6.453 Quantum Optical Communication
Spring 2009

Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science

6.453 Quantum Optical Communication

Lecture Number 8
Fall 2008

Jeffrey H. Shapiro

Date: Thursday, October 2, 2008

Reading: For eigenkets of $\hat{y} \equiv \hat{a}_A \otimes \hat{I}_B + \hat{I}_A \otimes \hat{a}_B^\dagger$:

- J.H. Shapiro and S.S. Wagner, "Phase and amplitude uncertainties in heterodyne detection," IEEE J. Quantum Electron. **QE-20**, 803–813 (1984).

For quantum photodetection:

- C.C. Gerry and P.L. Knight, *Introductory Quantum Optics* (Cambridge University Press, Cambridge, 2005) Sect. 5.2.
- L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995), Sects. 12.1–12.4, 12.9, 12.10.
- H.P. Yuen and J.H. Shapiro, "Optical communication with two-photon coherent states—part III: quantum measurements realizable with photoemissive detectors," IEEE Trans. Inform. Theory **IT-26**, 78–92 (1980).

For semiclassical photodetection:

- L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995), Sects. 9.1–9.8.
- R.M. Gagliardi and S. Karp, *Optical Communications* (New York, Wiley, 1976) Chap. 2.

Introduction

In this lecture we shall complete our treatment of the $\hat{a}$ positive operator-valued measurement by connecting the POVM approach to the more familiar case of observables. As a prelude to that work, however, we shall cast the observables with which we are already acquainted—the number operator and the quadrature operators—into the POVM mold. This is a worthwhile exercise because quantum information science typically uses POVMs for its characterization of measurements, as many of these measurements do *not* correspond to observables on the state space of the system being measured. Thus it behooves us to become comfortable with POVMs. Once we complete this POVM work, we'll move on to contrasting the semiclassical and quantum theories of ideal photodetection for single-mode fields.

Positive Operator-Valued Measurements

If $\hat{O}$ is an observable with discrete, distinct eigenvalues $\{o_n\}$ and an associated complete orthonormal set of eigenkets $\{|o_n\rangle\}$, then it has the eigenket-eigenvalue expansion

$$\hat{O} = \sum_n o_n |o_n\rangle\langle o_n|. \tag{1}$$

According to Axiom 3, measurement of $\hat{O}$ when the system is in state $|\psi\rangle$ yields an outcome that is one of the $\{o_n\}$ with the probability that the outcome is o_n being given by

$$\Pr(\,\text{outcome} = o_n \mid \text{state} = |\psi\rangle\,) = |\langle o_n|\psi\rangle|^2. \tag{2}$$

There is an equivalent POVM description of this measurement, as we now will provide. Define the collection of operators $\{\hat{\Pi}_n\}$ by

$$\hat{\Pi}_n = |o_n\rangle\langle o_n|. \tag{3}$$

Physically, the $\{\hat{\Pi}_n\}$ are *projectors*, i.e., for any state $|\psi\rangle$ we have that

$$\hat{\Pi}_n|\psi\rangle = (\langle o_n|\psi\rangle)|o_n\rangle \tag{4}$$

is the ket vector obtained by projecting $|\psi\rangle$ into the $|o_n\rangle$ direction. The $\{\hat{\Pi}_n\}$ have two properties that make them a positive operator-valued measurement:

- The $\{\hat{\Pi}_n\}$ are Hermitian operators, $\hat{\Pi}_n^\dagger = \hat{\Pi}_n$, as is self-evident from their definition.

- The $\{\hat{\Pi}_n\}$ resolve the identity, i.e.,

$$\hat{I} = \sum_n \hat{\Pi}_n, \tag{5}$$

  which follows immediately from the $\{|o_n\rangle\}$ being a complete orthonormal set of kets for the system.

We can now define the POVM $\{\hat{\Pi}_n\}$: the outcome of this measurement is n, the projector index, and the probability for getting the outcome n when the system is in the state $|\psi\rangle$ is

$$\Pr(\,\text{outcome} = n \mid \text{state} = |\psi\rangle\,) = \langle\psi|\hat{\Pi}_n|\psi\rangle. \tag{6}$$

Clearly, except for labeling the outcomes n instead of o_n, the POVM description is fully equivalent to the observable description of this measurement. So, for example, measurement of the number operator, $\hat{N} = \sum_n n|n\rangle\langle n|$, is equivalent to the POVM $\{\hat{\Pi}_n \equiv |n\rangle\langle n|\}$.

A similar relationship between observables and POVMs prevails when the former have a continuum of eigenvalues, as we will now show. Suppose that $\hat{O}$ is an observable with distinct eigenvalues $-\infty < o < \infty$ and an associated complete orthonormal (in the delta-function sense) set of eigenkets $\{|o\rangle\}$. This observable has the eigenket-eigenvalue expansion

$$\hat{O} = \int_{-\infty}^{\infty} \mathrm{d}o\, o|o\rangle\langle o|, \tag{7}$$

and, according to Axiom 3a, measurement of $\hat{O}$ yields an outcome that is a continuous random variable o with probability density function

$$p(o) = |\langle o|\psi\rangle|^2, \quad \text{for } -\infty < o < \infty, \tag{8}$$

when the system is in the state $|\psi\rangle$. The equivalent POVM formulation for this measurement is provided by the collection of operators $\{\hat{\Pi}(o)\}$, which are defined as follows,

$$\hat{\Pi}(o) \equiv |o\rangle\langle o|, \quad \text{for } -\infty < o < \infty. \tag{9}$$

The properties of these operators are like those we found for the discrete-eigenvalue case.

- The $\{\hat{\Pi}(o)\}$ are projectors, i.e.,

$$\hat{\Pi}(o)|\psi\rangle = (\langle o|\psi\rangle)|o\rangle, \tag{10}$$

  projects $|\psi\rangle$ along the $|o\rangle$ direction.[1]

- The $\{\hat{\Pi}(o)\}$ are Hermitian operators, i.e., $\hat{\Pi}(o)^\dagger = \hat{\Pi}(o)$, as is self-evident from their definition.

- The $\{\hat{\Pi}(o)\}$ resolve the identity, viz.,

$$\hat{I} = \int_{-\infty}^{\infty} \mathrm{d}o\, \hat{\Pi}(o), \tag{11}$$

  because $\{|o\rangle\}$ is a complete orthonormal set of eigenkets.

In light of the preceding properties, and in analogy with what we have stated for the case of an observable with discrete eigenvalues, we say that the POVM $\{\hat{\Pi}(o)\}$ has as its outcome a continuous random variable, o, whose probability density function is

$$p(o) = \langle\psi|\hat{\Pi}(o)|\psi\rangle, \quad \text{for } -\infty < o < \infty, \tag{12}$$

when the system is in state $|\psi\rangle$. It should be clear that the POVM description, $\{\hat{\Pi}(o)\}$, is fully equivalent, in this case, to that of the observable, $\hat{O}$. Thus, for example, measurement of the quadrature operator

$$\hat{a}_1 = \int_{-\infty}^{\infty} \mathrm{d}\alpha_1\, \alpha_1|\alpha_1\rangle_{11}\langle\alpha_1|, \tag{13}$$

[1] Due care must be taken here in interpreting this is as a projection. This is because a finite-energy $|\psi\rangle$ has unit length, but $\hat{\Pi}(o)|\psi\rangle$ will have infinite length unless $\langle o|\psi\rangle = 0$.

is equivalent to that of the POVM

$$\hat{\Pi}(\alpha_1) \equiv |\alpha_1\rangle_{11}\langle\alpha_1|, \quad \text{for } -\infty < \alpha_1 < \alpha_1. \tag{14}$$

Were we only to be concerned with measurements that are observables, then the value of POVM representations would be relatively low, as they are merely reformulations of what we have already characterized. The power of POVMs comes from their being able to describe measurements that *cannot* be characterized as an observable on the state space of the quantum system that is being measured. A prime example of such behavior is the $\hat{a}$ POVM, introduced last time, whose key properties are reviewed on Slide 2. To put this example into a more general setting, let's *define* discrete and continuous POVMs *without* the prior assumption that these measurements can also be represented by observables on the state space of the quantum system that is being measured.[2]

Discrete Outcome POVM

A POVM with a discrete outcome, $\{\hat{\Pi}_n\}$, is defined as follows.

- The $\{\hat{\Pi}_n\}$ are Hermitian operators that are positive semi-definite and resolve the identity,

$$\hat{\Pi}_n^\dagger = \hat{\Pi}_n, \quad \langle\psi|\hat{\Pi}_n|\psi\rangle \geq 0 \text{ for all } |\psi\rangle, \qquad \hat{I} = \sum_n \hat{\Pi}_n. \tag{15}$$

- When $\{\hat{\Pi}_n\}$ is measured, the outcome n occurs with probability

$$\Pr(\,\text{outcome} = n \mid \text{state} = |\psi\rangle\,) = \langle\psi|\hat{\Pi}_n|\psi\rangle, \tag{16}$$

  when the system is in the state $|\psi\rangle$.

If the $\{\hat{\Pi}_n\}$ are orthonormal projectors, so that

$$\hat{\Pi}_n\hat{\Pi}_m = \hat{\Pi}_n\delta_{nm}, \tag{17}$$

then this POVM is equivalent to an observable, but otherwise it is not. However, the POVM definition does *not* require that (17) hold. Indeed, it is easy to show that the preceding definition is fully consistent with classical probability theory, viz., it predicts probabilities that are non-negative and sum to one, with the former property being guaranteed by the positive semi-definite nature of the $\{\hat{\Pi}_n\}$, and the latter being due to the identity resolution property,

$$\sum_n \Pr(\,\text{outcome} = n \mid \text{state} = |\psi\rangle\,) = \langle\psi|\left(\sum_n \hat{\Pi}_n\right)|\psi\rangle = \langle\psi|\hat{I}|\psi\rangle = 1. \tag{18}$$

[2]Saying "on the state space of the quantum system that is being measured" is important because, as we will show later in this lecture, POVMs that are *not* observables on the system's state space *can* be represented as observables on an *enlarged* state space consisting of the original system adjoined with a (quantum-mechanical) measurement apparatus.

Continuous Outcome POVM
A POVM with a continuous outcome, $\{\hat{\Pi}(x)\}$, is defined as follows.

- The $\{\hat{\Pi}(x)\}$ are Hermitian operators that are positive semi-definite and resolve the identity,

$$\hat{\Pi}^\dagger(x) = \hat{\Pi}(x), \quad \langle\psi|\hat{\Pi}(x)|\psi\rangle \geq 0 \text{ for all } |\psi\rangle, \quad \hat{I} = \int \mathrm{d}x\, \hat{\Pi}(x). \tag{19}$$

- When $\{\hat{\Pi}(x)\}$ is measured, the outcome x occurs with probability density function

$$p(x) = \langle\psi|\hat{\Pi}(x)|\psi\rangle, \tag{20}$$

  when the system is in the state $|\psi\rangle$.

If the $\{\hat{\Pi}(x)\}$ are orthonormal projectors, so that

$$\hat{\Pi}(x)\hat{\Pi}(y) = \hat{\Pi}(x)\delta(x-y), \tag{21}$$

then this POVM is equivalent to an observable, but otherwise it is not. Once again, the POVM description does *not* require that the projection property—here, (21)—be satisfied. Note that the tenets of classical probability theory are obeyed by this prescription: non-negativity of a probability density is ensured by the positive semi-definite nature of the $\{\hat{\Pi}(x)\}$, and total probability equalling one follows from the identity resolution via

$$\int \mathrm{d}x\, p(x) = \langle\psi|\left(\int \mathrm{d}x\, \hat{\Pi}_x\right)|\psi\rangle = \langle\psi|\hat{I}|\psi\rangle = 1. \tag{22}$$

It should now be completely apparent that the POVM we introduced last time—as summarized on Slide 2—is

$$\hat{\Pi}(\alpha) \equiv \frac{|\alpha\rangle\langle\alpha|}{\pi}, \quad \text{for } \alpha \in \mathcal{C}, \tag{23}$$

where the $\{|\alpha\rangle\}$ are the coherent states. That this POVM does *not* correspond to an observable on the oscillator's state space is seen from

$$\hat{\Pi}(\alpha)\hat{\Pi}(\beta) = \frac{(\langle\alpha|\beta\rangle)|\alpha\rangle\langle\beta|}{\pi^2} = \frac{e^{-|\alpha|^2/2-|\beta|^2/2+\alpha^*\beta}|\alpha\rangle\langle\beta|}{\pi^2} \neq \hat{\Pi}(\alpha)\delta(\alpha-\beta), \tag{24}$$

where $\delta(\alpha-\beta) \equiv \delta(\alpha_1-\beta_1)\delta(\alpha_2-\beta_2)$.

Complete Statistics of the $\hat{a}$ POVM

Our next task will be to explore the statistics of the $\hat{a}$ POVM, i.e., $\hat{\Pi}(\alpha)$, from a transform (characteristic function) domain perspective. The outcome of this measurement is a pair of real numbers, α_1 and α_2, whose joint probability density is

$$p(\alpha) = \frac{|\langle\psi|\hat{\Pi}(\alpha)|\psi\rangle}{\pi} = \frac{|\langle\alpha|\psi\rangle|^2}{\pi}, \quad \text{for } \alpha \equiv \alpha_1 + j\alpha_2 \in \mathcal{C}. \tag{25}$$

when the system's state is $|\psi\rangle$. The classical joint characteristic function associated with this joint pdf is then

$$M_{\alpha_1,\alpha_2}(jv_1, jv_2) \equiv \int \mathrm{d}^2\alpha \, e^{jv_1\alpha_1 + jv_2\alpha_2} p(\alpha). \tag{26}$$

Consider the anti-normally ordered characteristic function associated with the state $|\psi\rangle$, i.e.,

$$\chi_A(\zeta^*, \zeta) \equiv \langle\psi|e^{-\zeta^*\hat{a}} e^{\zeta\hat{a}^\dagger}|\psi\rangle. \tag{27}$$

Inserting an identity operator that is resolved into coherent states, we find that

$$\chi_A(\zeta^*, \zeta) = \langle\psi|e^{-\zeta^*\hat{a}} \left(\int \frac{\mathrm{d}^2\alpha}{\pi} |\alpha\rangle\langle\alpha|\right) e^{\zeta\hat{a}^\dagger}|\psi\rangle \tag{28}$$

$$= \langle\psi| \left(\int \frac{\mathrm{d}^2\alpha}{\pi} |\alpha\rangle e^{-\zeta^*\alpha + \zeta\alpha^*} \langle\alpha|\right) |\psi\rangle = \int \mathrm{d}^2\alpha \, e^{-\zeta^*\alpha + \zeta\alpha^*} \frac{|\langle\alpha|\psi\rangle|^2}{\pi}, \tag{29}$$

where the second equality follows because $|\alpha\rangle$ is an eigenket of $e^{-\zeta^*\hat{a}}$ with eigenvalue $e^{-\zeta^*\alpha}$, as can be verified by Taylor series expansions of the operator and classical exponentials. From $-\zeta^*\alpha + \zeta\alpha^* = 2j\zeta_2\alpha_1 - 2j\zeta_1\alpha_2$, we then see that

$$M_{\alpha_1,\alpha_2}(jv_1, jv_2) = \chi_A(\zeta^*, \zeta)|_{\zeta = jv/2}, \quad \text{where } v \equiv v_1 + jv_2, \tag{30}$$

so that—unlike what we found last time for the Wigner characteristic function—the anti-normally ordered characteristic function *always* can be taken (with appropriate argument scaling) to be the joint characteristic function of two real-valued, classical random variables.

Let's use the preceding result to obtain the $\hat{a}$ POVM statistics when the oscillator is in the squeezed state $|\beta; \mu, \nu\rangle$, with, for simplicity, μ and ν real. For this state we have that

$$\chi_A(\zeta^*, \zeta) = \langle\beta; \mu, \nu|e^{-\zeta^*(\mu\hat{b} - \nu\hat{b}^\dagger)} e^{\zeta(\mu\hat{b}^\dagger - \nu\hat{b})}|\beta; \mu, \nu\rangle, \tag{31}$$

where $\hat{b} \equiv \mu\hat{a} + \nu\hat{a}^\dagger$ defines $|\beta; \mu, \nu\rangle$ via the eigenket-eigenvalue relation $\hat{b}|\beta; \mu, \nu\rangle = \beta|\beta; \mu, \nu\rangle$. Repeated use of the Baker-Campbell-Hausdorff theorem then leads to

$$\chi_A(\zeta^*, \zeta) = \langle\beta; \mu, \nu|e^{\zeta^*\nu\hat{b}^\dagger} e^{-\zeta^*\mu\hat{b}} e^{-\zeta^{*2}\mu\nu/2} e^{\zeta\mu\hat{b}^\dagger} e^{-\zeta\nu\hat{b}} e^{-\zeta^2\mu\nu/2}|\beta; \mu, \nu\rangle \tag{32}$$

$$= \langle\beta; \mu, \nu|e^{(\zeta\mu + \zeta^*\nu)\hat{b}^\dagger} e^{-(\zeta^*\mu + \zeta\nu)\hat{b}}|\beta; \mu, \nu\rangle e^{-|\zeta|^2\mu^2 - \mathrm{Re}(\zeta^2)\mu\nu}. \tag{33}$$

Using the fact that $|\beta;\mu,\nu\rangle$ is an eigenket of $e^{\xi\hat{b}}$, with eigenvalue $e^{\xi\beta}$, for ξ a complex number, gives us

$$\chi_A(\zeta^*,\zeta) = e^{(\zeta\mu+\zeta^*\nu)\beta^*-(\zeta^*\mu+\zeta\nu)\beta-|\zeta|^2\mu^2-\mathrm{Re}(\zeta^2)\mu\nu}. \tag{34}$$

Introducing $\zeta = jv/2$ leads to the final form,

$$M_{\alpha_1,\alpha_2}(jv_1,jv_2) = \chi_A(\zeta^*,\zeta)|_{\zeta=jv/2} = e^{jv_1(\mu-\nu)\beta_1-v_1^2\sigma_1^2/2}e^{jv_2(\mu+\nu)\beta_2-v_2^2\sigma_2^2/2}, \tag{35}$$

where β_1 and β_2 are the real and imaginary parts of β, and

$$\sigma_1^2 \equiv \mu^2-\mu\nu = \frac{(\mu-\nu)^2+1}{4} \quad \text{and} \quad \sigma_2^2 \equiv \mu^2+\mu\nu = \frac{(\mu+\nu)^2+1}{4}, \tag{36}$$

with the second equalities in the σ_k^2 expressions following from $\mu^2-\nu^2=1$. Equation (35) shows that, when the oscillator is in the squeezed state $|\beta;\mu,\nu\rangle$ with μ and ν real, the real and imaginary parts of the $\hat{a}$ POVM are statistically independent Gaussian random variables with mean values

$$\langle\alpha_1\rangle = (\mu-\nu)\beta_1 \quad \text{and} \quad \langle\alpha_2\rangle = (\mu+\nu)\beta_1, \tag{37}$$

and variances

$$\langle\Delta\alpha_1^2\rangle = \sigma_1^2 \quad \text{and} \quad \langle\Delta\alpha_2^2\rangle = \sigma_2^2. \tag{38}$$

Thus the $\hat{a}$ POVM gives information about the mean values of *both* quadratures, because $\langle\alpha_k\rangle = \langle\hat{a}_k\rangle$ for $k=1,2$, and, as we have shown earlier, it does so *without* violating the Heisenberg uncertainty principle. Indeed, the only new thing that we have learned here is that the $\hat{a}$ POVM statistics for the squeezed state—at least the one with μ, ν real—are Gaussian, and hence completely characterized by first and second moments.

Reconciling the $\hat{a}$ POVM with Observables

Positive operator-valued measurements entered into quantum information science in the 1960's, when the question of classical information transmission over quantum channels received its first thorough theoretical studies. Whether for digital or analog information transmission, deriving an optimum quantum receiver required that an optimization be performed over all possible quantum measurements. Restricting the set of possible measurements to observables on the state space of the quantum system whose state was modulated by the transmitter turned out to be too restrictive. Better performance could be obtained, in important cases, by using a positive operator-valued measurement on that system that was *not* an observable. Yet, as we will now show explicitly for the case of the $\hat{a}$ POVM, it is possible to construct an observables scheme that is equivalent—in the sense that it gives the same measurement statistics—as a POVM that cannot be represented by observables on the state space of the original quantum system. The key to this demonstration is to adjoin an ancilla quantum system and measure observables on the joint structure comprising the original quantum system and the ancilla.

On Problem Set 3 you have done the heavy lifting for this demonstration, by developing the notion of tensor product spaces. So, let us directly build on that background now. Consider *two* quantum harmonic oscillators, the system (S) and the ancilla (A), with state spaces $\mathcal{H}_S$ and $\mathcal{H}_A$, respectively. The annihilation operators for these oscillators will be denoted $\hat{a}_S$ and $\hat{a}_A$, and their coherent states will be denoted $|\alpha_S\rangle_S$ and $|\alpha_A\rangle_A$. If we measure the $\hat{a}_S$ POVM when the S system is in state $|\psi\rangle_S$, then we know that the outcome will be a complex-value random variable α_S with probability density function $p(\alpha_S) = |_S\langle\alpha_S|\psi\rangle_S|^2/\pi$. Now, consider the operator

$$\hat{y} \equiv \hat{a}_S \otimes \hat{I}_A + \hat{I}_S \otimes \hat{a}_A^\dagger, \tag{39}$$

where $\hat{I}_A$ and $\hat{I}_S$ are the identity operators on $\mathcal{H}_A$ and $\mathcal{H}_S$. We know, from Problem Set 3, that $\hat{y}$ is an operator on the tensor product state space $\mathcal{H} \equiv \mathcal{H}_S \otimes \mathcal{H}_A$ of the two oscillators. Because

$$\hat{y}^\dagger = \hat{a}_S^\dagger \otimes \hat{I}_A + \hat{I}_S \otimes \hat{a}_A, \tag{40}$$

we see that $\hat{y}$ is *not* Hermitian. Note, however, that $\hat{y}$ does commute with its adjoint, viz.,

$$\begin{aligned}
[\hat{y}, \hat{y}^\dagger] &= [\hat{a}_S \otimes \hat{I}_A + \hat{I}_S \otimes \hat{a}_A^\dagger, \hat{a}_S^\dagger \otimes \hat{I}_A + \hat{I}_S \otimes \hat{a}_A] && (41)\\
&= [\hat{a}_S \otimes \hat{I}_A, \hat{a}_S^\dagger \otimes \hat{I}_A] + [\hat{I}_S \otimes \hat{a}_A^\dagger, \hat{I}_S \otimes \hat{a}_A] = 1 - 1 = 0. && (42)
\end{aligned}$$

It follows that the real and imaginary parts of $\hat{y}$,

$$\hat{y}_1 \equiv \hat{a}_{S_1} \otimes \hat{I}_A + \hat{I}_S \otimes \hat{a}_{A_1} \quad \text{and} \quad \hat{y}_2 \equiv \hat{a}_{S_2} \otimes \hat{I}_A - \hat{I}_S \otimes \hat{a}_{A_2}, \tag{43}$$

are commuting observables that can be measured simultaneously. The supplementary reading for this lecture provides a reference in which the eigenkets of $\hat{y}$ are derived. We shall not take that route to quantifying the statistics of the simultaneous $\hat{y}_1$ and $\hat{y}_2$ measurements. For our needs, it is simpler to employ a characteristic function derivation.

Suppose that the state of the two oscillators is the product state $|\psi\rangle = |\psi\rangle_S \otimes |0\rangle_A$, where $|\psi\rangle_S$ is an arbitrary unit-length ket in $\mathcal{H}_S$ and $|0\rangle_A$ is the vacuum state in $\mathcal{H}_A$. Let y_1 and y_2 be the real-valued, random-variable outcomes obtained from the simultaneous measurement of $\hat{y}_1$ and $\hat{y}_2$ when the two oscillators are in this product state. We have that the joint characteristic function for these random variables is

$$M_{y_1,y_2}(jv_1, jv_2) = \langle e^{jv_1y_1+jv_2y_2}\rangle = \langle e^{jv_1\hat{y}_1+jv_2\hat{y}_2}\rangle \tag{44}$$

$$= \langle e^{(jv_1+v_2)\hat{y}/2}e^{(jv_1-v_2)\hat{y}^\dagger/2}\rangle \tag{45}$$

$$= \langle e^{(jv_1+v_2)\hat{a}_S\otimes\hat{I}_A}e^{(jv_1-v_2)\hat{a}_S^\dagger\otimes\hat{I}_A}\rangle\langle e^{(jv_1-v_2)\hat{I}_S^\dagger\otimes\hat{a}_A^\dagger}e^{(jv_1-v_2)\hat{I}_S\otimes\hat{a}_A}\rangle \tag{46}$$

$$= [{}_S\langle\psi|e^{-\zeta^*\hat{a}_S}e^{\zeta\hat{a}_S^\dagger}|\psi\rangle_S\,{}_A\langle 0|e^{-\zeta^*\hat{a}_A^\dagger}e^{\zeta\hat{a}_A}|0\rangle_A]|_{\zeta=jv/2} \tag{47}$$

$$= \chi_{A_S}(\zeta^*,\zeta)|_{\zeta=jv/2}, \tag{48}$$

where

$$\chi_{A_S}(\zeta^*,\zeta) = {}_S\langle\psi|e^{-\zeta^*\hat{a}_S}e^{\zeta\hat{a}_S^\dagger}|\psi\rangle_S \tag{49}$$

is the anti-normally ordered characteristic function of the signal oscillator's quantum state. Comparing Eqs. (30) and (48) shows that the $\hat{a}$ POVM has the same measurement statistics as the simultaneous measurement of $\hat{y}_1$ and $\hat{y}_2$ when the ancilla oscillator is in its vacuum state. When we study single-mode optical heterodyne detection—which realizes the $\hat{a}$ POVM—we will be able to identify a physical locus for the ancilla oscillator. There is, however, one more thing worth doing while we are comparing the $\hat{a}$ POVM with its commuting observables equivalent, and that is to see where the extra noise in the $\langle\Delta\alpha_1^2\rangle$ and $\langle\Delta\alpha_2^2\rangle$ comes from.

Because the α_1 and α_2 outcomes from the $\hat{a}$ POVM have the same statistics as the simultaneous measurement of $\hat{y}_1$ and $\hat{y}_2$ when the ancilla oscillator is in its vacuum state, we see that

$$\langle\Delta\alpha_k^2\rangle = \langle\Delta\hat{y}_k^2\rangle = \langle\Delta\hat{a}_{S_k}^2\rangle + \langle\Delta\hat{a}_{A_k}^2\rangle = \langle\Delta\hat{a}_{S_k}^2\rangle + 1/4, \quad \text{for } k = 1, 2, \tag{50}$$

where the second equality follows from the signal and ancilla being in a product state, and the last equality follows from the ancilla being in the vacuum state. Thus, the extra noise that appeared in Table 4 of Lecture 7 is due to the zero-point fluctuations of the ancilla.

Single-Mode Photodetection

We are now done with our development of the quantum harmonic oscillator. It is time to turn that knowledge into results for single-mode photodetection, and to contrast the quantum theory of photodetection with the semiclassical (shot-noise) theory of photodetection. Before doing so, we pause for a quick phenomenological discussion of photodetection, as it will make clear what idealizations underlie the quantum and semiclassical models that we will be presenting.

A Real Photodetector

Slide 8 shows a theorist's cartoon of a real photodetector. The two large blocks on this slide are the photodetector and the post-detection preamplifier. The smaller blocks within the two large blocks are phenomenological, i.e., they do not represent discrete components out of which the larger entities are constructed. Nevertheless, it is instructive to walk our way through this photodetection system by means of these phenomenological blocks. Incoming light—whether we model it in classical or quantum terms—illuminates an optical filter that models the wavelength dependence of the photodetector's sensitivity. The light emerging from this filter then strikes the core of the photodetector, i.e., the block that converts light into a light-induced current, which we call the photocurrent. Photodetectors have some current flow in the absence of illumination, and this *dark* current adds to the photocurrent within the detector. High-sensitivity photodetectors—such as avalanche photodiodes and photomultiplier tubes—have internal mechanisms that amplify (multiply) the initial photocurrent (and the dark current), and we have shown that on Slide 8 as a current multiplication block.[3] This current multiplication in general has some randomness associated with it, imposing an excess noise on top of any noise already inherent in the photocurrent and dark current. The electrical filter that is next encountered models the electrical bandwidth of the photodetector's output circuit, and the thermal noise generator models the noise associated with the dissipative elements in the detector. Because the output current from a photodetector may not be strong enough to regard all subsequent processing as noiseless, we have included the preamplifier block in Slide 8. Its filter, noise generator, and gain blocks model the bandwidth characteristics, noise figure, and gain of a real preamplifier. Ordinarily, the output from such a preamplifier is strong enough that any further signal processing can be regarded as noise free.

An Ideal Photodetector

Because we are interested in the *fundamental* limits of photodetection—be they represented in semiclassical or quantum terms—we will strip away all the inessential elements of the real photodetection system shown on Slide 8, and restrict our attention to the ideal photodetector shown on Slide 9. Be warned, however, that in experimental work we cannot always ignore the phenomena cited in our discussion of Slide 8. Nevertheless, it does turn out that there are photodetection systems that can approach the ideal behavior under some circumstances.

Our ideal photodetector is a perfect version of the photocurrent generator block from Slide 8. In particular, our ideal photodetector has these properties.

- Its optical sensitivity covers all frequencies.

[3] We have shown the photocurrent and dark current as undergoing the *same* multiplication process. In real detectors, these two currents may encounter different multiplication factors.

- Its conversion of light into current is perfectly efficient.
- It does not have any dark current.
- It does not have any current multiplication.
- It has infinite electrical bandwidth.
- Its subsequent preamplifier has infinite bandwidth and no noise, so it need not be considered as it does not degrade the photodetection performance.

As a result, the photocurrent takes the form of a random train of area-q impulses, where q is the electron charge,[4] and a counting circuit driven by this photocurrent will produce, as its output, a staircase function of unit-height steps which increments when each impulse occurs, i.e.,

$$N(t) = \frac{1}{q}\int_0^t \mathrm{d}u\, i(u), \quad \text{for } 0 \leq t \leq T, \tag{51}$$

as shown on Slide 9.

Single-Mode Fields: Classical and Quantum

Suppose that our ideal photodetector has an active region—the photosensitive region over which light is converted into current—$(x, y) \in \mathcal{A}$ in some constant-z plane. Also suppose that we make our photodetection measurement over the time interval $0 \leq t \leq T$. Later this semester we will consider semiclassical and quantum photodetection when the illuminating field can have arbitrary spatio-temporal behavior. For today, however, we will restrict our attention to single-mode fields.

Classical Single-Mode Field:
For semiclassical photodetection, we will take this single-mode field to be

$$E_z(x, y, t) = \frac{ae^{-j\omega t}}{\sqrt{AT}}, \quad \text{for } (x, y) \in \mathcal{A} \text{ and } 0 \leq t \leq T, \tag{52}$$

on the detector's photosensitive region. Here, a is dimensionless—it is the field equivalent of the classical harmonic oscillator initial time phasor—and A is the area of the the region $\mathcal{A}$. It follows that $E_z(x, y, t)$ has the units $\sqrt{\text{photons/m}^2\text{s}}$. Physically, $E_z(x, y, t)$, is the positive-frequency part[5] of a monochromatic (frequency ω), $+z$-

[4]This random train of impulses is known as shot noise, because individual current "shots" are discernible and randomly located in time.

[5]Even though there is a factor of $e^{-j\omega t}$ here, it is conventional to refer to this field as being the positive-frequency field. The real-valued field is then $\mathrm{Re}[E_z(x, y, t)]$. In what follows, we will work almost exclusively with positive-frequency fields (in the semiclassical theory of photodetection), and the corresponding positive-frequency field operators (in the quantum theory of photodetection).

going, plane-wave pulse impinging on the photodetector.[6]

Quantum Single-Mode Field:
For quantum photodetection, the single-mode classical field from the previous subsection becomes a quantum field operator with only one excited—one non-vacuum state—mode:

$$\hat{E}_z(x,y,t) = \underbrace{\frac{\hat{a}e^{-j\omega t}}{\sqrt{AT}}}_{\text{excited mode}} + \underbrace{\text{other terms}}_{\text{unexcited modes}}, \quad \text{for } (x,y) \in \mathcal{A} \text{ and } 0 \le t \le T. \tag{53}$$

Here, the complex phasor a from the classical field has become a photon annihilation operator $\hat{a}$. Thus, the field operator $\hat{E}_z(x,y,t)$ has the units $\sqrt{\text{photons/m}^2\text{s}}$. The "other terms" are operator-valued modes that are needed to ensure that $\hat{E}_z(x,y,t)$ has the proper commutator with its adjoint, as required for a free-space propagating wave with the units as given. In classical physics, saying that a field mode is unexcited means that its value is zero. This is why there are no "other terms" in our representation of the single-mode classical field. In quantum physics, however, saying that a field mode is unexcited means that it is in its vacuum state. Although the vacuum state contains no photons, it does possess zero-point fluctuations. So, depending on what measurement is made on that quantum field, the unexcited modes *may* contribute noise to the observations.

Direct Detection

Slide 11 shows the semiclassical and quantum models for direct detection of a single-mode field.[7] Both descriptions show that photodetectors are intrinsically square-law devices. In particular, if we integrate the photocurrent from $t = 0$ to $t = T$, and divide by the electron charge q, we get a counting variable—the photon count—N, given by

$$N = \frac{1}{q}\int_0^T \mathrm{d}u\, i(u), \tag{54}$$

cf. Slide 9.

According to semiclassical photodetection theory, the conditional probability mass function for N, given $a = \alpha$, is

$$\Pr(\, N = n \mid a = \alpha\,) = \frac{|\alpha|^{2n}e^{-|\alpha|^2}}{n!}, \quad \text{for } n = 0,1,2,\ldots, \tag{55}$$

[6]We have chosen to use scalar field notation, because almost all of our photodetection work this semester does not require us to consider the vector properties of the field. For consistency with the vector theory, we could say that this single-mode field is linearly polarized in, say, the x direction.

[7]Direct detection means that the field to be measured impinges on the photodetector *without* being combined with any additional field.

i.e., given knowledge of a, semiclassical theory says that N is Poisson distributed with mean $|a|^2$. Thus the variance of N equals its mean value, and the only way to have a variance of zero is to have $a = 0$, i.e., no illumination, in which case $N = 0$ with probability one.

In contrast, according to the quantum theory of photodetection, the conditional probability mass function for N given that the $\hat{a}$ mode is in the state $|\psi\rangle$ is

$$\Pr(\, N = n \mid \text{state } = |\psi\rangle\,) = |\langle n|\psi\rangle|^2, \quad \text{for } n = 0, 1, 2, \ldots,, \tag{56}$$

where $|n\rangle$ is n-photon number state. This shows that ideal direct detection of the single-mode quantum field realizes the number operator measurement, $\hat{N} = \hat{a}^\dagger\hat{a}$. From what we have already done, we know that putting the $\hat{a}$ mode into the coherent state $|\alpha\rangle$ will yield a Poisson distribution for N, with mean $|\alpha|^2$, in quantum photodetection theory, *exactly* matching the statistics predicted for a classical field with $a = \langle\hat{a}\rangle = \alpha$. On the other hand, if the $\hat{a}$ mode is in the number state $|k\rangle$ with $k > 0$, then quantum photodetection theory predicts

$$\Pr(\, N = n \mid \text{state } = |k\rangle\,) = \delta_{nk}, \tag{57}$$

so that $\langle N\rangle = k > 0$ and $\langle \Delta N^2\rangle = 0$, something which is *impossible* in semiclassical theory.

If a quantum state is such that *all* its possible quantum photodetection measurements (direct detection, homodyne detection, and heterodyne detection) have statistics that are identical to those obtained from the corresponding semiclassical (shot noise) model, then that state is said to be "classical." A purely quantum, or "non-classical" state is one for which at least one of the quantum photodetection measurements has statistics that *cannot* be explained by semiclassical theory. So far, we know that the number state is non-classical and the coherent state appears classical insofar as direct detection is concerned.

Homodyne Detection

Slide 12 shows the block diagram for balanced homodyne detection of a single-mode field. In semiclassical photodetection, the incoming positive-frequency signal field is

$$E_S(x, y, t) = \frac{a_S e^{-j\omega t}}{\sqrt{AT}}, \tag{58}$$

and the incoming positive-frequency local-oscillator field is

$$E_{\text{LO}}(x, y, t) = \frac{a_{\text{LO}} e^{-j\omega t}}{\sqrt{AT}}. \tag{59}$$

The signal field is quite weak compared to that of the local oscillator, $|a_S| \ll |a_{\text{LO}}|$, and the latter is assumed to be of the form $a_{\text{LO}} = \sqrt{N_{\text{LO}}}\, e^{j\theta}$. The 50/50 beam splitter is such that the positive-frequency fields reaching the two photodetectors are

$$E_\pm(x,y,t) = \frac{a_\pm e^{-j\omega t}}{\sqrt{AT}}, \quad \text{where} \quad a_\pm \equiv \frac{a_S \pm a_{\text{LO}}}{\sqrt{2}}. \tag{60}$$

The output of the balanced homodyne system is then

$$\alpha_\theta = \frac{qN_+ - qN_-}{K}, \tag{61}$$

where $K = 2q\sqrt{N_{\text{LO}}}$ and $N_\pm$ are the output counts from the two photodetectors. Shot noises from physically different detectors are statistically independent random variables, so we have that N_+ and N_- are statistically independent Poisson random variables with mean values $|a_+|^2$ and $|a_-|^2$, respectively. It is then a simple matter to find the classical characteristic function of α_θ in the limit of a very strong local oscillator field, i.e., when $N_{\text{LO}} \to \infty$. We have that

$$M_{\alpha_\theta}(jv) = \langle e^{jv(qN_+ - qN_-)/K}\rangle = \langle e^{jvqN_+/K}\rangle\langle e^{-jvqN_-/K}\rangle \tag{62}$$

$$= \exp[|a_+|^2(e^{jvq/K} - 1)]\exp[|a_-|^2(e^{-jvq/K} - 1)]. \tag{63}$$

As $N_{\text{LO}} \to \infty$ we have $K \to \infty$, so that

$$e^{\pm jvq/K} - 1 \approx \pm jvq/K - v^2q^2/2K^2 \tag{64}$$

to second order. Also, we have that

$$|a_\pm|^2 = \frac{N_{\text{LO}} \pm 2\text{Re}(a_S\sqrt{N_{\text{LO}}}\,e^{-j\theta}) + |a_S|^2}{2} \approx \frac{N_{\text{LO}} \pm 2\text{Re}(a_S\sqrt{N_{\text{LO}}}\,e^{-j\theta})}{2}, \tag{65}$$

as $N_{\text{LO}} \to \infty$. Substituting these two approximations into our expression for $M_{\alpha_\theta}(jv)$ and letting $N_{\text{LO}} \to \infty$ then yields

$$M_{\alpha_\theta}(jv) = e^{jv\text{Re}(a_S e^{-j\theta}) - v^2/8}, \tag{66}$$

which implies that α_θ is a Gaussian random variable with mean $\text{Re}(a_S e^{-j\theta})$ and variance $1/4$. When $\theta = 0$, the mean of the semiclassical homodyne outcome is $a_{S_1} = \text{Re}(a_S)$. Likewise, when $\theta = \pi/2$, the mean of the semiclassical homodyne outcome is $a_{S_2} = \text{Im}(a_S)$. Evidently, ideal semiclassical homodyne detection yields a signal-field quadrature—whose phase shift is determined by the relative phase, θ, between the signal and the local oscillator—embedded in an additive zero-mean, variance $1/4$ noise. Because the local oscillator is so strong, this noise is the shot noise that it creates.

The quantum theory of homodyne detection is quite different. Now the positive-frequency signal and local oscillator field operators are

$$\hat{E}_S(x,y,t) = \underbrace{\frac{\hat{a}_S e^{-j\omega t}}{\sqrt{AT}}}_{\text{excited mode}} + \underbrace{\text{other terms}}_{\text{unexcited modes}}, \tag{67}$$

and

$$\hat{E}_{\rm LO}(x,y,t) = \underbrace{\frac{\hat{a}_{\rm LO}e^{-j\omega t}}{\sqrt{AT}}}_{\text{excited mode}} + \underbrace{\text{other terms}}_{\text{unexcited modes}}, \tag{68}$$

respectively. The signal field is assumed to be quite weak compared to the local oscillator, i.e., $\langle \hat{a}_S^\dagger \hat{a}_S\rangle \ll \langle \hat{a}_{\rm LO}^\dagger \hat{a}_{\rm LO}\rangle$, and the local oscillator is assumed to be in the coherent state $|\sqrt{N_{\rm LO}}\,e^{j\theta}\rangle$. The 50/50 beam splitter combines the signal and local oscillator fields so that the positive-frequency field operators that illuminate the two detectors are

$$\hat{E}_\pm(x,y,t) = \underbrace{\frac{\hat{a}_\pm e^{-j\omega t}}{\sqrt{AT}}}_{\text{excited mode}} + \underbrace{\text{other terms}}_{\text{unexcited modes}}, \quad \text{where } \hat{a}_\pm \equiv \frac{\hat{a}_S \pm \hat{a}_{\rm LO}}{\sqrt{2}}. \tag{69}$$

From quantum photodetection theory we know that

$$\alpha_\theta = \frac{N_+ - N_-}{2\sqrt{N_{\rm LO}}} \longleftrightarrow \frac{\hat{a}_+^\dagger \hat{a}_+ - \hat{a}_-^\dagger \hat{a}_-}{2\sqrt{N_{\rm LO}}}, \tag{70}$$

where $\longleftrightarrow$ indicates that the statistics of the classical random variable on the left-hand side coincides with those of the quantum measurement (observable) on the right-hand side. Then, writing $\hat{a}_\pm$ in terms of $\hat{a}_S$ and $\hat{a}_{\rm LO}$, and using $\langle \hat{a}_S^\dagger \hat{a}_S\rangle/2\sqrt{N_{\rm LO}} \to 0$ as $N_{\rm LO}\to\infty$ to justify dropping the term $\hat{a}_S^\dagger \hat{a}_S/2\sqrt{N_{\rm LO}}$, we get

$$\alpha_\theta = \frac{N_+ - N_-}{2\sqrt{N_{\rm LO}}} \longleftrightarrow \frac{{\rm Re}(\hat{a}_S \hat{a}_{\rm LO}^\dagger)}{\sqrt{N_{\rm LO}}}, \quad \text{as } N_{\rm LO}\to\infty. \tag{71}$$

Finally, because the local oscillator is in the coherent state $|\sqrt{N_{\rm LO}}\,e^{-j\theta}\rangle$, we can say that $\hat{a}_{\rm LO}^\dagger/\sqrt{N_{\rm LO}} \to e^{-j\theta}$ as $N_{\rm LO}\to\infty$. Thus we get our the quantum photodetection description of balanced homodyne detection in the infinite local oscillator limit:

$$\alpha_\theta \longleftrightarrow {\rm Re}(\hat{a}_S e^{-j\theta}), \tag{72}$$

i.e., the setup shown on Slide 12 measures a field quadrature whose phase shift is determined by the relative phase between the excited signal mode and that of the local-oscillator's coherent state. When $\theta = 0$, the quantum homodyne outcome is that of the $\hat{a}_{S_1} = {\rm Re}(\hat{a}_S)$ observable. Likewise, when $\theta = \pi/2$, the quantum homodyne outcome is that of the $\hat{a}_{S_2} = {\rm Im}(\hat{a}_S)$ observable. If the $\hat{a}_S$ mode is in the coherent state $|a_S\rangle$, then the statistics predicted by quantum photodetection theory coincide with those of the semiclassical theory. But, if the $\hat{a}_S$ mode is in the squeezed state $|\beta;\mu,\nu\rangle$, with $|\nu|>0$, then there will be a θ value for which quantum photodetection theory predicts a Gaussian probability density with a variance that is less than $1/4$, something that is *impossible* in the semiclassical theory. Squeezed states, therefore are non-classical.

The Road Ahead

Next time we will continue our development of single-mode photodetection by studying balanced heterodyne detection. Here we will find measurement statistics—in the quantum case—that correspond to those of the $\hat{a}$ POVM. Moreover, we will find a physical locus for the ancilla mode that injects the extra noise in the commuting observables—on a larger state space—explanation of the $\hat{a}$ POVM.