at the receiver-end just before detection, or "in-line," where the optical signal is passed on without regeneration to another fiber. This is the only strategy that can, in principle, recover the original optical signal completely. However, higher-order PMD compensation requires a device with many degrees of freedom.

Among optical compensators, there is an interesting discussion about how many degrees of freedom are required for a given level of performance. Among the first-order compensators, there are two varieties. Full first-order compensators have three degrees of freedom, so that they can fully cancel the three components of the the PMD vector at ω_0. That is, one can choose $\vec{\Omega}_c$ so that[1]

$$\vec{\Omega}(\omega_0) = \mathcal{R}_{\text{comp}}(\omega_0)\vec{\Omega}_{\text{fiber}}(\omega_0) + \vec{\Omega}_{\text{comp}}(\omega_0) = 0. \tag{8.26}$$

However, the component of $\vec{\Omega}$ along the signal polarization actually has no effect to first order. For this reason, fixed-DGD compensators are an interesting, simpler alternative. For these, only the direction of $\vec{\Omega}_{\text{comp}}$ can be controlled. If the fixed magnitude is large enough, these two degrees of freedom are sufficient to align $\vec{\Omega}(\omega_0)$ with an arbitrary signal polarization,

$$\hat{s} \times \vec{\Omega}(\omega_0) = 0. \tag{8.27}$$

The field is rapidly progressing. All basic strategies have been demonstrated in field experiments, and there are active efforts to implement compensation in commercial systems. Notable recent updates on compensator performance include:

- Electronic compensation of a 10Gb/s signal was demonstrated for DGDs exceeding the bitrate: 3dB of penalty was reported at 150ps [81].

- Alcatel demonstrated a fixed-DGD first-order compensator on 40Gb/s NRZ data, bringing the PMD limit from 4ps to 8ps [53].

---

[1] In fact, the optimal choice for reducing PMD-induced distortion may differ from this choice slightly, as shown in [42]

- Theoretical and experimental comparisons show that variable DGD compensators outperform fixed-DGD compensators [42, 85].

- A number of partial second-order compensators are being discussed. We are not currently aware of any that completely cancel the second-order PMD distortion.

- Multiple WDM channels would need one compensator per channel for full compensation at a particular order. However, outages on any channel are rare, and so it is possible that a single compensator shared by many channels could significantly improve the outage probability. This approach has been investigated by [45]

Chapter 9

PMD compensation in a network

A network is composed not only of fibers, but includes regenerators, PMDCs, switches, etc. So far, the limitations of first-order compensators due to residual, higher-order PMD have been studied for point-to-point lines [54, 74]. If we imagine that networks will soon be composed of compensated fibers cascaded without regeneration, then the obvious next step is to understand the accumulation of higher-order PMD in such systems. We have calculated the relevant statistics and derived a simple scaling law for system design [23]. Multiple first-order compensators substantially increase the usable propagation length. As bandwidth is increased, however, the benefit of this strategy diminishes.

9.1 Compensated Fiber

Consider now a single compensated fiber, composed of a standard fiber and a compensator, as shown in Fig. 9-1. In particular, we will consider only *complete* first-order compensation, in contrast to fixed-delay compensators, for example [27]. The fiber PMD is described by

$$\vec{\Omega}^a = \vec{\Omega}^a(\omega_0) + \Delta\omega \vec{\Omega}^a_\omega(\omega_0) + \ldots \tag{9.1}$$

The random parameters $\vec{\Omega}^a(\omega_0)$, etc., have well known statistics [26]. The first-order compensator may have a number of implementations. Often (and in the best

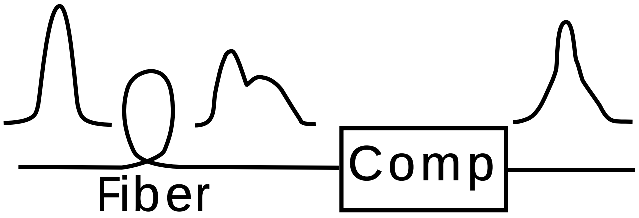


Figure 9-1: A first order compensator removes the primary component of PMD distortion, delay between the principal states, but leaves residual higher-order PMD .

case), it has little higher-order PMD of its own. We represent the compensator then by a *constant* PMD $\vec{\Omega}^b$. In general, the transformation described by the PMD equation (8.8) with constant $\vec{\Omega}^b$ can be written

$$\mathcal{R}^b = e^{\Delta\omega[\vec{\Omega}^b\times]}\mathcal{R}^b(\omega_0), \tag{9.2}$$

where $\mathcal{R}^b(\omega_0)$ is an arbitrary frequency-independent rotation. By design, a compensator adjusts $\vec{\Omega}^b$ to cancel the lowest-order effects of PMD. In [23], we assumed that a reasonable controller would choose $\vec{\Omega}^b = -\mathcal{R}^b(\omega_0)\vec{\Omega}^a(\omega_0)$; plugging into Eq. (8.16), we have, to lowest order,

$$\begin{aligned}
\vec{\Omega}^c &= \vec{\Omega}^b + e^{\Delta\omega[\vec{\Omega}^b\times]}\mathcal{R}^b(\omega_0)[\vec{\Omega}^a(\omega_0) + \Delta\omega\vec{\Omega}^a_\omega(\omega_0)] && (9.3)\\
&= \vec{\Omega}^b + e^{\Delta\omega[\vec{\Omega}^b\times]}\vec{\Omega}^b + \Delta\omega e^{\Delta\omega[\vec{\Omega}^b\times]}\vec{\Omega}^a_\omega(\omega_0) && (9.4)\\
&= \Delta\omega\mathcal{R}^b\vec{\Omega}^a_\omega(\omega_0) && (9.5)
\end{aligned}$$

That is, the residual second-order PMD $\vec{\Omega}^c_\omega(\omega_0)$ is simply a rotated copy of $\vec{\Omega}^a_\omega(\omega_0)$. It thus has the same magnitude as the second-order PMD of the uncompensated fiber.

This result does little more than formalize what is intuitively clear: when a fiber is compensated to first-order, residual second-order PMD remains. Although not surprising, this gives us a compact description of residual PMD to be used in calculations below. In fact, a recent paper [42] suggests that a "first-order" compensator with three degrees of freedom will actually do somewhat better than this. The above argument assumes that a control system will select $\vec{\Omega}^b = -\mathcal{R}^b(\omega_0)\vec{\Omega}^a(\omega_0)$ to cancel the PMD vector. However, one component of $\vec{\Omega}$, the component parallel to the signal polarization $\hat{s}$, actually does not contribute any distortion to lowest order. So cancellation of lowest-order effects determines only two of the three degrees of freedom. Karlsson has calculated the optimum of the third degree of freedom by using pulse broadening as the measure of distortion to be minimized [42]. We present our original analysis, and later provide some comments on the impact of [42].

The statistics of the compensated fiber can now be derived. An uncompensated fiber of length l has [26] $(\tau^u)^2 \equiv \langle|\vec{\Omega}^a|^2\rangle = Dl$, and $\langle|\vec{\Omega}^a_\omega|^2\rangle = D^2l^2/3$. The single parameter D characterizes the growth of PMD along a fiber. It can be larger than 1 ps^2/km for some older installed fibers. For the compensated fiber, statistics follow from Eq. (9.5). Since a rotation does not affect the vector magnitude, we have

$$\langle|\vec{\Omega}^c(\omega)|^2\rangle = (\omega-\omega_0)^2\langle|\vec{\Omega}^u_\omega|^2\rangle = (\omega-\omega_0)^2(\tau^u)^4/3 \tag{9.6}$$

The magnitude is explicitly dependent on ω (and ω_0, as discussed below). Ultimately, we will need only this expectation, and no other statistics of $\vec{\Omega}^c$, in our calculations.

Not surprisingly, the effect of a first-order compensator is a reduction of PMD on the order of the expansion parameter, $\tau^u\Delta\omega$. With $(\tau^c)^2 \equiv \langle|\vec{\Omega}^c|^2\rangle$, we have

$$\tau^c/\tau^u = \tau^u\Delta\omega/\sqrt{3} \tag{9.7}$$

9.2 Cascade of Fibers

A cascade of compensated fibers is depicted in Fig. 9-2. Consider, for simplicity, N statistically identical and independent segments of length l and total length $L = Nl$. The total PMD vector $\vec{\Omega}^L$, is the sum of N contributions, $\vec{\Omega}^L = \sum_{m=1}^{N}\mathcal{R}_m\vec{\Omega}^c_m$, each involving a rather complicated rotation $\mathcal{R}_m$. Regardless, interesting statistics can be calculated in just a few steps using the above framework.

The sum $\vec{\Omega}^L$ is of independent contributions randomly oriented in Stokes space. Thus, the average magnitude is unaffected by the rotations, and follows from Eq.

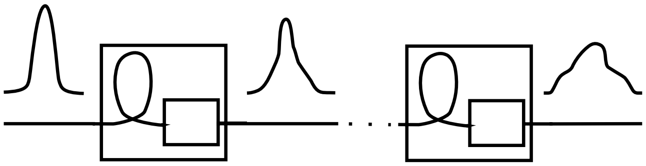


Figure 9-2: For a cascade of compensated fibers, residual, higher-order PMD accumulates with random-walk scaling.

(9.6):

$$\langle|\vec{\Omega}^L(\omega)|^2\rangle = N\langle|\vec{\Omega}^c(\omega)|^2\rangle = (\omega-\omega_0)^2 N(\tau^u)^4/3. \tag{9.8}$$

Using the temporal moments such as Eq. (8.25), we now calculate the PMD-induced pulse spreading. The basic idea is to separate PMD-induced spreading $\Delta t^2_{\mathrm{PMD}}$ from non-PMD contributions Δt_0^2 to the pulse width:

$$\Delta t^2 \equiv \langle \overline{t^2} - \overline{t}^2 \rangle = \Delta t_0^2 + \Delta t^2_{\mathrm{PMD}}. \tag{9.9}$$

In fact, the moment integrals separate cleanly under quite reasonable assumptions. Our first assumption is that a compensator automatically makes an "ideal" choice of center frequency ω_0: that is, it adjusts ω_0 to minimize distortions. Mathematically, one can show that by choosing $\omega_0 = \int d\omega |f|^2 \omega/2\pi$, the center of the signal spectrum, we simultaneously minimize the lowest-order pulse distortion and zero the polarization-dependent time delay. The assumption is natural—by design, realistic compensators will adjust to meet these physical requirements, at least to lowest order. Under these conditions, one can show that $\overline{t}$ has no PMD contribution. The delay $\overline{t}$ is simply a constant.

We turn to the second moment, Eq. (8.25). One can clearly identify the PMD and non-PMD terms in the integrand: PMD involves an $|\vec{\Omega}|^2$ term and an $\vec{s}\cdot\vec{\Omega}$ cross-term. Our second assumption is that $\vec{\Omega}^L$ is statistically independent of the signal polarization $\vec{s}$. This is true unless the transmitter receives feedback control. In this case, the cross term contributes nothing to the expectation (9.9), since $\langle\vec{s}\cdot\vec{\Omega}\rangle = 0$.

We are left with a single PMD-related term from $\langle\overline{t^2} - \overline{t}^2\rangle$,

$$\Delta t^2_{\mathrm{PMD}} = \frac{1}{4}\int \frac{d\omega}{2\pi}|f|^2 \left\langle|\vec{\Omega}^L|^2\right\rangle. \tag{9.10}$$

Interestingly, $\vec{\Omega}^L$ represents higher-order PMD. This calculation is an example of how Fourier integrals can handle higher-order PMD in a natural way. In this calculation, we simply include the explicit frequency dependence (9.8) of $\vec{\Omega}^L$ in the integral. The result is proportional to the pulse bandwidth, defined by $\Delta\omega^2_{\mathrm{pulse}} \equiv 1/4\tau^2_{\mathrm{pulse}} \equiv \int d\omega|f|^2(\omega-\omega_0)^2/2\pi$. Our result,

$$\Delta t^2_{\mathrm{PMD}} = \frac{1}{4}\frac{N(\tau^u)^4}{3}\Delta\omega^2_{\mathrm{pulse}} = \frac{D^2Ll}{12}\Delta\omega^2_{\mathrm{pulse}}. \tag{9.11}$$

agrees with [74] in the relevant limit, $N = 1$ and τ^u small.

9.3 System Requirements

The above results imply practical limitations on system design. Our first requirement is implicit: our approach assumes the parameter $\tau^u\Delta\omega = \Delta\omega\sqrt{Dl}$ is small. We further require that the total signal distortion be small enough to avoid errors. Accurately predicting bit-error rates is quite complicated in general (see, for example, [59, 8]). We can obtain a simple guideline easily by adopting PMD-induced spreading (9.9) as a metric of total distortion. A reasonable requirement is that $\Delta t^2_{\mathrm{PMD}}$ be less than some ratio R_{tol} of the pulse width (and thus the bit slot):

$$\Delta t^2_{\mathrm{PMD}} < R_{\mathrm{tol}}\tau^2_{\mathrm{pulse}}. \tag{9.12}$$

In general the tolerance ratio R_{tol} must be consistent with the pulse shape and bit error-rate, and may differ for different orders of PMD; while the correct choice is system-dependent, we can infer a rough estimate from the analysis of outage probabilities in [8], $R_{\mathrm{tol}} \sim .014$ (for 80ps FWHM Gaussian input, 10 ps PMD leads to acceptable probability of outage, BER $< 10^{-12}$).

Along with Eq. (9.11), this requirement gives us the total distance that pulses can propagate without regeneration,

$$L < \frac{12 R_{\text{tol}}}{D^2 l \Delta\omega_{\text{pulse}}^4}. \tag{9.13}$$

It is interesting to compare this distance to the maximum length allowed for an uncompensated fiber, $L_0 = 3R_{\text{tol}}/2D\Delta\omega_{\text{pulse}}^2$. The improvement in total distance is related to the expansion parameter, as we would expect:

$$\sqrt{L/L_0} \sim 1/\tau^u \Delta\omega_{\text{pulse}}. \tag{9.14}$$

The scaling of maximum distance given by Eqs. (9.13) and (9.14) demonstrates both the effectiveness and the limitation of a multiple first-order compensator strategy. At a fixed bandwidth, L is improved over L_0 by the large factor $(\tau^u \Delta\omega_{\text{pulse}})^{-2}$. For example, with $L_0 = 100$km and a expansion parameter ~ 0.1 , $L \sim 10,000$km could span a large terrestrial network. On the other hand, the scaling of L with bandwidth (9.13) is very strict. A doubling of bandwidth for fixed l requires a reduction of the product lL by a factor of 16. Thus, either the total distance or the spacing between compensators must drastically decrease. As bitrates increase, lumped compensation will offer limited improvement.

9.4 Discussion

As compensation is incorporated into existing networks, limitations due to the residual PMD of compensated fibers must be understood. We have set up a convenient description of residual PMD for a compensated fiber, and analyzed the accumulation of signal distortion in cascaded fibers. Our analysis gives a quick guideline (9.13) for the practical limit imposed on propagation distance and bandwidth. This simplified analysis can naturally be extended to more realistic systems. This might include non-identical fibers, better modeling of how distortions lead to system outage, and generalization to fixed-DGD compensators.

Our results highlight the challenge of increasing bandwidth. On the one hand, we have seen that the strategy works: by including multiple intermediate compensators, we can reduce all orders of PMD, and thus increase the usable propagation length L beyond that allowed by receiver-end first-order compensation alone. However, the strategy exhibits a scaling $L \propto \Delta\omega^{-4}$ for constituent fibers of a fixed length. As we push data rates higher, the improvement obtained by first-order compensation diminishes. This suggests that all-optical regenerators or closely spaced compensators may be required for sustained growth of large terrestrial networks.

Our basic starting point was the expression for a single compensated segment, which we assumed followed complete first-order compensation, Eq. 9.5. This is certainly not the case for a fixed-DGD, 2 degree-of-freedom compensator, and the calculation must be amended to include residual first-order PMD component parallel to the input polarization. There is also evidence that a realistic 3 degree-of-freedom compensator under feedback control does not perform "complete" first-order compensation. The basic argument is that only two components of the PMD vector contribute to the distortion to lowest order. So to minimize the distortion, a feedback controller will use two degrees of freedom to cancel these two orthogonal components (approximately). The third degree of freedom will be used to remove part of the higher-order distortion, and has a more complicated form [42]. For such a compensator, extending the analysis of this chapter is not trivial. The residual first-order PMD component must be retained, and further, we can no longer assume that $\langle \vec{\Omega} \cdot \hat{s} \rangle = 0$, since now Ω_c is not independent of $\hat{s}$.

Fixed-DGD compensators

We now outline an extension of our cascading analysis to fixed-DGD compensators. Consider again the fiber PMD

$$\vec{\Omega}^a = \vec{\Omega}^a(\omega_0) + \Delta\omega \vec{\Omega}^a_\omega(\omega_0) + \Delta\omega^2 \vec{\Omega}^a_{\omega\omega}(\omega_0)/2 + \ldots \tag{9.15}$$

For simplicity, we let the nominal rotation of the compensator be unity $\mathcal{R}^b(\omega_0) = 1$. The orthogonal component $\vec{\Omega}_{a,\perp}(\omega_0) \perp \hat{s}(\omega_0)$ of the PMD contributes the lowest-order distortion. Thus, a good approximation for a feedback-controlled fixed-DGD compensator is that it cancels this orthogonal component. We can achieve $\vec{\Omega}^c(\omega_0) \perp \hat{s}(\omega_0)$ with $\vec{\Omega}^b = -\vec{\Omega}^{a,\perp}(\omega_0) + B\hat{s}_0$. Then,

$$\begin{aligned} \vec{\Omega}^c &= \vec{\Omega}^b + e^{-\Delta\omega[\vec{\Omega}^b \times]}[\vec{\Omega}^a(\omega_0) + \Delta\omega\vec{\Omega}^a_\omega(\omega_0) + \ldots] \\ &\approx (\Omega^a_{||} + B)\hat{s}(\omega_0) + \Delta\omega[\vec{\Omega}^a_\omega(\omega_0) + (\Omega^a_{||} + B)(\vec{\Omega}^a_\perp \times \hat{s}(\omega_0))] \end{aligned} \tag{9.16}$$

Since the DGD of the compensator is fixed, $B = \pm\sqrt{|\vec{\Omega}^b|^2 - |\vec{\Omega}^a_\perp|^2}$. The compensator can choose only the sign.

To assess the residual distortion, we look at the pulse broadening. For simplicity, we use the case of no pulse chirp or chromatic dispersion:

$$\overline{t^2} - \bar{t}^2 = \Delta t_0^2 + \overline{|\vec{\Omega}|^2} - [\overline{\hat{s} \cdot \vec{\Omega}}]^2 \tag{9.17}$$

We identify components perpendicular and parallel to $\hat{s}(\omega)$,

$$\overline{t^2} - \bar{t}^2 = \Delta t_0^2 + \overline{|\vec{\Omega}^c_{||}|^2} + \overline{|\vec{\Omega}^c_\perp|^2} - [\overline{\hat{s} \cdot \vec{\Omega}^c_{||}} + \overline{\hat{s} \cdot \vec{\Omega}^c_\perp}]^2. \tag{9.18}$$

The zeroth order PMD distortion arises only from $|\vec{\Omega}_\perp|^2$, but is zero when the compensator cancels this component. The residual broadening is more complicated than for "complete first-order" compensation; Combining Eqs. (9.16) and (9.18) results in a number of terms of order $\Delta\omega^2\tau_{\mathrm{PMD}}$. Statistically, these term are all of the same order, and we might expect a similar scaling to that obtained in the simpler case, Eq. (9.5), but with different numerical constants. Naturally, we expect the fixed DGD compensator to do somewhat worse, because of the presence of additional terms in the residual PMD.

We would similarly argue that Karlsson's analysis of a three degree-of-freedom compensator again has residual PMD that is statistically of the same order. In this case, the optimal choice should perform somewhat better than the trends we gave above. The complexity of these other cases highlights the fortunate choice of complete compensation in simplifying the analysis. However, it cautions us to support analytic estimates with accurate simulations.

Chapter 10

Feedforward compensation of PMD

Virtually all optical PMD compensators have been implemented in a feedback configuration, as in Figure 10-1. The controllable PMD element is adjusted to optimize the signal monitor. The signal monitor measures some indicator of the residual PMD, for example degree-of-polarization, eye opening, etc. In general, a feedback system does not need to measure the PMD of the fiber. All the feedback circuitry needs to know is how "good" the signal looks, in order to find an optimum.

The feedforward configuration in Figure 10-2 is very different. The fiber PMD must now be accurately measured and the response of the control element must be predictable. This difference between feedforward and feedback control is common to other applications: in some sense, one always has to work harder to implement a feedforward control system because it is sensitive to deviations of the model from the real system.

We share the enthusiasm that engineers have in general for feedback. However, a number of important issues specific to PMD compensation remain unresolved. Their practical impact on feedback and feedforward implementations is different, suggesting some possible advantages for feedforward. For first-order compensation, progress in feedback schemes has been substantial, and it would be surprising if feedforward could displace the existing, successful technology. For the more difficult problem of higher-order compensation, we feel that feedforward or hybrid strategies may be optimal. The issues include:

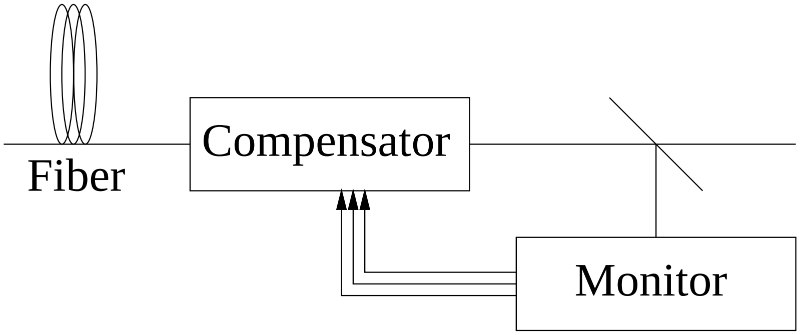


Figure 10-1: In a feedback configuration, the compensating element is controlled to optimize the signal monitor. For example, a polarization controller in a fixed DGD-compensator might be adjusted to maximize the eye opening of the bitstream.

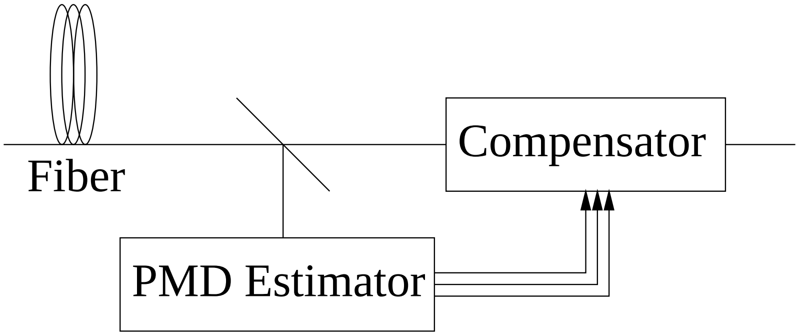


Figure 10-2: A feedforward compensator estimates the fiber PMD vector, and aligns the compensating element to cancel its effects.

- Cascading of controlled PMD segments: Most higher-order compensator architectures employ cascaded PMD elements. Adjusting one element in a cascade changes the orientation of all elements behind it. This creates a substantial challenge for either architecture, since changing one control variable can misalign other elements, leading to a system outage. For the feedback configuration, there is also the problem of the feedback loops being unable to track such variations, and requiring frequent reset.

- Local optima of the signal monitor: The measured signal monitor is often a complicated function of the PMD parameters. Local optima can cause the feedback controller to get stuck at an undesirable setting.

- Control resets Some practical difficulty has been added by the finite bounds on control voltages allowed by polarization rotators. While feedback compensators have been successful at overcoming this issue, they require additional complexity and knowledge of the state of the compensating element, detracting from the simplicity that feedback can often afford.

- Complexity: Higher-order PMD unavoidably involves many degrees of freedom, and thus many control variables. We expect that efficient control by either method will be challenging for second-order compensators, and perhaps impractical beyond second order.

- Other physics: In a system where higher-order PMD is significant, we very well may find that other physics plays a role as well. For example, compensation should tolerate interactions that may exist between PMD, PDL, and fiber nonlinearity. Feedback will generally have an advantage in cases where the fiber is not well characterized.

The issue of controlling cascaded PMD elements is considered in [73]. Consider the compensator of Figure 10-3. It consists of many cascaded PMD elements with controllable orientation and DGD. This is in some sense the ideal compensator—its structure resembles that of the fiber, and we have many degrees of freedom with

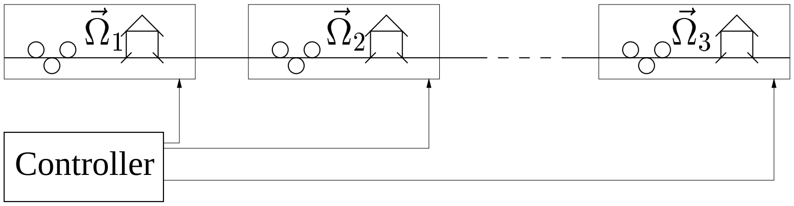


Figure 10-3: Here we see a compensator with many segments of controllable PMD $\vec{\Omega}_j$ in cascade. This strategy is attractive because it mimics the natural accumulation of PMD in the fiber. However, it has the undesirable property that tiny variations in any $|\vec{\Omega}_j|$ cause large polarization rotations.

which we can make its PMD precisely cancel that of the fiber. In principle, one can simply plug such a device into a generic feedback controller, and find an optimum of the signal monitor.

The practical problem of utilizing the degrees of freedom is difficult, because of the cascading rules. Tiny changes in the DGD of one element can drastically effect how the remaining elements combine. One proposal is to vary each DGD on a coarse scale, but keep its wavelength-scale variation tightly locked with a separate control loop. This is a step away from a simple feedback architecture, where the polarization transformation $\boldsymbol{T}(\omega)$ of the compensator does not need to be explicitly known. It may be a step towards a more feed-forward system, where the control system is able to predictably set the transformation $\boldsymbol{T}(\omega)$. Naturally, there may be some feedback involved in the workings of the controller. One justification for working on feedforward compensation is this: by thinking about how to efficiently implement feedforward control of a compensator, we learn a lot about practical feedback controllers.

10.1 Theory of polarimeter measurements

Various techniques have been developed for characterizing the PMD of a communications line. Most techniques do not allow on-line characterization; they require that a communications system be taken off-line to perform the measurement. We are exploring on-line PMD characterization of a communications link using a polarimeter. This method has been the subject of concurrent research in two other groups as well [15, 14, 24, 67, 75]. The applicability of these estimates for compensation in a feedforward [14] and feedback [67] configuration has been demonstrated.

Polarimeter measurements of broadband signals are also important in a much wider range of applications. In this chapter we develop compact, general expressions for the time-averaged state of polarization measured on a polarimeter. Our frequency-domain expressions allow easy calculations for an exact, frequency-dependent system response matrix $\boldsymbol{T}(\omega)$, or for approximate, narrowband parameters, such as the first- and second-order PMD vectors.

Mapping the polarization-dependent response

The high-level strategy is to characterize PMD by measuring the system's response to a variety of inputs. The more distinct inputs we provide, the more information we obtain about the propagation medium. Traditionally this is done by providing a variety of narrowband frequency inputs with a tunable laser [3]. Equivalently, one can select out a desired frequency at the receiver end using a narrow-band tunable filter [68].

An alternative is to provide finite-bandwidth input signals with a variety of different polarizations. This strategy can be implemented using a polarization scrambler at the receiver, depicted in Figure 10-4. Polarimeter measurements are taken for a sampling of input polarizations, and reveal that the signal is *depolarized* more for some input polarizations than others. The pattern of this depolarization gives the principal state axis and the total amount of PMD. In fact, these measurements are sufficient to characterize a first-order PMD element to within a sign.

The basic concept is illustrated in Figure 10-5. An input pulse has components along the two principal polarizations of the fiber. A polarimeter measures vectors in the three-dimensional Stokes space, in which the principal states define the $\pm\hat{\Omega}$ directions. An input with all energy in one principal-state polarization is not distorted to lowest order, and gives fully polarized polarimeter measurements along the principal axis. Inputs with an even split of energy between principal states are depolarized,

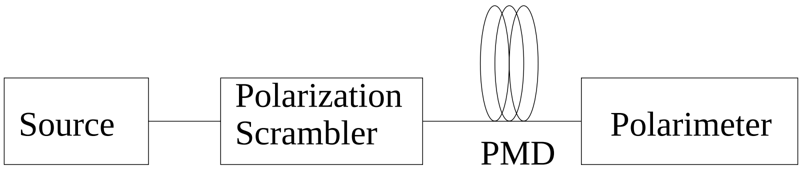


Figure 10-4: Schematic system where a diversity of input polarizations are sampled using a scrambler. In this way the systems response to all polarizations is probed.

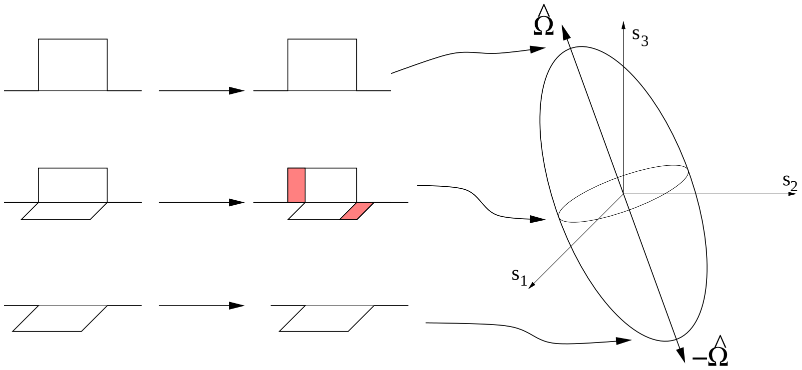


Figure 10-5: Output polarimeter measurements show the effects of system PMD. Signals falling near the principal state axis will be undistorted, and thus highly polarized. Those which have even splitting of energy between the principal states have equal and opposite contributions to the Stokes vector from the skewed components (shaded in the figure), leading to depolarization. The oblong shape has an orientation and minimum degree of polarization which reflect the principal states and DGD. If a diversity of input polarizations are used, the depolarization of the output signal can be used to estimate the PMD parameters.

since there is cancellation between the time-skewed contributions of the principal states. Polarimeter measurements then fall on an oblong surface, which can give us an estimate of the principal state axis and the time skew, or DGD. In this chapter, we analyze PMD-induced depolarization in detail. We show that the oblong shape of the polarimeter measurements, for a given (higher-order) PMD $\vec{\Omega}(\omega)$ and signal spectrum, is in fact an ellipsoid.

Here, we outline the theory of this "ellipsoid" estimation technique. The simplest version extracts PMD parameters from samples of a fixed ellipsoid. The formalism we set up lends itself nicely to generalizations of this scheme. For example, one could augment the ellipsoid measurement with optical filters or known-PMD elements. We briefly describe recent progress [63] on one such generalized strategy, using additional degrees of freedom. The framework we outline below is readily extended to estimates of higher-order PMD, variable-bandwidth input signals, and polarization-dependent loss (PDL).

10.2 Theory of polarimeter measurements

A polarimeter generates currents proportional to the three Stokes parameters of the detected light. The response of the photodetectors and electronics is generally much slower than the communications data rate. Thus, what is measured is not the instantaneous Stokes vector $\boldsymbol{b}^\dagger\vec{\boldsymbol{\sigma}}\boldsymbol{b}$, but the time-averaged Stokes vector,

$$\vec{r} = \int dt \boldsymbol{b}^\dagger(t)\vec{\boldsymbol{\sigma}}\boldsymbol{b}(t) \tag{10.1}$$

In Figure 10-5, we saw conceptually how time averaging leads to cancellation of the currents, and thus depolarization of the signal. The vector $\vec{r}$ allows us to describe this effect quantitatively for general signals.

If the effect of PMD in the time domain is already known, one might plug directly into Eq. 10.1. For example, an exact analysis of a purely first-order PMD system can be performed by writing the signal as two delayed components and evaluating $\vec{r}$. However, the bulk of PMD analysis is based on frequency expansions, and so it is often more natural to work in the frequency domain. The formulation we now develop allows us to keep track of the approximate nature of the PMD "orders." For example, the formulation allows us to test the first-order approximation of a real fiber by including the effects of higher orders.

Plugging the Fourier relation (8.2) into the above expression, we get the general result,

$$\vec{r} \;=\; \int dt \int \frac{d\omega}{2\pi} \int \frac{d\omega'}{2\pi} f^*(\omega')f(\omega)e^{-i(\omega-\omega')t}\boldsymbol{a}_0^\dagger \boldsymbol{T}^\dagger(\omega')\vec{\boldsymbol{\sigma}}\boldsymbol{T}(\omega)\boldsymbol{a}_0. \tag{10.2}$$

$$= \int \frac{d\omega'}{2\pi} |f(\omega)|^2 \boldsymbol{a}_0^\dagger \boldsymbol{T}^\dagger(\omega) \vec{\boldsymbol{\sigma}} \boldsymbol{T}(\omega) \boldsymbol{a}_0. \tag{10.3}$$

$$= \int \frac{d\omega'}{2\pi} |f(\omega)|^2 \hat{s}(\omega) \tag{10.4}$$

Thus the polarimeter measurement is a weighted average of the Stokes vector across the spectrum of the signal. It is convenient to normalize the measurement to the total power entering the polarimeter. This is easy experimentally and described simply by

$$\int \frac{d\omega'}{2\pi} |f(\omega)|^2 = 1, \tag{10.5}$$

so that in the limit of a narrowband signal, $|f(\omega)|^2 \to \delta(\omega - \omega_0)$,

$$\vec{r} \to \hat{s}(\omega_0) \tag{10.6}$$

Finally we define the degree of polarization (DOP) as the fraction of the signal power which is polarized.

$$\text{DOP} \equiv \frac{\left| \int dt \boldsymbol{b}^\dagger(t) \vec{\boldsymbol{\sigma}} \boldsymbol{b}(t) \right|}{\int dt \boldsymbol{b}^\dagger(t) \boldsymbol{b}(t)} = |\vec{r}| \tag{10.7}$$

As the input is scanned through all polarizations, the depolarization is quantified using the DOP.

10.2.1 Purely first-order system

In the next section, we derive approximate, narrowband ellipsoid expressions for a general narrowband system. Before moving on to this important practical problem, we solve simplest case: If the system consists of only a single birefringent element (that is, if higher-order PMD is identically zero) then there is no need for a Taylor expansion in frequency. One can perform the analysis to all orders, in either the time or frequency domain.

We describe the time-domain calculation in detail in Appendix E. Here we simply summarize the results. The input signal is assumed to have a pure polarization with

Jones vector $\boldsymbol{a}_0$ and Stokes vector $\hat{s}_{\text{in}} \equiv \boldsymbol{a}_0^\dagger \vec{\boldsymbol{\sigma}} \boldsymbol{a}_0$,

$$\boldsymbol{a}(t) = f(t)\boldsymbol{a}_0. \tag{10.8}$$

The polarimeter output is derived directly from Eq. (10.1), where the output signal has experienced a differential delay τ,

$$\boldsymbol{b}(t) = Uf(t - \tau/2)\boldsymbol{e}_1 + Vf(t + \tau/2)\boldsymbol{e}_2, \tag{10.9}$$

where U and V are the amplitudes of the input signal along the principal state polarizations.

The input Stokes vector is composed of components parallel $\hat{s}_{||}$ and perpendicular $\hat{s}_\perp$ to the principal state axis of the PMD,

$$\hat{s}_{\text{in}} \quad = \quad \hat{s}_{||} + \hat{s}_\perp \tag{10.10}$$

The action of the PMD is to map this input polarization vector onto an output polarization $\vec{r}$. The PMD causes a rotation about the principal state axis, and also a depolarization effect. Our calculation gives the form of the time-average output Stokes vector,

$$\vec{r} = \hat{s}_{||} + |R_{ff}(\tau)|\hat{s}'_\perp. \tag{10.11}$$

The depolarization is polarization dependent. It leaves the parallel component unchanged but scales down the orthogonal component by the factor $|R_{ff}(\tau)|$, where R_{ff} is the autocorrelation of the input waveform,

$$R_{ff}(\tau) = \int dt f(t - \tau/2)^* f(t + \tau/2). \tag{10.12}$$

The rotation about the principal state axis $\hat{s}_\perp \to \hat{s}'_\perp$ is simply the rotation of the polarization at the center frequency ω_0, and not of great physical importance.

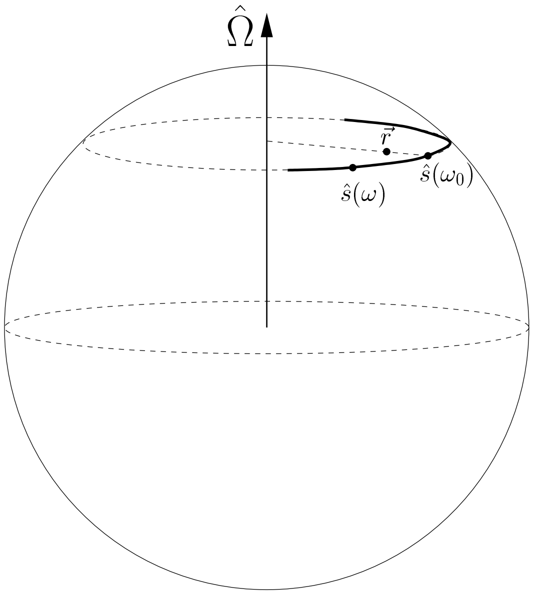


Figure 10-6: The polarimeter measurement of a finite-bandwidth signal is the average of the frequency components $\hat{s}(\omega)$ weighted by the signal spectrum $|f(\omega)|^2$. For a first-order PMD system, $\hat{s}(\omega)$ traces out a circular arc about the principal state axis, $\hat{\Omega}$. The weighted average of points on the circle is pulled in towards the axis, leading to depolarization, $|\vec{r}| < 1$.

Frequency-domain analysis

Frequency domain expressions allow us to approach first-order and more complicated estimations on the same footing. The basic starting point uses Eqs. (10.4) and (8.8):

$$\vec{r} = \int \frac{d\omega}{2\pi} |f(\omega)|^2 \hat{s}(\omega) \tag{10.13}$$

$$\frac{d}{d\omega}\hat{s} = \vec{\Omega} \times \hat{s}. \tag{10.14}$$

For a strictly first-order PMD system, the PMD vector is a constant. The trajectory of $\hat{s}(\omega)$ is then exactly solvable, corresponding to the intuitive geometrical picture of Figure 10-6. We can write the solution as

$$\hat{s}(\omega) = e^{\Delta\omega\vec{\Omega}\times}\hat{s}(\omega_0) = \hat{s}_{||}(\omega_0) + \cos(|\vec{\Omega}|\Delta\omega)\hat{s}_{\perp}(\omega_0) + \sin(|\vec{\Omega}|\Delta\omega)\hat{\Omega} \times \hat{s}_{\perp}(\omega_0). \tag{10.15}$$

Again, the polarization at the center wavelength is not physically interesting. What we are interested in is the depolarization that results from averaging $\hat{s}$ over the signal bandwidth. Substituting the DGD τ for $|\vec{\Omega}|$

$$\vec{r} = \hat{s}_{||}(\omega_0) + \int \frac{d\omega'}{2\pi} |f(\omega)|^2 [\cos(\tau\Delta\omega)\hat{s}_\perp(\omega_0) + \sin(\tau\Delta\omega)\hat{\Omega} \times \hat{s}_\perp(\omega_0)] \qquad (10.16)$$

The form is simplest if the signal spectrum is symmetric,

$$\vec{r} = \hat{s}_{||}(\omega_0) + \left[\int \frac{d\omega'}{2\pi} |f(\omega)|^2 \cos(\tau\Delta\omega)\right] \hat{s}_\perp(\omega_0) \qquad (10.17)$$

As we should expect, for a symmetric spectrum we recover the (real) autocorrelation function,

$$\int \frac{d\omega'}{2\pi} |f(\omega)|^2 \cos(\tau\Delta\omega) = \int \frac{d\omega'}{2\pi} |f(\omega)|^2 e^{i\tau\Delta\omega} = R_{ff}(\tau) \qquad (10.18)$$

A non-symmetric spectrum causes only a slight complication, corresponding to a nontrivial phase variation in the time-domain. The result, Eq. (10.11), is naturally the same for the time-domain and frequency-domain calculations. As before, the maximum DOP is unity (fully polarized) along the principal states. With energy evenly split among the principal states, we reach a minimum DOP of $|R_{ff}(\tau)|$,

$$\mathrm{DOP}^2 \equiv \vec{r} \cdot \vec{r} = 1 - [1 - |R_{ff}(\tau)|] |\hat{s}_{0,\perp}|^2 \qquad (10.19)$$

10.2.2 Higher-order PMD

We have seen that our basic approach efficiently deals with simple first-order systems. We now show that it is useful for more general systems. We first focus on the general form of the time-averaged output Stokes vector: A given signal spectrum and (higher-order) PMD characteristic $\vec{\Omega}(\omega)$ lead to a polarization ellipsoid. The parameters of this ellipsoid are related to the functions $f(\omega)$ and $\vec{\Omega}(\omega)$. We then turn to frequency-expansion techniques. Since PMD is typically defined in terms of such an expansion, these are more relevant to the communications problem than an exact analysis of a PMF done in the last section. Our analysis identifies limitations of the technique:

the ellipsoid can be used to extract most but not all of the information on first and second order PMD.

General relations

The starting point is again Eqs. (10.13-10.14). We can write an exact formal solution to Eq. (8.8), using the matrix A to represent the cross product operation:

$$\hat{s}(\omega) = R(\omega)\hat{s}_{\text{in}} \tag{10.20}$$

$$\frac{d}{d\omega}R(\omega) = A(\omega)R(\omega) \tag{10.21}$$

$$A = [\vec{\Omega}(\omega)\times] = \dots \tag{10.22}$$

In general, there is a nominal transformation $R(\omega_0) \neq 1$ of the polarization at ω_0:

$$\hat{s}(\omega_0) = R(\omega_0)\hat{s}_{\text{in}}. \tag{10.23}$$

This formal solution is simply a way of rewriting the problem, yet from it we can get some understanding about the general input-output relation,

$$\vec{r} = \left[\int \frac{d\omega}{2\pi}|f(\omega)|^2 R(\omega)\right]\hat{s}_{\text{in}} \equiv B\hat{s}_{\text{in}}. \tag{10.24}$$

This says that a single 3 × 3 real matrix B maps the input polarization state to the polarization measurement, even for an arbitrary, frequency-dependent $\vec{\Omega}(\omega)$. The conclusion is very important for the problem of PMD parameter estimation. A real 3 × 3 matrix B has at most 9 degrees of freedom. Thus no matter how many data points are taken, the simple measurement outlined in [14] is fundamentally limited to at most nine independent PMD parameters.

In fact, we can go further. Three degrees of freedom represent an arbitrary rotation of the input polarization, $\hat{s}_{\text{in}}$. Such a rotation is generally not observed in a communications system, and does not relate to pulse-distortion. Only frequency-dependent polarization effects, not an overall polarization transformation, distort the signal, but are reflected by only six independent parameters of the ellipsoid.

A singular value decomposition of the matrix B is one useful way to separate out the input-output transformation $\hat{s}_{\text{in}} \rightarrow \vec{r}$ into three simple geometrical parts. This gives us better intuition for what the nine degrees of freedom of B are doing. The singular value decomposition of B is

$$B = USV^T \tag{10.25}$$

where U and V are orthogonal rotation matrices and S is a diagonal scaling matrix. Thus the rotation V only serves to change the reference frame of the input:

$$\begin{aligned} \vec{r} &= B\vec{s}_{\text{in}} = USV^T\vec{s}_{\text{in}} && (10.26) \\ &\equiv US\vec{s}\,'_{\text{in}} && (10.27) \end{aligned}$$

Since the degrees of freedom associated with this rotation are unimportant, the fiber is characterized only by U and S. That is, the input $\vec{s}\,'_{\text{in}}$ is scaled differently in three orthogonal directions by matrix S and then rotated by matrix U. These operations each have 3 degrees of freedom: three real numbers scale the x, y, and z axes, and three angles rotate these axes to an arbitrary orientation.

The number of degrees of freedom gives us an important conclusion that polarimeter measurements in this simple configuration. Since each order of PMD includes three independent degrees of freedom (corresponding to the components of the vectors $\vec{\Omega}$, $\vec{\Omega}_\omega$, etc.), at most two orders of PMD parameters can be independently determined from the configuration of Figure 10-4. In fact, we show below that some second-order PMD information is unobservable as well. Modifications of the basic configuration are required for independent estimation of first- and second-order PMD.

Since the measurements are normalized to total power, $|\hat{s}_{\text{in}}| = 1$ and Eq. (10.24) is a convenient way of parameterizing a general ellipsoid. Thus the "ellipsoid" often referred to in our earlier work is in fact the most general shape possible without including non-PMD effects. Random unit input vectors *generally* give polarimeter measurement vectors $\vec{r}$ on the surface of an ellipsoid in the presence of higher-order PMD. Naturally, effects such as PDL and nonlinearities have not been included above.

Frequency expansion and PMD parameters

To relate to practical compensation schemes, we connect with the narrowband expansion in "orders" of PMD. The first, second, and higher-order PMD parameters are defined by

$$\vec{\Omega}(\omega) = \underbrace{\vec{\Omega}(\omega_0)}_{1^{\text{st}}\text{order}} + \Delta\omega \underbrace{\vec{\Omega}_\omega(\omega_0)}_{2^{\text{nd}}\text{order}} + \ldots \tag{10.28}$$

The natural approach is to plug a frequency expansion for $\hat{s}(\omega)$ into Eq. (10.4),

$$\vec{r} = \int \frac{d\omega}{2\pi}|f(\omega)|^2[\hat{s}(\omega_0) + \Delta\omega\hat{s}_\omega(\omega_0) + \Delta\omega^2\hat{s}_{\omega\omega}(\omega_0)/2 + \ldots] \tag{10.29}$$

$$= \hat{s}(\omega_0) + \hat{s}_\omega(\omega_0)\left(\int \frac{d\omega}{2\pi}|f(\omega)|^2\Delta\omega\right) + \hat{s}_{\omega\omega}(\omega_0)\left(\int \frac{d\omega}{2\pi}|f(\omega)|^2\Delta\omega^2/2\right) + \ldots \tag{10.30}$$

We adopt the shorthand $\overline{\Delta\omega^k}$ for the weighted averages on the spectrum,

$$\overline{\Delta\omega^k} \equiv \int \frac{d\omega}{2\pi}|f(\omega)|^2\Delta\omega^k \tag{10.31}$$

so that

$$\vec{r} = \hat{s}(\omega_0) + \hat{s}_\omega(\omega_0)\overline{\Delta\omega} + \hat{s}_{\omega\omega}(\omega_0)\overline{\Delta\omega^2}/2 + \ldots \tag{10.32}$$

Applying the basic equation (8.8) recursively, we can obtain derivatives of any order, for example,

$$\hat{s}_\omega(\omega) = \vec{\Omega}(\omega) \times \hat{s}(\omega) \tag{10.33}$$

$$\hat{s}_{\omega\omega}(\omega) = \frac{d}{d\omega}[\vec{\Omega}(\omega) \times \hat{s}(\omega)] \tag{10.34}$$

$$= \vec{\Omega}_\omega(\omega) \times \hat{s}(\omega) + \vec{\Omega}(\omega) \times [\vec{\Omega}(\omega) \times \hat{s}(\omega)] \tag{10.35}$$

Combining the above equations, we obtain an expression for $\vec{r}$ in terms of the first and second order PMD,

$$\begin{aligned}\vec{r} \;\approx\; & \hat{s}(\omega_0) + \overline{\Delta\omega}\vec{\Omega}(\omega_0) \times \hat{s}(\omega_0) + \frac{1}{2}\overline{\Delta\omega^2}\vec{\Omega}_\omega(\omega_0) \times \hat{s}(\omega_0) \\ & + \frac{1}{2}\overline{\Delta\omega^2}\vec{\Omega}(\omega_0) \times [\vec{\Omega}(\omega_0) \times \hat{s}(\omega_0)] \end{aligned} \qquad (10.36)$$

Estimation of first-order PMD

The experimental estimation of PMD parameters using polarimeter measurements is described in [14, 24]. It consists of fitting the measurements $\vec{r}$ to the model (10.36).

For that specific setup, the signal spectrum was fixed, and we assumed that $\overline{\Delta\omega} = 0$. In fact, this is not an assumption, but rather a definition: if the signal spectrum is not changing, we can define the center frequency ω_0 as [23]

$$\omega_0 = \int \frac{d\omega}{2\pi} |f(\omega)|^2 \omega, \qquad (10.37)$$

which is equivalent to $\overline{\Delta\omega} = 0$. For realistic fiber communications, the "orders" of PMD are parameters of an approximation. It is important to consider these approximations in the context of a measurement. For example, with $\overline{\Delta\omega} = 0$, the first- and second-order PMD vectors enter to order $\Delta\omega^2$ in the *measured* $\vec{r}$.

We can clarify this further by looking at the DOP, resulting from Eq. (10.36)

$$\mathrm{DOP}^2 = \vec{r} \cdot \vec{r} = |\hat{s}(\omega_0)|^2 + \overline{\Delta\omega^2}\hat{s}(\omega_0) \cdot \{\vec{\Omega}(\omega_0) \times [\vec{\Omega}(\omega_0) \times \hat{s}(\omega_0)]\} + \mathcal{O}(\Delta\omega^3) \quad (10.38)$$

To this order of approximation, only the first-order PMD effects the DOP ellipse. The second-order PMD effects the measurement only by creating a discrepancy between the direction of $\hat{r}$ and the direction of $\hat{s}(\omega_0)$. Since $\hat{s}(\omega_0)$ is not directly observable, this can lead to errors in the estimation of of the PSP axis.

Evaluating the vector products, we can simplify the form of the DOP

$$\mathrm{DOP}^2 \approx 1 - \overline{\Delta\omega^2}\tau^2|\hat{s}_\perp(\omega_0)|^2 \qquad (10.39)$$

where $\hat{s}_\perp$ is the component of the polarization perpendicular to the principal states. If θ is the angle between $\hat{s}(\omega_0)$ and $\hat{\Omega}$, we can write this as

$$\mathrm{DOP}^2 \approx 1 - \sin^2(\theta)\overline{\Delta\omega^2}\tau^2 \tag{10.40}$$

The first-order PMD can be estimated by varying the input polarization and measuring the variation of DOP. The principal axes occur at the maximum DOP, and τ can be computed from the minimum DOP,

$$\min_\theta(\mathrm{DOP}) \approx 1 - \frac{1}{2}\overline{\Delta\omega^2}\tau^2. \tag{10.41}$$

The result is the same ellipsoid described by Eq. (10.19), but with the autocorrelation function truncated at the $\Delta\omega^2$ term,

$$R_{ff}(\tau) = \int \frac{d\omega'}{2\pi}|f(\omega)|^2 e^{i\tau\Delta\omega} \approx 1 - \frac{1}{2}\overline{\Delta\omega^2}\tau^2. \tag{10.42}$$

The truncation emphasizes the importance of the narrowband approximation: for a real fiber with unknown, higher-order PMD, the first-order approximation is only valid when $\overline{\Delta\omega^2}\tau^2$ is small. One should be cautious about using Eq. (10.19) in the analysis of a real fiber; using Eq. (10.39) instead makes the approximations explicit.

DOP vs. observed direction In [14], we state without derivation the functional form of $|\vec{r}|$ *vs.* $\hat{r}$. This is derived as follows. Defining $P_{||}$ as the projection operator for the direction $\hat{\Omega}$, the effect of first-order PMD can be written,

$$\vec{r} \approx [1 - (1 - P_{||})\frac{1}{2}\overline{\Delta\omega^2}\tau^2]\hat{s}(\omega_0) \tag{10.43}$$

We can now use the fact that $\hat{s}(\omega_0)$ is normalized as a constraint on $\vec{r}$,

$$|[1 - (1 - P_{||})\frac{1}{2}\overline{\Delta\omega^2}\tau^2]^{-1}\vec{r}|^2 = |\hat{s}(\omega_0)|^2 = 1. \tag{10.44}$$

Then

$$
\begin{aligned}
|\vec{r}|^2 &= \frac{1}{|[1-(1-P_{\|})\frac{1}{2}\overline{\Delta\omega^2}\tau^2]^{-1}\hat{r}|^2} && (10.45) \\
&= \frac{(1-\overline{\Delta\omega^2}\tau^2)^2}{|\hat{r}-\overline{\Delta\omega^2}\tau^2 P_{\|}\hat{r}|^2} && (10.46)
\end{aligned}
$$

We can now rewrite this as

$$
\mathrm{DOP} = |\vec{r}| = \frac{1-|\vec{a}|^2}{|\hat{r}-(\vec{a}\cdot\hat{r})\vec{a}|} \qquad (10.47)
$$

where $\vec{a} \equiv \pm\sqrt{\overline{\Delta\omega^2}}\tau\hat{\Omega}$. This is exactly the expression used in [14].

Higher order parameters We would like to extract some PMD information for broad spectra, when the above approximation starts to break down. The basic approach is to relate the general DOP ellipsoid,

$$
\mathrm{DOP}^2 = \vec{r}\cdot\vec{r} = \hat{s}_{\mathrm{in}}^T B^T B \hat{s}_{\mathrm{in}} \qquad (10.48)
$$

to the PMD parameters.

We have also outlined the calculation of the DOP to order $\Delta\omega^4$ for a symmetric spectrum. The expressions are much more complicated, but can be treated systematically. It is possible to do this kind of calculation in the time-domain as well for special cases; The case of two PMF segments has been carried out exactly in [28].

The result is, of course, that additional information about the PMD can be determined using the same setup, where the input is polarization scrambled and has fixed spectrum. However, it is not possible in general to determine the first- and second-order PMD. In fact, the issue of various PMD order being entangled together becomes more severe for the higher-order terms. It remains to be determined, perhaps in simulations, whether the additional information can be utilized by a compensator.

Special case: unobservable 2nd order Unfortunately, the six ellipsoid parameters *do not* map to the six real-number parameters of the first and second order PMD.

This is demonstrated simply by the following example: Assume the PMD is

$$\vec{\Omega}(\omega) = \alpha(\omega)\vec{\Omega}_0 \tag{10.49}$$

The solution is perhaps more intuitive using the matrix representation, following Eqs. (10.21 - 10.22)

$$\begin{aligned} A(\omega) &= \alpha(\omega)A_0 && (10.50)\\ \frac{dR}{d\omega} &= \alpha(\omega)A_0R(\omega) && (10.51) \end{aligned}$$

which has the exact solution,

$$R(\omega) = e^{A_0\int_{\omega_0}^{\omega} d\omega'\alpha(\omega')}R(\omega_0). \tag{10.52}$$

The matrix B which determines the polarimeter measurements is

$$B = \int \frac{d\omega}{2\pi}|f(\omega)|^2R(\omega) = \int \frac{d\omega}{2\pi}|f(\omega)|^2 e^{A_0\int_{\omega_0}^{\omega} d\omega'\alpha(\omega')} \tag{10.53}$$

where we have substituted $R(\omega_0) = 1$ for simplicity. We now argue that the frequency-dependent α cannot be observed by the ellipsoid measurement. That is, the resulting ellipsoid cannot distinguish $\alpha(\omega)\vec{\Omega}_0$ from an effective first-order PMD $\vec{\Omega}'$ which we will construct.

The net effect of a frequency-dependent α can be thought of as distorting the spectrum. To show this, we define the modified frequency $\nu(\omega)$ so that $d\nu/d\omega = \alpha$. Then $\vec{r}$ has the same form as the first-order PMD case, but with a modified "effective" spectrum $g(\nu)$

$$B = \int \frac{d\nu}{2\pi}\frac{1}{\alpha}|f(\omega)|^2 e^{A_0\Delta\nu} = \int \frac{d\nu}{2\pi}|g(\nu)|^2 e^{A_0\Delta\nu}. \tag{10.54}$$

The ellipsoid is thus identical to the case of purely first-order PMD $\vec{\Omega}'$ with a modified magnitude

$$R_{ff}(|\vec{\Omega}'|) = R_{gg}(|\vec{\Omega}_0|), \tag{10.55}$$

and the same direction

$$\hat{\Omega}' = \hat{\Omega}_0. \tag{10.56}$$

Polarization-dependent loss

In general, a communications system will have some polarization-dependent loss (PDL). This also changes the surface of the polarimeter measurements, and must be taken into account or eliminated if accurate PMD estimates are to be made. Other signal degradations such as PDL will interfere with any PMD mitigation scheme, but are especially problematic for a feedforward scheme. any significant effect can cause deviations from the model that effect both the parameter estimation and the appropriate compensation. A feedback scheme, on the other hand, can often partially mitigate even non-PMD distortions, simply by optimizing the signal quality with respect to any available degrees of freedom. Having said that, we would like to include as many significant effects in our model so that our feedforward controller is not fooled by these effect. We outline a crude analysis of PDL below. Further theory will be required to see if the analysis applies to relevant system parameters. Within the simple model, the original PMD estimation can be improved to independently determine the PMD and PDL.

In Jones space, PDL can be included with a non-unitary transformation matrix,

$$\boldsymbol{b}(\omega) = \boldsymbol{U}(\omega)\boldsymbol{a}(\omega). \tag{10.57}$$

The communications link is a cascade of many independent fiber segments, each with small PMD and PDL,

$$\boldsymbol{U}(\omega) = \boldsymbol{T}_1(\omega)\boldsymbol{S}_1 \ldots \boldsymbol{T}_M(\omega)\boldsymbol{S}_M \tag{10.58}$$

$$\boldsymbol{T}_j(\omega) \approx (1 + i\Delta\omega\boldsymbol{M}_j)\boldsymbol{T}_j(\omega_0) \tag{10.59}$$

$$\boldsymbol{S}_j(\omega) \approx 1 + \delta\boldsymbol{L}_j. \tag{10.60}$$

Clearly the simplest approximation when both δ and $\Delta\omega$ are small is to allow only first-order terms. In this case, we neglect any product terms that might lead to an interaction of PMD and PDL. It is not clear that this is a good approximation; in fact interaction of the effects has been demonstrated in realistic systems.

If the distortions *are* sufficiently small and we neglect product terms of $\boldsymbol{M}_j$'s and $\boldsymbol{L}_k$'s, then we can reorder the PMD and PDL transformations for convenience. That is, $\boldsymbol{U}$ can be decomposed into a frequency-dependent, unitary part and a frequency-independent, non-unitary part:

$$\boldsymbol{b}(\omega) = \boldsymbol{T}(\omega)\boldsymbol{L}\boldsymbol{a}(\omega). \tag{10.61}$$

In this case, we define

$$\vec{s} = \boldsymbol{b}^\dagger \vec{\boldsymbol{\sigma}} \boldsymbol{b} = \boldsymbol{a}_0^\dagger \boldsymbol{U}^\dagger \vec{\boldsymbol{\sigma}} \boldsymbol{U} \boldsymbol{a} \tag{10.62}$$

and the evolution equation is still valid

$$\frac{d}{d\omega}\vec{s} = \vec{\Omega} \times \vec{s} \tag{10.63}$$

and so the transformation of the PMD has the same structure as before,

$$\vec{s}(\omega) = R_{\mathrm{PMD}}(\omega)\vec{s}(\omega_0) \tag{10.64}$$

$$\frac{d}{d\omega}R_{\mathrm{PMD}} = \vec{\Omega} \times R_{\mathrm{PMD}}. \tag{10.65}$$

An important difference is that now $\vec{s}(\omega)$ is not simply a rotation of the input polarization $\hat{s}_{\mathrm{in}}$. In fact, it is no longer normalized. Then $\vec{r}$ is no longer normalized to the power measured at the polarimeter $s_0 = \boldsymbol{b}^\dagger \boldsymbol{b}$. Instead, we assume the input power is constant and use this as our normalization $|\hat{s}_{\mathrm{in}}| = 1$. Additional information is derived from the polarization dependence of s_0 and of $|\vec{r}|$, (which are not equal for depolarized signals).

Instead, the transformation $\hat{s}_{\mathrm{in}} \longrightarrow \hat{s}(\omega)$ consists of PDL distortion followed by a rotation (the zeroth-order PMD). The distortion has a characteristic axis $\hat{p}$ and is given by an affine relation:

$$\hat{s}_{\mathrm{in}} \longrightarrow B_{\mathrm{PDL}}\hat{s}_{\mathrm{in}} + c\hat{p} \tag{10.66}$$

The matrix B_{PDL} and c can easily be derived from the four parameters of the polarization-dependent loss (the maximum and minimum loss, and the direction of $\hat{p}$). The overall transformation maps the normalized input $\hat{s}_{\mathrm{in}}$ to an offset ellipsoid. As long as the input power is constant, the offset $c\hat{p}$ and power measurements s_0 determine B_{PDL}, and so the ellipsoid can be adjusted to reflect only the PMD.

10.3 PMD estimation algorithms

Naturally the "ellipsoid" PMD estimation requires some numerical algorithm to obtain the ellipsoid parameters given a set of N measured points $\vec{r}_1, \ldots, \vec{r}_N$. A natural starting point is the nonlinear least squares estimator,

$$\vec{a}_{\mathrm{LS}} = \underset{\vec{b}}{\arg\min} \sum_k [|\vec{r}_k| - f(\hat{r}_k, \vec{b})]^2, \tag{10.67}$$

where the function f gives the theoretical degree of polarization (DOP), or magnitude of $\vec{r}$. For first-order PMD, the function f has three degrees of freedom, and so $\vec{b}$ is a 3-vector

$$|\vec{r}_k|_{\mathrm{theory}} = f(\hat{r}_k, \vec{b}) = \frac{1 - |\vec{b}|^2}{|\hat{r}_k - (\vec{b} \cdot \hat{r}_k)\vec{b}|}. \tag{10.68}$$

Using this strategy, we have obtained repeatable experimental estimates of the principal states. The variance of these estimates was small even in the most difficult case, when PMD is small and few measurements are needed for each estimate, as seen in the results section below.

This may in fact be the most sensible type of algorithm to use in practice. However, we have pursued another type of stripped-down algorithms, which may be easier to implement on a minimal processor and may be more insensitive to deviations from the model (due to fiber nonlinearity, for example). One interesting class of estimates use a quadratic optimization function [14],

$$\hat{a}_{\mathrm{q}} = \underset{\hat{b}}{\arg\min} \sum_k (\vec{q}_k \cdot \hat{b})^2 \quad ; \quad \vec{q}_k = g(\vec{q}_k) \tag{10.69}$$

We have shown that versions of this can yield accurate estimates, and require very little computation per measurement. However, the accuracy of estimates is considerably lower than that of the least-squares method for a given number of measurements.

10.4 PMD experiments

We have done a simple experimental demonstration of feedforward PMD compensation [61] using the ellipsoid estimation method, also described in the thesis of Patrick Chou [61]. The setup is shown in Figure 10-7. A crude "transmitter" is composed of a DFB laser at 1543nm and a lithium niobate Mach-Zehnder type modulator. The polarization is scrambled using a 4-element liquid crystal polarization controller. Polarimeter measurements are processed in Labview to obtain PMD vector estimates and compute control voltages.

Compensation is performed with a second polarization controller (a squeezed fiber, General Photonics device) and a PMD element (either a PM fiber or a JDS device composed of a polarizing beamsplitter and motorized stage). Since the polarization controller does not have a predictable response to control voltages, a calibration arm was included using a polarization-maintaining splitter (fiber coupler). The calibration device consists of a polarizing beamsplitter and matched detectors. The beamsplitter has the same alignment as the compensator, so that the output of the calibration arm is the projection of the Stokes vector on the axis of the compensator. This is exactly the projection we need to align the fast axis of the fiber with the slow axis of the compensator, thus cancelling the PMD. A lookup table stored this mapping at all sets of control voltages.

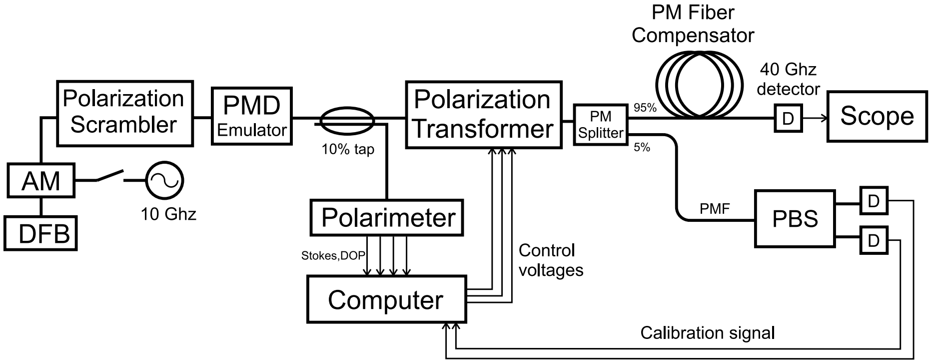


Figure 10-7: Experimental setup of a feedforward PMD compensation demo. PMD parameters are estimated by the ellipsoid method, and the principal states are aligned with a compensating element.

10.4.1 Results

We now present results confirming that PMD parameters can be accurately estimated and used for compensation. For each of a set of fixed, emulated PMDs, we took many polarimeter measurements so that several independent estimates of the PMD could be compared. Plotting each estimate of the principal state *direction* as a point in Stokes space, we obtain Figure 10-8. We see that the points are clustered tightly together, indicating that the independent measurements are consistent, or repeatable.

The results of several data sets such as that in Figure 10-8 have been condensed into Figure 10-9. Here the consistency of the PSP estimates is quantified by the angular variance from the mean. This is plotted for several values of PMD. As PMD gets larger, the ellipsoid becomes more oblong, and its axis is easier to identify. One expects that better estimates can be obtained if more polarimeter measurements are used in the ellipsoid fit. This is the case, as seen in the decrease in angular variance with the number of measurements, N. Interestingly, however, even in the worst case of small PMD and small N, the measurements are quite accurate.

Similarly, our measurements show that the magnitude of the PMD vector can be determined, as expected, from the amount of depolarization along the waist of the ellipsoid. In Figure 10-10, measured points are plotted along with the theoretical curve. The bandwidth of the input signal is used as a single fitted parameter.

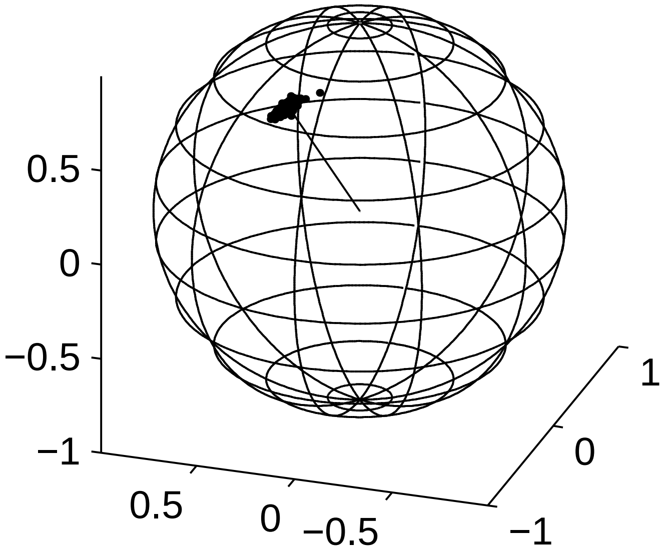


Figure 10-8: Multiple estimates of the same PSP direction are plotted in Stokes space to demonstrate repeatability of the ellipsoid estimates.

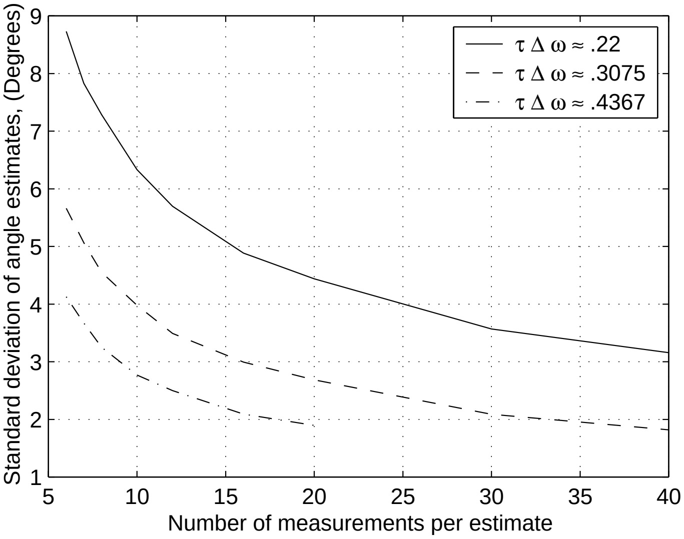


Figure 10-9: Consistency of PSP estimates are quantified by the angular deviation from the mean. This is plotted versus the number of measurements used per estimate, for several different values of DGD, τ.

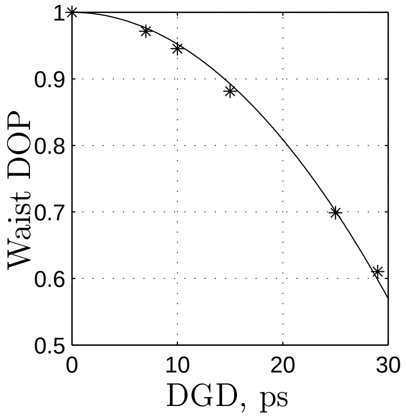


Figure 10-10: The experimental measurements of minimum DOP *vs.* DGD of the emulator are in good agreement with the theory. This confirms that we can use DOP to determine the magnitude of the PMD vector.

We should mention that determination of PSPs and magnitude only determines $\vec{\Omega}$ to within a sign. In fact, this is a fundamental ambiguity; the ellipsoid cannot distinguish $\vec{\Omega}$ from $-\vec{\Omega}$. We see several ways to get around this "fundamental" limitation in practice. For one thing, once a compensator is working, it can track slow changes in $\vec{\Omega}$ with no ambiguity. A limited degree of feedback could be introduced to help initially lock the estimate on the correct sign. A more robust approach involves using a fixed, known PMD in cascade with the unknown fiber PMD, and before the polarimeter. With this configuration, the sign of the total PMD,

$$\vec{\Omega} = \vec{\Omega}_{rmknown} + \vec{\Omega}_{rmunknown} \tag{10.70}$$

is known *a priori*, as long as the known part is larger than the unknown part. In this configuration, it is important that the fixed PMD component be stable so that it does not introduce varying rotations of the polarization.

Finally, we have set up a semi-automated demonstration of compensation using the PSP estimates. Three digital scope traces have been combined on a single axis in Figure 10-11. Clearly, the PMD distortion is quite substantial, but the original pulse shape is restored by the compensation. We lacked equipment to easily measure

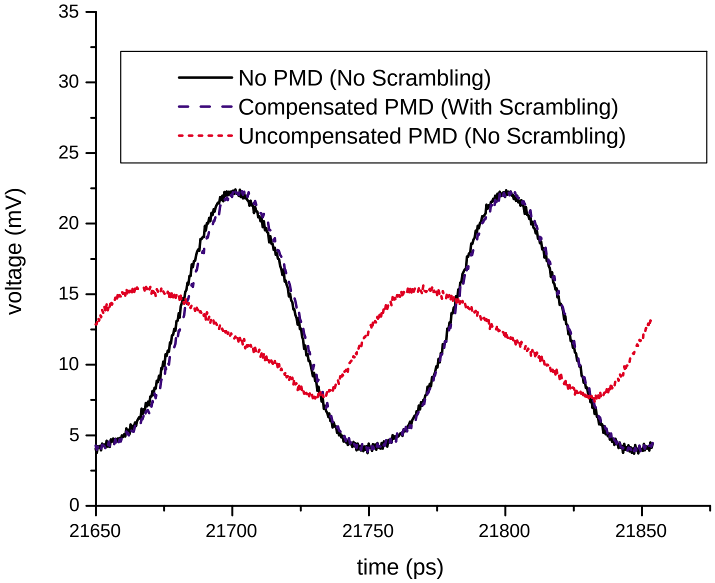


Figure 10-11: Three digital scope traces are shown here on the same axes. A large amount of PMD leads to a clearly visible distortion in the uncompensated signal. The compensated signal and undistorted signal are essentially identical.

standard metrics of the compensator performance, such as bit-error rate. To quantify the performance, we were able to estimate the residual PMD from the scrambler-induced timing jitter, depicted in Figure 10-12. That is, unless the PMD is perfectly compensated, there will still be a residual delay between the fast axis and slow axis of the total system. The polarization scrambler causes the signal to jitter back and forth between the fast and slow axes, as illustrated in Figure 10-12. For an uncompensated DGD of 40ps, the residual DGD was around 5ps.

Our setup was not fully automated and falls far short of a real system demonstration: we have not used a realistic bitpattern, a higher-order PMD emulator, or explicitly demonstrated the important known-offset variation of our method. The speed of the estimation and compensation was not pushed anywhere near its limits. However, we feel that all components needed for such a full demonstration are currently becoming commercially available.

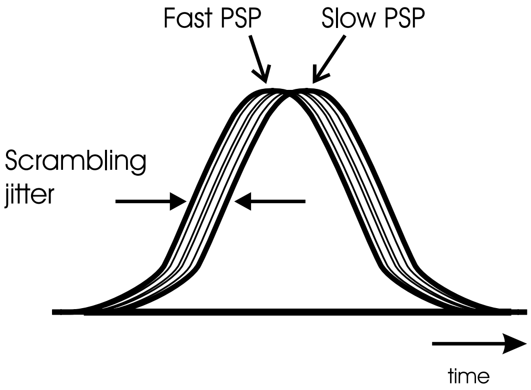


Figure 10-12: Performance of the compensator can be quantified by residual PMD, which can be inferred from the scrambler-induced timing jitter. If the PMD is not perfectly compensated, the composite system (fiber plus compensator) has a residual PMD, with fast and slow polarizations. As the signal is scrambled over all polarizations, it's time delay jitters back and forth in proportion to the residual DGD.

10.4.2 Future Work

The method has a natural extension to systems where signal bandwidth is varied. This could be implemented using a set of "test signals" with different bandwidths, or by using filters at the receiver end. This family of techniques is related to other frequency-scanning and frequency filtering methods. Our contribution includes a more general understanding of PMD and its effect on polarimeter measurements.

To implement the more general method, consider using various filters in combination with a polarimeter. One can then obtain various values of $\overline{\Delta\omega}$, $\overline{\Delta\omega^2}$, etc. With enough variation of the spectrum, one can independently determine the components in the moment expansion of $\vec{r}$ (10.32)

$$\vec{r} \;=\; \hat{s}(\omega_0) + \hat{s}_\omega(\omega_0)\overline{\Delta\omega} + \hat{s}_{\omega\omega}(\omega_0)\overline{\Delta\omega^2}/2 + \ldots \tag{10.71}$$

In the simplest case, two filters generate different $\overline{\Delta\omega}$ values, and one can independently determine $\hat{s}(\omega_0)$ and $\hat{s}_\omega(\omega_0) = \vec{\Omega}(\omega_0) \times \hat{s}(\omega_0)$. This would allow first-order, feedforward PMD compensation without a scrambler (since the parallel component of $\vec{\Omega}$ does not contribute to first order. We are considering a number of generalizations of this scheme with and without a polarization controller (scrambler) at the transmitter.

Chapter 11

Conclusions

In this thesis, we have investigated some fascinating ideas in quantum optics. Our basic goal is to find a realm beyond the semi-classical, a regime in which an optical system shows quantum dynamics qualitatively unlike the corresponding classical system. Our investigations have shown that such effects are possible in principle. Interesting physics such as multi-photon tunneling follow directly from the basic models we have studied, as confirmed by theory and simulations.

The models that we have used are also quite reasonable and standard, except that they neglect decoherence effects and device imperfections. This suggests that interesting physics may be hidden within many of the models and calculations that we study routinely in traditional quantum optics. They are obscured because remarkable quantum states are fragile.

One conclusion we can draw is a familiar one: many quantum effects seem counterintuitive because they are very difficult to observe directly. Our semiclassical intuition is formed by studying problems with typical, real world material parameters. For these problems, nonlinearities are small, and so the effect of many photons are generally seen collectively. Loss and other decoherence mechanisms are hard to neglect in real systems. By the time nonlinear phase shifts accumulate that might lead to novel quantum effects, a system is generally pushed back into the semiclassical regime because of decoherence.

The notion that photon DeBroglie effects are counterintuitive, then, does not come from nowhere. However, by playing with these mathematical models, we can develop a sense for what could be possible in this different and counterintuitive regime. One can imagine that this will be realized by discovery of a new material system, or by clever design. The rewards would be great, as the effects are strikingly different and potentially useful.

For the particular designs we have looked at, the NLDC interferometer and the momentum squeezer, the requirements of loss resistance and adiabaticity seem impossible to meet. While our simulations consistently show tunneling effects and λ/n interferometery, the delicate effects that we calculate would require physical couplers of fantastic lengths and impossible fabrication tolerances. Beyond this, various decoherence effects come in to play. These fundamentally push us towards a more classical regime, by breaking superpositions.

These results seem discouraging. However, there is reason to continue the investigation, as many questions have only been given preliminary answers. A separate analysis of the problem, using different methods to analyze similar types of devices, revealed strategies which may drastically improve the performance of the nonlinear coupler. In addition, some of the arguments that we have given against a real-world design have not been fully tested. We look forward to future investigations of the coupler design, as well as clever detectors and other elements that will bring these striking effects closer to a real-world implementation.

If fundamentally new effects are difficult, new intuition apparently comes much more easily. The configuration-space, or “particle,” perspective on quantum optics problems is very useful for understanding the normal semiclassical regime, even if the holy grail of this investigation has not been attained. Specifically, we found a new mechanism for loss-induced noise in soliton systems. The particle approach gives us a different description, a new set of intuitive pictures, but describes precisely the same physics. The equivalence of the models is confirmed by the quantitative agreement with the previous second-quantized calculation of the Gordon-Haus effect. Similarly, the particle approach suggested a simple method of generating momentum-squeezing, which we feel is currently realizable.

Appendix A

Spin and Related Quantum Systems

A.1 Introduction

A number of problems have been considered in our research related to spin systems. We attempt to summarize some of their similarities, differences, and basic relations. These physical systems all draw on the geometrical intuition of three-dimensional spin vectors. In this appendix we review the important important mathematical connection between the SU(2) and SO(3) algebraic groups, which underlies the spin representation. The related physical problems are as follows:

True two-state systems If the entire system has only two states of interest, life is simple. By the time one has identified two basis states, many complex issues have already been resolved (for example, indistinguishability and sepration of unimportant degrees of freedom of the system). The physical details may vary, but ultimately, *a two-level system is a two-level system.* The dynamics of all such system are essentially the same. The descriptions in many quantum books, such as Shankar [72], are somewhat terse, but essentially complete.

Polarization of photons The polarization of a photon can be thought of as the photon spin. This raises the confusing notion that photons are spin-1 particles and should somehow behave like three-level systems. As is commonly known, a photon in a plane-wave mode with given $\vec{k}$ has two, not three, independent polarizations. While it is worthwhile to ponder this discrepancy, does not present any difficulty in practical calculations. One can identify all classical electromagnetic modes *including polarization*, and then quantize in the usual way. Photons, from our point of view, are simply bosonic excitations of this multi-mode field. Their inherant spin does not play an important role in the analysis of our main chapters. In single-mode fiber, for instance, the polarization states of light are states of a two-mode boson problem. This approach allows us to study effects like polarization-squeezed states without treating polarization modes any differently from other second-quantized modes.

Two-mode boson problem In these problems, two interesting bosonic modes have been identified. These may be states of orthogonal polarizations, or of different spatial profiles (or both). Such a system *is not* a two-mode system, except in the limit that exactly one photon is present. More generally, N photons lead to an $N + 1$-dimensional state space, and superpositions of states with different photon numbers are of course possible as well. Calculations can be done using familiar tools of second-quantization.

System including many two-state subsystems A more general problem is the quantum system composed of two-state subsystems. The crucial question here is whether the sub-systems are distinguishable or not. The subsystems may be fundamentally indistinguishable: the individual photons in a two-mode photon problem mentioned above are indistinguishable, but each has two states it can be in. Or one can have manifestly distinguishable subsystems, like the two-state polarizations of photons in separate waveguides, or the states of two atoms of different species.

Other interesting cases are governed by pseudo-spin dynamics: a superradiant system, for example, consists of two-level atoms that are distinguishable: each atom sits at a given position, it's internal state could in principle be measured distinctly from the state of another atom at a different position. However, if the atoms interact with a field in such a way that they are indistinguishable to the field, the resulting dynamics resembles that of a spin system.

Spin of an atom, molecule, etc. The spin state of a molecule involves spins of many particles, and the state of each particle has degrees of freedom other than spin. Often, a spin system with total spin number S is extracted from the larger problem. The number of states for the spin problem is $2S + 1$ (for example, for spin $1/2$ there are two states). The spin formalism is remarkably useful for a variety of systems with a finite number of levels.

Entangled-photon problems Entangled-photon states are important in quantum computation and related fields. There are two distinct usages of "entanglement" of photons: On the one hand, many nonclassical states of a light field can loosely be thought of as having photons that are in some way correlated, or "entangled." Since the photons are indistinguishable, in some cases we could argue that there are no clear, separate subsystems to be "entangled," and so the word should not be used.

In the stricter usage, entangled photons have two or more degrees of freedom which are entangled. For example, two photons may appear in two spatially separated waveguides and have anti-correlated polarizations. In this case, there are position and polarization degrees of freedom. This is important: the positions of the photons renders them *distinguishable* in a sense, so that we can reasonably talk about the two entangled polarization states of two distinct photons: the one in waveguide A and the one in waveguide B, even though photons are indistinguishable on a more fundamental level. Since more degrees of freedom are generally included in entangled-photon systems (typically, four electromagnetic modes), one typically restricts the number of particles to make the problem manageable.

A.2 Two-state systems

Very often, a physical system is reduced to two quantum states of interest. While the physics leading up to this approximation can be incredibly varied and complex, the dynamics of all two-state systems are essentially the same, and reasonably simple.

The two states can represent anything: two atomic levels (that is, internal electronic states of an atom), two polarization states of a photon, the state of some complex quantum object being on one side or the other of a tunneling barrier, the before-and-after states of some nuclear reaction, etc. Ultimately, the states of the system are represented in terms of some basis:

$$|\Psi\rangle = \alpha|a\rangle + \beta|b\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \boldsymbol{v}. \tag{A.1}$$

The state vector is a complex 2-vector and the Hamiltonian is represented by a 2×2 Hermitian matrix:

$$\boldsymbol{H} = \begin{pmatrix} H_{aa} & H_{ab} \\ H_{ba} & H_{bb} \end{pmatrix} \tag{A.2}$$

$$H_{ab} = \langle a|\hat{H}|b\rangle, \quad \text{etc.} \tag{A.3}$$

There are four real degrees of freedom in $\hat{H}$ once one has enforced Hermiticity (H_{aa}, H_{bb} real and $H_{ab} = H_{ba}$).

If a constant Hamiltonian for a particular two-state problem is known, one can plug the values into well-known formulas for the Rabi oscillations of such a system. One can solve analytically for the eigenstates of the system,

$$\hat{H}|g\rangle = (E_0 - \hbar\Omega)|g\rangle \tag{A.4}$$

$$\hat{H}|e\rangle = (E_0 + \hbar\Omega)|e\rangle, \tag{A.5}$$

and write the evolution of an arbitrary state accordingly,

$$|\Psi(t)\rangle = e^{-iE_0t/\hbar}\left[e^{i\Omega t}\langle g|\Psi(0)\rangle|g\rangle + e^{-i\Omega t}\langle e|\Psi(0)\rangle|e\rangle\right] \tag{A.6}$$

Problems with time-varying Hamiltonians are similarly well understood. In the next section, we discuss some useful relations in dealing with such systems.

A.3 SU(2) and SO(3) Algebras: Jones and Stokes space

The linear evolution of a two-state quantum system is a unitary transformation of 2D complex vectors. Over the years, physicists have made use of a connection between this group of transformations, called the SU(2) group, and the rotations (orthogonal transformations) of real 3D vectors, the SO(3) group. In this section, we outline some useful relations connecting the two.

A.3.1 Basic properties: the static problem

Expansion of the Hamiltonian We mentioned above that the Hamiltonian has four degrees of freedom. We can expand a static Hamiltonian using the Pauli spin matrices:

$$\boldsymbol{H} = H_0\boldsymbol{I} + H_x\boldsymbol{\sigma_x} + H_y\boldsymbol{\sigma_y} + H_z\boldsymbol{\sigma_z} \tag{A.7}$$

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad ; \quad \boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad ; \quad \boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{A.8}$$

The component H_0 is clearly just an energy offset and does not effect the energy eigenstates or Rabi oscillations. We adopt a convenient vector notation for the other components of the expansion:

$$\vec{\boldsymbol{\sigma}} \equiv \begin{pmatrix} \boldsymbol{\sigma}_x \\ \boldsymbol{\sigma}_y \\ \boldsymbol{\sigma}_z \end{pmatrix}$$

$$\vec{a} \cdot \vec{\boldsymbol{\sigma}} = \sum_{j=x,y,z} a_j \boldsymbol{\sigma}_j .$$

so that the Hamiltonian can be written

$$\boldsymbol{H} = H_0\boldsymbol{I} + \vec{a}\cdot\vec{\boldsymbol{\sigma}} \tag{A.9}$$

This representation is not only compact and convenient, but improves our understanding of the properties and symmetry of the system.

Basic algebraic properties The Pauli matrices $\boldsymbol{\sigma}_j$ with $j \in \{x, y, z\}$ have the basic properties,

$$\mathrm{Tr}\{\boldsymbol{\sigma}_j\} = 0$$

$$\boldsymbol{\sigma}_j^2 = 1$$

$$\boldsymbol{\sigma}_x\boldsymbol{\sigma}_y = -\sigma_y\sigma_x = i\sigma_z$$

and cyclic permutations.

Using these, we can also derive properties of linear combinations,

$$\mathrm{Tr}\{\vec{a}\cdot\vec{\boldsymbol{\sigma}}\} = 0$$

$$(\vec{a}\cdot\vec{\boldsymbol{\sigma}})(\vec{b}\cdot\vec{\boldsymbol{\sigma}}) = \vec{a}\cdot\vec{b} + i(\vec{a}\times\vec{b})\cdot\vec{\boldsymbol{\sigma}}$$

$$[\vec{a}\cdot\vec{\boldsymbol{\sigma}}, \vec{b}\cdot\vec{\boldsymbol{\sigma}}] = 2i(\vec{a}\times\vec{b})\cdot\vec{\boldsymbol{\sigma}}$$

Using these identities, we can calculate the Rabi oscillation frequency in a simple and elegant way:

$$(\boldsymbol{H} - H_0)^2 = (\vec{a}\cdot\vec{\boldsymbol{\sigma}})(\vec{a}\cdot\vec{\boldsymbol{\sigma}}) = |\vec{a}|^2\boldsymbol{I} \tag{A.10}$$

Comparing this with the eigenvalue expressions (A.4- A.5) above, we see that,

$$\hbar\Omega = |\vec{a}|. \tag{A.11}$$

Symmetry between the matrices The above identities and eigenvalue calculation begin to reveal the geometrical intuition and symmetry of the SO(3) representation. By symmetry, we mean that any orientation in the three-space is just like any other orientation. Said another way, any three matrices $\hat{a} \cdot \vec{\boldsymbol{\sigma}}$, $\hat{b} \cdot \vec{\boldsymbol{\sigma}}$, $\hat{c} \cdot \vec{\boldsymbol{\sigma}}$ will behave exactly like the standard spin matrices as long as $\hat{a}$, $\hat{b}$, and $\hat{a}$ form a right-handed orthonormal set. They correspond to the spin matrices in a different basis, or a different frame of reference in the Stokes space.

The Pauli matrices transform to one another under change of basis. For example, the unitary matrix

$$\boldsymbol{T} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

leads to the permutation,

$$\boldsymbol{T}^{\dagger}\boldsymbol{\sigma}_x\boldsymbol{T} = -\boldsymbol{\sigma}_z \ \ ; \ \ \boldsymbol{T}^{\dagger}\boldsymbol{\sigma}_y\boldsymbol{T} = \boldsymbol{\sigma}_y \ \ ; \ \ \boldsymbol{T}^{\dagger}\boldsymbol{\sigma}_z\boldsymbol{T} = \boldsymbol{\sigma}_x$$

A system with $\hat{H} = \boldsymbol{\sigma}_x$ and one with $\hat{H} = \boldsymbol{\sigma}_z$ have essentially the same evolution, but in a different basis. The transformation above diagonalizes $\boldsymbol{\sigma}_x$. To further appreciate this point, we now study the connection between a change of basis in SU(2) and the rotations in SO(3).

Matrix Exponent The matrix exponent of an Hermitian spin matrix is a unitary transformation. The following formula is useful, for example, in analyzing the evolution of a two-mode system

$$e^{-i\vec{a}\cdot\vec{\boldsymbol{\sigma}}} = \cos(|\vec{a}|) - i\sin(|\vec{a}|)\hat{a} \cdot \vec{\boldsymbol{\sigma}} \tag{A.12}$$

Transformations: the connection A change of basis in a vector space is accomplished by a unitary transformation $\boldsymbol{U}$. The transformation acts on state vectors,

$$\tilde{\boldsymbol{v}} = \boldsymbol{U}\boldsymbol{v} \tag{A.13}$$

and on operators

$$\tilde{\boldsymbol{B}} = \boldsymbol{U}\boldsymbol{B}\boldsymbol{U}^{\dagger} \tag{A.14}$$

in a consistent way,

$$\boldsymbol{w} = \boldsymbol{B}\boldsymbol{v} \quad \longrightarrow \quad \tilde{\boldsymbol{w}} = \boldsymbol{U}\boldsymbol{w} = \boldsymbol{U}\boldsymbol{B}\boldsymbol{U}^{\dagger}\boldsymbol{U}\boldsymbol{v} \tag{A.15}$$

A unitary transformation can be written in exponential form (neglecting the offset phase),

$$\hat{U} = e^{-i\vec{a}\cdot\vec{\boldsymbol{\sigma}}}. \tag{A.16}$$

We would like to understand the transformation of a Hermitian $\boldsymbol{B} = \vec{b}\cdot\vec{\boldsymbol{\sigma}}$,

$$\boldsymbol{U}\boldsymbol{B}\boldsymbol{U}^{\dagger} = e^{-i\vec{a}\cdot\vec{\boldsymbol{\sigma}}}\vec{b}\cdot\vec{\boldsymbol{\sigma}}e^{-i\vec{a}\cdot\vec{\boldsymbol{\sigma}}} \tag{A.17}$$

We start by plugging in the exponential expression above,

$$\begin{aligned} \boldsymbol{U}\boldsymbol{B}\boldsymbol{U}^{\dagger} \;\; = \;\; & \cos(|a|)\vec{b}\cdot\vec{\boldsymbol{\sigma}}\cos(|a|) + i\cos(|a|)\sin(|a|)[(\vec{b}\cdot\vec{\boldsymbol{\sigma}})(\hat{a}\cdot\vec{\boldsymbol{\sigma}}) - (\hat{a}\cdot\vec{\boldsymbol{\sigma}})(\vec{b}\cdot\vec{\boldsymbol{\sigma}})] \\ & + \sin^2(|a|)(\hat{a}\cdot\vec{\boldsymbol{\sigma}})(\vec{b}\cdot\vec{\boldsymbol{\sigma}})(\hat{a}\cdot\vec{\boldsymbol{\sigma}}) \end{aligned} \tag{A.18}$$

after a few algebraic steps, we obtain

$$\boldsymbol{U}\boldsymbol{B}\boldsymbol{U}^{\dagger} = \vec{b}'\cdot\vec{\boldsymbol{\sigma}} \tag{A.19}$$

$$\vec{b}' = (\hat{a}\cdot\vec{b})\hat{a} + \cos(2|a|)\vec{b}_{\perp} + \sin(2|a|)\hat{a}\times\vec{b}_{\perp} \tag{A.20}$$

where

$$\vec{b}_{\perp} \equiv \vec{b} - (\hat{a}\cdot\vec{b})\hat{a} \tag{A.21}$$

is the component of $\vec{b}$ orthogonal to $\vec{a}$. This expression corresponds to a simple geometrical construction: the vector $\vec{b}'$ is obtained by taking $\vec{b}$ and rotating it about the axis $\hat{a}$ by an angle $2|\vec{a}|$.

Transformations and symmetry This connection allows us to complete the analysis of a system with Hamiltonian (A.9). The eigenvalues are $H_0 \pm |\vec{a}|$. The eigenvectors are obvious if we move into a rotated frame where $\vec{a}$ is aligned in the z-direction.

Thus the change-of-basis that diagonalizes the Hamiltonian can be constructed from the geometrical rotation in SO(3). This completes our understanding of the symmetry of the system—all directions are essentially the same, only in a different reference frame.

A.4 Photon spin and spin-1

The spin of the photon is one, which can seem confusing in light of only two possible polarizations. We offer a few comments here.

A.4.1 Second-quantized boson operators

As a practical matter, we can do calculations of nonclassical polarization states without ever considering what spin number the photon has. We simply take the spatial modes, labeled j, along with the polarization, for example labeled $+$ or $-$, to define a total mode, and then let the photon operator algebra do the rest. That is, $\hat{a}^\dagger_{j,+}$ creates a photon of one polarization, $\hat{a}^\dagger_{j,-}$ creates an orthogonally polarized photon, etc.

One can make an interesting connection between true boson operators and spin-1/2 Fermi-like operators by considering a two-mode field with only one photon on it. The basic physical intuition is that, since there is only one photon and two modes, there are only two orthogonal states. The system is thus a two-mode like like all others, and must map onto a spin-1/2 system.

Mathematically, we start with a complete basis of the two-mode field states, $\{|n_1, n_2\rangle\}$ and restrict ourselves to the states with total photon number one. The dimension of this subspace is two. We can rename the states however we like:

$$|\uparrow\rangle \equiv |1,0\rangle \quad ; \quad |\downarrow\rangle \equiv |0,1\rangle \tag{A.22}$$

We now introduce the operator

$$\hat{c} \equiv \hat{a}^\dagger_2 \hat{a}_1. \tag{A.23}$$

The operator $\hat{c}^\dagger$ then takes the state from $|\downarrow\rangle$ to $|\uparrow\rangle$. We can think of this action as "creating" an excitation, just as moving an electron is often regarded as creating an excitation of the medium.

We will now show that this operator actually has Fermi properties on the 1-photon subspace. First of all, we have the transitions

$$\hat{c}^\dagger|\downarrow\rangle = |\uparrow\rangle \quad ; \quad \hat{c}|\uparrow\rangle = |\downarrow\rangle. \tag{A.24}$$

We also have the Pauli exclusion principle, which does not allow more than one excitation,

$$\hat{c}^\dagger\hat{c}^\dagger|\downarrow\rangle = 0 \qquad ; \qquad \hat{c}^\dagger\hat{c}^\dagger|\uparrow\rangle = 0 \tag{A.25}$$

$$\hat{c}\hat{c}|\downarrow\rangle = 0 \qquad ; \qquad \hat{c}\hat{c}|\uparrow\rangle = 0. \tag{A.26}$$

Finally, we can prove that the Fermi anti-commutation relation holds on the subspace,

$$\hat{c}\hat{c}^\dagger + \hat{c}^\dagger\hat{c} = 1. \tag{A.27}$$

To see this, we go back to the defining boson operators,

$$\hat{c}\hat{c}^\dagger + \hat{c}^\dagger\hat{c} = \hat{a}_2^\dagger\hat{a}_1\hat{a}_1^\dagger\hat{a}_2 + \hat{a}_1^\dagger\hat{a}_2\hat{a}_2^\dagger\hat{a}_1 \tag{A.28}$$

Using the boson commutation rules, this becomes,

$$\hat{c}\hat{c}^\dagger + \hat{c}^\dagger\hat{c} = \hat{a}_2^\dagger\hat{a}_2\hat{a}_1\hat{a}_1^\dagger + \hat{a}_2\hat{a}_2^\dagger\hat{a}_1^\dagger\hat{a}_1 = 2\hat{a}_2^\dagger\hat{a}_2\hat{a}_1^\dagger\hat{a}_1 + \hat{a}_2^\dagger\hat{a}_2 + \hat{a}_1^\dagger\hat{a}_1 = 2\hat{n}_2\hat{n}_1 + \hat{n}_1 + \hat{n}_2 \tag{A.29}$$

On the one-photon subspace, we have

$$\hat{n}_1 + \hat{n}_2 = 1 \tag{A.30}$$

and we always have either n_1 or n_2 zero,

$$\hat{n}_2\hat{n}_1 = 0. \tag{A.31}$$

Combining the last three equations, we get the desired anti-commutation relation (A.27).

A.4.2 Angular momentum and a photon "wavefunction"

Formally, one can reconcile the spin-one character of the photon with its two-state spin as outlined in Cohen-Tannoudji [17]. Their quantum electrodynamics description involves breaking up the field into transverse modes including the polarization. The equation of motion for the "normal variables" without sources becomes

$$i\hbar\frac{d}{dt}\vec{\alpha}(\vec{k},t) = \hbar\omega\vec{\alpha}(\vec{k},t).$$

These normal variables are complex 3-vectors, much like the plane-wave amplitudes of the electric field. They are constructed so that the simple harmonic evolution of the transverse field evolution is apparent in the above equation.

Spin comes in when we ask for the angular momentum of the transverse field. This can be written as

$$\vec{J}_{\text{trans}} = \hbar\int d^3k\left[\sum_a \alpha_a^*(-i\vec{k}\times\nabla)\alpha_a - i\alpha^*\times\alpha\right].$$

The form of this expression is suggestive of a particle wavefunction: the first term is exactly what we would expect for the orbital angular momentum of a particle with wavefunction α_a. The second term, when α is a 3-vector, is the spin contribution that we would expect for a particle of spin 1. Thus, the photon is regarded as a spin-1 particle, consistent with the boson statistic of light fields.

However, the field modes were assumed to be restricted to transverse modes when the normal variables were defined. For various reasons, one cannot think of the photon as having an unrestricted wavefunction $\vec{\alpha}$. When all is said and done, the polarization must of course be transverse to $\vec{k}$. So while the spin of the photon is one, it does not exhibit three spin states as other spin-1 particles, but rather the 2 states characteristic of spin-1/2 particles. The absence of a third spin state is difficult to fully reconcile with our intuition. It is related to the relativistic idea that a photon does not have a rest frame, and thus does not have a spin that is clearly defined independent of its momentum. A more intuitive approach is presented in [37]. Here the quantization is done not of a plane wave, but of a wave with finite transverse extent. The angular momentum of the field arises in a natural way even for a *classical* field, and can then be quantized.

A.5 Two-mode boson problems

Two-mode problems come up often in quantum optics. The situation is similar to the two-state system: once two optical modes have been identified, all problems of this type basically work the same way. The physics and approximations leading up to this point may be different, but the two-mode formalism is the same.

We start with the usual states and operators. We have the vacuum $|0\rangle$ and the photon operators for the two modes $\hat{a}_1$ and $\hat{a}_2$. A perfectly reasonable basis to work in is the number states:

$$|n_1, n_2\rangle = \frac{(\hat{a}_1^\dagger)^n_1}{\sqrt{n_1!}} \frac{(\hat{a}_2^\dagger)^n_2}{\sqrt{n_2!}} |0\rangle \tag{A.32}$$

These are normalized and, of course, obey

$$\hat{a}_1^\dagger \hat{a}_1 |n_1, n_2\rangle = n_1 |n_1, n_2\rangle \tag{A.33}$$

and the other standard relations.

The Hamiltonian is not quite as simple as the two-state problem. Naturally, the state space is infinite, and so enumerating all matrix elements H_{n_1', n_2', n_1, n_2} is not possible. However, optical nonlinearities are generally weak and of low order. In the absence of optical nonlinearities,

$$\hat{H} = H_0 + \sum_{j,l=1}^{2} E_{jl}\hat{a}_j^{\dagger}\hat{a}_l \tag{A.34}$$

Again, the dynamics are clear once we diagonalize the 2×2 matrix $\boldsymbol{E}$. After diagonalization, we have two uncoupled modes. The eigenstates are simply the number states of these modes.

The case of low-order nonlinearities can lead to interesting and manageable dynamics. For example, we recently looked at systems of the type

$$\hat{H} = H_0 + \sum_{j,l=1}^{2} E_{jl}\hat{a}_j^{\dagger}\hat{a}_l + \sum_{j,k,l,m=1}^{2} V_{j,k,l,m}\hat{a}_j^{\dagger}\hat{a}_k^{\dagger}\hat{a}_l\hat{a}_m. \tag{A.35}$$

If the Hamiltonian conserves photon number, the dynamics can basically be understood by looking at the fixed N-photon number problem. This amounts to the diagonalization of an $(N+1) \times (N+1)$ matrix.

A.5.1 Spin formalism for two-mode problems

The spin formalism is essentially just like the angular momentum formalism in quantum mechanics. The basic operators are $\hat{J}_x$, $\hat{J}_y$, and $\hat{J}_z$. The total spin number S is conserved by these operators, and is often a conserved quantity of the Hamiltonian. The commutation relations are:

$$[\hat{J}_x, \hat{J}_y] = i\hat{J}_z \;\; ; \;\; [\hat{J}_y, \hat{J}_z] = i\hat{J}_x \;\; ; \;\; [\hat{J}_z, \hat{J}_x] = i\hat{J}_y \tag{A.36}$$

The eigenstates of the system are labeled by

$$M \in [-S, -S+1, \ldots, S]$$

and defined by

$$\hat{J}_z|S, M\rangle = M|S, M\rangle. \tag{A.37}$$

These $2S + 1$ states form a complete basis on the sector with total spin S. The spin vector is a set of three non-commuting variables with inherent uncertainty. The uncertainties can be related to the commutators (A.36). They can also be considered the result of quantizing a single component of the spin: for large spin numbers, the discrete nature of spin components is less noticeable, and one can construct states with well-defined spin vectors. We return to this idea below in the context of polarization states.

Properties of these operators and states can be found in the literature. For example, we have the matrix elements,

$$\hat{J}_x|S,M\rangle = [(S+1+M)(S-M)]^{1/2}/2|S,M+1\rangle + [(S+M)(S+1-M)]^{1/2}/2|S,M-1\rangle$$

All of these properties are shared exactly by two-mode boson problem once we make the following definitions:

$$S = N/2$$

$$\hat{N} = \hat{n}_1 + \hat{n}_2 = \hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2$$

$$\hat{J}_x = \frac{1}{2}(\hat{a}_1^\dagger \hat{a}_2 + \hat{a}_1^\dagger \hat{a}_2)$$

$$\hat{J}_y = \frac{1}{2i}(\hat{a}_1^\dagger \hat{a}_2 - \hat{a}_1^\dagger \hat{a}_2)$$

$$\hat{J}_z = \frac{1}{2}(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2)$$

For example, we can define a basis of states on the N-photon space $|S,M\rangle$ that follow Eq. (A.37). In fact these are simply the $|n_1, n_2\rangle$ states with

$$S = (n_1 + n_2)/2 \tag{A.38}$$

$$M = (n_1 - n_2)/2. \tag{A.39}$$

Clearly, the range of possible M is from $-S$ to S in increments of 1, just as for the spin states. We can derive the commutation relations and matrix element directly from the properties of $\hat{a}_1$, $\hat{a}_2$, and $|n_1, n_2\rangle$. Since the states and matrix elements behave identically, the models are mathematically equivalent.

The formalis lends us the geometrical intuition of magnetic/spin systems. It can also allow more compact, elegant form of a linear or non-linear Hamiltonian. For example, the Kerr component of a Hamiltonian can be written as

$$\hat{H}_{\mathrm{NL}} = \sum_{j,k,l,m=1}^{2} V_{j,k,l,m}\hat{a}_j^\dagger \hat{a}_k^\dagger \hat{a}_l \hat{a}_m = \sum_{a,b=x,y,z} M_{a,b} : \hat{J}_a \hat{J}_b : \tag{A.40}$$

Where :: is used to obtain normal-ordered expressions, for simplicity, and

$$V_{j,k,l,m} = \sum_{a,b=x,y,z} M_{a,b}[\sigma_a]_{j,m}[\sigma_b]_{k,l}. \tag{A.41}$$

Evolution of a system according to a spin Hamiltonian can then be visualized in three dimensional space; the corresponding classical evolution is simply the trajectory $\vec{J}(t)$.

Finally, we not that a spin system has a finite number of excitation levels for a given S. The level-raising operator is then neither simple bosonic nor fermionic (except on the one-photon subspace). It is a different algebra, but also well-known.

Appendix B

External Soliton Variables

In this appendix, we review the definitions of soliton variables, and make connections between momentum, group velocity, and soliton energy. The soliton variables [51] become

$$\hat{X} = \frac{1}{n}\int dx \hat{\phi}^{\dagger}(x) x \hat{\phi}(x)$$

$$\hat{P} = -i\hbar \int dx \hat{\phi}^{\dagger}(x) \frac{d}{dx}\hat{\phi}(x)$$

$$\hat{H}_{\mathrm{A}} = \hbar \int dx \hat{\phi}^{\dagger}(x) \left[\omega_0 - i\omega' \frac{\partial}{\partial x} - \frac{\omega''}{2}\frac{\partial^2}{\partial x^2}\right] \hat{\phi}(x) + \hat{H}_{nonlinear}$$

when we use a stationary frame. Position and momentum are of course canonically conjugate variables [51],

$$[\hat{X}, \hat{P}] = i\hbar.$$

The speed at which a soliton is traveling is important in analyzing timing jitter. This speed can be formalized as a group velocity operator, and can be evaluated using photon-operator algebra or using the configuration-space separation of variables described in Section 3.1.2:

$$\hat{H}_{\mathrm{sol}} = \hat{H}_{\mathrm{rel}} + \omega_0 \hat{N} + \omega' P + \frac{\omega''}{2\hbar}\hat{P}^2 \hat{N}^{-1} \tag{B.1}$$

The velocity operator is

$$\hat{v}_g \equiv \frac{1}{i\hbar}[\hat{X}, \hat{H}] = \omega' + \frac{\omega''}{\hbar}\hat{P}\hat{N}^{-1}. \tag{B.2}$$

In a fiber, the group velocity is nearly constant (that is $\omega' \gg \omega''\Delta k$) and so $\hat{v}_g \approx \omega' \equiv v_{g0}$. In this case, $\hat{H}_X \approx \omega'\hat{P}$, and the soliton frequency, velocity, and momentum are all related approximately linearly.

Finally, we can relate the group velocity to the total soliton energy,

$$\begin{aligned} \hat{v}_g &\approx v_{g0} + \frac{\omega''}{\hbar}\hat{P}\hat{N}^{-1} && \text{(B.3)} \\ &\approx v_{g0} + \frac{\omega''}{\hbar v_g 0}(\hat{H}_A - E_s n)\hat{N}^{-1} && \text{(B.4)} \end{aligned}$$

When uncertainty in photon number is small, this means

$$\langle \Delta\hat{v}_g^2 \rangle \approx \left(\frac{\omega''}{\hbar\omega'}\right)^2 \langle (\Delta\hat{H}\hat{N}^{-1})^2 \rangle.$$

What all of this means is that group velocity $\hat{v}_g$, momentum $\hat{P}$ and soliton energy $\hat{H}_{\mathrm{A}}$ are all closely related variables. For a given photon number, they all represent precisely the same degree of freedom. Indeed, they all are approximately linearly related to one another. The timing jitter arising from velocity noise is then closely connected to energy disturbances.

Appendix C

Perturbation of localized state

One strength of our formalism is its ability to treat states with general momentum wavefunctions. Given an initial pure quantum wavefunction for momentum (implying states of uncertain energy), we will find a perturbed soliton in a statistical mixture of pure quantum states. There is then a coherent perturbation and a statistical broadening associated with a given physical expectation.

We start with

$$\sin(\theta)|\Psi_1(t_f)\rangle = \int dh_i a(h_i) \int d\epsilon T(\epsilon; h_i)|nh_i + \epsilon, E_{Bi} - \epsilon\rangle \tag{C.1}$$

$$\sin(\theta)|\Psi_1(t_f)\rangle = \int d\epsilon |E_{Bi} + \epsilon\rangle \otimes \int dh_i a(h_i) T(\epsilon; h_i)|nh_i - \epsilon\rangle \tag{C.2}$$

In the limit that $a(h_i)$ is infinitely narrow, we start with an energy eigenstate, and end up with a statistical mixture of energy eigenstates,

$$\sin(\theta)|\Psi_1(t_f)\rangle \quad \overset{|a|^2 \longrightarrow \delta(h_i - \bar{h})}{\longrightarrow} \quad \int d\epsilon T(\epsilon; \bar{h})|E_{Bi} + \epsilon\rangle \otimes |n\bar{h} - \epsilon\rangle \tag{C.3}$$

This is equivalent to a density operator representation of the soliton subsystem which is a statistical mixture of energy eigenstates.

For cases of practical interest, $a(h_i)$ is very narrowly confined about $h_i = \bar{h}$ compared to the slow variation of T with h_i. However, the width of a is physically important, and so the simplification (C.3) is undesirable. It is useful to reshape the expression (C.2)

$$\sin(\theta)|\Psi_1(t_f)\rangle = \underbrace{\int d\epsilon T(\epsilon;\bar{h})|E_{Bi}+\epsilon\rangle}_{\text{statistical spread}} \otimes \underbrace{\int dh_i \tilde{a}_\epsilon(h_i)|nh_i-\epsilon\rangle}_{\text{coherent wavefunction}} \tag{C.4}$$

where the new wavefunction,

$$\tilde{a}_\epsilon(h_i) = a(h_i)\frac{T(\epsilon;h_i)}{T(\epsilon;\bar{h})} \tag{C.5}$$

will be a slightly perturbed version of a in the limit that T is slowly varying compared to a. This is the usual situation for physical states: for coherent-state solitons [35]: T will have a width corresponding to the classical spectrum $\hbar\Delta\omega_{\mathrm{clas}}$, but $a(h_i)$ has a spread on the order $\hbar\Delta\omega_{\mathrm{clas}}/\sqrt{n}$, which is often smaller by a factor of a thousand or more.

Equation (C.4) corresponds to a very interesting picture of an input and output soliton. Imagine a is a Gaussian momentum wavefunction for a localized pulse in the fiber. After losing a photon, the soliton is in a mixture of quantum states, each having only slight distortion in its momentum wavefunction. In fact,

$$a(h_i) = e^{-(h_i-\bar{h})^2/4\sigma_a^2} \tag{C.6}$$

gives rise to

$$\begin{aligned} \tilde{a}_\epsilon(h_i) &= e^{-(h_i-\bar{h})^2/4\sigma_a^2}\frac{\operatorname{sech}\left(\frac{\epsilon-h_i}{\Delta E_{clas}}\right)}{\operatorname{sech}\left(\frac{\epsilon-\bar{h}}{\Delta E_{clas}}\right)} && \text{(C.7)} \\ &\approx e^{-r(h_i-\bar{h}-\delta h)^2/4\sigma_a^2}. && \text{(C.8)} \end{aligned}$$

That is, as long as a is narrow compared to T, an initially Gaussian wavefunction remains Gaussian, but with small changes in its parameters. The perturbations to parameters $\bar{h}$ and σ_a are small, and given by

$$\frac{r}{\sigma_a^2} = \frac{1}{\sigma_a^2} + \frac{2}{\Delta E_{\text{clas}}^2}\text{sech}\left(\frac{\epsilon - \bar{h}}{\Delta E_{\text{clas}}}\right) \tag{C.9}$$

$$\delta h = \frac{2\sigma_a^2}{\Delta E_{\text{clas}}}\tanh\left(\frac{\epsilon - \bar{h}}{\Delta E_{\text{clas}}}\right) \tag{C.10}$$

There are several interesting things to note. first of all δh is indeed small compared to σ_a for $\sigma_a \ll \Delta E_{\text{clas}}$. Similarly, the new Gaussian width is close to the unperturbed width. Finally, we recognize that the center-of-mass position spread $\langle \Delta \hat{X}^2 \rangle$ for each coherent component is proportional to r/σ_a^2. Specifically,

$$\Delta X_{\tilde{a}}^2 = \int dh_i \tilde{a}^\dagger(h_i) \Delta \hat{X}^2 \tilde{a}(h_i) \approx \Delta X_a^2 + \frac{\Delta X_{\text{clas}}^2}{n^2}\text{sech}(\ldots) \tag{C.11}$$

Statistical mixtures and entangled states

Above, we have referred to an entanglement as a statistical mixture. We here clarify this statement. If systems A and B are entangled and not interacting, then the entanglement with B is effectively the same as a statistical mixing *with regard to expectations in operators on system A.*

Take the entangled state,

$$|\Psi\rangle = \int d\mu b(\mu) |\mu\rangle \otimes |\psi_\mu\rangle, \tag{C.12}$$

where the kets $|\mu\rangle$ are an orthogonal continuum of states for an independent, non-interacting system. Then if the operator $\hat{\Omega}$ acts on the ψ states only,

$$\langle \Psi | \hat{\Omega} | \Psi \rangle = \int d\mu |b(\mu)|^2 \langle \psi_\mu | \hat{\Omega} | \psi_\mu \rangle. \tag{C.13}$$

Here $|b|^2$ plays the role of $p(\mu)$, and essentially represents a classical probability for the statistical mixture. In the more typical density operator representation, we can write,

$$\hat{\rho}_{\text{A}} = \int d\mu |b(\mu)|^2 |\psi_\mu\rangle\langle\psi_\mu|. \tag{C.14}$$

Appendix D

Internal state of a perturbed soliton

The action of $\hat{\phi}$ on an n-photon soliton state generates components of the field state with perturbed internal coordinates. It can be shown, however, that the loss of a single photon leaves a dominant component in the internal ground state of an $n-1$ soliton. In this paper, we have used this fact to neglect the excited internal-state contributions. The proof is taken with minor modifications from [32].

The state with one photon removed is denoted $|\bar{\Phi}_n\rangle$, and the true $n-1$ ground-state is written $|\Phi_{n-1}\rangle$. We would like to show that these state have an overlap of very nearly one:

$$\langle\Phi_{n-1}|\bar{\Phi}_n\rangle \approx 1.$$

An extremely similar overlap integral is performed in [32, Appendix A]. In that case, the photon was not removed locally; an additional correction can account for the shift in X by our local annihilation operator. The result is qualitatively the same,

$$\langle\Phi_{n-1}|\bar{\Phi}_n\rangle = 1 - O(1/n).$$

The soliton term dominates as n becomes large.

Appendix E

Time-domain ellipsoid calculation

In this appendix, we go through the complete time-domain analysis of a strictly first-order PMD system. This model could apply, for example to a single polarization-maintaining fiber with negligible PDL. Since the system has no higher-order PMD, the narrowband expansion in frequency is unnecessary: as we see below, the problem is tractable with no further approximations.

In general, a single birefringent element has a DGD τ and orthogonal principal states $\boldsymbol{e}_k$. The effect of the birefringence is a differential linear phase

$$\boldsymbol{T}(\omega) = e^{i\omega\tau/2}\boldsymbol{e}_1\boldsymbol{e}_1^\dagger + e^{-i\omega\tau/2}\boldsymbol{e}_2\boldsymbol{e}_2^\dagger \tag{E.1}$$

or, more simply, a differential time delay,

$$\boldsymbol{b}(t) = f(t-\tau/2)\boldsymbol{e}_1(\boldsymbol{e}_1^\dagger\boldsymbol{a}_0) + f(t+\tau/2)\boldsymbol{e}_2(\boldsymbol{e}_2^\dagger\boldsymbol{a}_0) \tag{E.2}$$

where the input field has been rewritten as

$$\boldsymbol{a}(t) = f(t)\boldsymbol{a}_0 \tag{E.3}$$

$$f(t) \equiv \int d\omega f(\omega)e^{-i\omega t}/2\pi. \tag{E.4}$$

The polarimeter measurement can now be derived by plugging into (10.1)

$$\begin{aligned} \vec{r} &= \int dt[f^*(t-\tau/2)\boldsymbol{e}_1^\dagger(\boldsymbol{e}_1^\dagger\boldsymbol{a}_0)^* + f^*(t+\tau/2)\boldsymbol{e}_2^\dagger(\boldsymbol{e}_2^\dagger\boldsymbol{a}_0)^*]\vec{\boldsymbol{\sigma}} && \text{(E.5)} \\ & [f(t-\tau/2)\boldsymbol{e}_1(\boldsymbol{e}_1^\dagger\boldsymbol{a}_0) + f(t+\tau/2)\boldsymbol{e}_2(\boldsymbol{e}_2^\dagger\boldsymbol{a}_0)] && \text{(E.6)} \end{aligned}$$

The four resulting terms are actually quite manageable when we recognize the form of the time integrals: we have the normalization integral,

$$\int dt|f(t-\tau/2)|^2 = \int dt|f(t+\tau/2)|^2 = 1 \tag{E.7}$$

and the autocorrelation,

$$R_{ff}(\tau) = \int dtf(t-\tau/2)^* f(t+\tau/2) = \left[\int dtf(t+\tau/2)^* f(t-\tau/2)\right]^* \tag{E.8}$$

Then,

$$\begin{aligned} \vec{r} &= |\boldsymbol{e}_1^\dagger\boldsymbol{a}_0|^2\boldsymbol{e}_1^\dagger\vec{\boldsymbol{\sigma}}\boldsymbol{e}_1 + |\boldsymbol{e}_2^\dagger\boldsymbol{a}_0|^2\boldsymbol{e}_2^\dagger\vec{\boldsymbol{\sigma}}\boldsymbol{e}_2 && \text{(E.9)} \\ & +R_{ff}(\tau)^*(\boldsymbol{e}_2^\dagger\boldsymbol{a}_0)^*(\boldsymbol{e}_1^\dagger\boldsymbol{a}_0)\boldsymbol{e}_2^\dagger\vec{\boldsymbol{\sigma}}\boldsymbol{e}_1 + R_{ff}(\tau)(\boldsymbol{e}_1^\dagger\boldsymbol{a}_0)^*(\boldsymbol{e}_2^\dagger\boldsymbol{a}_0)\boldsymbol{e}_1^\dagger\vec{\boldsymbol{\sigma}}\boldsymbol{e}_2 && \text{(E.10)} \end{aligned}$$

The geometrical interpretation of this expression is actually quite clear once we give the various terms a Stokes-domain interpretation. The principal state axis in Stokes space is

$$\hat{i}_x = \boldsymbol{e}_1^\dagger\vec{\boldsymbol{\sigma}}\boldsymbol{e}_1 = -\boldsymbol{e}_2^\dagger\vec{\boldsymbol{\sigma}}\boldsymbol{e}_2 \tag{E.11}$$

For a normalized $\boldsymbol{a}_0$, the power splitting along the two polarizations can be expressed as an angle,

$$\cos^2(\theta/2) = |\boldsymbol{e}_1^\dagger\boldsymbol{a}_0|^2 \tag{E.12}$$

$$\sin^2(\theta/2) = |\boldsymbol{e}_2^\dagger\boldsymbol{a}_0|^2 \tag{E.13}$$

The other axes of the coordinate system can be defined as

$$\hat{i}_y = \mathcal{R}e\{\boldsymbol{e}_2^\dagger\vec{\boldsymbol{\sigma}}\boldsymbol{e}_1\} \tag{E.14}$$

$$\hat{i}_z = \mathcal{I}m\{\boldsymbol{e}_2^\dagger \vec{\boldsymbol{\sigma}} \boldsymbol{e}_1\} \tag{E.15}$$

Finally, the time-averaged Stokes vector becomes

$$\vec{r} \;=\; \cos(\theta)\hat{i}_x + |R_{ff}(\tau)| \left[\sin(\theta)\cos(\phi)\hat{i}_y + \sin(\theta)\sin(\phi)\hat{i}_z\right] \tag{E.16}$$

where ϕ is defined by

$$\phi = \arg\{R_{ff}(\tau)(\boldsymbol{e}_2^\dagger \boldsymbol{a}_0)(\boldsymbol{e}_1^\dagger \boldsymbol{a}_0)^*\} \tag{E.17}$$

The geometrical importance becomes transparent when we compare $\vec{r}$ with the Stokes vector of the input signal, $\hat{s}_0$. The calculation is trivial, we simply set the PMD to zero, $R_{ff}(0) = 1$, $\phi_0 = \arg\{\boldsymbol{e}_2^\dagger \boldsymbol{a}_0(\boldsymbol{e}_1^\dagger \boldsymbol{a}_0)^*\}$, so that

$$\hat{s}_0 \;=\; \underbrace{\cos(\theta)\hat{i}_x}_{\hat{s}_{0,\|}} + \underbrace{\sin(\theta)\cos(\phi_0)\hat{i}_y + \sin(\theta)\sin(\phi_0)\hat{i}_z}_{\hat{s}_{0\perp}} \tag{E.18}$$

The transformation of input Stokes vector $\hat{s}_0$ to output $\vec{r}$ by a first-order PMD system is very simple: It consists of a rotation about the principal state axis $\phi_0 \rightarrow \phi$ and a scaling down of the Stokes components orthogonal to the principal axis. If we assume for simplicity that R_{ff} is real (essentially, this means removing the carrier frequency component of the phase), the polarization-dependent scaling can be written as

$$\vec{r} = \hat{s}_{0,\|} + R_{ff}(\tau)\hat{s}_{0,\perp}. \tag{E.19}$$

This describes an ellipsoid with major axis of length one along the principal states. The two minor axis are of equal length, representing the minimum degree of polarization measured:

$$\min(\text{DOP}) = R_{ff}(\tau) \tag{E.20}$$

Bibliography

[1] Agrawal. *Nonlinear fiber optics.* Academic, New York, 1995.

[2] A. P. Alodzhants and S. M. Arakelyan. *JETP*, 81:34, 1995.

[3] D. Andresciani, F. Curti, et al. Measurement of the group-delay difference between the principal states of polarization on a low-birefringence terrestrial fiber cable. *Opt. Lett.*, 12:844, October 1988.

[4] S. M. Barnett, N Imoto, and B. Huttner. Photonic de Broglie wave interferometers. *J. Mod. Opt.*, 45:2217, 1998.

[5] G. Bjork, A. Trifonov, et al. *Quantum Semiclass. Opt.*, 10:705, 1998.

[6] Roy S Bondurant and J. H. Shapiro. Squeezed states and phase-sensing interferometers. *Phys. Rev. D*, 30:2548, 1984.

[7] A. N. Boto, P. Kok, et al. Quantum interferometric optical lithography: exploring entanglement to beat the diffraction limit. *Phys. Rev. Lett.*, 85:2733, 2000.

[8] H. B. Bülow. System outage probability due to first- and second-order PMD. *IEEE Photon. Tech. L.*, 10(5):696, May 1998.

[9] R. A. Campos, B. E. A. Saleh, and M. C. Teich. Quantum-mechanical lossless beamsplitter: SU(2) symmetry and photon statistics. *Phys. Rev. A*, 40:1371, 1989.

[10] C. M. Caves. *Phys. Rev. D*, 23:1693, 1981.

[11] A. Chefles and S. M. Barnett. Quantum theory of two-mode nonlinear directional couplers. *J. Mod. Opt.*, 43:709, 1996.

[12] S. V. Chernikov and P. V. Mamyshev. Femtosecond soliton propagation in fibers with slowly decreasing dispersion. *J. Opt. Soc. B*, 8:1633, August 1991.

[13] A. S. Chirkin, A. A. Orlov, and D. Y. Parashchuk. *Quantum Electron.*, 23:870, 1993.

[14] P. C. Chou, J. M. Fini, and H. A. Haus. Real-time principal state characterization technique for feedback-free PMD compensation. *IEEE Photon. Tech. Lett.*, 13:568, June 2001.

[15] P. C. Chou and H. A. Haus. patent, filed February 2000.

[16] Claude Cohen-Tannoudji. *Atom-Photon Interactions*. Wiley, New York, 1992.

[17] Claude Cohen-Tannoudji. *Photons and atoms*. Wiley, New York, 1997.

[18] G. M. D'Ariano and C. Macchiavello. *Phys. Rev. A*, 57:3131, 1998.

[19] P. D. Drummond, R. M. Shelby, et al. Quantum solitons in optical fibers. *Nature*, 365:307, 1993.

[20] M. Enz and R. Schilling. *J. Phys. C.*, 19:L711–5, 1765–70, 1986.

[21] J. M. Fini, P. L. Hagelstein, and H. A. Haus. *Phys. Rev. A*, 57:4842, 1998.

[22] J. M. Fini, P. L. Hagelstein, and H. A. Haus. *To appear in Phys. Rev. A*, 2001.

[23] J. M. Fini and H. A. Haus. Accumulation of polarization mode dispersion in cascades of compensated optical fibers. *Photon. Tech. Lett.*, 13:124, February 2001.

[24] John M. Fini, Patrick C. Chou, and Hermann A. Haus. Estimation of polarization dispersion parameters for compensation with reduced feedback. In *Optical Fiber Communications Conference (OFC'01)*, page WAA6, March 2001.

[25] E. J. S. Fonseca, C. H. Monken, and S. Padua. Measurement of the DeBroglie wavelength of a multi-photon wavepacket. *Phys. Rev. Lett.*, 82:2868, 1999.

[26] G. J. Foschini and C. D. Poole. Statistical theory of polarization dispersion in single mode fibers. *J. Lightwave Technol.*, 9(11):1439, November 1991.

[27] C. Francia, F Bruyere, et al. Simple dynamic polarisation mode dispersion compensator. *Electron. Lett.*, 35(5):414, March 1999.

[28] Goldhar et al. Submitted to Photon. Tech. L.

[29] J. P. Gordon and H. A. Haus. *Opt. Lett.*, 11:665, 1986.

[30] J. P. Gordon and H. Kogelnik. PMD fundamentals: Polarization mode dispersion in optical fibers. *Proc. Nat. Acad. Sciences*, 97(9):4541, April 2000.

[31] P. L. Hagelstein. unpublished.

[32] P. L. Hagelstein. *Phys. Rev. A.*, 54:2426, 1996.

[33] H. A. Haus. *J. Opt. Soc. Am. B.*, 12:2019, 1995.

[34] H. A. Haus. *Electromagnetic noise and quantum optical measurements.* Springer, New York, 2000.

[35] H. A. Haus and Y. Lai. *J. Opt. Soc. Am. B*, 7:386, 1990.

[36] H. A. Haus and J. A. Mullen. *Phys. Rev.*, 128:2407, 1962.

[37] H. A. Haus and J. L. pan. *Am. J. Phys.*, 61:818, 1993.

[38] Helstrom. *Quantum Detection and Estimation Theory.* Academic, 1976.

[39] R. Horák. *Optics Comm.*, 72:239, 1989.

[40] L. G. Joneckis and J. H. Shapiro. Quantum propagation in a kerr medium: lossless, dispersionless fiber. *J. Opt. Soc. Am. B*, 1102:2, 1993.

[41] M. Karlsson. Polarization mode dispersion-induced pulse broadening in optical fibers. *Opt. Lett.*, 23(9):688, May 1998.

[42] M. Karlsson, C. Xie, et al. Higher order polarization mode dispersion compensator with three degrees of freedom. In *Optical Fiber Communications Conference (OFC'01)*, number MO1, 2001.

[43] F. X. Kartner and L. Boivin. *Phys. Rev. A*, 53:454, 1996.

[44] D. J. Kaup. Perturbation theory for solitons in optical fibers. *Phys. Rev. A*, 42:5689, 1990.

[45] R. Khosravani, S. A. Havstad, et al. Simultaneous PMD compensation of multiple WDM channels using a single compensator. In *ECOC'00*, 2000.

[46] R. Khosravani, Y. Xie, et al. Limitations to first-order pmd compensation in wdm systems due to xpm-induced pspchanges. In *Optical Fiber Communications Conference (OFC'01)*, page WAA5, March 2001.

[47] J. A. Kong. *Electromagnetic Wave Theory.* Wiley, New York, 1990.

[48] N. Korolkova and J. Perina. Quantum statistics and dynamics of Kerr nonlinear couplers. *Opt. Comm.*, 136:135, 1997.

[49] H. H. Kuehl. Solitons on an axially nonuniform optical fiber. *J. Opt. Soc. B*, 5:709, March 1988.

[50] P. G. Kwiat, W. A. Vareka, et al. Correlated two-photon interference in a dual-beam Michaelson interferometer. *Phys. Rev. A*, 41:2910, 1990.

[51] Y. Lai and H. A. Haus. *Phys. Rev. A*, 40:844, 1989.

[52] Y. Lai and H. A. Haus. *Phys. Rev. A*, 40:854, 1989.

[53] S. Lanne, W. Idler, et al. Demonstration of adaptive pmd compensation at 40gb/s. In *Optical Fiber Communications Conference (OFC'01)*, number TuP3, 2001.

[54] D Mahgerefteh and C. R. Menyuk. Effect of first-order pmd compensation on the statistics of pulse broadening in a fiber with randomly varying birefringence. *IEEE Photon. Tech. Lett.*, 11:340, March 1999.

[55] Y Matsui, M. D. Pelusi, and A Suzuki. Generation of 20-fs optical pulses from a gain-switched laser diode by a four-stage soliton compression scheme. *IEEE Photon. Tech. Lett.*, 11(10):1217, October 1999.

[56] A. Mecozzi et al. *Opt. Lett.*, 16:1841, 1991.

[57] Messiah. *quantum mechanics.* Wiley, New York, 1976.

[58] A Mostofi, H. Hatami-Hanza, and P. L. Chu. Optimum dispersion profile for compression of fundamental solitons in dispersion decreasing fibers. *IEEE J. Quant. Elec.*, 33:620, April 1997.

[59] L. E. Nelson, R. M. Jopson, and H. Kogelnik. Polarization mode dispersion penalties associated with rotation of principal states of polarization in optical fiber. In *Optical Fiber Communications Conference (OFC'00)*, 2000.

[60] Z. Y. Ou. *Phys. Rev. Lett.*, 77:2352, 1996.

[61] Patrick Chien pang Chou. *Optical pulse distortion and manipulation through polarization effects and chromatic dispersion.* PhD thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, February 2001.

[62] J. Perina, B. E. A. Saleh, and M. C. teich. Multiphoton absorption cross section nad virtual-state spectroscopy for the entangled n-photon state. *Phys. Rev. A*, 57:3972, 1998.

[63] P. B. Phua, J. M. Fini, and H. A. Haus. *to be submitted.*

[64] C. D. Poole, S. Bergano, et al. Polarizaton dispersion and principle states in a 147-km undersea lightwave cable. *J. Lightwave Technol.*, 6(7):1185, July 1988.

[65] J. G. Rarity, P. R. Tapster, et al. Two-photon interference in a Mach-Zehnder interferometer. *Phys. Rev. Lett.*, 65:1348, 1990.

[66] M. Romagnoli, P. Franco, et al. Time-domain fourier optics for polarization-mode dispersion compensation. *Opt. Lett.*, 24:1197, September 1999.

[67] H. Rosenfeldt, Ch. Knothe, et al. Automatic PMD compensation at 40Gbit/s and 80Gbit/s using a 3-dimensional DOP evaluation for feedback. In *Optical Fiber Communications Conference (OFC'01)*, page PD27, March 2001.

[68] H. Rosenfeldt, R. Ulrich, et al. PMD compensation in 10Gbit/s NRZ field experiment using polarimetric error signal. *Elec. Lett.*, 36:448, March 2000.

[69] G. Scharf, W. F. Wreszinski, and J. L. van Hemmen. Tunnelling of a large spin: mapping onto a particle problem. *J. Phys. A.*, 20:4309, 1987.

[70] B. L. Schumaker. Noise in homodyne detection. *Opt. Lett.*, 9:189, 1984.

[71] A. C. Scott and J. C. Eilbeck. *Phys. Lett. A*, 119:60, 1986.

[72] Shankar. *Principles of quantum mechanics*. Plenum, 1994.

[73] A. Shtaif, A. Mecozzi, et al. A compensator for the effects of high-order polarization mode dispersion in optical fibers. *IEEE Photonic. Tech. L.*, 12(4):434, April 2000.

[74] H. Sunnerud, M. Karlsson, and P. A. Andrekson. Analytic theory for PMD-compensation. *IEEE Photon. Tech. Lett.*, 12:50, January 2000.

[75] P. M. Sylla, C. J. K. Richardson, et al. A technique for characterization and visualization of PMD. In *Conference on Lasers and Electro-Optics (CLEO'01)*, page CThL55, May 2001.

[76] K. Tajima. Compensation of soliton broadening in nonlinear optical fibers with loss. *Opt. Lett.*, 12:54, January 1987.

[77] M. C. Teich and B. E. A. Saleh. Entangled photon microscopy, spectroscopy, and display. U. S. Patent no. 5,796,477, August 1998.

[78] J. L. van Hemmen and A Sütõ. Semiclassical quantization and resonance in spin tunnelling. *J. Phys. A.*, 31:10029, 1998.

[79] P. L. K. Wa, J. E. Sitch, et al. All-optical multiple-quantum-well waveguide switch. *Electron. Lett.*, 21:26, 1985.

[80] D Wang and C. R. Menyuk. Calculation of penalties due to polarization effects in a long-haul wdm system using a stokes-parameter model. *J. Lightwave Tech.*, 19(4):487, April 2001.

[81] B. Wedding, A. Chiarotto, et al. Fast adaptive control for electronic equalization of pmd. In *Optical Fiber Communications Conference (OFC'01)*, number TuP4, 2001.

[82] M. Xiao, L. A. Wu, and H. J. Kimble. Precision measurement beyond the shot-noise limit. *Phys. Rev. Lett.*, 59:278, 1987.

[83] B. Yoon and J. W. Negele. *Phys. Rev. A*, 16:1451, 1977.

[84] C. X. Yu, H. A. Haus, and E. P. Ippen. *Optics Letters*, 26, May 2001.

[85] Q. Yu and A. E. Willner. Comparison of optical pmd compensation using a variable and fixed differential group delays. In *Optical Fiber Communications Conference (OFC'01)*, number MO2, 2001.

[86] H. P. Yuen. *Phys. Rev. Lett.*, 56:2176, 1986.

[87] B. Yurke, S. L. McCall, and J. R. Klauder. *Phys. Rev. A*, 33:4033, 1986.