

Quantum theory of angle and relative-phase measurement

Scott Roger Shepard*

College of Engineering and Science, Louisiana Tech University, Ruston, Louisiana 71272, USA

(Received 3 October 2014; published 12 December 2014)

The complementarity between time and energy, as well as between an angle and a component of angular momentum, is described at three different layers of understanding. The phenomena of super-resolution are readily apparent in the quantum phase representation which also reveals that entanglement is not required. We modify Schwinger’s harmonic oscillator model of angular momentum to include the case of photons. Therein, the quantum angle measurement is shown to be equivalent to the measurement of the relative phase between the two oscillators. Two reasonable ways of dealing with degeneracy are shown to correspond to a conditional measurement which takes a snapshot in absolute time (corresponding to adding probability amplitudes) and a marginal measurement which takes an average in absolute time (corresponding to adding probabilities). The sense in which distinguishability is a “matter of how long we look” is discussed and the meaning of the general theory is illustrated by taking the two oscillators to be circularly polarized photons. It is shown that an odd number of x -polarized photons will never have an angle in correspondence with the y axis, but an even number of x -polarized photons always can! The behavior of an x -polarized coherent state is examined and the snapshot angular distributions are seen to evolve into two counter-rotating peaks resulting in considerable correspondence with the y axis at the time for which a classical linear polarization vector would shrink to zero length. We also demonstrate how the probability distribution of absolute time (herein a measurable quantity, rather than just a parameter) has an influence on how these snapshot angular distributions evolve into a quantum version of the polarization ellipse.

DOI: [10.1103/PhysRevA.90.062117](https://doi.org/10.1103/PhysRevA.90.062117)

PACS number(s): 03.65.Ta, 42.50.Tx, 42.50.St, 42.50.Lc

I. SINGLE-OSCILLATOR PHASE STATISTICS

A. Introduction

Dirac [1], in 1927, postulated the existence of a Hermitian phase operator $\hat{\phi}$ as part of a polar decomposition of the annihilation operator $\hat{a} = e^{i\hat{\phi}}(\hat{a}^\dagger\hat{a})^{1/2}$ for a single quantum harmonic oscillator, where $\hat{a}^\dagger\hat{a} = \hat{n}$ is the photon-number operator, the eigenvalues of which correspond to the energy levels of the oscillator. In this ordering, the operator $e^{i\hat{\phi}} = \hat{a}(\hat{n})^{-1/2}$ lowers the photon number (like \hat{a} but without any \sqrt{n} factor)

$$e^{i\hat{\phi}}|n\rangle = |n-1\rangle \quad (\forall n \geq 1), \quad (1)$$

however, the action on the vacuum state is undefined. In 1964, Susskind and Glogower [2] demonstrated that no such Hermitian operator exists on the denumerably infinite-dimensional space [3] spanned by the number-kets $\{|n\rangle : n = 0, 1, 2, \dots, \infty\}$. They proposed, instead, the use of a polar decomposition of \hat{a} with the opposite ordering; the Susskind-Glogower (SG) phase operator $\widehat{e^{i\hat{\phi}}} \equiv (\hat{a}\hat{a}^\dagger)^{-1/2}\hat{a}$, so that the SG operator is a pure lowering operator which stops (i.e., yields the null ket) at the vacuum:

$$\hat{A}|n\rangle = |n-1\rangle \quad (\forall n \geq 1) \quad \text{and} \quad \hat{A}|0\rangle = 0, \quad (2)$$

where to simplify the notation we let $\hat{A} \equiv \widehat{e^{i\hat{\phi}}}$. Thus, the SG operator has a number representation given by $\hat{A} = \sum_{n=0}^{\infty} |n\rangle\langle n+1|$. Due to the bound on photon-number eigen-spectra (i.e., the absence of negative-energy eigenstates) this translation, however, cannot be unitary. It can only be one-sided unitary, i.e., $\hat{A}\hat{A}^\dagger = \hat{I}$ but $\hat{A}^\dagger\hat{A} = \hat{I} - \hat{V}$ where \hat{I} is

the identity operator on the state space of a single harmonic oscillator and $\hat{V} \equiv |0\rangle\langle 0|$ is the vacuum projector. Thus, pure translation on a bounded state space cannot be unitary; therefore, it cannot be expressed as the power series of a Hermitian (i.e., self-adjoint) operator.

In light of an ever-improving understanding of what it means to associate a measurement with an operator which does not commute with its adjoint, we [4,5] demonstrated a connection between the SG operator and Helstrom’s maximum likelihood (ML) quantum phase estimator [6]. Helstrom was not concerned with polar decompositions of the annihilation operator, nor with obtaining a description of a phase measurement that is complementary to that of photon counting. The ML measurement is based on the state-dependent kets

$$|\phi, \psi\rangle = \sum_{n=0}^{\infty} e^{i(n\phi + \chi_n)} |n\rangle, \quad (3)$$

where the χ_n are the phases of the number-ket expansion coefficients $\psi_n \equiv \langle n|\psi\rangle$. Thus, the ML phase estimation procedure can be decomposed into two steps: the first being to effectively remove the phases of the input quantum state’s number-ket expansion coefficients ψ_n , i.e., these phases are all effectively set to be equal to zero. This first step can only be omitted if the ψ_n are already real. The second step turns out to be equivalent to performing a relative phase measurement between two oscillators (complementary to the measurement of the difference in their photon numbers) when one of those oscillators is in the vacuum state [7,8], resulting in the single-oscillator (or single-mode) phase statistics, which are *fuzzy* in a sense that we will elucidate herein. Thus, the ML statistics correspond to the single-mode statistics for the case of states with real ψ_n , as depicted in the Venn diagram of Fig. 1.

*sshepard@latech.edu

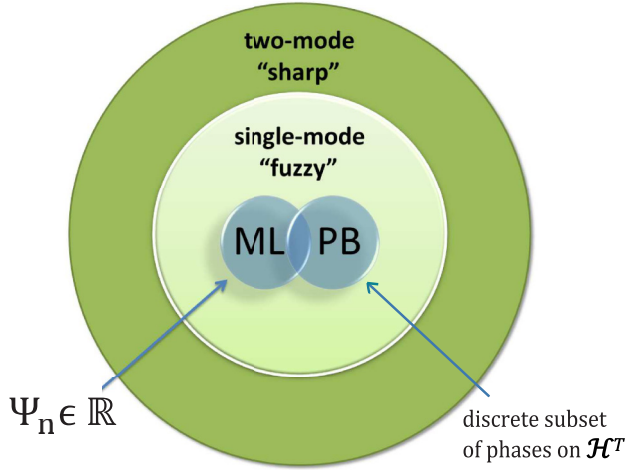


FIG. 1. (Color online) Hierarchy of some quantum phase measurements.

Concurrently (with respect to [4,5]) and independently, an alternate method for obtaining the single-oscillator statistics was derived by Pegg and Barnett [9]. Their approach requires the truncation of the infinite-dimensional state space of a harmonic oscillator to one of finite but arbitrarily large dimension. This subspace, denoted $\mathcal{H}^T(s)$, is spanned by the number-kets $\{|n\rangle : 0 \leq n \leq s\}$. When s is finite, the resulting Pegg-Barnett (PB) discrete-phase kets are an orthogonal subset of the single-mode continuous phase-kets, as depicted in Fig. 1. This approach is described in Appendix A but it is important to note in passing that *when s is finite* the discrete-phase measurement is “sharp” (i.e., it permits wave-function collapse via projections onto the eigenkets of a Hermitian operator). Yet, we also have that *when instead* the limit $s \rightarrow \infty$ is taken, these discrete-phase statistics converge (in distribution, i.e., in as much as rational numbers can converge to real numbers) to those of a fundamentally “fuzzy” measurement, the single-mode continuous phase statistics (i.e., the continuous envelope to which they converge cannot collapse to a delta function). Thus, unless s is left to be finite, this is an alternate means of calculating the same statistics via a limiting procedure. Note in passing, upon review it was pointed out that a discrete-phase operator on a truncated space was also described by Popov and Yarunin in [10].

The connection to the SG operator is as follows. Although the SG operator has other eigenkets (the coherent phase states [11]) we restrict our attention to the infinite energy subset of these, the continuous phase-kets, of number representation:

$$|\phi\rangle = \sum_{n=0}^{\infty} e^{in\phi} |n\rangle, \quad (4)$$

which are normalizable in the continuous ϕ domain [12]. The restriction to these was made not only to connect with the essence of ML phase estimation, but (more significantly) to obtain complementarity to the measurement of photon number. These kets have been said [13] to be of fundamental importance because they underlie ML phase estimation. It was also noted [7,8], however, that these phase-kets are of even greater significance because they provide the first of three layers of

understanding complementarity in a more general context. Complementarity at the first layer corresponds to the fuzzy single-mode case (depicted in Fig. 1, which generalizes the ML measurement and to which the PB measurement can converge). The second and third layers of understanding can only be achieved when we extend beyond the fuzzy measurement by working in a state space larger than that of a single oscillator, as we will in the next section.

The SG operator does not commute with its adjoint, hence, it is not comprised of a set of commuting Hermitian operators. Therefore, the measurement statistics we associate with it cannot be calculated via the familiar Hermitian operator rules (e.g., moments calculated via $\langle \psi | \hat{A}^k | \psi \rangle$ for $k = 1, 2, \dots$ do *not* correspond to the single-mode statistics).

Quantum measurements can be described in terms of wave functions (as well as operators) and perhaps the simplest path to the single-mode statistics is to form the phase wave function

$$\psi(\phi) \equiv \langle \phi | \psi \rangle, \quad (5)$$

from which the phase probability distribution $P(\phi) = |\psi(\phi)|^2 / (2\pi)$ and its associated moments follow directly. This will be formally justified below but it stems from the fact that the (nonorthogonal) phase-kets are complete:

$$\int_{-\pi}^{\pi} \frac{d\phi}{2\pi} |\phi\rangle \langle \phi| = \sum_{n=0}^{\infty} |n\rangle \langle n| = \hat{I}. \quad (6)$$

The resolution of the identity by the phase-kets in (6) permits the extremely useful phase representation of an arbitrary quantum state:

$$|\psi\rangle = \hat{I}|\psi\rangle = \int_{-\pi}^{\pi} \frac{d\phi}{2\pi} \langle \phi | \psi \rangle |\phi\rangle, \quad (7)$$

analogous to the familiar number-ket expansion of a state:

$$|\psi\rangle = \hat{I}|\psi\rangle = \sum_{n=0}^{\infty} \langle n | \psi \rangle |n\rangle. \quad (8)$$

Just as the number-ket expansion coefficients $\psi_n \equiv \langle n | \psi \rangle$ may be viewed as a wave function in discrete n space, the inner product $\psi(\phi) \equiv \langle \phi | \psi \rangle$ is a wave function in continuous ϕ space. The Fourier transform between the number and phase wave functions

$$\psi(\phi) = \sum_{n=0}^{\infty} \psi_n e^{-in\phi} \leftrightarrow \psi_n = \int_{-\pi}^{\pi} \frac{d\phi}{2\pi} \psi(\phi) e^{in\phi} \quad (9)$$

demonstrates the complementarity of photon number and quantum phase.

Thus, to reveal the phase properties of an arbitrary state, we can simply take the familiar number-ket expansion coefficients ψ_n to be the Fourier series coefficients for $\psi(\phi)$, which underlies the harmonic oscillator’s continuous and periodic phase distribution. Several relations among ψ_n and $\psi(\phi)$ are reminiscent of those encountered in Schrödinger’s wave mechanics. Analogous to the position representation of the momentum operator $\hat{p} \rightarrow -i\hbar \frac{d}{dx}$, for example, we have a phase representation of the number operator $\hat{n} \rightarrow -i \frac{d}{d\phi}$:

$$\langle (\hat{n})^k \rangle = \int_{-\pi}^{\pi} \frac{d\phi}{2\pi} \psi^*(\phi) \left(-i \frac{d}{d\phi} \right)^k \psi(\phi) \quad (\forall \text{ integer } k). \quad (10)$$

We note in passing, however, that differentiation with respect to a discrete variable is an undefined operation.

Alternatively, we might have arrived at this, (4)–(10), without any reference to the SG operator, via a general theory of complementarity (as is done in Appendix B, and as is *required* for the fuzzy description of the angle of a particle of finite angular momentum since then there is also an upper bound on the complementary eigenspectra and that prohibits a lowering operator from having any eigenkets at all). Historically, such an obvious approach (resulting in a Fourier transform between complementary wave functions) has been slow to gain acceptance, leading to a lack of time and angle operators in quantum theory, due to the mathematical subtleties which arise from our perfectly valid predilection for Hermitian operators (and/or equivalently for wave functions that can collapse). One might argue that these can now be accepted as “fuzzy” measurements, i.e., nonprojection-valued POMs (probability-operator measures) [14] so that, at layer 1, this *is* a solution since these measurements do exist. The author, however, views fuzzy measurements as “incomplete descriptions” [7,8] of a realizable measurement, and would argue that one still needs to achieve a (sharp) “complete description” of such measurements in terms of sets of commuting Hermitian operators and their associated collapsible wave functions to fully understand the complementarity alluded to at this first layer.

A nonprojection-valued POM is simply a resolution of the identity operator by nonorthogonal eigenkets [14]. The bound in their complementary eigenspectra prevents orthogonality for the single-mode phase-kets of (4). Equivalently, (9) describes a one-sided Fourier series, but it would take a two-sided Fourier series to represent a delta function. The limited dimensionality of the space in one domain cannot support such sharp behavior in the complementary wave function. This limitation does more than prohibit delta functions, it restricts the class of single-mode phase statistics to those which must satisfy a Paley-Weiner theorem [15] which can be expressed as

$$\int_{-\pi}^{\pi} \frac{d\phi}{2\pi} |\ln|\psi(\phi)|| < \infty. \quad (11)$$

This theorem demonstrates, for example, that $|\psi(\phi)|$ cannot vanish over an interval of nonzero width. The limited dimensionality of the underlying wave functions simply cannot support such sharp behavior. Thus, the single-mode phase distribution $P(\phi)$ can vanish (equal zero) only at isolated points in ϕ .

B. Examples of Naimark’s extension theorem

The fact that nonprojection-valued POMs do correspond to realizable quantum measurements can be made more palatable via Naimark’s extension theorem [14] which (to paraphrase) states that these correspond to the measurement of sets of commuting Hermitian operators defined on a larger state space when subsets of that larger space are not entangled with the state of the original space prior to the measurement. In this section, we describe two examples from quantum optics in which the original space \mathcal{H}_s is that of a single mode (i.e., a single harmonic oscillator, the original system

of interest); the larger space is the product space $\mathcal{H}_s \otimes \mathcal{H}_a$ of the original mode with some additional (a.k.a. auxiliary) mode on \mathcal{H}_a which fundamentally *must* be a part of the physical apparatus which realizes the quantum measurement, and the additional mode is “off” (i.e., placed in the vacuum state) prior to the measurement. It is important to note from the onset, however, that Naimark’s extension theorem provides a means of describing a measurement which is *still* fuzzy in the aforementioned senses. It does not provide a general means of extending to a sharp measurement (as we will in the next section) although it can sometimes give clues as to how that might proceed.

Clearly, the operator $\hat{o} = \hat{x} + i\hat{y}$, where \hat{x} and \hat{y} are the commuting (hence simultaneously measurable) x and y position operators, is measurable in the sharp sense (since it commutes with its adjoint [16]). For an operator such as $\hat{a} = \hat{x} + i\hat{p}$, however, where \hat{p} is the x component of the momentum operator, we certainly cannot directly associate a sharp measurement since a perfectly precise simultaneous measurement of its real and imaginary components would constitute a violation of the uncertainty principle (i.e., \hat{a} does not commute with its adjoint). We can, however, associate a quantum measurement with such an operator in a fuzzy sense: a simultaneous measurement of its real and imaginary components which is not perfectly precise in either. That such a measurement exists stems from the fact that the eigenkets of this annihilation operator (the coherent states $|\alpha\rangle$ [17]) are complete, i.e., they resolve the identity operator, and completeness alone is sufficient to guarantee that $|\langle\alpha|\psi\rangle|^2$ is a perfectly valid PDF (probability distribution function) which must therefore, in some sense, describe a realizable quantum measurement. The fact that the coherent states are not orthogonal is a reflection of the fact that this is a fuzzy measurement.

Fundamental to the realizable measurement of any operator which does not commute with its adjoint is the existence of an auxiliary noise source. Zero-point fluctuations [18] from this auxiliary mode prevent a perfectly precise simultaneous measurement of the noncommuting real and imaginary parts of the original operator (so that the uncertainty principle is not violated). We can see this more clearly by extending the operator of interest to a larger space. Any quantum measurement described by a nonprojection-valued POM on \mathcal{H}_s can be represented by a collection of commuting observables on a larger Hilbert space. The utility of this representation lies in the identification of the aforementioned noise source. This auxiliary system is an integral part of the physical apparatus which realizes the quantum measurement.

Formally, the procedure is to find the Naimark extension, on $\mathcal{H} = \mathcal{H}_s \otimes \mathcal{H}_a$, of the desired POM on \mathcal{H}_s . For our purpose, this amounts to finding an operator on \mathcal{H} , say $\hat{E}_{s\otimes a}$, which commutes with its adjoint such that its real and imaginary parts form a pair of commuting observables (a.k.a. Hermitian operators). Furthermore, we require that the measurement statistics of $\hat{E}_{s\otimes a}$ reproduce those of the original operator on \mathcal{H}_s (associated with the desired POM) when the auxiliary system is in some appropriate quantum state.

We consider now a product space description of the measurement associated with the annihilation operator \hat{a}_s , for a single quantum harmonic oscillator, defined on the space

\mathcal{H}_s . When this oscillator is used to model a single mode of the electromagnetic field, this measurement can be realized by the heterodyne detection process and the noncommuting real and imaginary parts of \hat{a}_s , denoted $\hat{\chi}_s$ and $\hat{\rho}_s$, respectively, represent the in-phase and quadrature field components. It has been shown [19] that one extension of \hat{a}_s onto $\mathcal{H} = \mathcal{H}_s \otimes \mathcal{H}_a$ is

$$\hat{y} \equiv \hat{a}_s \otimes \hat{I}_a + \hat{I}_s \otimes \hat{a}_a^\dagger, \quad (12)$$

where $\hat{a}_a = \hat{\chi}_a + i\hat{\rho}_a$ is the annihilation operator for the auxiliary mode. Since operators on different spaces commute, we have $[\hat{y}, \hat{y}^\dagger] = 0$ so that the real and imaginary parts of \hat{y} , denoted as \hat{X} and \hat{P} , comprise a pair of commuting observables.

Notice that when the auxiliary mode is in the vacuum state the expected values of \hat{X} and \hat{P} , in this realizable measurement of \hat{a}_s , are the same as those of two *distinct* measurements of the noncommuting operators $\hat{\chi}_s$ and $\hat{\rho}_s$ on two identically prepared systems:

$$\langle \hat{X} \rangle_{s \otimes a} = {}_s \langle \psi | \hat{\chi}_s | \psi \rangle_s + {}_a \langle 0 | \hat{\chi}_a | 0 \rangle_a = {}_s \langle \psi | \hat{\chi}_s | \psi \rangle_s \quad (13)$$

and

$$\langle \hat{P} \rangle_{s \otimes a} = {}_s \langle \psi | \hat{\rho}_s | \psi \rangle_s - {}_a \langle 0 | \hat{\rho}_a | 0 \rangle_a = {}_s \langle \psi | \hat{\rho}_s | \psi \rangle_s, \quad (14)$$

where we used $\hat{X} \equiv (\hat{y} + \hat{y}^\dagger)/2 = \hat{\chi}_s \otimes \hat{I}_a + \hat{I}_s \otimes \hat{\chi}_a$ and $\hat{P} \equiv (\hat{y} - \hat{y}^\dagger)/2i = \hat{\rho}_s \otimes \hat{I}_a - \hat{I}_s \otimes \hat{\rho}_a$.

For the second moments, however, we find

$$\langle \hat{X}^2 \rangle_{s \otimes a} = {}_s \langle \psi | \hat{\chi}_s^2 | \psi \rangle_s + {}_a \langle 0 | \hat{\chi}_a^2 | 0 \rangle_a = {}_s \langle \psi | \hat{\chi}_s^2 | \psi \rangle_s + 1/4 \quad (15)$$

and

$$\langle \hat{P}^2 \rangle_{s \otimes a} = {}_s \langle \psi | \hat{\rho}_s^2 | \psi \rangle_s + {}_a \langle 0 | \hat{\rho}_a^2 | 0 \rangle_a = {}_s \langle \psi | \hat{\rho}_s^2 | \psi \rangle_s + 1/4. \quad (16)$$

The variance of a measurement of $\hat{\chi}_s$ on the original system (of state space \mathcal{H}_s) is $\langle \Delta \hat{\chi}_s^2 \rangle \equiv {}_s \langle \psi | \hat{\chi}_s^2 | \psi \rangle_s - ({}_s \langle \psi | \hat{\chi}_s | \psi \rangle_s)^2$. The variance of a *separate* or independent measurement of $\hat{\rho}_s$, on an identically prepared system, is $\langle \Delta \hat{\rho}_s^2 \rangle \equiv {}_s \langle \psi | \hat{\rho}_s^2 | \psi \rangle_s - ({}_s \langle \psi | \hat{\rho}_s | \psi \rangle_s)^2$. Therefore, the variances of the outcomes X and/or P , for the simultaneous measurement of \hat{X} and \hat{P} on \mathcal{H} , are larger than $\langle \Delta \hat{\chi}_s^2 \rangle$ and/or $\langle \Delta \hat{\rho}_s^2 \rangle$ by $\frac{1}{4}$. The additive terms of $\frac{1}{4}$ arise from the zero-point fluctuations of the auxiliary system. Similar terms appeared in the Arthurs and Kelly derivation of an uncertainty principle for the simultaneous measurement of position and momentum [20].

In the measurement apparatus of heterodyne detection, the auxiliary noise source is the “image” electromagnetic mode (characterized by annihilation operator \hat{a}_a) which resides at the same frequency displacement from the classical local oscillator (LO) frequency as does the original mode of interest (characterized by annihilation operator \hat{a}_s) as in Fig. 2. Even a classical treatment of heterodyning reveals that the beat of the image band with the local oscillator is mapped onto the same detector frequency as the beat of the local oscillator with the



FIG. 2. (Color online) Heterodyne detection.

original signal. Quantum mechanically, however, we note that we cannot turn the image mode off in the sense that when it is in the vacuum state we have zero-point fluctuations which contribute noise to the detected signal. This noise (essential for preventing the unrealizable perfectly precise measurement of $\hat{\chi}_s$ and $\hat{\rho}_s$) is seen to be irrevocably imbedded in the measurement apparatus of the heterodyne detection process.

Extensions of this type are not unique and Naimark’s theorem is only meant to recover the fuzzy statistics (not go beyond them). Algebraically, we see that for any operator \hat{O}_s whose commutator with its adjoint is $\hat{\zeta}_s \equiv [\hat{O}_s, \hat{O}_s^\dagger]$ one such extension is $\hat{E} = \hat{O}_s \otimes \hat{\zeta}_a + \hat{\zeta}_s^\dagger \otimes \hat{O}_a^\dagger$, provided that $\hat{\zeta}^2 = \hat{\zeta}$ (as it does for \hat{a} and \hat{A}).

Let \hat{A}_s be the SG operator for the original system of interest; an extension of this form onto $\mathcal{H} = \mathcal{H}_s \otimes \mathcal{H}_a$ is

$$\hat{Y} \equiv \hat{A}_s \otimes \hat{V}_a + \hat{V}_s \otimes \hat{A}_a^\dagger. \quad (17)$$

We can also see explicitly from the number representations of \hat{A} and \hat{A}^\dagger , and recalling the definition of the vacuum projector $\hat{V} = |0\rangle\langle 0|$, we have

$$\hat{A} \hat{V} = 0 = \hat{V} \hat{A}^\dagger \quad (18)$$

so that

$$[\hat{Y}, \hat{Y}^\dagger] = [\hat{A}_s, \hat{A}_s^\dagger] \otimes \hat{V}_a + \hat{V}_s \otimes [\hat{A}_a^\dagger, \hat{A}_a] = 0. \quad (19)$$

Next, by solving for the eigenkets of \hat{Y} , we obtain the \hat{Y} measurement statistics for an arbitrary state $|\psi\rangle_{s \otimes a} \equiv \sum_{n_s, n_a} \psi_{n_s, n_a} |n_s, n_a\rangle$ on \mathcal{H} . Setting $\hat{Y}|Y\rangle = Y|Y\rangle$ yields

$$|Y\rangle = \sum_{n_s=0}^{\infty} e^{in_s \phi} |n_s, 0\rangle + \sum_{n_a=1}^{\infty} e^{-in_a \phi} |0, n_a\rangle, \quad (20)$$

where $Y = e^{i\phi}$. The eigenvalue Y could also be zero, in which case the corresponding eigenket can be any superposition of “off-axis” number states, i.e., those which do not involve $|n_s, 0\rangle$ or $|0, n_a\rangle$. When the auxiliary mode is in the vacuum state $|\psi_a\rangle_a = |0\rangle_a$, the general result reduces to the following: the outcome of the \hat{Y} measurement is a complex number of unit magnitude $Y = e^{i\phi}$ and the probability distribution for its phase is

$$p(\phi) = \frac{1}{2\pi} |\langle \phi | \psi \rangle_s|^2, \quad (21)$$

identical to that from the POM description of the measurement associated with \hat{A} .

The single complex-valued outcome $Y = e^{i\phi}$ of the \hat{Y} measurement may also be viewed as a pair of real-valued results Y_1 and Y_2 , where $Y = Y_1 + iY_2$. The fact that \hat{Y} commutes with its adjoint implies that its real and imaginary parts $\hat{Y}_1 \equiv (\hat{Y} + \hat{Y}^\dagger)/2$ and $\hat{Y}_2 \equiv (\hat{Y} - \hat{Y}^\dagger)/2i$ are commuting observables. These can also be written in the form $\hat{Y}_1 = \hat{C}_s \otimes \hat{V}_a + \hat{V}_s \otimes \hat{C}_a$ and $\hat{Y}_2 = \hat{S}_s \otimes \hat{V}_a - \hat{V}_s \otimes \hat{S}_a$ (where \hat{C} and \hat{S} are the real and imaginary parts of \hat{A} [21]). The expected values of these, from a single \hat{Y} measurement with the auxiliary mode in the vacuum state, are identical to those of the two distinct measurements of \hat{C} and \hat{S} , i.e.,

$$\begin{aligned} \langle \hat{Y}_1 \rangle_{s \otimes a} &= {}_s \langle \psi | \hat{C}_s | \psi \rangle_s {}_a \langle 0 | \hat{V}_a | 0 \rangle_a + |\psi_0|^2 {}_a \langle 0 | \hat{C}_a | 0 \rangle_a \\ &= {}_s \langle \psi | \hat{C}_s | \psi \rangle_s \end{aligned} \quad (22)$$

and similarly

$$\langle \widehat{Y}_2 \rangle_{s \otimes a} = {}_s \langle \psi | \widehat{S}_s | \psi \rangle_s. \quad (23)$$

For the second moments, first note that

$$\widehat{Y}_1^2 = \widehat{C}_s^2 \otimes \widehat{V}_a + \widehat{V}_s \otimes \widehat{C}_a^2 + \frac{\widehat{A}_s^\dagger \widehat{V}_s}{2} \otimes \frac{\widehat{V}_a \widehat{A}_a}{2} + \frac{\widehat{V}_s \widehat{A}_s}{2} \otimes \frac{\widehat{A}_a^\dagger \widehat{V}_a}{2} \quad (24)$$

and

$$\widehat{Y}_2^2 = \widehat{S}_s^2 \otimes \widehat{V}_a + \widehat{V}_s \otimes \widehat{S}_a^2 - \frac{\widehat{A}_s^\dagger \widehat{V}_s}{2} \otimes \frac{\widehat{V}_a \widehat{A}_a}{2} - \frac{\widehat{V}_s \widehat{A}_s}{2} \otimes \frac{\widehat{A}_a^\dagger \widehat{V}_a}{2}. \quad (25)$$

Furthermore, since $\widehat{V}_a \widehat{A}_a = |0\rangle_s \langle 1|$ and $\widehat{A}_a^\dagger \widehat{V}_a = |1\rangle_s \langle 0|$, the second moments of \widehat{Y}_1 and \widehat{Y}_2 in a \widehat{Y} measurement with the auxiliary mode in the vacuum state are

$$\begin{aligned} \langle \widehat{Y}_1^2 \rangle_{s \otimes a} &= {}_s \langle \psi | \widehat{C}_s^2 | \psi \rangle_s + |\psi_0|^2 {}_a \langle 0 | \widehat{C}_a^2 | 0 \rangle_a \\ &= {}_s \langle \psi | \widehat{C}_s^2 | \psi \rangle_s + \frac{|\psi_0|^2}{4} \end{aligned} \quad (26)$$

and

$$\langle \widehat{Y}_2^2 \rangle_{s \otimes a} = {}_s \langle \psi | \widehat{S}_s^2 | \psi \rangle_s + \frac{|\psi_0|^2}{4}. \quad (27)$$

To interpret the real and imaginary components let us return momentarily to the original notation of the SG operator $\widehat{A} \rightarrow e^{i\phi}$. The fact that the SG operator does not commute with its adjoint implies that its Hermitian real and imaginary parts $\widehat{C} \equiv (e^{i\phi} + e^{-i\phi})/2$ and $\widehat{S} \equiv (e^{i\phi} - e^{-i\phi})/2i$ do not commute: $[\widehat{C}, \widehat{S}] = i\widehat{V}/2$. Classically, we have the trigonometric identity

$$\begin{aligned} 1 &= (e^{i\phi})(e^{-i\phi}) = (\cos \phi + i \sin \phi)(\cos \phi - i \sin \phi) \\ &= \cos^2 \phi + \sin^2 \phi. \end{aligned} \quad (28)$$

Quantum mechanically, we have $\widehat{I} = (e^{i\phi})(e^{-i\phi})$, so

$$\begin{aligned} 1 &= \langle (\widehat{C} + i\widehat{S})(\widehat{C} - i\widehat{S}) \rangle = \langle \widehat{C}^2 \rangle + \langle \widehat{S}^2 \rangle - i \langle [\widehat{C}, \widehat{S}] \rangle \\ &= \langle \widehat{C}^2 \rangle + \langle \widehat{S}^2 \rangle + \frac{|\psi_0|^2}{2}. \end{aligned} \quad (29)$$

The moments of \widehat{A} , or equivalently those of \widehat{C} and \widehat{S} , do not correspond to a realizable quantum measurement, but the moments of \widehat{Y}_1 and \widehat{Y}_2 do and from (25), (26), and (29) we obtain

$$\langle \widehat{Y}_1^2 \rangle_{s \otimes a} + \langle \widehat{Y}_2^2 \rangle_{s \otimes a} = 1 \quad (30)$$

in accordance with the assertion that the outcome of the \widehat{Y} measurement is a complex number of unity magnitude: $Y = e^{i\phi}$.

The physical apparatus which realizes the \widehat{Y} measurement has not yet been identified. The presence of the auxiliary mode, however, *must* (on physical grounds) be an inextricable part of this apparatus, as it is in heterodyne detection.

II. BEYOND NAIMARK: COMPLEMENTARY PHASE AT THE SECOND LAYER OF UNDERSTANDING

In this section, we provide an example of how one can extend the SG operator onto a subset of the two-mode space in order to describe a sharp quantum phase measurement which

yields the next layer of understanding complementarity (be it for phase, time, or an angle). There are an infinite number of such subsets that can be defined on a two-mode space, but a general theory (and final layer of understanding) on an unrestricted two-mode space will be discussed in Sec. IV.

Historically, \widehat{Y} was derived on physical (rather than algebraic) grounds to achieve a sharp measurement, rather than just recover the fuzzy statistics [7,8] as follows. Clearly, the problems in formulating a time or angle operator stem from the bounded eigenspectra of the complementary quantity (energy or angular momentum). For example, for a phase operator it is the absence of negative-energy states for the quantum harmonic oscillator which leads to the SG operator not commuting with its adjoint. The SG operator cannot lower below the vacuum (2) since there are no ‘‘negative-number’’ (negative-energy) states for the oscillator. The extension \widehat{Y} , however, lowers the original system mode photon number

$$\widehat{Y}|n_s\rangle_s |0\rangle_a = |n_s - 1\rangle_s |0\rangle_a \quad (n_s \geq 1), \quad (31)$$

then *continues* through the vacuum

$$\widehat{Y}|0\rangle_s |0\rangle_a = |0\rangle_s |1\rangle_a \quad (32)$$

and raises the auxiliary mode photon number

$$\widehat{Y}|0\rangle_s |n_a\rangle_a = |0\rangle_s |n_a + 1\rangle_a. \quad (33)$$

Topologically, it is as if \widehat{Y} continues to lower below the vacuum into the auxiliary (negative-number) mode. The visualization of this behavioral aspect can be facilitated by simply relabeling the number states according to the value of $m = n_s - n_a$. One might anticipate that these translations in energy difference will lead to complementarity between m and the ‘‘relative phase’’ between the two oscillators. Indeed, let us relabel the \widehat{Y} eigenkets as

$$|\phi\rangle' = \sum_{n_s=0}^{\infty} e^{in_s\phi} |n_s\rangle_s |0\rangle_a + \sum_{n_a=1}^{\infty} e^{-in_a\phi} |0\rangle_s |n_a\rangle_a. \quad (34)$$

These reside on a subset \mathcal{H}' of $\mathcal{H}_s \otimes \mathcal{H}_a$ which is spanned by $|n_s\rangle_s |n_a\rangle_a : n_s n_a = 0$. When the auxiliary mode is in the vacuum state ($n_a = 0$), the \widehat{Y} measurement yields the single-mode statistics and their attendant Paley-Wiener restriction.

We can go *beyond* these fuzzy statistics by exciting the auxiliary mode to create an arbitrary state on \mathcal{H}' :

$$|\psi\rangle = \sum_{n_s=0}^{\infty} \psi_{n_s,0} |n_s\rangle_s |0\rangle_a + \sum_{n_a=1}^{\infty} \psi_{0,n_a} |0\rangle_s |n_a\rangle_a. \quad (35)$$

Let $\psi_m \equiv \psi_{m,0}$ ($\forall m \geq 0$) and $\psi_m \equiv \psi_{0,-m}$ ($\forall m < 0$). The generalized phase wave function

$$\psi'(\phi) \equiv \langle \phi | \psi \rangle = \sum_{m=-\infty}^{\infty} \psi_m e^{-im\phi} \quad (36)$$

is a two-sided Fourier series. The Paley-Wiener restriction is removed and the $\psi'(\phi)$ can now ‘‘collapse’’ to a delta function.

Commensurate with its negative-number behavioral aspect, the auxiliary mode can be interpreted as a phase-reversed mode in the following sense. Consider the case of when the auxiliary mode is in the vacuum state ($n_a = 0$) and denote an initial state by $|\psi\rangle_0$. The state (in the Schrödinger picture) after time

evolution of an amount τ is

$$|\psi_\tau\rangle = e^{-i(\hat{n}_s + \hat{n}_a)\omega\tau} |\psi_0\rangle|_{n_a=0} = e^{-i\hat{n}_s\omega\tau} |\psi_0\rangle \quad (37)$$

so that the relation between the generalized phase representations of the initial and delayed states is simply

$$\psi'_\tau(\phi) = \psi'_0(\phi + \omega\tau) \quad (n_a = 0). \quad (38)$$

Now, consider the case of the original system being in the vacuum state ($n_s = 0$). The Schrödinger picture of the delayed version of an initial state $|\psi_0\rangle$ is

$$|\psi_\tau\rangle = e^{-i(\hat{n}_s + \hat{n}_a)\omega\tau} |\psi_0\rangle|_{n_s=0} = e^{-i\hat{n}_a\omega\tau} |\psi_0\rangle. \quad (39)$$

The initial and delayed generalized phase representations for this case are related by

$$\psi'_\tau(\phi) = \psi'_0(\phi - \omega\tau) \quad (n_s = 0). \quad (40)$$

Thus, the two modes are phase reversed in that, under time evolution, the $n_a \geq 1$ portion of the generalized phase wave function moves backwards with respect to the $n_s \geq 1$ portion.

Of course, there is nothing mysterious about the fact that the energy difference m can be negative. Nor does the phase-reversed aspect of the auxiliary mode imply a violation of temporal causality since ϕ (which is complementary to m) turns out to be the relative phase between the two modes, as we now demonstrate, which will also start to define what we mean by “relative phase” which of course implies a quantum measurement. In so doing we shall also set the stage for the general two-mode relative phase representation (which does not require restriction of the state to \mathcal{H}').

Generalizing complementarity to be viewed as Fourier relations among wave functions, we anticipate that what might be of use here is to start with the two-dimensional Fourier transform of the $\{\psi_{n_s, n_a}\}$. Indeed, $\Psi(\phi_s, \phi_a) \equiv {}_s\langle\phi_s| {}_a\langle\phi_a|\psi\rangle$ provides this:

$$\Psi(\phi_s, \phi_a) = \sum_{n_s=0}^{\infty} \sum_{n_a=0}^{\infty} \psi_{n_s, n_a} e^{-in_s\phi_s} e^{-in_a\phi_a}, \quad (41)$$

where $|\psi\rangle$ is now an arbitrary state on $\mathcal{H}_s \otimes \mathcal{H}_a$ with number-ket expansion coefficients $\psi_{n_s, n_a} \equiv {}_s\langle n_s| {}_a\langle n_a|\psi\rangle$ and $|\Psi(\phi_s, \phi_a)|^2 / (2\pi)^2$ is the probability density function for the simultaneous measurement of ϕ_s and ϕ_a .

Under the change of variables

$$\Phi_\Sigma \equiv (\phi_s + \phi_a)/2, \quad \Phi_\Delta \equiv (\phi_s - \phi_a)/2, \quad (42)$$

we map to a different wave function

$$\psi(\Phi_\Sigma, \Phi_\Delta) = \sum_{j=0}^{\infty} \sum_{m=-j}^j \psi_{n_s, n_a} e^{-i(j\Phi_\Sigma)} e^{-i(m\Phi_\Delta)}, \quad (43)$$

where $j \equiv n_s + n_a$, so that $n_s \rightarrow (j + m)/2$ and $n_a \rightarrow (j - m)/2$ in the above. Notice that by making the change of variables in the wave function (rather than in the PDF) we have also changed the quantum measurement. If instead we made a similar change of variables in the PDF, it would correspond to measuring both ϕ_s and ϕ_a first and then adding and/or subtracting the results; but, that is not what Φ_Σ and Φ_Δ represent in (43). Since j is bounded from below, the sum phase Φ_Σ is not measurable in the sharp sense. In Sec. IV, two reasonable ways of dealing with Φ_Σ and hence defining a

direct measurement of the relative phase ϕ_Δ on $\mathcal{H}_s \otimes \mathcal{H}_a$ are presented. For states restricted to \mathcal{H}' , one of these ways can be obtained from (43) and we find

$$\psi(\Phi_\Sigma = 0, \Phi_\Delta = \phi)|_{\text{on } \mathcal{H}'} = \psi'(\phi), \quad (44)$$

which demonstrates that the argument in $\psi'(\phi)$ is a relative phase as asserted. We can define a relative phase measurement between any two modes we wish. If, however, we choose two modes which are already time or phase reversed in some physical sense (e.g., an electromagnetic mode of wave vector \vec{k} and another of wave vector $-\vec{k}$), then the formalism can lead to more physical insight. For example, if the two modes are the right and left circular polarizations of an electromagnetic plane wave, then the relative phase measurement is equal to (not just isomorphic to) the quantum angle measurement.

As clarified in the next section, for such photons our definitions of j and m and the change of variables in (42) are appropriate. If instead of such photons our harmonic oscillators (modeling angular momenta) are Schwinger's [22] fermionic primitives, then the factors of $\frac{1}{2}$ should go elsewhere [we would divide j and m by two in the above definitions and multiply by two in (42) for the appropriate change of variables].

III. HARMONIC OSCILLATOR MODELS OF ANGULAR MOMENTA

In 1952, Schwinger [22] demonstrated a connection between the algebra of two uncoupled harmonic oscillators and the algebra of angular momenta. In quantum optics, this connection has proved useful in the analysis of optical beam splitters [23] although a beam splitter does not actually perform a rotation in physical space, and the connection is merely within the mathematics. This connection has also proved useful in calculating the effects of actual rotations on systems, but the oscillators (which behave like spin- $\frac{1}{2}$ bosons) are deemed unphysical [24]. We put more physics into this connection by considering rotations of the electromagnetic field. This leads to a subtle but surprisingly significant modification of Schwinger's model. Also, by describing the oscillator states in the phase representation, we will be led to insights on the angles themselves (rather than their conjugate momenta). We begin with a brief summary of the key points of Schwinger's model. Let \hat{a}_u and \hat{a}_d denote the annihilation operators for two harmonic oscillators which are uncoupled (i.e., independent) so that

$$[\hat{a}_u, \hat{a}_d^\dagger] = 0 = [\hat{a}_u, \hat{a}_d]. \quad (45)$$

Defining

$$\hat{J}_+ \equiv \hbar \hat{a}_u^\dagger \hat{a}_d, \quad \hat{J}_- \equiv \hbar \hat{a}_d^\dagger \hat{a}_u, \quad \text{and} \quad \hat{J}_z \equiv \frac{\hbar}{2} (\hat{n}_u - \hat{n}_d), \quad (46)$$

it is easy to show that

$$[\hat{J}_+, \hat{J}_-] = 2\hbar \hat{J}_z \quad \text{and} \quad [\hat{J}_\pm, \hat{J}_\pm] = \pm \hbar \hat{J}_\pm, \quad (47)$$

which are the fundamental commutation relations of angular momentum. The raising and lowering operators \hat{J}_\pm can of course be alternatively expressed in terms of the x - and y -component angular momentum operators as

$$\hat{J}_\pm = \hat{J}_x \pm i \hat{J}_y. \quad (48)$$

Since, from (43), \hat{J}_\pm raise and lower the eigenvalue of \hat{J}_z/\hbar (i.e., m) by one, it is as though we have one spin- $\frac{1}{2}$ particle, with spin up (or down) associated with each quanta of the u (or d) oscillators. The z component of angular momentum is then simply $\hbar/2$ times the difference in the number of up and down quanta, commensurate with (43). Therefore, in terms of eigenvalues we have

$$m = (n_u - n_d)/2, \quad (49)$$

from which we anticipate $j = (n_u + n_d)/2$, where $\hbar^2 j(j+1)$ is the eigenvalue of $\hat{J}^2 \equiv \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$. Indeed,

$$\begin{aligned} \hat{J}^2 &= \frac{1}{2}(\hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+) + \hat{J}_z^2 \\ &= \hbar^2 \left(\frac{\hat{n}_u + \hat{n}_d}{2} \right) \left(\frac{\hat{n}_u + \hat{n}_d}{2} + 1 \right) \end{aligned} \quad (50)$$

so that $j = (n_u + n_d)/2$ as expected.

In Schwinger's scheme, however, a total of two up and down quanta will always yield $j = 1$, i.e., the antisymmetric singlet state (of $j = 0$) would never occur. As Sakurai [24] puts it as follows: "Only totally symmetrical states are constructed by this method. The primitive spin- $\frac{1}{2}$ particles appearing here are actually *bosons*! This method is quite adequate if our purpose is to examine the properties under rotations of states characterized by j and m without asking how such states are built up initially." Since spin- $\frac{1}{2}$ particles must obey Fermi statistics, we shall not attempt to make physical sense of these *spin- $\frac{1}{2}$ primitives which act like bosons* and shall instead simply refer to them as fermionic primitives (mathematical entities which need not be represented in the physical world).

The photon is a boson which nonetheless resembles a fermion in the sense that its spin space is two dimensional, i.e., it is "spin-1 with $m = 0$ missing" [25]. Indeed, two harmonic oscillator modes are sufficient to describe the polarization state of a single k -vector component of an electromagnetic wave since we need only consider the transverse components of its vector potential. Therefore, it seems reasonable to attempt to reconstruct the algebra of angular momenta from these physically significant photonic primitives. We pursue this by considering rotations of the electromagnetic field.

Let \hat{A} be the vector potential operator for an electromagnetic plane wave comprised of the two circularly polarized, \bar{z} propagating, same frequency modes. By requiring that the expected value of this vector operator transforms like a classical vector under rotation about \bar{z} , we obtain the well-known [25] results

$$\hat{R}_z(\phi) \hat{a}_r^\dagger \hat{R}_z^\dagger(\phi) = \hat{a}_r^\dagger e^{-i\phi} \quad \text{and} \quad \hat{R}_z(\phi) \hat{a}_l^\dagger \hat{R}_z^\dagger(\phi) = \hat{a}_l^\dagger e^{i\phi}. \quad (51)$$

From this Heisenberg picture of a rotation about the z axis by an amount ϕ , denoted by $\hat{R}_z(\phi)$, we have in the Schrödinger picture that

$$e^{-i\hat{J}_z\phi/\hbar}|1\rangle_r = e^{-i\phi}|1\rangle_r, \quad (52)$$

i.e., a right-handed circularly polarized photon is an eigenstate of \hat{J}_z/\hbar with eigenvalue $m = +1$, where we used the assumption that the vacuum is rotationally invariant $\hat{R}_z^\dagger(\phi)|0\rangle = |0\rangle$, and $|1\rangle_r$ is $|1,0\rangle \equiv \hat{a}_r^\dagger|0,0\rangle$ in $|n_r, n_l\rangle$ notation. Similarly, a left-handed circularly polarized photon is associated with

$m = -1$ and therefore photons are said to be particles of spin 1 with $m = 0$ missing. Furthermore, from (51) we find that the Schrödinger picture of a rotation about the z axis for an arbitrarily polarized field (expressed in the circularly polarized basis) is

$$\hat{R}_z(\phi)|\psi\rangle = \sum_{n_r, n_l} \psi_{n_r, n_l} e^{-i(n_r - n_l)\phi} |n_r, n_l\rangle, \quad (53)$$

which we notice is physically indistinguishable from a differential phase shift of the two circularly polarized modes. Therefore, when we utilize this rotation to derive its complementary angle-kets, as in Appendix B, for this choice of modes we would also obtain the relative phase-kets (similarly derivable under differential phase shift). Thus, for this particular choice of mode set, the angle and phase measurements are exactly identical (rather than merely isomorphic). In any event, the connection between angular momentum and these photonic primitives is clearly

$$\hat{J}_z = \hbar(\hat{n}_r - \hat{n}_l). \quad (54)$$

Note the absence of the factor of $\frac{1}{2}$ which was present in the case of fermionic primitives. As (54) leads to $m = n_r - n_l$, we expect $j = n_r + n_l$. However, because we have scaled up \hat{J}_z (by 2) we find that we must also scale up \hat{J}_x and \hat{J}_y (by 2) in order to make sense of \hat{J}^2 . We can introduce the scaling either in the relation between $\hat{J}_{x,y}$ and $\hat{J}_{+,-}$ or in the definitions of \hat{J}_+ and \hat{J}_- . Choosing the latter,

$$\hat{J}_+ \equiv 2\hbar \hat{a}_r^\dagger \hat{a}_l \quad \text{and} \quad \hat{J}_- \equiv 2\hbar \hat{a}_l^\dagger \hat{a}_r \quad (55)$$

and (48) still holds. This leads to

$$\hat{J}^2 = \hbar^2(\hat{n}_r + \hat{n}_l)(\hat{n}_r + \hat{n}_l + 1) \quad (56)$$

so that $j = n_r + n_l$ as desired. An unavoidable consequence of this scaling is the appearance of a factor of 2 in the commutators

$$[\hat{J}_i, \hat{J}_j] = 2i\hbar \epsilon_{ijk} \hat{J}_k \quad (i, j, k \in \{x, y, z\}). \quad (57)$$

This, however, is still the same group (although the structure constant has doubled), i.e., we *have* reproduced the algebra of angular momenta with these physically interpretable photonic primitives. What has happened is perhaps more clear when viewed in terms of the commutators involving \hat{J}_\pm .

We now have

$$[\hat{J}_+, \hat{J}_-] = 4\hbar\hat{J}_z, \quad (58)$$

where the 4 comes from scaling up the \hat{J}_\pm , but there is no way to scale these to alter the fact that we now have

$$[\hat{J}_z, \hat{J}_\pm] = \pm 2\hbar\hat{J}_\pm. \quad (59)$$

The factor of 2 that has to appear in (59) is a solution rather than a problem, however, since it means that \hat{J}_\pm will raise and lower the eigenvalue of \hat{J}_z/\hbar by two rather than one, which is exactly what we want! We can also see this from (55) which indicates that \hat{J}_- (for example) will annihilate one right-handed photon (with z -component angular momentum of \hbar) and create one left-handed photon (with z -component angular momentum of $-\hbar$) thereby lowering the value of m by two. Thus, \hat{J}_- lowers the $m = +1$ state immediately to the

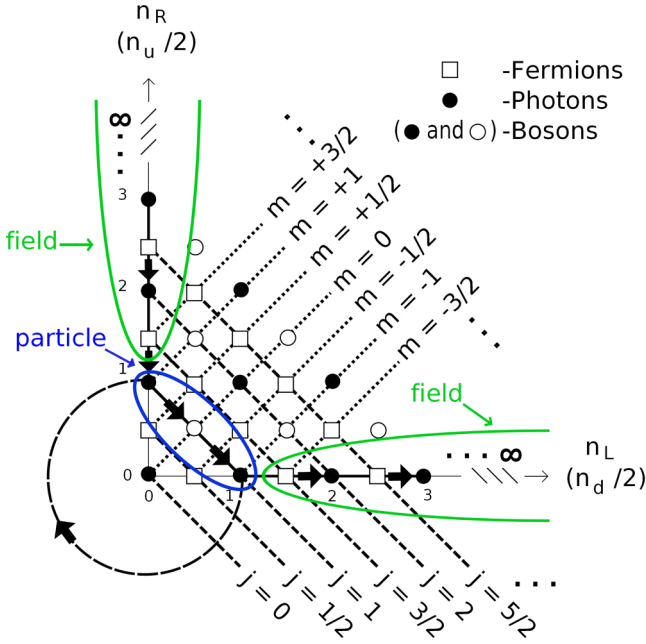


FIG. 3. (Color online) Harmonic oscillator models of angular momentum.

$m = -1$ state while automatically skipping over the $m = 0$ case which does not exist for a (plane-wave) photon.

In Fig. 3, we indicate the allowed photonic states by points (solid circles) in the n_r, n_l plane, which is also labeled by j and m . Notice that for an odd (or even) total number of photons, m must also be odd (or even). The so-called missing states are indicated by empty circles. For ordinary bosons, we would use *all* of the circles and as we can still use the angle representation to describe their polarization state we maintain the connection with harmonic oscillators, but then those would have to be the presumed unphysical fermionic primitives (for which we would then use the $n_u/2$ and $n_d/2$ axes).

In hindsight [26], we see why a similar factor of 2 has to occur in the algebra of the Stokes operators:

$$\begin{aligned} \hat{S}_0 &= \hat{a}_H^\dagger \hat{a}_H + \hat{a}_V^\dagger \hat{a}_V, & \hat{S}_2 &= \hat{a}_H^\dagger \hat{a}_V e^{i\theta} + \hat{a}_V^\dagger \hat{a}_H e^{-i\theta}, \\ \hat{S}_1 &= \hat{a}_H^\dagger \hat{a}_H - \hat{a}_V^\dagger \hat{a}_V, & \hat{S}_3 &= i \hat{a}_V^\dagger \hat{a}_H e^{-i\theta} - i \hat{a}_H^\dagger \hat{a}_V e^{i\theta}, \end{aligned} \quad (60)$$

presumably originating in [27] where H and V are the two *linear* polarization modes of relative phase shift θ . These operators satisfy the commutation relations of angular momentum $\forall \theta$ “apart from a factor of 2” [27] yet, they are not the three components of angular momentum for arbitrary θ [28].

Returning to the issue of angle, consider for example the quantum angle representation of the state of a single particle (labeled “particle” in Fig. 3). Formally, the angle of a particle requires a field for its measurement (else the finite-dimensional state space would not permit wave function collapse [29]). When the field (also labeled in Fig. 3) is comprised of photons (rather than up down oscillators) the state space for the complete, sharp, description of this measurement is a composite of the state space of the particle and that of the field (for which we could also include the vacuum state in order to turn the field “off”). For this field-particle system

to go beyond the fuzzy statistics, the state of the field would have to be entangled with that of the particle prior to the measurement. Otherwise (e.g., if the field was “off” prior to the measurement), the outcomes will be Paley-Weiner limited and these “single-particle” fuzzy statistics can be obtained via a simple Fourier transform $\psi(\varphi) = \sum_{m=-j}^j \psi_{j,m} e^{-im\varphi}$, where φ is the angle about the z axis. The fact that the quantum angle distribution cannot vanish over an interval of nonzero width has some interesting physical consequences such as *spin up really does point up* in [30]. These phenomena and the necessity of a field, etc., would not be revealed if we restricted our attention to a different measurement: the discrete-angle measurement [31] which can be described by a Hermitian operator on the space of a single-particle via a wraparound term (depicted by the semicircle in Fig. 3) akin to that of Pegg and Barnett *et al.* [9,10]. Unlike the case of discrete-phase convergence to the fuzzy single-mode phase statistics, however, the discrete-angle statistics cannot converge to the fuzzy single-particle angle statistics since we cannot just take the limit $j \rightarrow \infty$ and still be referring to a spin- j particle. Moreover, the discrete-angle statistics are complementary to a periodically replicated version of the angular momentum spectra. But, real angles are continuous, just as real angular momenta are not periodic.

IV. A GENERAL QUANTUM THEORY OF ANGLE AND RELATIVE PHASE MEASUREMENT

There are an infinite number of subsets such as \mathcal{H}' , or the field particle space in Fig. 3, that can be defined on a two-mode space wherein each value of m also corresponds to a single value of j . The general theory (and final layer of understanding) on an unrestricted two-mode space, presented herein, can cover all of those infinite number of possible measurements and it demonstrates two reasonable ways of dealing with degeneracy in j . These also correspond to the two reasonable ways of dealing with the uncollapsible Φ_Σ and hence defining a direct measurement of the relative phase ϕ_Δ on $\mathcal{H}_s \otimes \mathcal{H}_a$.

In probability theory, to eliminate one variable from a two-dimensional PDF, there are two reasonable choices: form a conditional PDF or form a marginal PDF. Rather than doing this on the PDF, we extrapolate and apply the concepts to a POM to form either a “conditional measurement,” i.e., a snapshot in absolute time of the relative phase, or a “marginal measurement,” i.e., a time average of the relative phase distribution.

In quantum theory, there are two reasonable alternatives: if the final states of a measurement are distinguishable we add probabilities; if the final states of a measurement are indistinguishable we add probability amplitudes [32]. It is perhaps initially surprising that these two things reasonably (from two totally different perspectives) coincide, i.e., we will show that taking a snapshot corresponds to adding amplitudes and taking a time average corresponds to adding probabilities.

If states of different j correspond to different (distinguishable) particles, then we could impose the “add probabilities” constraint and equivalently argue that the snapshot measurement cannot be realized for such a system. Otherwise, we could argue that as designers of quantum measurements we are free

to choose to do a measurement that yields final states which are distinguishable (in j) or not. Specifically, let the system be right and left circularly polarized photons so that the oscillators have physical significance and the relative phase measurement is equivalent to the quantum angle measurement. In a snapshot measurement of this angular distribution (say at $\Phi_\Sigma = 0$), how are we to tell which j branch contributed a result? The snapshot measurement has a connection to (43) from which we see that information on j vanishes when we take $\Phi_\Sigma = 0$. It is as if “we do not take enough time” to distinguish the different j . On the other hand, the time-average measurement is a marginal POM (rather than a marginal PDF) so the connection with (43) is not as direct but one can see that when Φ_Σ varies, the differences in j can have an effect, making them distinguishable (even if the entire system is comprised of indistinguishable photons) and so it is palatable that we end up adding probabilities. It is as if “distinguishability is a matter of how long we look.” Examples will clarify this after we first present the details of the formalism.

Let

$$|\Phi_\Sigma, \Phi_\Delta\rangle \equiv \sum_{j=0}^{\infty} \sum_{m=-j}^j |n_s, n_a\rangle e^{i(j\Phi_\Sigma)} e^{i(m\Phi_\Delta)}, \quad (61)$$

where for photons $n_s \rightarrow (j+m)/2$ and $n_a \rightarrow (j-m)/2$ in the above and m increments by two in the sum. We can eliminate Φ_Σ to obtain a *marginal measurement* of Φ_Δ on $\mathcal{H}_s \otimes \mathcal{H}_a$ by applying an “absolute time average” to $|\Phi_\Delta, \Phi_\Sigma\rangle \langle \Phi_\Delta, \Phi_\Sigma|$, resulting in the marginal POM:

$$\begin{aligned} (2\pi) d\hat{\Pi}_M(\Phi_\Delta) &\equiv \int_{-\pi}^{+\pi} \frac{d\Phi_\Sigma}{2\pi} |\Phi_\Delta, \Phi_\Sigma\rangle \langle \Phi_\Delta, \Phi_\Sigma| \\ &= \sum_{j=0}^{\infty} \left[\left(\sum_{m=-j}^{+j} |j, m\rangle e^{im\Phi_\Delta} \right) \left(\sum_{m'=-j}^{+j} \langle j, m'| e^{-im'\Phi_\Delta} \right) \right]. \end{aligned} \quad (62)$$

Because both of the inner sums use the same value of j , interference among states of different j is excluded and we have (for pure states) the following probability distribution function:

$$P_M(\Phi_\Delta) = \text{Tr}[\hat{\rho} d\hat{\Pi}_M(\Phi_\Delta)] = \left(\frac{1}{2\pi} \right) \sum_{j=0}^{\infty} |\Psi^{(j)}(\Phi_\Delta)|^2, \quad (63)$$

where Tr denotes trace; $\hat{\rho}$ is the density matrix; and

$$\Psi^{(j)}(\Phi_\Delta) \equiv \sum_{m=-j}^{+j} \Psi_{j,m} e^{-im\Phi_\Delta} \quad (64)$$

is the quantum angle representation for each j branch of expansion coefficients $\Psi_{j,m} = \langle j, m | \psi \rangle$. Thus, in this marginal or time-averaged measurement of Φ_Δ , the results from states of different j are distinguishable and we are led to adding probabilities in (63).

We also might eliminate Φ_Σ to obtain a *conditional measurement* of Φ_Δ on $\mathcal{H}_s \otimes \mathcal{H}_a$ by taking “snapshot in absolute time” via conditioning $|\Phi_\Delta, \Phi_\Sigma\rangle \langle \Phi_\Delta, \Phi_\Sigma|$ to $\Phi_\Sigma = 0$

(for example) resulting in the conditional POM:

$$\begin{aligned} (2\pi C) d\hat{\Pi}_C(\Phi_\Delta) &\equiv |\Phi_\Delta, \Phi_\Sigma = 0\rangle \langle \Phi_\Delta, \Phi_\Sigma = 0| \\ &= \left(\sum_{j=0}^{\infty} \sum_{m=-j}^{+j} |j, m\rangle e^{im\Phi_\Delta} \right) \left(\sum_{j'=0}^{\infty} \sum_{m'=-j'}^{+j'} \langle j', m'| e^{-im'\Phi_\Delta} \right), \end{aligned} \quad (65)$$

where the renormalization constant C is

$$C \equiv P(\Phi_\Sigma = 0) = \sum_{m=-\infty}^{\infty} \left\{ \left| \left(\sum_{j=0}^{\infty} \Psi_{j,m} \right) \right|^2 \right\} \quad (66)$$

the probability of the conditioning event.

In contrast to the case in (62), the sums over m in (65) now use different values of j , thereby permitting interference among the states of different j . Therefore, we have (for pure states) the probability distribution function

$$P_C(\Phi_\Delta) = \text{Tr}[\hat{\rho} d\hat{\Pi}_C(\Phi_\Delta)] = \left(\frac{1}{2\pi C} \right) \left| \sum_{j=0}^{\infty} \Psi^{(j)}(\Phi_\Delta) \right|^2 \quad (67)$$

so that in this *conditional* “snapshot” measurement of Φ_Δ we are adding amplitudes, the $\Psi^{(j)}(\Phi_\Delta)$, before taking the magnitude square in (67).

Note that the snapshot measurement recovers the \hat{Y} measurement on \mathcal{H}' , or any other of the infinite number of operators that could be defined on subsets of $\mathcal{H}_s \otimes \mathcal{H}_a$ in which each value of m corresponds to a unique value of j (for these nondegenerate cases $C = 1$ so that renormalization would not be required). For the degenerate cases, we could equivalently form an amplitude for being in a state of m , independent of j via $\Psi_m \equiv \sum_j \Psi_{j,m}$ and Fourier transform these to form the wave function underlying the snapshot PDF of (67). We could similarly define a conditional measurement of Φ_Δ by taking a snapshot at some other value of Φ_Σ but that would be equivalent to applying the above procedure to a state that has undergone time evolution of that same (conditioning) amount $\Phi_\Sigma/\omega = t$ — which is computationally easier (where ω is the radian frequency of our oscillator, which we now set equal to one for simplicity and to emphasize that Φ_Σ corresponds to absolute time).

V. QUANTUM ANGLE REPRESENTATIONS OF OPTICAL POLARIZATION

Applying the general formalism to a case that can naturally support a time-reversal symmetry (as in Sec. II) should yield more physical insights. Herein, we take the original system (s) and the auxiliary mode (a) to be the right (R) and left (L) circularly polarized modes of a single k -vector z -propagating plane wave so that the relative phase measurement is the quantum angle measurement in a system comprised of physically realizable primitives. The conditional measurement now corresponds to taking a snapshot of the angular distribution of the electric field about the z axis (at some angle $\phi = \Phi_\Delta$ with

respect to the x axis) taken at some time $t = \Phi_\Sigma$. The marginal measurement forms a properly weighted time average of these angular snapshots which trace out, and thereby result in, a quantum version of the polarization ellipse.

When the state of the field is comprised of a single value of total angular momentum (i.e., when the only nonzero probability amplitudes have one unique value of $j = n_R + n_L$), the polarization ellipse is the same distribution as a snapshot taken at any point in absolute time. Moreover, in the case of an x -polarized number state, with the y -polarized mode in the vacuum state, one would expect these identical distributions to peak at both $\phi = 0$ and $\pm\pi$. For example, one x -polarized photon, or two x -polarized photons, should have a polarization ellipse along the x axis with “up along x ” and “down along x ” being equally most likely; the identical snapshot distribution at any time must follow suit. Indeed, this physically reasonable result holds in the quantum angle representation.

However, if we have a superposition of one x -polarized photon and two x -polarized photons, then the value of j is not unique and the snapshot distributions are not identical to the polarization ellipse. For example, the superposition

$$\begin{aligned} \sqrt{2} |\psi\rangle &= |1\rangle_x |0\rangle_y + |2\rangle_x |0\rangle_y = (|1\rangle_R |0\rangle_L + |0\rangle_R |1\rangle_L) / \sqrt{2} \\ &+ (|2\rangle_R |0\rangle_L + |0\rangle_R |2\rangle_L + \sqrt{2} |1\rangle_R |1\rangle_L) / 2 \end{aligned} \quad (68)$$

has an angle representation comprised of $j = 1$ and 2 components given by $\sqrt{2} \psi(\phi) = \Psi^{(1)}(\phi) + \Psi^{(2)}(\phi)$ where

$$\Psi^{(1)}(\phi) = \sqrt{2} \cos(\phi) \quad \text{and} \quad \Psi^{(2)}(\phi) = \cos(2\phi) + 1/\sqrt{2}. \quad (69)$$

If we magnitude square either of these it will have a peak at $\phi = \pi$ and the sum of those probabilities would yield the polarization ellipse. If we add amplitudes instead, then although both j components have an amplitude for being “down along x ” at this time (when $\Phi_\Sigma = 0$), those amplitudes are out of phase at this time so they cancel and the snapshot distribution (when $\Phi_\Sigma = 0$) is only “up along x ” (i.e., peaks only at $\phi = 0$) as shown in Fig. 4, where the x axis is absolute time (Φ_Σ from 0 to π); the y axis is the angle (ϕ from $-\pi$ to π); and the contours indicate the *angular* probability density (from 0.1 to 0.8), i.e., each “slice in x ” is a (normalized) snapshot PDF. Later (when $\Phi_\Sigma = \pi$), the snapshot distribution must reflect an angular distribution of electric field vectors that is primarily “down along x ” (i.e., the snapshot distribution, when $\Phi_\Sigma = \pi$, must peak at $\phi = \pm\pi$, as it does).

Consider now how “up along x ” can evolve into “down along x .” Classically, the two (right- and left-handed) circular polarization vectors counter-rotate in time and their sum creates a linearly polarized vector that shrinks to zero halfway between the time it is “up along x ” and the time it is “down along x .” Quantum mechanically, the renormalized snapshot PDFs cannot vanish (although without renormalization these can shrink towards zero, reflecting an unlikelihood of taking a snapshot at that time). Moreover, these PDFs are Paley-Weiner restricted in shape (for finite j) so delta functions are not allowed and the PDFs cannot equal zero over angular intervals of nonzero length. It is analogous to considering how one might deform a balloon from one pointing up to one pointing down without breaking the balloon, while also maintaining

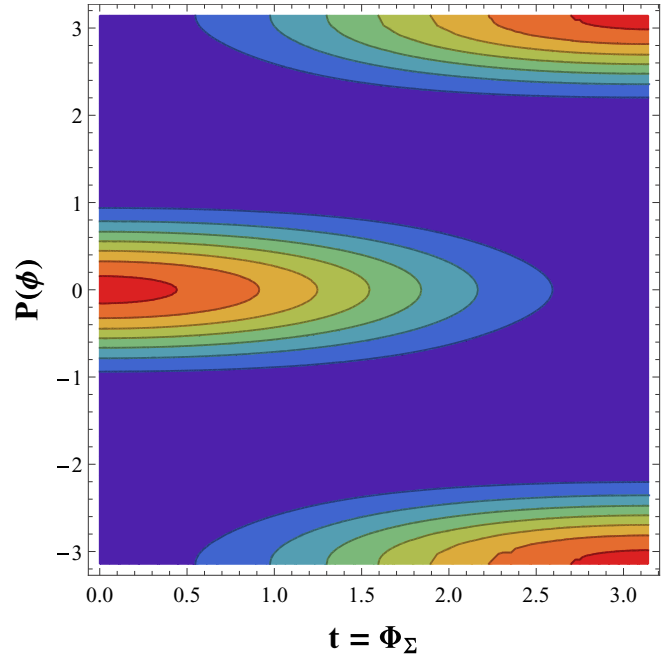


FIG. 4. (Color online) A sequence (at times t) of snapshot PDFs (along y) for a superposition of one and two x -polarized photons.

symmetry in ϕ . Two possibilities are to have a peak at $\phi = 0$ come down and spread out into a more uniform shape while the peak at $\phi = \pm\pi$ comes up or the peak at $\phi = 0$ can come down and spread out into two or more discernible counter-rotating peaks, which then recombine into a peak at $\phi = \pm\pi$. The superposition of one and two x -polarized photons of (68) gives an example of the former in Fig. 4, and an x -polarized coherent state of $N = 1$ gives an example of the latter in Fig. 5.

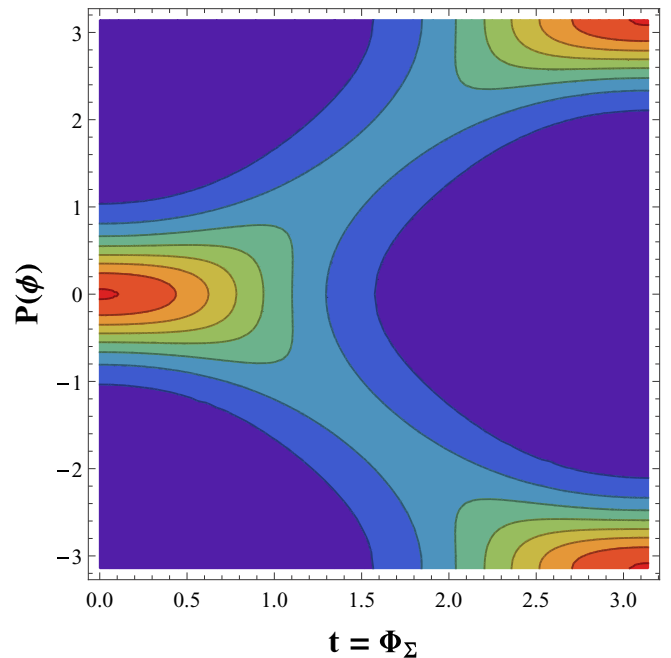


FIG. 5. (Color online) A sequence (at times t) of snapshot PDFs (along y) for a coherent State of $N = 1$.

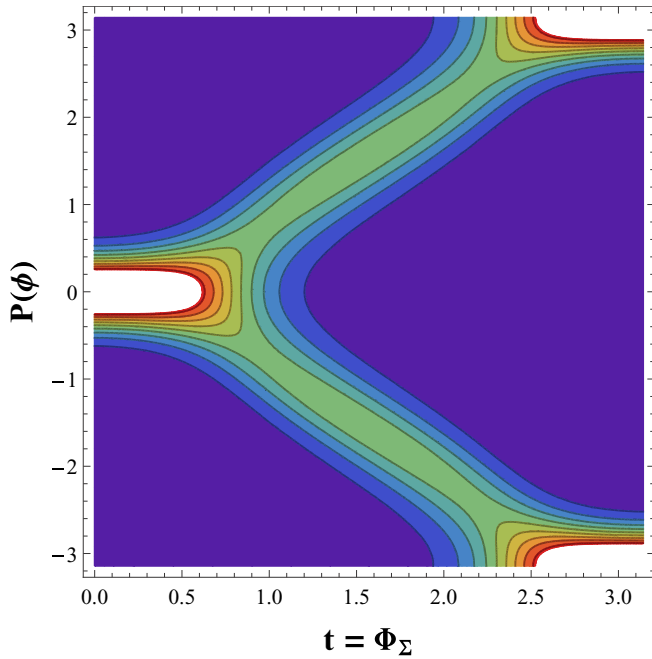


FIG. 6. (Color online) A sequence (at times t) of snapshot PDFs (along y) for a coherent state of $N = 4$.

Figures 6 and 7 similarly depict the evolution of the snapshot PDFs for x -polarized coherent states of $N = 4$ and 9, respectively. Both demonstrate again the splitting of the initial “up along x ” peak into two discernible peaks (which would be counter-rotating in a polar plot) which then again have to recombine into a single “down along x ” peak (at $\phi = \pm\pi$) when $\Phi_\Sigma \rightarrow \pi$. For $N = 9$, Fig. 7 also reveals that when Φ_Σ

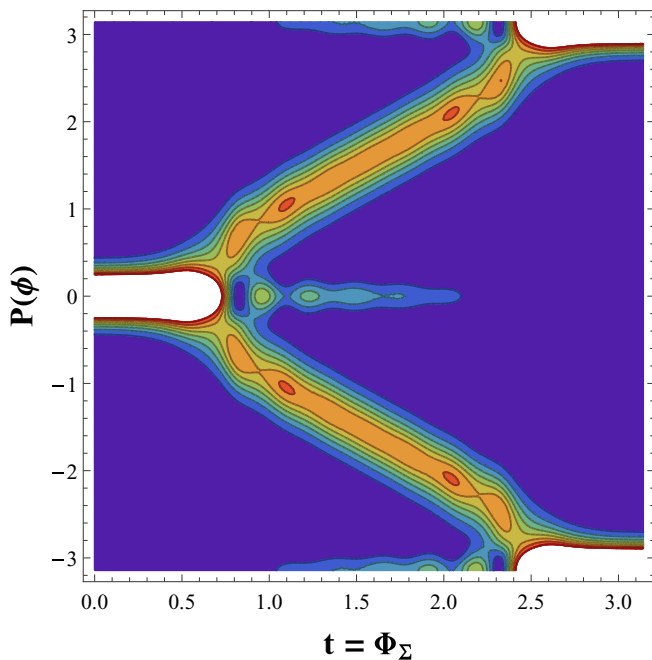


FIG. 7. (Color online) A sequence (at times t) of snapshot PDFs (along y) for a coherent state of $N = 9$.

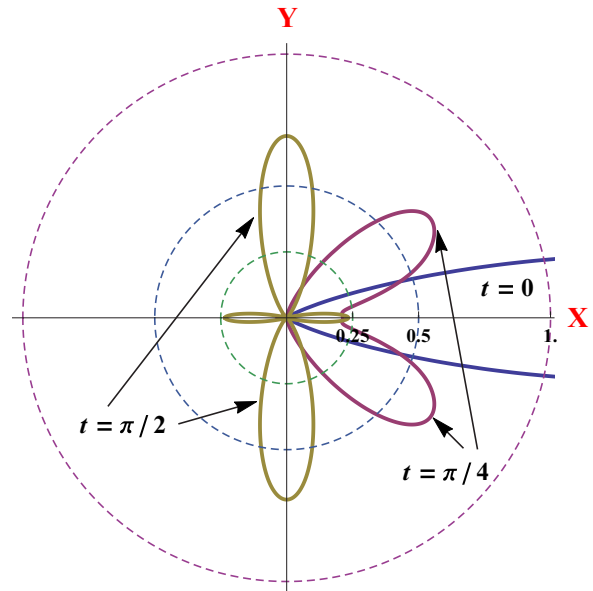


FIG. 8. (Color online) Three snapshot PDFs from Fig. 7: presented in a polar plot on a linear scale.

is near $\pi/2$, two more peaks or side lobes (at $\phi = 0$ and at $\phi = \pm\pi$) become visible.

Before taking up the issue of the behavior at $\phi = \pi/2$, it is useful to consider the role that the probability of Φ_Σ plays in the relation between a sequence of snapshots and the quantum polarization ellipse. For example, the snapshot PDF for an x -polarized coherent state of $N = 9$ reveals an appreciable $P(\phi = \pi/2)$ when the sum phase is near $\pi/2$, as shown in Fig. 8 which presents three snapshot PDFs in a polar plot on a linear scale. The peak of the snapshot PDF at time $t = \pi/2$ is down from the peak of the snapshot PDF at time $t = 0$ by less than a factor of 4. However, the probability of taking a snapshot at time $t = \pi/2$ is over three orders of magnitude smaller than the probability of taking a snapshot at time $t = 0$. Thus, when we take the time average inherent in forming the polarization ellipse the influence of the $t = \pi/2$ snapshot is greatly diminished.

The renormalization constant C is the probability of the conditioning event, i.e., the constant is itself also a PDF, the PDF for the measurement of absolute phase which (like absolute time) is measurable in the fuzzy, albeit not in the sharp, sense. In considering a sequence of snapshot PDFs for various Φ_Σ we should dispense with the notion of absolute time marching along uniformly (as it has historically, as a parameter in quantum theory rather than an operator). Instead, when we realize that the probability of the conditioning event is telling us that taking a snapshot at one time is not as probable as at some other time, then we see more clearly how these snapshots turn into the quantum polarization ellipse. It is as if the polarization vectors spend more time in snapshots where $P(\Phi_\Sigma)$ is large and zip past (or even skip) snapshots taken at Φ_Σ of small (or even zero) $P(\Phi_\Sigma)$. The time averaging inherent in the marginal POM naturally incorporates this weighting.

The quantum polarization ellipses for an x -polarized coherent state of $N = 1, 4,$ and 9 are presented on a dB scale in the polar plots of Fig. 9. To better reveal the side-lobe

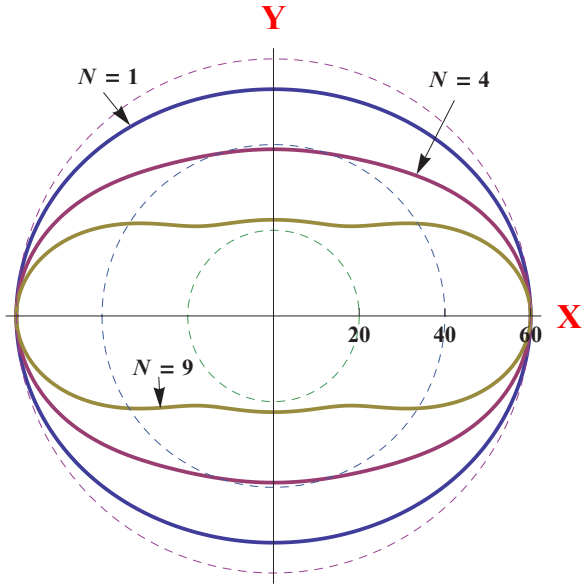


FIG. 9. (Color online) Quantum polarization ellipses on a dB scale for x -polarized coherent states of $N = 1, 4$, and 9 .

structure it is preferable to plot the log of probability (rather than the probability itself) and to avoid a negative radius in a polar plot we add a scaling constant. In the graphics, we arbitrarily scale each of these so that the peak of each of the PDFs corresponds to 60 dB (setting the peaks to some fixed reference level facilitates comparison of the underlying shapes). We see, for $N = 9$, the probability of being “up along the y axis” is almost 40 dB below the probability of being “up along the x axis.” For $N = 4$ we see this ratio is over 20 dB and for $N = 1$ we are approaching the $N = 0$ case of a uniform distribution (a circular polarization ellipse) since the vacuum state is rotationally invariant.

Of course, an x -polarized coherent state is not exactly orthogonal to a y -polarized coherent state since they share the same vacuum state. But, an x -polarized number state can also have an appreciable $P(\phi = \pi/2)$ although these are exactly orthogonal to any y -polarized number state (when that number is not zero). More generally, within the phase representation one can readily see that only terms of the form $\cos(m\phi)$ will arise for states of $\psi_{j,m}$ that are symmetric under $m \rightarrow -m$ [33–35]. Any state for which the y -polarization mode is in the vacuum state will have this symmetry. For photons, m will be odd when j is odd (m increments by two for photons, see Fig. 3) and in that case the probability of $\phi = \pm\pi/2$ vanishes (for all Φ_Σ). Thus, a single x -polarized photon will never have an electric field pointing along the y axis [i.e., $P(\phi = \pm\pi/2) = 0$] when its quantum angle is measured, but a two-photon x -polarized field can! Indeed, any x -polarized field with nonvanishing components of even j (such as a coherent state) will exhibit such behavior. In the classical limit $P(\phi = \pi/2)$ rapidly diminishes (even for N only equal to 9 it is already less than 0.0002) but for a weak coherent state of $N = 1$ it is slightly over 6%. Since weak coherent states are sometimes used in polarization-based quantum communication systems to mimic single-photon states, such an effect might merit consideration [36].

VI. CONCLUDING REMARKS

The complementarity between time and energy, as well as between an angle and a component of angular momentum, was described at three layers of understanding complementarity in a more general context. The first layer, comprised of a simple Fourier transform of the complementary wave function, amounts to a nonprojection-valued probability-operator measure and we elucidated ways in which these can be interpreted as fuzzy measurements. The phase of a single harmonic oscillator and the angle of a single particle are examples in which the limited dimensionality of the state space was shown to prevent wave-function collapse in the phase or angle wave functions and further restrict the class of their realizable fuzzy measurement statistics. Such measurements can, however, be described via sets of commuting observables on a larger state space which includes an auxiliary system (which *must* be a part of the apparatus which realizes the measurement) when the auxiliary and original systems are not entangled prior to the measurement.

Therein, the auxiliary system is shown to function as a noise source and the fuzzy statistics manifest. Such extensions to larger state spaces are not unique and are only intended to recover the fuzzy statistics. To go beyond these, we must also extend the *meaning* of what it is that we wish to measure. Clearly, no general way of doing that can exist but using complementarity as a guide we were led to the conclusion that formally only relative phase can be measured and the angle of a particle will require a field for its measurement. The meaning of relative phase gleaned at the second layer of understanding was achieved on an important subspace of the more general two-mode (two-oscillator) state space. This subspace has some interesting physical properties (automatically entangling the two modes) which are also useful for quantum noise reduction and it is the space in which the celebrated NOON states reside. The phenomena of super-resolution are readily apparent in the quantum phase representation which also reveals that entanglement is not required and the NOON state performance can be identically reproduced on the single-mode space: simply notice that the periodicity of the magnitude square of a Fourier series is set by the minimal distance in m between nonvanishing Fourier coefficients.

In preparation for the final layer of understanding complementarity, Schwinger’s harmonic oscillator model of angular momentum was modified to include the case of photons (instead of only the unrealizable fermionic primitives). Therein, the quantum angle measurement (complementary to the measurement of a component of angular momentum) was shown to be equivalent to the relative phase measurement between those two oscillators.

The meaning of relative phase was finalized at the third layer. At the second layer, there are an infinite number of subspaces that could be defined on a two-mode space wherein each value of photon-number difference also corresponds to a unique value of number sum. The general theory (and final layer of understanding) on an unrestricted two-mode space can cover all of those infinite number of possible measurements and it demonstrates two reasonable ways of dealing with the degeneracy in number sum. These also correspond to two reasonable ways of eliminating absolute time (which is measurable in a fuzzy albeit not in a sharp sense) in

order to define a direct measurement of the relative phase: a conditional measurement which takes a snapshot in absolute time (corresponding to adding probability amplitudes); and a marginal measurement which takes an average in absolute time (corresponding to adding probabilities). The sense in which distinguishability is a “matter of how long we look” was discussed and the meaning of the general theory was illustrated by taking the two oscillators to model the right and left circularly polarized modes of an electromagnetic plane wave so that the conditional measurement reveals a snapshot of the angular distribution of the electric field vector and the marginal measurement corresponds to a quantum version of the polarization ellipse.

The quantum angle representation demonstrated that any excitation of an odd number of x -polarized photons will *never* have an angle in correspondence with the y axis, but that of an even number of x -polarized photons *always* can! The behavior of an x -polarized coherent state was examined and the snapshot angular distributions were seen to evolve into two counter-rotating peaks resulting in considerable correspondence with the y axis (particularly for weaker coherent states) at the time for which a classical linear polarization vector would shrink to zero length. Such an effect could be of significance for polarization-based quantum communication systems since weak coherent states are sometimes used to mimic single-photon states. We also demonstrated how the probability distribution of absolute time (now treated as a measurable quantity, rather than just a parameter) has an influence on how these snapshot angular distributions trace out and evolve into the quantum polarization ellipse.

ACKNOWLEDGMENT

The author would like to acknowledge many useful discussions with J. H. Shapiro on quantum measurements and support from NASA EPSCoR (Grant No. NNX13AB14A) and the Louisiana Board of Regents.

APPENDIX A: ALTERNATE PATH TO THE SINGLE-MODE STATISTICS

Concurrent to (and independent of) the development of the continuous single-mode phase representation, an alternate method for obtaining the single-oscillator statistics was derived by Pegg and Barnett [9] (see also [10]). Their approach requires the truncation of the infinite-dimensional state space of a harmonic oscillator to one of finite but arbitrarily large dimension. This subspace, denoted $\mathcal{H}^T(s)$, is spanned by the number-kets $\{|n\rangle : 0 \leq n \leq s\}$. Furthermore, their formalism relies on an ordering of terms, in a polar decomposition of the annihilation operator, which is akin to Dirac’s ordering [1] rather than Susskind and Glogower’s [2]. This ordering permits them to impose an “additional condition,” which specifies the action on the vacuum state, whereas the action of the SG operator on the vacuum state is uniquely determined by that of \hat{a} , as mentioned.

For any number state $|n\rangle$, with $n \in \{1, 2, \dots, s\}$ but $n \neq 0$, the unitary power series of Pegg and Barnett’s operator, $\exp(i\hat{\phi}_{PB})$ defined on $\mathcal{H}^T(s)$, is a lowering operator. The action of this operator on the vacuum state is then defined to be a

“wraparound” term

$$e^{i\hat{\phi}_{PB}}|0\rangle \equiv e^{i(s+1)\theta_0}|s\rangle, \tag{A1}$$

where θ_0 is the location of the branch cut for phase (which is $-\pi$ in our formalism). This cyclic behavior, possible only in a truncated space such as $\mathcal{H}^T(s)$, is essential for their definition of a Hermitian phase operator on $\mathcal{H}^T(s)$. Unitarity is accomplished by not having to “stop” at the vacuum state, but the “wraparound to the top of the stack” term causes the discrete-phase wave functions to be complementary to a periodically replicated (and hence truncated) version of the $\{\psi_n\}$.

This operator can be expressed in terms of an orthogonal subset $\{|\theta_m\rangle\}$ of truncated phase-kets as

$$\hat{\phi}_{PB} = \sum_{m=0}^s \theta_m |\theta_m\rangle \langle \theta_m|, \tag{A2}$$

where

$$|\theta_m\rangle \equiv (s+1)^{-1/2} \sum_{n=0}^s e^{in\theta_m} |n\rangle \quad \text{and} \quad \theta_m \equiv \theta_0 + 2\pi \left(\frac{m}{s+1} \right). \tag{A3}$$

A measurement of $\hat{\phi}_{PB}$ on $\mathcal{H}^T(s)$ will yield one of its discrete eigenvalues θ_m , which are rational multiples of 2π plus θ_0 , with probability

$$P_r(\theta_m) = |\langle \theta_m | \psi \rangle|^2, \tag{A4}$$

where $|\psi\rangle^T$ is a truncated state $\sum_{n=0}^s \psi_n |n\rangle$ on $\mathcal{H}^T(s)$.

As the truncation point s goes to infinity, Pegg and Barnett’s discrete phase eigenspectra converge to our phase continuum (in as much as rational numbers can converge to real numbers), i.e., their probability mass function (A4) converges in distribution to our probability density. It is this $s \rightarrow \infty$ limit in which they argue that the discrete phase statistics make physical sense. It should be emphasized that although Pegg and Barnett indeed have a Hermitian phase operator on $\mathcal{H}^T(s)$ with s finite, when $s \rightarrow \infty$ this is an alternative means of calculating the fuzzy single-mode continuous phase measurement statistics via a limiting procedure. If s is left to be finite, then their discrete phase statistics correspond to a discrete Fourier transform, complementary to a truncated (and periodically replicated) version of the $\{\psi_n\}$. In the limit $s \rightarrow \infty$ these converge (in distribution) to the continuous phase statistics which we can obtain directly via the Fourier transform of the actual $\{\psi_n\}$. Note that by accepting the non-Hermitian nature of phase, we were led to the necessity of a larger state space which emphasizes the existence of an auxiliary noise source.

APPENDIX B: GENERAL THEORY OF COMPLEMENTARITY

Herein, we analyze complementarity at the first (fuzzy) layer, which is general in the sense that it can cover an angle, phase, or time (in the case of a time-independent Hamiltonian) when there is *no degeneracy in the complementary eigenspectra*. To cover cases of degenerate eigenspectra, one must use procedures akin to (62) and/or (65). At layer one, the proof can be made to follow (almost verbatim) the case which stems

from linear momentum being the generator of translations in space [24].

Postulate the existence of a set of eigenstates $|\chi\rangle$ which resolve the identity operator so that these correspond to a realizable measurement. Furthermore, let the eigenvalues χ be continuous, nondegenerate, and real valued. Denote the operator which effects an infinitesimal translation in χ by $\hat{T}(d\chi)$:

$$\hat{T}(d\chi)|\chi\rangle = |\chi + d\chi\rangle. \quad (\text{B1})$$

The following physically reasonable properties

$$\hat{T}^\dagger(d\chi) \hat{T}(d\chi) = \hat{I}, \quad (\text{B2})$$

$$\hat{T}(d\chi_1) \hat{T}(d\chi_2) = \hat{T}(d\chi_1 + d\chi_2), \quad (\text{B3})$$

$$\hat{T}(-d\chi) = \hat{T}^{-1}(d\chi) \quad \text{and} \quad \lim_{d\chi \rightarrow 0} \hat{T}(d\chi) = \hat{I} \quad (\text{B4})$$

are obtained (for *infinitesimal* $d\chi$) when

$$\hat{T}(d\chi) = \hat{I} - i\hat{G}d\chi, \quad (\text{B5})$$

where \hat{G} is said to be the generator of translations in χ . An example of this is when χ is space and \hat{G} is proportional to linear momentum with Planck's constant incorporated into its definition so that (B5) is dimensionless. To progress from this to the Fourier transform between representations (and hence achieve complementarity) we will assume that \hat{G} is Hermitian (self-adjoint) so that its eigenspectra are real. If however, we try to take \hat{G} to be an angle operator, for example, generating finite (not infinitesimal) translations in m then complications arise (differentiation not being defined for a discrete parameter being the least of them) and indeed these complications are trying to tell us something "is wrong" here which then forces us to a higher-dimensional state space in order to achieve a complete description (i.e., a sharp measurement) in terms of sets of commuting observables. We can, however, remain at the first layer and obtain an incomplete description (i.e., a fuzzy measurement) if instead we take an angle to be χ . Note, however, that in (B1) there can be no "stopping," as in the sense of the SG operator stopping at the vacuum, i.e., the eigenspectra of χ must range from $-\infty$ to ∞ else (B1) cannot hold $\forall d\chi$ and $\forall \chi$, which would preclude the definition of a derivative in what follows, i.e., (B7). *Later*, when we find the angle distribution to be periodic mod 2π (although clearly fermions can exhibit mod 4π behavior the observation of such requires their interference with another system) *then* we can restrict our attention to one of these identically distributed 2π intervals.

To be sure, avoidance of stopping is what leads us to (at the second layer) extend the SG operator to one on \mathcal{H}' , or to define a lowering of m on a field or particle system. Rather than taking a phase operator to generate translations in photon number, we can take \hat{n} as the generator of translations in ϕ (as indeed it is already accepted that the Hermitian Hamiltonian

generates translations in time, and \hat{J}_z generates translations in the angle about the z axis, etc.) and therein we can remain at the first layer and simply justify the Fourier transform that leads to the fuzzy complementary measurement statistics.

Let $\psi(\chi) \equiv \langle \chi | \psi \rangle$, from (B1) and (B5) we have

$$\langle \chi | \hat{T}^\dagger(d\chi) | \psi \rangle = \psi(\chi + d\chi) = \psi(\chi) + i\langle \chi | \hat{G}^\dagger | \psi \rangle d\chi \quad (\text{B6})$$

so that

$$\frac{d\psi}{d\chi} = i\langle \chi | \hat{G} | \psi \rangle \rightarrow \frac{d}{d\chi} \langle \chi | G \rangle = iG \langle \chi | G \rangle \quad (\text{B7})$$

when we take $|\psi\rangle$ to be an eigenket of \hat{G} (and we used $\hat{G}^\dagger = \hat{G}$.) The solution to this differential equation is $\langle \chi | G \rangle = Ne^{iG\chi}$, the kernel of the Fourier transform, where N is a normalization constant. The presumed completeness of the $|\chi\rangle$ then leads to

$$\langle G | \psi \rangle = \int d\chi \langle G | \chi \rangle \langle \chi | \psi \rangle, \quad (\text{B8})$$

i.e., with $\psi_G(G) \equiv \langle G | \psi \rangle$ we have

$$\psi_G(G) = N \int d\chi e^{-iG\chi} \psi(\chi) \quad (\text{B9})$$

from which the inverse Fourier transform follows. Then, Rayleigh's energy theorem (or Parseval's power theorem) [15] proves that a normalized distribution in one domain will have a complementary distribution which is also normalized (in the complementary domain). Thus, if the eigenkets of the generator $|G\rangle$ are complete in some space, then the eigenkets $|\chi\rangle$ are complete in that same space, proving the assumption, which concludes the proof (at the first layer).

If the original distribution, i.e., that of the Hermitian generator of translations, such as a Hamiltonian, is continuous and aperiodic (i.e., not periodic), then the complementary distribution is also continuous and aperiodic, and the two wave functions are related by the Fourier integral transform. If the original distribution is "rationally discrete" (i.e., in correspondence with numbers whose ratios are rational numbers) and aperiodic, then the complementary distribution will be continuous and periodic (the Fourier series transform relationship). If the original distribution is aperiodic and in correspondence with a discrete set of numbers whose ratios are *not* rational numbers, then the complementary distribution will be continuous and "quasiperiodic." Lastly, if the original distribution is discrete and periodic, then the complementary distribution will be also (the discrete Fourier transform). Thus, for any system with a rationally discrete energy spectrum, the temporal distribution will be periodic. Likewise, the only system which exhibits truly discrete temporal behavior is one in which the energy distribution is truly periodic. Similarly, the quantization of angular momentum (projected onto an axis) is the simple and immediate consequence of the periodicity of the angle about that axis.

[1] P. A. M. Dirac, Proc. R. Soc. London, **Ser. A** **114**, 243 (1927).

[2] L. Susskind and J. Glogower, Physics **1**, 49 (1964).

[3] Formally, a Hilbert space corresponds to a nondenumerably infinite vector space (in correspondence with the set of real numbers) rather than a denumerably infinite space (e.g., in

correspondence with the set of positive-semidefinite integers). The latter is occasionally referred to as a rigged Hilbert space, but it has become common practice in the literature to refer to even a finite-dimensional set of vectors (e.g., the eigenkets of the z component of spin for a particle of finite total spin) as

a Hilbert space. Since the dimensionality of the space in what follows is more than a matter of semantics, we shall refer to these as a state space and reserve the phrase “Hilbert space” to those spanned by eigenkets of unbounded real eigenvalues.

- [4] S. R. Shepard and J. H. Shapiro, *Proceedings of the Annual Meeting of the Optical Society of America*, Santa Clara, CA, 1988.
- [5] J. H. Shapiro, S. R. Shepard, and N. C. Wong, *Phys. Rev. Lett.* **62**, 2377 (1989).
- [6] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976), Chap. 8.
- [7] S. R. Shepard, Ph.D. thesis, Massachusetts Institute of Technology, 1990.
- [8] S. R. Shepard, Workshop on Squeezed States and Uncertainty Relations, NASA-CP-3135, 177 (1991).
- [9] D. T. Pegg and S. M. Barnett, *Europhys. Lett.* **6**, 483 (1988); *Phys. Rev. A* **39**, 1665 (1989).
- [10] V. N. Popov and V. S. Yarunin, *Vestn. Leningr. Univ.* **22**, 7 (1973); *Theor. Math. Phys.* **89**, 1292 (1991).
- [11] J. H. Shapiro, S. R. Shepard, and N. C. Wong, in *Coherence and Quantum Optics VI*, edited by L. Mandel, E. Wolf, and J. H. Eberly (Plenum, New York, 1990), p. 1077.
- [12] Historical note: These kets were first used by Shepard and Shapiro in [4] and later in [5], for their noise reduction potential. But, it was also realized (and finally documented by Shepard in [7,8]) that they held a fundamental significance for describing a quantum phase measurement complementary to the measurement of photon number as in (9). One might argue that the phase-kets first appeared as far back as 1976 in Helstrom’s [6] ML POM, but one would have to remove the “first step” component of that POM to achieve equivalence between (3) and (4), as was done in [4,7]. Interestingly (while I was trying to find a review article to reference on the approach of others), I came across M. M. Nieto, *Phys. Scripta* **T48**, 5 (1993); in which he mentions their appearance in [2] and in Carruthers and Nieto (1968) followed by “To explain why this is wrong. . .” so it seems that [4,5] and perhaps arguably [6] were the first documentations of “how to use them right” (i.e., to describe a fuzzy measurement at the first layer). Anyway, [7] appears to be the first documentation of going from the fuzzy into the sharp at the second layer in 1990 (as in Sec. II), and the third layer (as in Sec. IV) was documented in [30,31] in terms of distinguishability in 1992, whereas the notion that distinguishability can be a matter of what we chose to measure was not documented until 1993 by Shepard in NASA-CP-3270.
- [13] J. H. Shapiro and S. R. Shepard, *Phys. Rev. A* **43**, 3795 (1991).
- [14] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976), Chap. 5.
- [15] A. Papoulis, *The Fourier Integral and its Applications* (McGraw-Hill, New York, 1962).
- [16] The notion that quantum measurements correspond to Hermitian operators (which have real eigenvalues) originally stemmed from the notion that all measurements yield real numbers. Yet, a complex number is merely a way of representing two real numbers simultaneously, and no one would argue that we could not measure the operator $\hat{o} = \hat{x} + i\hat{y}$, where \hat{x} and \hat{y} are the commuting (hence simultaneously measurable) x and y position operators. Therein, we have simply labeled the y axis as imaginary and reported the outcome as a complex number, rather than a vector. Thus, we might restrict the class of realizable measurements to sets of operators which commute with their adjoint, rather than to operators which are equal to their adjoint, since then their real and imaginary components (which are Hermitian by construction) will also commute. We have referred to such measurements as sharp and for these we have a complete description of their measurement in terms of sets of commuting Hermitian operators. Therein, all of the conventional descriptions of quantum theory hold and the predilection for Hermitian operators rightfully stems from the fact that their eigenkets are orthogonal (hence providing the sharpness which permits wave-function collapse).
- [17] R. J. Glauber, *Phys. Rev.* **131**, 2766 (1963).
- [18] In general, these could be quantum fluctuations, but in our quantum optics examples, these are zero-point fluctuations since therein the auxiliary system is a mode which is in the vacuum state prior to the measurement.
- [19] S. D. Personick, *Bell Syst. Tech. J.* **50**, 213 (1971).
- [20] E. Arthurs and J. L. Kelly Jr., *Bell Syst. Tech. J.* **44**, 725 (1965).
- [21] P. Carruthers and M. M. Nieto, *Phys. Rev. Lett.* **14**, 387 (1965).
- [22] J. Schwinger, U.S. Atomic Energy Report No. NYO-3071 (1952).
- [23] B. Yurke, S. L. McCall, and J. R. Klauder, *Phys. Rev. A* **33**, 4033 (1986).
- [24] J. J. Sakurai, *Modern Quantum Mechanics* (Benjamin/Cummings, Menlo Park, CA, 1985).
- [25] V. B. Berestetskii *et al.*, *Quantum Electrodynamics* (Permagon, Elmsford, NY, 1982).
- [26] The author was unaware of [27] when he used the two circularly polarized modes to study angular momentum in [34,35].
- [27] J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (Addison-Wesley, Reading, MA, 1955).
- [28] Moreover, in [27] we find “. . . it would be entirely wrong to associate these . . . with any rotations in ordinary space.” The operators in (60), however, do reproduce the algebra of angular momentum because they are comprised of a photon-number difference and linear combinations of terms isomorphic to \hat{J}_{\pm} and because the linear and circular polarization modes are related via a linear transformation. In the algebra of the Stokes operators, Planck’s constant is eliminated, but we now see that the factor of 2 had to reappear because of the nature of the circularly polarized modes’ angular momenta (which are associated with rotations in ordinary space) per the discussion in Sec. III.
- [29] Equivalently, if we wish to describe the angle measurement in terms of operators rather than wave functions, a field is required so that we can generate translations in an unbounded m space (as exemplified by the arrows in Fig. 3 within the field-particle system). Only then can such translations be both-sided unitary and thereby define a Hermitian angle operator. If we instead introduce a wraparound term (depicted by the semicircle) and stay on the finite-dimensional state space of the particle, then we lose complementarity.
- [30] S. Shepard, *Ann. NY Acad. Sci.* **755**, 812 (1995).
- [31] A. Vourdas, *Phys. Rev. A* **43**, 1564 (1991).
- [32] R. P. Feynmann, R. B. Leighton, and M. Sands, *The Feynmann Lectures on Physics*, Vol. III (Addison Wesley, Reading, MA, 1965).
- [33] Such states are also symmetric under $\phi \rightarrow -\phi$, the unitary time-reversal eigenkets, which constitute a generalization of

- “*x*-polarized” states, in which the *y*-polarization mode need not be in the vacuum state.
- [34] S. R. Shepard, Workshop on Harmonic Oscillators, NASA-CP-3197, 109 (1992).
- [35] S. R. Shepard, Ph.D. thesis, Massachusetts Institute of Technology, 1992.
- [36] An apparatus which realizes the quantum angle measurement has yet to be realized so one must carefully consider whatever apparatus (polarization beam splitter, etc.) is actually used. Nevertheless, the quantum phase and angle representations are so fundamental that their predictions are often discernible in other quantum measurements.