QUANTUM ESTIMATION THEORY AND OPTICAL DETECTION

G. M. D'ARIANO
Dipartimento di Fisica "A. Volta",
via Bassi 6, I-27100 Pavia, Italy

1. Introduction

In quantum mechanics we call "observable" any physical quantity that can be represented by numbers. An observable is associated in a one-to-one way with a selfadjoint operator $\hat{X}$ acting on the Hilbert space $\mathcal{H}_S$ of the quantum system S, and the spectrum of $\hat{X}$ represents the set of all possible readings from the measurement. Let us consider, for example, an observable with spectrum equal to whole real line $\mathbf{R}$, and with spectral decomposition

$$\hat{X} = \int x \, \mathrm{d}\hat{E}(x) \,. \tag{1}$$

If the operator $\hat{X}$ is non degenerate, the spectral measure $\mathrm{d}\hat{E}(x)$ is simply the projector on the eigenvector $|x\rangle$ of $\hat{X}$, namely

$$\mathrm{d}\hat{E}(x) = \mathrm{d}x|x\rangle\langle x| \,, \qquad \hat{X}|x\rangle = x|x\rangle \,, \qquad \langle x|x'\rangle = \delta(x - x') \,. \tag{2}$$

Eqs. (1) and (2) supply the physical observable with the minimal mathematical outfit that is needed for stating the basic rule of quantum mechanics—the Born's rule—which at the same time provides the probabilistic interpretation of physical "state". The Born's rule can be enunciated as follows: "If we know in advance that the system is in a (pure) state described by the vector $|\psi\rangle \in \mathcal{H}_S$, we can predict *a priori* the probability $\mathrm{d}P(x)$ that the experimental reading will fall in the range $[x, x + \mathrm{d}x)$ by means of the formula

$$\mathrm{d}P(x) = \langle\psi|\mathrm{d}\hat{E}(x)|\psi\rangle \,. \tag{3}$$

The *Born's statistical formula* (3) can be further generalized in two ways: i) considering a prior undetermined "mixed" state described by a density operator $\hat{\rho}$; ii) embracing also the description of joint measurements of compatible observables. Compatible observables correspond to commuting operators $\hat{X}_i$ $(i = 1, \ldots, n)$ that share an orthogonal spectral decomposition $\mathrm{d}\hat{E}(\mathbf{x}) \equiv |\mathbf{x}\rangle\langle\mathbf{x}|\mathrm{d}\mathbf{x}$ as follows

$$\hat{X}_i = \int x_i \, \mathrm{d}\hat{E}(\mathbf{x}) \,, \tag{4}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ denotes the vector of simultaneous eigenvalues x_i of $\hat{X}_i$ with common eigenvector $|\mathbf{x}\rangle$. Including both generalizations, the statistical formula (3) now

T. Hakioğlu and A. S. Shumovsky (eds.), Quantum Optics and the Spectroscopy of Solids, 139–174.

reads

$$dP(\mathbf{x}) = \langle \mathbf{x}|\hat{\rho}|\mathbf{x}\rangle \equiv \mathrm{Tr}[\hat{\rho}\, d\hat{E}(\mathbf{x})] \,. \tag{5}$$

The Born's rule is very basic: it provides only the interpretation of "observables" and "states" in quantum mechanics. It assumes that one knows in advance what a measuring instrument is and which observables are measured. However, despite any experimental evidence, this assumption cannot be granted from start, because the measuring instrument is a special physical system and, as such, it is itself submitted to the laws of quantum mechanics. In most cases the measuring apparatus is a very complicated system, and some interpretation is already needed to understand what it is and how it works. The Born's rule makes no attempt to provide answers to more "operational" issues as: 1) Given a physical parameter—on the basis of its classical definition, or of the procedure for measuring it[22]—which selfadjoint operator describes the measurement? 2) How to describe the measurement of a physical quantity that apparently does not match any selfadjoint operator? [this is the case of the phase of the electromagnetic field]; 3) How to describe joint measurements of non compatible observables? 4) How to describe instrumental precision/resolution? 5) In which way the state of the system changes after the measurement?

The above issues urge a further generalization of the Born's rule (5) in a way that can be easily recognized at the mathematical level. If a quantum mechanical instrument is to provide information about a physical system S, the probability $dP(\mathbf{x})$ must be governed only by the state of the system, which is represented by a density operator $\hat{\rho}$. However complex the system-apparatus interaction is, quantum mechanics must provide a prevision of the result of the measurement in terms of operators acting on the Hilbert space $\mathcal{H}_S$ of the system S only. Depending on the measurement result $\mathbf{x}$, an operator $d\hat{\Pi}(\mathbf{x})$ will furnish the required probability $dP(\mathbf{x})$ through a rule of the general form

$$dP(\mathbf{x}) = \mathrm{Tr}[\hat{\rho}\, d\hat{\Pi}(\mathbf{x})] \,. \tag{6}$$

In order to have $dP(\mathbf{x})$ as a genuine probability, the operators $d\hat{\Pi}(\mathbf{x})$ in Eq. (6) must be nonnegative (hence selfadjoint)

$$d\hat{\Pi}(\mathbf{x}) \geq 0 \,, \tag{7}$$

as a consequence of positivity of density operators $\hat{\rho}$. Normalization of probability $dP(\mathbf{x})$ is guaranteed by the completeness relation

$$\int d\hat{\Pi}(\mathbf{x}) = \hat{1} \,. \tag{8}$$

The trace rule (6) is intimately connected with the probabilistic interpretation of physical states and their description in terms of density operators. In fact, the linear functional "Tr" guarantees propagation of convex linear combinations from density operators toward probabilities. In mathematical terms, the set of operators $d\hat{\Pi}(\mathbf{x})$ form a mapping that is a *positive operator measure*—more precisely, a *probability-operator measure* (POM)—on $\mathbf{R}^n$. Generally speaking, if $\Delta, \Delta_i \subset \mathbf{R}^n$ denote possible experimental "events", the following map

$$\hat{\Pi}(\Delta) = \int_{\Delta} d\hat{\Pi}(\mathbf{x}) \tag{9}$$

satisfies the abstract axioms of POM

$$\hat{\Pi}(\emptyset) = 0\,, \qquad \hat{\Pi}(\Delta) \geq 0\,, \qquad \hat{\Pi}(\mathbf{R}^n) = \hat{1}\,,$$
$$\hat{\Pi}(\cup_i \Delta_i) = \sum_i \hat{\Pi}(\Delta_i) \qquad \text{for } \cap_i \Delta_i = \emptyset\,. \tag{10}$$

Axioms (10) can be stated more generally for a probability space Ω in place of $\mathbf{R}^n$, with Ω playing the role of the spectrum of a set of commuting selfadjoint operators in the old Born's rule (5). For simplicity of notation in the following I will consider the case $\Omega \equiv \mathbf{R}^n$, whereas the integration set will be not explicitly written when integrals are extended to the whole space, as in Eq. (8); it is implicit that integrals must be replaced by sums whenever Ω is a discrete set.

Eq. (5) is only a particular case of Eq. (6) with $d\hat{\Pi}(\mathbf{x}) \equiv d\hat{E}(\mathbf{x})$ *orthogonal* POM. In the following we will be interested mostly in *nonorthogonal* POM's $d\hat{\Pi}(\mathbf{x})d\hat{\Pi}(\mathbf{x}') \neq 0$ for $\mathbf{x} \neq \mathbf{x}'$. How nonorthogonal POM's enters the quantum description of a physical system? As we will see shortly, this happens when a part P of the apparatus—so called "probe"or "ancilla"—itself enters the quantum description of the measurement by its own Hilbert space $\mathcal{H}_P$. Then, if one considers the customary measurement of commuting observables $\hat{X}_i$ now acting on the extended Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_P$ in the uncorrelated joint state $\hat{\rho}_S \otimes \hat{\rho}_P$, the Born's rule reads

$$dP(\mathbf{x}) = \mathrm{Tr}_{S+P}[\hat{\rho}_S \otimes \hat{\rho}_P |\mathbf{x}\rangle\langle\mathbf{x}|]d\mathbf{x}\,. \tag{11}$$

The trace in (11) can be evaluated in two successive steps as follows

$$dP(\mathbf{x}) = \mathrm{Tr}_S\{\hat{\rho}_S \mathrm{Tr}_P[\hat{\rho}_P |\mathbf{x}\rangle\langle\mathbf{x}|]\}d\mathbf{x}\,. \tag{12}$$

From the point of view of an observer who ignores (deliberately or not) the apparatus P, the Born's rule (12) has to involve operators on the Hilbert space $\mathcal{H}_S$ of the system only, and hence it is written in the form (6) as follows

$$dP(\mathbf{x}) = \mathrm{Tr}_S[\hat{\rho}_S d\hat{\Pi}(\mathbf{x})]\,. \tag{13}$$

Comparing Eq. (12) with Eq. (11) leads to the following POM

$$d\hat{\Pi}(\mathbf{x}) = \mathrm{Tr}_P[\hat{\rho}_P |\mathbf{x}\rangle\langle\mathbf{x}|]\,. \tag{14}$$

The operator $d\hat{\Pi}(\mathbf{x})$ in Eq. (14) is the partial trace Tr_P over $\mathcal{H}_P$ of an operator acting on $\mathcal{H}_S \otimes \mathcal{H}_P$, an hence it is an operator acting on $\mathcal{H}_S$ only. By definition, $d\hat{\Pi}(\mathbf{x})$ in Eq. (14) is positive and normalized to identity, i.e. it satisfies the axioms of a POM. It is also clear that for a given probe state $\hat{\rho}_P$ the POM $d\hat{\Pi}(\mathbf{x})$ in Eq. (14) is generally not orthogonal.

This is the way in which nonorthogonal POM's enter the quantum mechanical description of a measurement, namely through the measuring apparatus. The POM $d\hat{\Pi}(\mathbf{x})$ depends on the considered experimental setup: for a fixed state $\hat{\rho}_S$ of the system one can have different probability distributions $dP(\mathbf{x})$ by changing the detector and/or on the detector preparation $\hat{\rho}_P$. As we will see in the following, the correspondence between detectors and POM's is not one-to-one, namely there are many detectors described by the same POM. The notion of POM provides a new concept of physical observable that is more "operational" than the original one, because it is based on the definition of the procedure for performing the measurement. We will examine POM's further in the following sections.

We are now in position to understand what is the meaning of the title of these lectures: "Quantum estimation theory" [1]. Quantum estimation theory analyzes POM's at a purely abstract level. With the purpose of seeking the best strategy for estimating one or more parameters of the system in a fixed state, the theory looks for the pertaining class of POM's, and then seeks the POM that is optimal according to some pre-chosen goodness criterion [for example: maximum likelihood, minimum r.m.s. noise, etc.] In this way the theory allows one to find the best or "ideal" detector for such measurement. Quantum optics is an ideal lab for testing the theory of quantum measurements: in these lectures we will examine some examples of application in this field. In the tool-box of quantum optics we can find simple yet concrete devices for measuring a variety of observables of the electromagnetic field: the homodyne detector, which measures any linear combination of a couple of canonically conjugated observables of the field—the socalled *quadratures*; the heterodyne detector, which jointly measures a couple of conjugated (hence non compatible) quadratures; finally, high-sensitive interferometry, which poses the problem of measuring the phase of the field, a quantity with well defined physical meaning, however with no corresponding selfadjoint operator.

After giving an introductory classification of different types of POM's, in Sect. 2, we will analyze the Naimark's theorem, which assures that every POM can be obtained from conventional observables that involve the measuring apparatus itself. Applications to quantum optics are analyzed in details in Sect. 3, with special emphasis on the heterodyne detector, which achieves the joint measurement of non-commuting observables. Joint measurements are then analyzed in Sect. 4, where general measurements in the phase space are studied, including the measurement of the phase of the field. Quantum estimation theory is reviewed in Sect. 5: here, as a relevant application, a long subsection is devoted to the method for finding the ideal measurement of the phase. The last section 6 analyzes the notion of "instrument", which is more powerful that the concept of POM, as it describes also the back-action on the system after the measurement: the so called "state reduction". Here we will analyze in detail the general scheme for indirect measurements, with two examples—the von Neumann and the Arthurs-Kelly measurements.

2. Probability operator measures (POM)

In this section I briefly analyze the different classes of POM's. The following classification is only for didactic reasons, with the purpose of understanding step by step the new conceptual issues in the generalization from orthogonal spectral resolutions to noncommuting POM's.

2.1. ORTHOGONAL POM'S

Within our general framework, conventional measurements correspond to orthogonal projection-valued measures

$$d\hat{\Pi}(x)d\hat{\Pi}(y) = d\hat{\Pi}(x)\delta(x-y)dy\,, \tag{15}$$

with $\int d\hat{\Pi}(x) = \hat{1}$. Ideal perfectly *resolved* measurements are non degenerate, namely $d\hat{\Pi}(x) = dx|x\rangle\langle x|$ is a one-dimensional projector on the Hilbert space. On the other hand, non-ideal *unresolved* measurements carry some degeneration: this is the case, for example, of a set of orthogonal operators $\hat{\Pi}_n \equiv \hat{\Pi}(\Delta_n)$ defined as in Eq. (9), with the subsets Δ_n exhaustive and disjoint.

Example
An obvious example of orthogonal POM is the spectral resolution of a selfadjoint operator $\hat{X}$ with spectrum $\mathbf{R}$. One has

$$\mathrm{d}\hat{\Pi}(x) = \delta(x - \hat{X})dx \equiv |x\rangle\langle x|\mathrm{d}x \ . \tag{16}$$

As a concrete example, in Sect. 3.2 we will analyze the homodyne detector, with $\hat{X}$ representing a quadrature of the electromagnetic field.

2.2. COMMUTING POM'S

A trivial generalization to nonorthogonal POM's is the case of commuting measures, namely

$$[\mathrm{d}\hat{\Pi}(\mathbf{x}), \mathrm{d}\hat{\Pi}(\mathbf{x}')] = 0 \ . \tag{17}$$

As the operators $\mathrm{d}\hat{\Pi}(\mathbf{x})$ are selfadjoint and commute at different $\mathbf{x}$, an orthogonal basis $|\mathbf{z}\rangle\langle\mathbf{z}|$ exists that diagonalizes all of them simultaneously for all $\mathbf{x}$ [here I distinguish between different sets of states only by changing their label, as, for example, $|\mathbf{x}\rangle$ and $|\mathbf{z}\rangle$]. Hence, $\mathrm{d}\hat{\Pi}(\mathbf{x})$ must be of the form

$$\mathrm{d}\hat{\Pi}(\mathbf{x}) = \int \mathrm{d}\mathbf{z}\, m(\mathbf{x}|\mathbf{z})|\mathbf{z}\rangle\langle\mathbf{z}| \ . \tag{18}$$

From the POM axioms it follows that the coefficients $m(\mathbf{x}|\mathbf{z})$ in Eq. (18) are probability densities, namely

$$m(\mathbf{x}|\mathbf{z}) \geq 0 \ , \qquad \int \mathrm{d}\mathbf{x}\, m(\mathbf{x}|\mathbf{z}) = 1 \ . \tag{19}$$

More precisely, $m(\mathbf{x}|\mathbf{z})$ can be interpreted as the conditional probability density of getting outcome $\mathbf{x}$ given that the system is known to be exactly in the state $|\mathbf{z}\rangle\langle\mathbf{z}|$. Therefore, the present kind of measurement describes again a conventional measurement, however with an additional *imprecision* (or extrinsic noise) that makes the outcome $\mathbf{x}$ unpredictable even when it is known a priori that the system is exactly in an eigenstate $|\mathbf{z}\rangle$ of the measured observable.

Example
As an example, consider the following function of the operator $\hat{X}$

$$\mathrm{d}\hat{\Pi}(x) = \frac{1}{\sqrt{2\pi\Delta^2}} \exp\left\{-\frac{(x - \hat{X})^2}{2\Delta^2}\right\} \mathrm{d}x \ , \tag{20}$$

It is easy to see that $\mathrm{d}\hat{\Pi}(x)$ in Eq. (20) is the Gaussian convolution of the orthogonal projector $|x\rangle\langle x|$. A concrete example will be given in Sect. 3.2, where we will analyze the homodyne detector with nonunit quantum efficiency.

2.3. NONCOMMUTING POM'S

The truly nontrivial generalization of the projector spectral decomposition $\mathrm{d}\hat{E}(x)$ is the case of a non commuting POM $\mathrm{d}\hat{\Pi}(\mathrm{x})$, namely

$$[\mathrm{d}\hat{\Pi}(\mathrm{x}), \mathrm{d}\hat{\Pi}(\mathrm{x}')] \neq 0 \ . \tag{21}$$

Here, there is no longer an orthogonal basis that diagonalizes all operators $\mathrm{d}\hat{\Pi}(\mathrm{x})$ simultaneously: hence, no interpretation is possible in terms of compatible observables, nor the noise can be considered as an additional instrumental imprecision that is added to an ideally sharp measurement. Due to nonorthogonality, the resulting probability distribution is always unsharp for any state $\hat{\rho}_S$ of the system. Therefore, the only possible interpretation of the noise from such measurement is as an "intrinsic unavoidable quantum-mechanical imprecision". As we will see soon, this is the case of the noise arising when one jointly measures two noncommuting observables, or, more generally, when the measuring procedure involves a joint measurement, as in the case of the phase of the electromagnetic field. However, there is no "canonical" measurement of noncommuting observables that corresponds to a given POM, and for this reason the noncommuting POM is usually referred to as *generalized observable.*

It is obvious that, similarly to the classification of commuting POM's, also in the case of noncommuting POM's one could distinguish between: *i)* ideal *resolved* measurements—when $\mathrm{d}\hat{\Pi}(\mathrm{x})$ is a 1-dim projector, now ranging over a nonorthogonal (overcomplete) set; *ii) unresolved* measurements—when the POM is degenerate; *iii)* measurements with instrumental imprecision, when the POM is convolved with a conditional probability density. However, in the present case, this classification is mostly academic.

At this point one could notice that POM's provide also new selfadjoint operators available to the theory. Hence, why we do not use them? For example, the following operator is manifestly selfadjoint

$$\hat{X} = \int x \, \mathrm{d}\hat{\Pi}(x) \ . \tag{22}$$

Hence, $\hat{X}$ admits also an orthogonal spectral resolution in terms of eigen-vectors $|x\rangle\langle x|$. The operator $\hat{X}$, however, does not describe the same measurement of $\mathrm{d}\hat{\Pi}(x)$, apart from giving the correct expectation value $\int x \, \mathrm{d}P(x) \equiv \mathrm{Tr}(\hat{\rho}\hat{X})$—and, in fact, the corresponding probability distribution $\langle x|\hat{\rho}|x\rangle$ is different from the experimental one $\mathrm{d}P(x) \doteq \mathrm{Tr}[\hat{\rho}\,\mathrm{d}\hat{\Pi}(x)]$. Differences are evident already from the second moment, where one has

$$\hat{X}^2 \doteq \left(\int x \, \mathrm{d}\hat{\Pi}(x)\right)^2 \neq \int x^2 \mathrm{d}\hat{\Pi}(x) \ , \tag{23}$$

as a consequence of non orthogonality of $\mathrm{d}\hat{\Pi}(x)$. From the following generalized Schwartz inequality

$$\int x^2 \mathrm{d}\hat{\Pi}(x) \geq \left(\int x \mathrm{d}\hat{\Pi}(x)\right)^2 \ , \tag{24}$$

one has

$$\overline{\Delta x^2} \geq \langle \Delta \hat{X}^2 \rangle \ , \tag{25}$$

where the over-bar denotes the experimental average $\overline{f} \doteq \int f(\mathbf{x})\mathrm{d}P(\mathbf{x})$, whereas brackets denote the ensemble average $\langle\hat{f}\rangle = \mathrm{Tr}[\hat{f}\hat{\rho}]$. It follows that the "true" variance $\overline{\Delta x^2}$ is larger than the one resulting from the selfadjoint operator $\hat{X}$, despite $\hat{X}$ provides the correct average $\overline{x} \equiv \langle\hat{X}\rangle$ for all states $\hat{\rho}$. In other words, nonorthogonal POM's introduce an additional noise that arises from violations of the operator function calculus, namely

$$\hat{f} = \int f(x)\,\mathrm{d}\hat{\Pi}(x) \neq f(\hat{X}) , \qquad \text{where } \hat{X} = \int x\,\mathrm{d}\hat{\Pi}(x) . \tag{26}$$

According to Eq. (26), in order to describe the experimental probability distribution $\mathrm{d}P(\mathbf{x}) = \mathrm{Tr}[\hat{\rho}\mathrm{d}\hat{\Pi}(\mathbf{x})]$ for any $\hat{\rho}$, one would need the infinite set of selfadjoint operators

$$\widehat{X^n} = \int x^n \mathrm{d}\hat{\Pi}(x) \neq \hat{X}^n . \tag{27}$$

But, why do we use the POM $\mathrm{d}\hat{\Pi}(x)$ instead of the selfadjoint operator $\hat{X}$, considering that the latter can have sharp probabilities? Simply because $\hat{X}$ does not describe the measurement of the physical parameter that we meant. Moreover, notice that the selfadjoint operators defined in Eq. (22) do not solve the problem of joint measurements, because they generally do not commute, namely

$$\hat{X}_i \doteq \int x_i \mathrm{d}\hat{\Pi}(\mathbf{x}) , \quad [\mathrm{d}\hat{\Pi}(\mathbf{x}), \mathrm{d}\hat{\Pi}(\mathbf{x}')] \neq 0 \Longrightarrow [\hat{X}_i, \hat{X}_j] \neq 0 . \tag{28}$$

Example

In the following a and $a^\dagger$ will represent the usual annihilation and creation operators of a selected mode of the electromagnetic field, with commutation relation $[a, a^\dagger] = 1$, and with vacuum vector $|0\rangle$, i.e. $a|0\rangle = 0$. It is convenient to adopt the complex notation $f = f(z, \overline{z})$ to denote generic functions of $z \in \mathbf{C}$ (z and $\overline{z}$ are treated as independent variables).

Consider the following POM

$$\mathrm{d}\hat{\Pi}(z, \overline{z}) = \frac{\mathrm{d}^2 z}{\pi}|z\rangle\langle z| , \quad z \in \mathbf{C} . \tag{29}$$

In Eq. (29) $|z\rangle$ denotes the customary coherent state

$$|z\rangle \doteq \hat{D}(z)|0\rangle \doteq \exp(za^\dagger - \overline{z}a)|0\rangle , \tag{30}$$

which is obtained by displacing the vacuum $|0\rangle$ by the operator $\hat{D}(z)$. It is obvious that $\mathrm{d}\hat{\Pi}(z, \overline{z})$ is not commutative, just because $\langle z|z'\rangle \neq \langle z'|z\rangle$. In Sect. 3.3 we will see that the POM (29) describes the ideal heterodyne detector, which provides the optimal joint measurement of a couple of conjugated quadratures—the optical equivalent of position and momentum of an harmonic oscillator. For the moment, just notice the following identities

$$\hat{X} = \int \mathrm{d}\hat{\Pi}(z, \overline{z})\,\mathrm{Re}z = \frac{1}{2}\left(a + a^\dagger\right) , \qquad \hat{Y} = \int \mathrm{d}\hat{\Pi}(z, \overline{z})\,\mathrm{Im}z = \frac{i}{2}\left(a^\dagger - a\right) , \tag{31}$$

with $\hat{X}$ and $\hat{Y}$ having commutation $[\hat{X}, \hat{Y}] = i/2$.

2.4. NAIMARK'S THEOREM

In the introductory section we have already seen how nonorthogonal/noncommuting POM's arise in a measurement description that involves the apparatus. In this case, the POM just plays the role of the customary projector in the Born's rule used by an observer who ignores the apparatus. Let us recall in formulas this Born's rule

$$\mathrm{d}P(\mathbf{x}) = \mathrm{Tr}_{S+P}[\hat{\rho}_S \otimes \hat{\rho}_P |\mathbf{x}\rangle\langle\mathbf{x}|] = \mathrm{Tr}_S[\hat{\rho}_S \mathrm{d}\hat{\Pi}(\mathbf{x})] \,, \tag{32}$$

with $\hat{\rho}_S$ and $\hat{\rho}_P$ denoting the states of the system S and the probe P respectively, and $\hat{X}_i$ denoting the observables that are measured, with $\hat{X}_i = \int x_i |\mathbf{x}\rangle\langle\mathbf{x}|\mathrm{d}\mathbf{x}$ acting on $\mathcal{H}_S \otimes \mathcal{H}_P$. The POM is given by

$$\mathrm{d}\hat{\Pi}(\mathbf{x}) = \mathrm{Tr}_P[\hat{1}_S \otimes \hat{\rho}_P |\mathbf{x}\rangle\langle\mathbf{x}|]\mathrm{d}\mathbf{x} \,, \tag{33}$$

namely, the POM is the partial trace over $\mathcal{H}_P$ of the probe preparation $\hat{\rho}_P$ with the projector of the $S+P$ observables. It is very remarkable that every POM can be always represented as in Eq. (33): this is the statement of the Naimark's theorem [2], namely "Given a POM $\mathrm{d}\hat{\Pi}(\mathbf{x})$ in the system Hilbert space $\mathcal{H}_S$, there is always an extension $\mathcal{H}_S \otimes \mathcal{H}_P$ of the Hilbert space, a pure state $|\psi_P\rangle$, and an orthogonal POM $|\mathbf{x}\rangle\langle\mathbf{x}|\mathrm{d}\mathbf{x}$, such that

$$\mathrm{d}\hat{\Pi}(\mathbf{x}) = \mathrm{Tr}_P[\hat{1}_S \otimes |\psi_P\rangle\langle\psi_P|\, |\mathbf{x}\rangle\langle\mathbf{x}|]\mathrm{d}\mathbf{x} \,. \tag{34}$$

As a consequence, using POM's in quantum mechanics is not in conflict with the dictum that "only observables can be measured", because every POM corresponds to a customary observable in a larger Hilbert space. But such an observable is not unique, and may have "unnatural" physical meaning, because it involves the measuring apparatus itself. For the proof of the Naimark theorem the reader is addressed to the original papers [2] or to Ref. [3] (a sketch of the proof is also reported in the Helstrom's book [1] and in the book of A. Peres [4]). In the following, I will illustrate the theorem on the basis of two examples (collected from the same Ref. [1]), which I think can be interesting for applications to quantum optics.

2.4.1. *Example 1: the quantum roulette wheel*

Consider the following (generally non commutative) POM

$$\hat{\Pi}_m = \sum_{i=1}^{M} \zeta_i \hat{E}_m^{(i)} \,, \qquad m = 1, \ldots, n \,, \tag{35}$$

where

$$\zeta_i \geq 0 \,, \quad \sum_{i=1}^{M} \zeta_i = 1 \,, \qquad \hat{E}_m^{(i)} \hat{E}_n^{(i)} = \delta_{mn} \hat{E}_m^{(i)} \,, \quad \sum_{m=1}^{n} \hat{E}_m^{(i)} = \hat{1}_S \,. \tag{36}$$

For fixed i the projectors $\hat{E}_m^{(i)}$ give an orthogonal resolution of the identity. The physical meaning of $\hat{\Pi}_m$ is clear: the POM (35) describes a measuring apparatus where one of M different observables is selected at random at every measurement step, with ζ_i as the probability of the i-th observable.

The Naimark's extension of the POM (35) can be obtained as follows. Consider an M-dimensional Hilbert space $\mathcal{H}_P$, with $\{|\omega_i\rangle\}_{i=1,\ldots,M}$ as an orthonormal basis spanning

$\mathcal{H}_P$. A set of orthogonal projectors $\hat{E}_m \hat{E}_n = \delta_{nm} \hat{E}_m$ in the extended Hilbert space is given by

$$\hat{E}_m = \sum_{i=1}^{M} \hat{E}_m^{(i)} \otimes |\omega_i\rangle\langle\omega_i| \,. \tag{37}$$

The Naimark extension of the POM (35) is given by the projectors in Eq. (37), with the following state preparation of the probe

$$|\psi_P\rangle = \sum_{i=1}^{M} \zeta_i^{1/2} |\omega_i\rangle \,. \tag{38}$$

In fact, one can immediately check that

$$\mathrm{Tr}_P[\hat{1} \otimes |\psi_P\rangle\langle\psi_P| \hat{E}_m] = \sum_{i=1}^{M} \zeta_i \hat{E}_m^{(i)} \equiv \hat{\Pi}_m \,. \tag{39}$$

In this example, the probe P plays the role of a random device corresponding to a sort of "quantum roulette" wheel.

2.4.2. Example 2: commuting POM's

Let us consider M elements $\hat{\Pi}_k$ of a finite commuting POM. Upon denoting by $\{|m\rangle\}$ their common orthonormal set of eigenvectors, one has

$$\hat{\Pi}_k = \sum_m \lambda_m^{(k)} |m\rangle\langle m| \,, \qquad k = 1, \ldots, M \,, \tag{40}$$

and, as already mentioned, $\lambda_m^{(k)} \geq 0$ are interpreted as conditional probabilities, with $\sum_{k=1}^{M} \lambda_m^{(k)} = 1$. As in the previous example, let us consider a probe in an M-dimensional Hilbert space $\mathcal{H}_P$ with $\{|\omega_k\rangle\}_{k=1,\ldots,M}$ denoting an orthonormal basis. The following linear combinations

$$|\Lambda_m\rangle = \sum_{k=1}^{M} [\lambda_m^{(k)}]^{1/2} |\omega_k\rangle \,, \tag{41}$$

can be written in terms of unitary transformations of a fixed probe vector $|\psi_P\rangle \equiv |\omega_1\rangle$ as follows

$$|\Lambda_m\rangle = \hat{U}_m |\psi_P\rangle \,, \tag{42}$$

where $\hat{U}_m$ are unitary operators. With the probe preparation $|\psi_P\rangle \equiv |\omega_1\rangle$, the Naimark extension of the POM (40) is given by the orthogonal projectors

$$\hat{E}_k = \sum_m |m\rangle\langle m| \otimes \hat{U}_m^\dagger |\omega_k\rangle\langle\omega_k| \hat{U}_m \,. \tag{43}$$

In fact, it is easy to check the following steps

$$\begin{aligned} \mathrm{Tr}_P[|\omega_1\rangle\langle\omega_1|\hat{E}_k] &= \sum_m |m\rangle\langle m| \mathrm{Tr}_P[|\omega_1\rangle\langle\omega_1|\hat{U}_m^\dagger|\omega_k\rangle\langle\omega_k|\hat{U}] \\ &= \sum_m |m\rangle\langle m| \mathrm{Tr}_P[\hat{U}_m|\omega_1\rangle\langle\omega_1|\hat{U}_m^\dagger|\omega_k\rangle\langle\omega_k|] = \sum_m |m\rangle\langle m| |\langle\Lambda_m|\omega_k\rangle|^2 \\ &= \sum_m |m\rangle\langle m| \lambda_m^{(k)} \equiv \hat{\Pi}_k \,. \end{aligned} \tag{44}$$

3. POM's in quantum optics

In this section I illustrate some applications of POM's to quantum optics, where we have detectors for measuring observables of the electromagnetic field, with spectrum either discrete—as for the number of photons—or continuous—as for the quadrature of the field. We will see that by homodyne detection we can measure any linear combination of a couple of canonically conjugated observables—the so called *quadratures* of the field: this is a fortunate situation, which does not occur in the quantum mechanics of massive particles, and that makes possible to detect even the state itself of the field [for this topic see my other set of lectures in this same book [5]]. A long subsection is devoted to the heterodyne detector, which jointly measures two conjugated quadratures of the field. Joint measurements will be analyzed in more detail in Sect. 4.

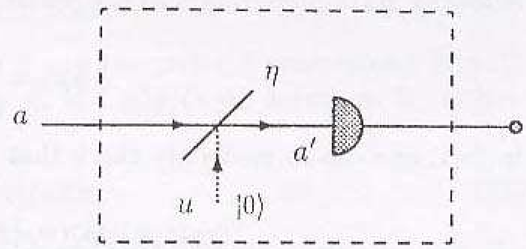

Figure 1. Equivalence of a nonideal ($\eta < 1$) detector with an ideal one preceded by a beam splitter of transmissivity η.

3.1. DIRECT DETECTION

The photon-count distribution for a photodetector (with a photo-tube small with respect to the coherence length of radiation) is given by the Mandel-Kelley-Kleiner formula [6, 7]

$$P_\eta(n) = \left\langle : \frac{(\eta a^\dagger a)^n}{n!} \exp(-\eta a^\dagger a) : \right\rangle , \tag{45}$$

where :: denotes normal ordering, and η is the overall quantum efficiency of the detector ($0 \leq \eta \leq 1$). For simplicity, I consider only monochromatic fields, with a denoting the annihilator of the nonvacuum mode: however, Eq. (45) can be written more generally in the wideband case, where instead of the operator $a^\dagger a$ one has the Pointing flux operator (with time-ordering and integration over the detector time). A simple derivation of Eq. (45) can be found in Ref. [8]. For $\eta = 1$ Eq. (45) gives the POM of ideal photon-number detection. In fact, from the identity

$$|0\rangle\langle 0| = \lim_{\epsilon \to 1} \sum_{l=0}^{\infty} \frac{(-\epsilon)^l}{l!} (a^\dagger)^l a^l \equiv \lim_{\epsilon \to 1} (1-\epsilon)^{a^\dagger a} . \tag{46}$$

and exploiting the recurrence

$$:(a^\dagger a)^n: = a^\dagger a (a^\dagger a - 1) \ldots (a^\dagger a - n + 1) , \tag{47}$$

one obtains

$$P_\eta(n) = \sum_{k=n}^{\infty} P(k) \binom{k}{n} \eta^n (1-\eta)^{k-n} , \tag{48}$$

where

$$P(n) \equiv P_1(n) = \langle\, |n\rangle\langle n|\, \rangle . \tag{49}$$

In Eq. (49) $|n\rangle$ denotes the photon-number eigenstate $a^\dagger a|n\rangle = n|n\rangle$. In other words, the POM for $\eta = 1$ is given by

$$\hat{\Pi}_1(n) = |n\rangle\langle n| , \tag{50}$$

whereas, more generally, for $\eta < 1$, using Eqs. (46) and (47) one can see that the probability distribution resulting from Eq. (45) is a Bernoulli convolution of the ideal probability (49), namely the detector POM is given by

$$\hat{\Pi}_\eta(n) = \sum_{k=n}^{\infty} \binom{k}{n} \eta^n (1-\eta)^{k-n} |k\rangle\langle k| . \tag{51}$$

Eq. (51) provides an example of a nonorthogonal commuting POM, with a form similar to Eq. (18) (here for a discrete spectrum) with conditional probability density $m(\mathrm{x}|\mathrm{z})$ given by a Bernoulli distribution.

Now I show that a detector with quantum efficiency $\eta < 1$ is equivalent to an ideal detector preceded by a beam splitter of transmissivity η. Such a "quantum-equivalence" between devices is schematically depicted in Fig. 1, and is relevant for detection theory in quantum optics. We have just to remind that, apart from trivial phase changes, a beam splitter of transmissivity η affects the unitary transformation of fields (see Fig. 2)

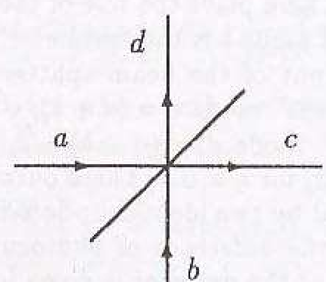

Figure 2. Field modes at a beam splitter.

$$\begin{pmatrix} c \\ d \end{pmatrix} = \hat{U}^\dagger \begin{pmatrix} a \\ b \end{pmatrix} \hat{U} = \begin{pmatrix} \eta^{1/2} & (1-\eta)^{1/2} \\ -(1-\eta)^{1/2} & \eta^{1/2} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} , \tag{52}$$

where all field modes are considered at the same frequency. Hence, the output mode a' in Fig. 1 is given by the linear combination

$$a' = \eta^{1/2} a + (1-\eta)^{1/2} u . \tag{53}$$

The mode a is entangled with the vacuum mode u, which here plays the role of a "probe" mode. The POM is obtained by partially tracing over the u mode as follows

$$\hat{\Pi}_\eta(n) = \mathrm{Tr}_u[\hat{1}_a \otimes |0\rangle\langle 0|_u |n\rangle\langle n|_{a'}] = {}_u\langle 0| : \exp(-a'^\dagger a')(-a'^\dagger a')^n/n! : |0\rangle_u . \tag{54}$$

Eq. (54) gives $\hat{\Pi}_\eta(n)$ in form of a normal-ordered expectation between coherent (vacuum) states. Here we can use the following identity valid for any function of a linear combination $Ka + Hb$ of two modes a and b

$$_b\langle\beta| : f\left(Ka + Hb, (Ka + Hb)^\dagger\right) : |\beta\rangle_b =: f(Ka + H\beta, \overline{K}a^\dagger + \overline{H\beta}) : , \tag{55}$$

with $|\beta\rangle_b$ denoting a coherent state for mode b only. Using Eq. (55) we immediately obtain

$$\hat{\Pi}_\eta(n) =: \exp(-\eta a^\dagger a)\frac{(\eta a^\dagger a)^n}{n!} : , \tag{56}$$

which gives the probability (45).

3.2. BALANCED HOMODYNE DETECTION

The scheme of a balanced homodyne detector is depicted in Fig. 3. The "signal" mode a is combined by means of a 50-50 beam splitter with a "local oscillator" (LO) mode b operating at the same frequency of a, and prepared in an "intense" coherent state $|z\rangle$. The signal mode a here plays the role of the "system" S, whereas mode b is the "probe" P. The field at the output of the beam splitter is described by a "sum" mode $c = (a + b)/\sqrt{2}$ and a "difference" mode $d = (a - b)/\sqrt{2}$, according to Eqs. (52) for $\eta = 0.5$. These output modes are detected by two identical photodetectors, and finally the difference of photocurrents (at zero frequency) is rescaled by $2|z|$. Thus, the output of the detector is given by the following operator

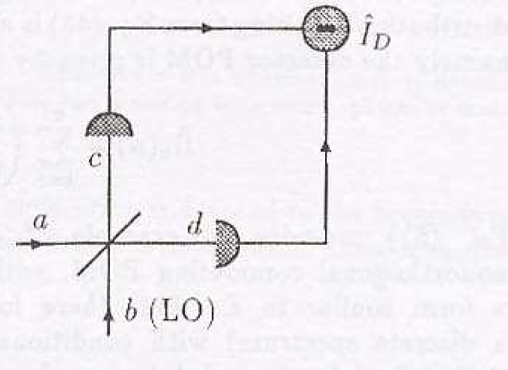

Figure 3. Scheme of the balanced homodyne detector.

$$\hat{I}_D = \frac{c^\dagger c - d^\dagger d}{2|z|} = \frac{a^\dagger b + b^\dagger a}{2|z|} . \tag{57}$$

Our intent is to evaluate the POM of the detector, or, in other words, to obtain the probability distribution of the output photocurrent $\hat{I}_D$ for any generic state $\hat{\rho}$ of the signal mode a. It is easier to evaluate the generating function of the moments of $\hat{I}_D$

$$\chi(\lambda) = \left\langle e^{i\lambda\hat{I}_D}\right\rangle_{ab} , \tag{58}$$

and then obtain the probability distribution of I_D as the Fourier transform of $\chi(\lambda)$, namely

$$\mathrm{d}P(I) = \mathrm{d}I \int_{-\infty}^{+\infty} \frac{d\lambda}{2\pi} e^{-i\lambda I}\left\langle e^{i\lambda\hat{I}_D}\right\rangle_{ab} . \tag{59}$$

Using the Backer-Campbell-Hausdorff (BCH) formula [9] for the $SU(2)$ group, namely

$$\exp\left(\xi a^\dagger b - \overline{\xi} b^\dagger a\right) = e^{\zeta b^\dagger a}\left(1 + |\zeta|^2\right)^{\frac{1}{2}\left(b^\dagger b - a^\dagger a\right)} e^{-\overline{\zeta} a^\dagger b} , \quad \zeta = \frac{\xi}{|\xi|}\tan|\xi| , \tag{60}$$

one can normal-order the exponential in Eq. (58) with respect to mode b as follows

$$\chi(\lambda) = \left\langle \exp\left[i\tan\left(\frac{\lambda}{2|z|}\right) b^\dagger a\right] \left[\cos\left(\frac{\lambda}{2|z|}\right)\right]^{a^\dagger a - b^\dagger b} \exp\left[i\tan\left(\frac{\lambda}{2|z|}\right) a^\dagger b\right]\right\rangle_{ab} . \tag{61}$$

The partial trace over b can be evaluated easily as follows

$$\begin{aligned}\chi(\lambda) &= \left\langle \exp\left[i\tan\left(\frac{\lambda}{2|z|}\right) \bar{z} a\right] \left[\cos\left(\frac{\lambda}{2|z|}\right)\right]^{a^\dagger a} \exp\left[i\tan\left(\frac{\lambda}{2|z|}\right) z a^\dagger\right]\right\rangle_a \\ &\times \left\langle z\left|\left[\cos\left(\frac{\lambda}{2|z|}\right)\right]^{-b^\dagger b}\right|z\right\rangle ,\end{aligned} \tag{62}$$

with the probe mode b is in the coherent state $|z\rangle$. Using now the customary BCH formula valid for $[\hat{A},[\hat{A},\hat{B}]] = [\hat{B},[\hat{A},\hat{B}]] = 0$

$$\exp\hat{A}\exp\hat{B} = \exp\left(\hat{A} + \hat{B} + \frac{1}{2}[\hat{A},\hat{B}]\right) , \tag{63}$$

one can recast Eq. (62) in normal-order with respect to a, namely [10]

$$\chi(\lambda) = \left\langle : \exp\left[iz\sin\left(\frac{\lambda}{2|z|}\right) a^\dagger\right] \exp\left[-2\sin^2\left(\frac{\lambda}{4|z|}\right)(a^\dagger a + |z|^2)\right] \exp\left[i\bar{z}\sin\left(\frac{\lambda}{2|z|}\right) a^\dagger\right] :\right\rangle_a . \tag{64}$$

Eq. (64) simplifies greatly in the strong-LO limit $z \to \infty$, where only the lowest order terms in $\lambda/|z|$ are retained, and $a^\dagger a$ is neglected with respect to $|z|^2$. One has

$$\lim_{z\to\infty} \chi(\lambda) = \left\langle : \exp\left[i\frac{\lambda}{2}e^{i\phi}a^\dagger\right] \exp\left[-\frac{\lambda^2}{8}\right] \exp\left[i\frac{\lambda}{2}e^{-i\phi}a^\dagger\right] :\right\rangle_a = \langle \exp[i\lambda\hat{a}_\phi]\rangle_a , \tag{65}$$

with $\phi = \arg z$ and $\hat{a}_\phi = \frac{1}{2}(a^\dagger e^{i\phi} + ae^{-i\phi})$ denoting the so called "quadrature" of the field mode a at phase ϕ with respect to the LO. The generating function in Eq. (65) is equivalent to the POM

$$d\hat{\Pi}(x) = dx \int_{-\infty}^{+\infty} \frac{d\lambda}{2\pi} e^{-i\lambda x} \exp[i\lambda\hat{a}_\phi] = dx\, \delta(x - \hat{a}_\phi) . \tag{66}$$

Hence, in conclusion, the balanced homodyne detector in the strong LO limit achieves the ideal measurement of the quadrature $\hat{a}_\phi$.

It is easy to take into account nonunit quantum efficiency at detectors. According to Eq. (53) one performs the substitutions

$$c \implies \eta^{1/2}c + (1-\eta)^{1/2}u , \qquad u, v \text{ vacuum modes} \tag{67}$$

$$d \implies \eta^{1/2}d + (1-\eta)^{1/2}v , \tag{68}$$

and now the output current is rescaled by $2|z|\eta$, thus obtaining

$$\hat{I}_D = \hat{a}_\phi + \sqrt{\frac{(1-\eta)}{2\eta}}(\hat{u}_\phi - \hat{v}_\phi) + \mathcal{O}(|z|^{-1}) , \tag{69}$$

with $\mathcal{O}(|z|^{-1})$ denoting terms vanishing as $|z|^{-1}$. Then, by tracing-out the vacuum modes u and v, one obtains

$$\begin{aligned} \mathrm{d}\hat{\Pi}_\eta(x) &= \mathrm{d}x \int \frac{\mathrm{d}\lambda}{2\pi} e^{-i\lambda x} e^{i\lambda\hat{a}_\phi} |\langle 0| e^{i\lambda\sqrt{\frac{1-\eta}{2\eta}}u_\phi}|0\rangle|^2 = \mathrm{d}x \int \frac{\mathrm{d}\lambda}{2\pi} e^{-i\lambda(x-\hat{a}_\phi)} e^{-\lambda^2\frac{1-\eta}{8\eta}} \\ &= \mathrm{d}x \frac{1}{\sqrt{2\pi\Delta_\eta^2}} \exp\left[-\frac{(x-\hat{a}_\phi)^2}{2\Delta_\eta^2}\right], \end{aligned} \tag{70}$$

where

$$\Delta_\eta^2 = \frac{1-\eta}{4\eta} . \tag{71}$$

Thus, in the nonideal case the POM is the convolution of the ideal POM with a Gaussian conditional probability: as in the case of photodetection, again nonunit quantum efficiency makes the POM nonorthogonal.

3.3. HETERODYNE DETECTOR

Heterodyne detection provides a method to perform joint measurements of two conjugated quadratures of the field [11]. The detector and the relevant field modes involved in the measurement are outlined in Fig. 4. The input field $\hat{E}_{in}$ impinges into a beam splitter with transmissivity θ, and has nonzero photon number only at frequency $\omega_0 + \omega_{IF}$. The LO works at a different frequency ω_0, and the output photocurrent $\hat{I}_{out}$ is measured at the intermediate frequency ω_{IF}. In the time-domain the measured photocurrent is given by

$$\hat{I}_{out}(t) = \hat{E}^-_{out}(t)\hat{E}^+_{out}(t) , \tag{72}$$

where $\hat{E}^\pm$ denote the usual positive and negative frequency components of the field, containing the annihilation and creation operators, respectively. The output photocurrent analyzed at frequency ω_{IF} is given by

$$\hat{I}_{out}(\omega_{IF}) = \int \mathrm{d}t\, \hat{I}_{out}(t) e^{i\omega_{IF}t} = \int \frac{\mathrm{d}\omega}{2\pi} \hat{E}^-_{out}(\omega+\omega_{IF})\hat{E}^+_{out}(\omega) . \tag{73}$$

The only field modes that are nonvacuum are the signal mode a_s at frequency $\omega + \omega_{IF}$ and the LO b_l at frequency ω_0. The integral in Eq. (73) involves modes at all frequencies: the terms that survive in the strong-LO limit are those linear in b_l or $b_l^\dagger$, namely $b_l^\dagger a_i$ and $b_l a_s^\dagger$, both having frequency difference equal to ω_{IF} (all other nonvacuum modes depicted in Fig. 4 do not involve the LO in Eq. (73)). The detector behaves ideally in the combined limits of strong-LO $z \to \infty$ and perfect transparency $\theta \to 1$, with $\gamma \doteq |z|\sqrt{\theta(1-\theta)}$ kept as constant. In fact, the output rescaled photocurrent $\hat{Z}$

$$\hat{Z} = \lim_{\substack{\theta \to 1, |z| \to \infty \\ \gamma = \mathrm{cost.}}} \gamma^{-1}\hat{I}_{out}(\omega_{IF}) , \tag{74}$$

is given by

$$\hat{Z} = |z|^{-1}(a_s^\dagger b_l + a_i b_l^\dagger) + \mathcal{O}(|z|^{-1}) , \tag{75}$$

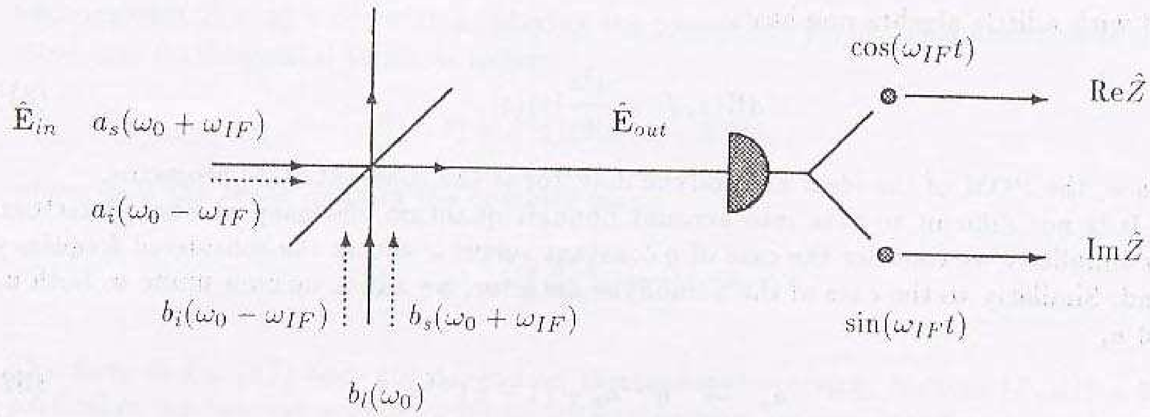

Figure 4. Scheme of the heterodyne detector and relevant field modes involved in the measurement. Dashed lines denote vacuum modes. Signal input modes are denoted by a, LO modes by b. The subindices s, l, and i refer to the frequency of modes: s is for signal band around $\omega_0+\omega_{IF}$, l is for LO band around ω_0, and i is for image band around $\omega_0-\omega_{IF}$. The output photocurrent is detected at the intermediate frequency ω_{IF}.

and after re-phasing the field modes $\hat{Z}$ becomes

$$\hat{Z} = a_s^\dagger + a_i \,. \tag{76}$$

The complex operator $\hat{Z}$ is equivalent to a couple of commuting selfadjoint operators

$$\hat{Z} = \hat{Z}_R + i\hat{Z}_I\,, \qquad [\hat{Z},\hat{Z}^\dagger] = [\hat{Z}_R,\hat{Z}_I] = 0\,, \tag{77}$$

and is described by a quantum mechanical probability density in the complex plane $p(z_R,z_I)\equiv p(z,\overline{z})$. The probability density is the Fourier transform of the generating function of the moments of $\hat{Z}$, and in complex notation is

$$p(z,\overline{z}) = \int \frac{d^2\lambda}{\pi^2}\langle e^{\lambda\hat{Z}^\dagger-\overline{\lambda}\hat{Z}}\rangle_{si}\, e^{\overline{\lambda}z-\lambda\overline{z}}\,, \tag{78}$$

where $\langle\ldots\rangle_{si}$ denotes the ensemble average on both modes a_s and a_i. In the present case the signal mode a_s represents the "system", whereas the image-band mode a_i is the "probe". The partial trace over the probe is carried out as follows

$$\begin{aligned}\left\langle e^{\lambda\hat{Z}^\dagger-\overline{\lambda}\hat{Z}}\right\rangle_{si} &= \mathrm{Tr}_s\left[\hat{\rho}_s\hat{D}_s(\lambda)\right]\langle 0|\hat{D}_i(-\lambda)|0\rangle_i = \mathrm{Tr}_s\left[\hat{\rho}_s\hat{D}_s(\lambda)\right]e^{-\frac{1}{2}|\lambda|^2}\\ &\equiv \mathrm{Tr}_s\left[\hat{\rho}_s : \hat{D}_s(\lambda) :_A\right]\,,\end{aligned} \tag{79}$$

where $\hat{D}(\alpha)=\exp(\alpha a^\dagger-\overline{\alpha}a)$ denotes the usual displacement operator, and sub-indices s,i pertain signal and image modes, respectively. Anti-normal ordering $::_A$ in Eq. (79) follows from the customary BCH formula (63) applied to the displacement operator. Comparing Eqs. (78) and (79) the POM of the detector is obtained in the form

$$d\hat{\Pi}(z,\overline{z}) = \int \frac{d^2\lambda}{\pi^2} e^{\overline{\lambda}z-\lambda\overline{z}} : \hat{D}_s(\lambda) :_A d^2z\,, \tag{80}$$

and with a little algebra one has[1]

$$d\hat{\Pi}(z,\bar{z}) = \frac{d^2 z}{\pi}|z\rangle\langle z| \,. \tag{81}$$

Hence, the POM of the ideal heterodyne detector is the coherent-state projector.

It is not difficult to take into account nonunit quantum efficiency at photodetectors. For simplicity we consider the case of η constant versus ω within the considered frequency band. Similarly to the case of the homodyne detector, we add a vacuum mode to both a_s and a_i

$$a_s \rightarrow \eta^{1/2}a_s + (1-\eta)^{1/2}u \,, \tag{82}$$
$$a_i \rightarrow \eta^{1/2}a_i + (1-\eta)^{1/2}v \,, \tag{83}$$

with u and v denoting the vacuum modes at frequencies $\omega_0+\omega_{IF}$ and $\omega_0-\omega_{IF}$ respectively. Upon rescaling the output photocurrent by an additional factor $\eta^{1/2}$, we obtain

$$\hat{Z} = a_s + \sqrt{\frac{1-\eta}{\eta}}u + a_i^\dagger + \sqrt{\frac{1-\eta}{\eta}}v^\dagger \,. \tag{84}$$

The two modes u and v enter the definition of the new enlarged "probe" of the detector, and must be traced out. In this way one obtains the POM

$$d\hat{\Pi}_\eta(\alpha,\bar{\alpha}) = d^2\alpha \int \frac{d^2\lambda}{\pi^2} e^{\alpha\bar{\lambda}-\bar{\alpha}\lambda} : \hat{D}_s(\lambda) :_A {}_{uv}\langle 0|\hat{D}_u(\lambda_\eta)\hat{D}_v(-\lambda_\eta)|0\rangle_{uv} \,, \tag{85}$$

where $\lambda_\eta = \sqrt{\frac{1-\eta}{\eta}}\lambda$. The POM (85) is the Gaussian convolution of the ideal POM

$$d\hat{\Pi}_\eta(\alpha,\bar{\alpha}) = d^2\alpha \int \frac{d^2\lambda}{\pi^2} e^{\alpha\bar{\lambda}-\bar{\alpha}\lambda} : \hat{D}_s(\lambda) :_A e^{-|\lambda_\eta|^2} = \int \frac{d^2 z}{\pi\Delta_\eta^2} e^{-\frac{|z-\alpha|^2}{\Delta_\eta^2}} d\hat{\Pi}(z,\bar{z}) \,, \tag{86}$$

where $\Delta_\eta^2 = \frac{1-\eta}{\eta}$.

Before continuing further, it is instructive to see an alternative derivation of the heterodyne POM (81). We have seen that the heterodyne detector measures the complex

[1] Here is the proof:

$$\int \frac{d^2\lambda}{\pi} : \hat{D}(\lambda) :_A = \int \frac{d^2\lambda}{\pi} e^{-|\lambda|^2} e^{\lambda a^\dagger} e^{-\bar{\lambda} a} = \int \frac{d|\lambda|^2}{2} e^{-|\lambda|^2} \sum_{n=0}^{\infty} \frac{(-)^n}{(n!)^2}|\lambda|^{2n}(a^\dagger)^n a^n = \sum_{n=0}^{\infty} \frac{(-)^n}{n!}(a^\dagger)^n a^n \,,$$

and using the identity (46) one has

$$\int \frac{d^2\lambda}{\pi} : \hat{D}(\lambda) :_A = |0\rangle\langle 0| \,.$$

Using the last equation we obtain

$$\int \frac{d^2\lambda}{\pi} e^{\alpha\bar{\lambda}-\bar{\alpha}\lambda} : \hat{D}(\lambda) :_A = \int \frac{d^2\lambda}{\pi} e^{-|\lambda|^2} e^{\lambda(a^\dagger-\bar{\alpha})} e^{-\bar{\lambda}(a-\alpha)} = \hat{D}^\dagger(-\alpha)|0\rangle\langle 0|\hat{D}(-\alpha) = |\alpha\rangle\langle\alpha| \,.$$

photocurrent $\hat{Z} = a_s + a_i^\dagger$, with a_i playing the role of the probe. This assertion is translated into mathematical terms as follow

$$d\hat{\Pi}(z,\bar{z}) = d^2z\,{}_i\langle 0|\delta_2(z-\hat{Z})|0\rangle_i\,, \tag{87}$$

where δ_2 is the Dirac delta in the complex plane

$$\delta_2(\alpha-\beta) = \int \frac{d^2\lambda}{\pi^2} e^{(\alpha-\beta)\bar{\lambda}-(\bar{\alpha}-\bar{\beta})\lambda}\,. \tag{88}$$

The form of Eq. (87) does not depend on the operator ordering, because $[\hat{Z},\hat{Z}^\dagger]=0$. In particular, we can use normal ordering with respect to a_i, corresponding to anti-normal ordering with respect to a_s. Then, the vacuum expectation is evaluated just upon setting a_i to zero, namely

$$d\hat{\Pi}(z,\bar{z}) = d^2z : \delta_2(z-a_s) :_A\ . \tag{89}$$

By definition, one has

$$: \delta_2(z-a_s) := \int \frac{d^2\lambda}{\pi^2} e^{(z-a_s)\bar{\lambda}-(\bar{z}-a_s^\dagger)\lambda} :_A = \int \frac{d^2\lambda}{\pi^2} : e^{z\bar{\lambda}-\bar{z}\lambda} : \hat{D}_s(\lambda) :_A\,, \tag{90}$$

namely Eq. (81).

4. Joint measurements

From the derivation of the POM of the heterodyne detector we can understand the basis of a joint measurement of two non commuting observables. The heterodyne detector performs a joint measurement of any couple of conjugated quadratures, say for example

$$\hat{X} = \frac{1}{2}(a_s + a_s^\dagger)\,, \qquad \hat{Y} = \frac{i}{2}(a_s^\dagger - a_s)\,, \tag{91}$$

with

$$[\hat{X},\hat{Y}] = \frac{i}{2}\,. \tag{92}$$

The method for jointly measuring $\hat{X}$ and $\hat{Y}$ consists of making a conventional measurement of two commuting currents $\hat{Z}_R$ and $\hat{Z}_I$ that have the same expectation values of $\hat{X}$ and $\hat{Y}$, namely

$$\langle\hat{X}\rangle = \langle\hat{Z}_R\rangle\,, \qquad \langle\hat{Y}\rangle = \langle\hat{Z}_I\rangle\,, \tag{93}$$

or, in complex notation

$$\langle a_s\rangle = \langle\hat{Z}\rangle\,. \tag{94}$$

Eq. (94) emphasizes the fact that measuring $\hat{X}$ and $\hat{Y}$ jointly is equivalent to "measuring the complex operator" a_s. Now we will see that the price to pay for jointly measuring non commuting observables is an additional noise. In fact, let us evaluate the r.m.s. fluctuations for $\hat{X}$, $\hat{Y}$, and for $\hat{Z}_R$, $\hat{Z}_I$, and then compare the respective results. One has

$$\langle\Delta\hat{Z}_R^2\rangle = \langle\Delta\hat{X}^2\rangle + \langle\Delta\hat{X}_i^2\rangle = \langle\Delta\hat{X}^2\rangle + \frac{1}{4}\,, \tag{95}$$

and, similarly

$$\langle\Delta\hat{Z}_I^2\rangle = \langle\Delta\hat{Y}^2\rangle + \langle\Delta\hat{Y}_i^2\rangle = \langle\Delta\hat{Y}^2\rangle + \frac{1}{4}\,, \tag{96}$$

where $\hat{X}_i$ and $\hat{Y}_i$ are the same quadratures as in Eq. (91), but for the image-band mode a_i. It follows that the experimental probability distribution of the photocurrent $\hat{Z}$ has the same average of the complex field a_s, but with an additional noise. From Eqs. (95) and (96) we deduce the "experimental" Heisenberg relation [12, 13]

$$\langle\Delta\hat{Z}_R^2\rangle\langle\Delta\hat{Z}_I^2\rangle \geq \frac{1}{4}\,, \tag{97}$$

which should be compared with the customary inequality

$$\langle\Delta\hat{X}^2\rangle\langle\Delta\hat{Y}^2\rangle \geq \frac{1}{4}|\langle[\hat{X},\hat{Y}]\rangle|^2 \geq \frac{1}{16}\,. \tag{98}$$

Notice that the usual Heisenberg relation pertains the intrinsic uncertainties of a couple of conjugated observables, and thus can be used only to analyze conventional measurements of one of the two observables at a time (the uncertainty of the other observable refers to a "preparation" before the measurement). On the other hand, the case of joint measurements is described by the new Heisenberg inequality (98): here, the "experimental" noise is double than the "theoretical" one, and their relative factor 2 is usually referred to as "the additional 3 decibels (3dB) noise due to the joint measure". Such noise is of quantum origin, and is unavoidable. This can be easily understood with the aid of the following argument. The 3dB noise originates from the vacuum fluctuations of the image-band mode, which is needed in order to have a commuting current $\hat{Z}$. For this purpose one needs to add $a_i^\dagger$—not a_i—to the signal annihilator a_s, and this produces the anti-normal ordering for the POM, which corresponds to the 3dB Gaussian convolution

$$\begin{aligned} \mathrm{d}\hat{\Pi}(z,\bar{z}) &= \mathrm{d}^2 z\, {}_i\langle 0|\delta_2(z-\hat{Z})|0\rangle_i = \mathrm{d}^2 z : \delta_2(z-a_s) :_A \\ &= \int \frac{\mathrm{d}^2\lambda}{\pi^2} e^{z\bar{\lambda}-\bar{z}\lambda} e^{-\frac{1}{2}|\lambda|^2}\hat{D}_s(\lambda)\,. \end{aligned} \tag{99}$$

In the following we will see that the 3dB additional noise is equivalent to measuring each quadrature with effective quantum efficiency $\eta = \frac{1}{2}$.

4.1. MARGINAL JOINT MEASUREMENTS

It is clear that once a method for measuring the complex field a is given, then any function of the field can be measured. Such measuring scheme resembles a "classical" measurement in the phase space, where one jointly measures the canonical pair and then evaluates functions of it. Operatively, the measurement works similarly to the classical case, namely, after detecting the complex photocurrent $\hat{Z}$ and obtaining the reading $z \in \mathbf{C}$, one evaluates the function $w = f(z, \bar{z})$. What is the POM that gives the probability distribution for w? The answer is simple: the probability density for w is just the marginal probability of $p(z, \bar{z})$, namely

$$p(w) = \int \mathrm{d}^2 z\, p(z, \bar{z})\, \delta(w - f(z, \bar{z}))\ . \tag{100}$$

Hence, the POM is the marginal POM of $\mathrm{d}\hat{\Pi}(z, \bar{z})$

$$\mathrm{d}\hat{\Pi}(w) = \mathrm{d}w \int \mathrm{d}\hat{\Pi}(z, \bar{z})\, \delta(w - f(z, \bar{z}))\ . \tag{101}$$

The Dirac delta function in Eq. (101) must be defined carefully on the complex plane, depending on the particular analytic form of the function f. Using Eq. (87) one obtains

$$\mathrm{d}\hat{\Pi}(w) = \mathrm{d}s\, \mathrm{d}w\ {}_i\langle 0|\delta(w - f(\hat{Z}, \hat{Z}^\dagger))|0\rangle_i = \mathrm{d}w : \delta(w - f(a, a^\dagger)) :_A\ . \tag{102}$$

In the following we will examine some relevant choices for the function f.

4.1.1. *Field quadrature*

Field quadrature corresponds to the function $f(z, \bar{z}) = \mathrm{Re}(z e^{-i\phi})$ of the field. In this way, from a joint measurement of any couple of conjugated quadratures, one obtains a marginal probability distribution for any desired single quadrature $\hat{a}_\phi$. In fact, from Eq. (102) one has

$$\begin{aligned} \mathrm{d}\hat{\Pi}(x) &= \mathrm{d}x : \delta(x - \hat{a}_\phi) :_A = \mathrm{d}x \int_{-\infty}^{+\infty} \frac{\mathrm{d}\lambda}{2\pi} e^{i\lambda x} e^{-\frac{1}{8}|\lambda|^2} e^{-i\lambda \hat{a}_\phi} \\ &= \mathrm{d}x\, \frac{2}{\sqrt{\pi}} \exp\left[-2(x - \hat{a}_\phi)\right]\ , \end{aligned} \tag{103}$$

which is a Gaussian with variance $\Delta^2 = \frac{1}{4}$, as expected from Eq. (95): this is just the 3dB noise due to the joint measure. Comparing Eq. (103) with Eq. (71) we immediately recognize that this noise corresponds to an effective quantum efficiency $\eta = \frac{1}{2}$.

4.1.2. *Field intensity*

This case corresponds to the function $f(z, \bar{z}) = |z|^2$. One has

$$\begin{aligned} \mathrm{d}\hat{\Pi}(w) &= \mathrm{d}w : \delta(w - a^\dagger a) :_A = \mathrm{d}w \int \frac{\mathrm{d}\lambda}{2\pi} e^{-i\lambda w} : e^{i\lambda a^\dagger a} :_A \\ &= \mathrm{d}w \int \frac{\mathrm{d}\lambda}{2\pi} e^{-i\lambda w} (1 - i\lambda)^{-a^\dagger a - 1} = \frac{w^{a^\dagger a}}{(a^\dagger a)!} e^{-w} \mathrm{d}w\ . \end{aligned} \tag{104}$$

The POM (104) is an unsharp version of the ideal POM (50). Notice that the function in Eq. (104) is not a Poisson, because here it is regarded as a function of w—not versus $a^\dagger a$.

4.1.3. *Phase of the field*

This case is particularly interesting, as in practice it is the only way to define a quantum mechanical measurement of the phase of the field, namely through the measurement of the polar angle of a complex photocurrent. It is instructive to analyze briefly the experimental procedure for obtaining the marginal phase distribution. This is illustrated in Fig. 5, where, as an example, a computer simulation of the experimental procedure is illustrated for a squeezed state. Each experimental event consists of a reading of the complex heterodyne photocurrent, which is represented by a point plotted in the complex plane of the field amplitude. The phase value inferred from the event is the polar angle of the point itself. The experimental histogram of the phase distributions is obtained upon dividing the plane into "infinitesimal" angular bins of equal width $\delta\phi$, from $-\pi$ to π, then counting the number of points which fall into each bin. In Fig. 5 the simulated experimental histogram (10^4 events) is compared with the theoretical probability as obtained from the marginal phase POM of the heterodyne detector. Formally, the marginal phase POM is given by

$$\mathrm{d}\hat{\Pi}(\phi) = \mathrm{d}\phi : \delta(\phi - \arg a) :_A , \tag{105}$$

where the meaning of the δ function is the marginal integral of the δ_2 distribution over the polar modulus on the complex plane, namely

$$\delta(\phi - \arg z) = \int_0^\infty \rho\, \mathrm{d}\rho\, \delta_2(\rho e^{i\phi} - z) = \int_0^\infty \mathrm{d}r \int_{\mathbf{C}} \frac{\mathrm{d}^2\lambda}{2\pi^2} \exp\left[\sqrt{r}(e^{i\phi} - z)\overline{\lambda} - c.c.\right] . \tag{106}$$

One has

$$\begin{aligned} \mathrm{d}\hat{\Pi}(\phi) &= \frac{\mathrm{d}\phi}{2\pi} \int_0^\infty \mathrm{d}r \int_{\mathbf{C}} \frac{\mathrm{d}^2\lambda}{\pi} e^{\sqrt{r}(e^{i\phi}\overline{\lambda} - e^{-i\phi}\lambda)} e^{-\overline{\lambda}a} e^{\lambda a^\dagger} \\ &= \frac{\mathrm{d}\phi}{2\pi} \int_{\mathbf{C}} \frac{\mathrm{d}^2\lambda}{\pi} e^{-\frac{1}{2}|\lambda|^2} \frac{\hat{D}(\lambda)}{[\mathrm{Im}(\lambda e^{-i\phi}) - i0^+]^2} . \end{aligned} \tag{107}$$

Equivalently, evaluating the marginal POM of the coherent-state projector, one has

$$\mathrm{d}\hat{\Pi}(\phi) = \frac{\mathrm{d}\phi}{2\pi} \int_0^\infty \mathrm{d}r |\sqrt{r}e^{i\phi}\rangle\langle\sqrt{r}e^{i\phi}| = \frac{\mathrm{d}\phi}{2\pi} \sum_{nm=0}^\infty e^{i(n-m)\phi} \frac{\Gamma[\frac{1}{2}(n+m)+1]}{\sqrt{n!m!}} |n\rangle\langle m| . \tag{108}$$

We will discuss this POM later and compare it with the ideal one coming from quantum estimation theory.

5. Quantum estimation theory

Quantum estimation theory analyzes POM's at a purely abstract level, with the purpose of seeking the best strategy for estimating one or more parameters of a quantum system. The theory looks for the general class of POM's that describe the specific measurement, then optimizes the POM according to some prefixed goodness criterion. In general, one can say that the problem resorts to seeking the best strategy for estimating a set of parameters $\theta = \{\theta_1, \theta_2, \ldots, \theta_m\}$ of the density operator $\hat{\rho}(\theta)$ of the system (for example the position and momentum of a particle, the amplitude of a field mode, etc). The observational strategy for estimating θ is expressed by a POM that pertains a generic apparatus along with its "data processing rule", i.e. the evaluation of a function of the experimental result. Let us denote by $\mathrm{d}\hat{\Pi}(\theta)$ the generic POM. Generally, the result of the measurement—i.e. the estimated values—are different from the "true" ones, and we will denote the true values by θ and the estimated values by θ_*. Then, the joint conditional probability density $p(\theta_*|\theta)$ of estimating θ_* for true values θ, is given by

$$p(\theta_*|\theta)\mathrm{d}^m\theta = \mathrm{Tr}[\hat{\rho}(\theta)\mathrm{d}\hat{\Pi}(\theta_*)] . \tag{109}$$

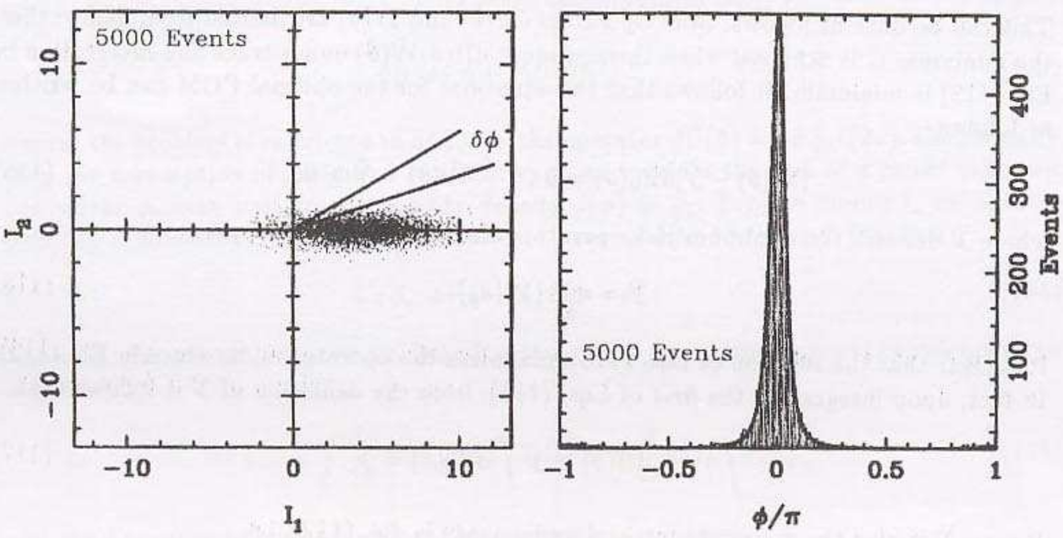

Figure 5. Computer simulation of a heterodyne phase detection experiment for a squeezed state with 4.53 squeezing photons and 20 photons in total. The histogram for 5000 events is compared with the theoretical result from the POM (108).

The goodness of the POM is considered on the basis of a cost function $C(\theta_*, \theta)$, which assesses the cost of errors in the estimates. Examples of cost functions are the delta-function cost

$$C(\theta_*, \theta) = -\prod_{k=1}^{m} \delta(\theta_{k*} - \theta_k) , \tag{110}$$

and the quadratic cost

$$C(\theta_*, \theta) = -\sum_{k=1}^{m} (\theta_{k*} - \theta_k)^2 . \tag{111}$$

More generally, in Eqs. (110) one could consider different weights for every component, or one can introduce a positive cost matrix in the quadratic case (111). We must provide also an *a priori* probability density $z(\theta)$ for the estimanda. Then, the average cost incurred when the strategy represented by $\mathrm{d}\hat{\Pi}(\theta)$ is employed, is given by

$$\overline{C} = \int \mathrm{d}^m\theta \, z(\theta) C(\theta_*, \theta) p(\theta_*|\theta) = \mathrm{Tr} \int \hat{\mathcal{W}}(\theta_*) \, \mathrm{d}\hat{\Pi}(\theta_*) , \tag{112}$$

where $\hat{\mathcal{W}}(\theta)$ denotes the selfadjoint risk operator

$$\hat{\mathcal{W}}(\theta_*) = \int \mathrm{d}^m\theta \, z(\theta) C(\theta_*, \theta) \, \hat{\rho}(\theta) . \tag{113}$$

We want now to select the POM that minimizes the average cost $\overline{C}$ under the constraints

$$\mathrm{d}\hat{\Pi}(\theta) \geq 0 , \qquad \int \mathrm{d}\hat{\Pi}(\theta) = \hat{1} . \tag{114}$$

This can be done as follows. Both operators $d\hat{\Pi}(\theta)$ and $\hat{\mathcal{W}}(\theta)$ are limited from below: then the minimum $\overline{C}$ is achieved when their product $d\hat{\Pi}(\theta)\hat{\mathcal{W}}(\theta)$ under trace and integration in Eq. (112) is minimum. It follows that the equations for the optimal POM can be written as follows

$$[\hat{\mathcal{W}}(\theta) - \hat{\mathcal{Y}}]d\hat{\Pi}_0(\theta) = 0 \ , \qquad \hat{\mathcal{W}}(\theta) - \hat{\mathcal{Y}} \geq 0 \ , \tag{115}$$

where $\hat{\mathcal{Y}}$ denotes the minimum risk operator (also called Lagrange operator)

$$\hat{\mathcal{Y}} = \min_{\theta}\{\hat{\mathcal{W}}(\theta)\} \ . \tag{116}$$

It is clear that the solution of Eqs. (115) minimizes the operator under trace in Eq. (112). In fact, upon integrating the first of Eqs. (115), from the definition of $\hat{\mathcal{Y}}$ it follows that

$$\int \hat{\mathcal{W}}(\theta)d\hat{\Pi}(\theta) \geq \hat{\mathcal{Y}} \int d\hat{\Pi}_0(\theta) = \hat{\mathcal{Y}} \ . \tag{117}$$

Hence, $\hat{\mathcal{Y}}$ is also the minimum integral under trace in Eq. (112), i.e.

$$\hat{\mathcal{Y}} = \int \hat{\mathcal{W}}(\theta)d\hat{\Pi}_0(\theta) \ , \qquad \overline{C}_{min} = \mathrm{Tr}\hat{\mathcal{Y}} \ . \tag{118}$$

Notice that the Lagrange operator is selfadjoint by definition, and hence

$$\hat{\mathcal{Y}} = \int \hat{\mathcal{W}}(\theta)d\hat{\Pi}_0(\theta) = \int d\hat{\Pi}_0(\theta)\hat{\mathcal{W}}(\theta) \ , \tag{119}$$

namely the optimal POM also satisfies the hermitian conjugated of equation (115).

In general, solving Eqs. (115) and (118) is a difficult task. In the following subsection we will analyze in some detail a relevant example: the canonical measurement of the phase.

5.1. CANONICAL MEASUREMENT OF THE PHASE

The estimation problem is the following: to estimate the phase-shift ϕ of a fixed density matrix $\hat{\rho}_0$ undergoing the unitary transformation

$$\hat{\rho}(\phi) = e^{-ia^\dagger a\phi}\hat{\rho}_0 e^{ia^\dagger a\phi} \ . \tag{120}$$

First, we observe that ϕ is defined on a circle (a 2π-window), because $a^\dagger a$ is an integer operator. Then, we notice that the family of states $\{\hat{\rho}(\phi)\}$ is "covariant", namely it is of the form

$$\hat{\rho}(\phi) = \hat{U}_\phi\hat{\rho}_0\hat{U}_\phi^\dagger \ , \tag{121}$$

where $\hat{U}_\phi$ are unitary operators representing a group—in the present case, the abelian group $U(1)$ of rotation along one axis. For a covariant estimation problem, the optimal POM must be itself covariant. This should be true at least if we want a likelihood $p(\phi|\phi)$ which is independent on ϕ [for a general study of covariant estimation problems see Ref. [14]] i.e.

$$p(\phi|\phi) = \mathrm{Tr}[\hat{U}_\phi\hat{\rho}_0\hat{U}_\phi^\dagger d\hat{\Pi}(\phi)] = \text{const.} \ , \tag{122}$$

which, due to invariance of trace under cyclic permutations leads to

$$\mathrm{d}\hat{\Pi}(\phi) = \hat{U}_\phi \mathrm{d}\hat{\Pi}(0)\hat{U}_\phi^\dagger \ . \tag{123}$$

Hence, the problem is restricted to find only the operator $\mathrm{d}\hat{\Pi}(0) \doteq \mathrm{d}\phi\,\hat{\xi}_0/(2\pi)$. Consistently with the assumption of likelihood constant vs ϕ, we consider the case of *a priori* unknown parameter ϕ, with uniform probability density $z(\phi) = \frac{1}{2\pi}$. For the moment, we address only the max-likelihood estimation problem, corresponding to the cost function

$$C(\phi_*, \phi) = -\delta_{2\pi}(\phi_* - \phi) \ , \tag{124}$$

where $\delta_{2\pi}$ denotes the 2π-periodic delta function. With the above choices, the risk operator is given by

$$\hat{\mathcal{W}}(\phi_*) = \int \frac{\mathrm{d}\phi}{2\pi} C(\phi, \phi_*)\, \hat{U}_\phi \hat{\rho}_0 \hat{U}_\phi^\dagger = -\frac{1}{2\pi}\hat{U}_{\phi_*}\hat{\rho}_0\hat{U}_{-\phi_*} \ , \tag{125}$$

and the Lagrange operator becomes diagonal with $a^\dagger a$, namely

$$\hat{\mathcal{Y}} \ = \ -\int \frac{\mathrm{d}\phi}{(2\pi)^2}\hat{U}_\phi \hat{\xi}_0 \hat{\rho}_0 \hat{U}_\phi^\dagger \ , \tag{126}$$

$$\langle k|\hat{\mathcal{Y}}|l\rangle \ = \ -\int \frac{\mathrm{d}\phi}{(2\pi)^2} e^{-i(k-l)\phi}\langle k|\hat{\xi}_0\hat{\rho}_0|l\rangle = -\delta_{kl}\frac{1}{2\pi}\langle k|\hat{\xi}_0\hat{\rho}_0|k\rangle \ . \tag{127}$$

Thanks to covariance, the estimation problem resorts to seeking the solution $\mathrm{d}\hat{\Pi}(0)$ of the following equations only

$$[\hat{\mathcal{W}}(0) - \hat{\mathcal{Y}}]\mathrm{d}\hat{\Pi}(0) = 0 \ , \qquad \hat{\mathcal{W}}(0) - \hat{\mathcal{Y}} \geq 0 \ . \tag{128}$$

Notice that from Eq. (125) one has $\hat{\mathcal{W}}(0) = -\frac{1}{2\pi}\hat{\rho}_0$: hence, Eqs. (128) can be written as follows

$$[\hat{\rho}_0 + 2\pi\hat{\mathcal{Y}}]\xi_0 = 0 \ , \qquad \hat{\rho}_0 + 2\pi\hat{\mathcal{Y}} \leq 0 \ . \tag{129}$$

The problem is still too difficult, and we restrict attention to the case of pure states $\hat{\rho}_0 = |\psi\rangle\langle\psi|$. We seek solutions of Eqs. (129) in the form

$$\mathrm{d}\hat{\Pi}(0) = \frac{\mathrm{d}\phi}{2\pi}|\gamma\rangle\langle\gamma| \ , \tag{130}$$

where $|\gamma\rangle$ is a (generally non normalizable) vector in the Hilbert space. The Lagrange operator has the following nonvanishing matrix elements

$$\langle k|\hat{\mathcal{Y}}|k\rangle = -\frac{1}{2\pi}\langle k|\gamma\rangle\langle\gamma|\psi\rangle\langle\psi|k\rangle \equiv -\frac{1}{2\pi}\gamma_k\psi_k^*\sum_{n=0}^{\infty}\gamma_n^*\psi_n \ , \tag{131}$$

where $\gamma_k \doteq \langle k|\gamma\rangle$ and $\psi_k \doteq \langle k|\psi\rangle$. Completeness of $\mathrm{d}\hat{\Pi}(\phi)$ implies that

$$\delta_{nm} = \langle n|\int \mathrm{d}\hat{\Pi}(\phi)|m\rangle = \int \frac{\mathrm{d}\phi}{2\pi} e^{-i\phi(n-m)}\langle n|\gamma\rangle\langle\gamma|m\rangle = \delta_{nm}|\langle n|\gamma\rangle|^2 \ , \tag{132}$$

which requires $|\gamma_k| = 1$. Moreover, reality of $\hat{\mathcal{Y}}$ needs $\arg(\gamma_k) = \arg(\psi_k)$ for $\psi_k \neq 0$. Hence, we write

$$\gamma_k = \begin{cases} \psi_k/|\psi_k| \doteq e^{i\chi_k}, & \psi_k \neq 0 \\ 1, & \psi_k = 0 \end{cases} \tag{133}$$

leading to

$$\langle k|\hat{\mathcal{Y}}|k\rangle = -\frac{1}{2\pi}|\psi_k| \sum_{n=0}^{\infty} |\psi_n| . \tag{134}$$

Now, we only need to check Eqs. (129). The second equation means that for any vector $|v\rangle$ in the Hilbert space, one has

$$0 \geq \langle v|\hat{\rho}_0 + 2\pi\hat{\mathcal{Y}}|v\rangle = \Big|\sum_{n=0}^{\infty} v_n^* \psi_n\Big|^2 - \sum_{n=0}^{\infty} |v_n|^2 |\psi_n| \sum_{k=0}^{\infty} |\psi_k| . \tag{135}$$

This bound is satisfied according to the Schwartz inequality

$$\begin{aligned} \Big|\sum_{n=0}^{\infty} v_n^* \psi_n\Big|^2 &= \Big|\sum_{n=0}^{\infty} v_n^* |\psi_n| \gamma_n\Big|^2 = \Big|\sum_{n=0}^{\infty} v_n^* |\psi_n|^{1/2} \gamma_n |\psi_n|^{1/2}\Big|^2 \\ &\leq \sum_{n=0}^{\infty} |v_n|^2 |\psi_n| \sum_{k=0}^{\infty} |\psi_k| . \end{aligned} \tag{136}$$

It remains to show that the first one of Eqs. (129) is also satisfied. One has

$$\begin{aligned} \langle k|[\hat{\rho}_0 + 2\pi\hat{\mathcal{Y}}]\xi_0|m\rangle &= \langle k|\psi\rangle\langle\psi|\gamma\rangle\langle\gamma|m\rangle + 2\pi\langle k|\hat{\mathcal{Y}}|k\rangle\langle k|\gamma\rangle\langle\gamma|m\rangle \\ &= [\langle k|\psi\rangle\langle\psi|\gamma\rangle - \langle k|\psi\rangle\langle\psi|\gamma\rangle\langle\gamma|k\rangle\langle k|\gamma\rangle]\langle\gamma|m\rangle \\ &= \langle k|\psi\rangle\langle\psi|\gamma\rangle[1 - |\gamma_k|^2]\langle\gamma|m\rangle = 0 , \end{aligned} \tag{137}$$

where we have considered that due to Eq. (119) also the following identity holds true

$$\hat{\mathcal{Y}} = -\frac{1}{2\pi}\hat{\xi}_0\hat{\rho}_0 . \tag{138}$$

In summary, we have proved the following assertion: the POM for estimating a phase shift of a pure state $|\psi\rangle\langle\psi|$ with max-likelihood cost-function is given by

$$d\hat{\Pi}(\phi) = \frac{d\phi}{2\pi} e^{-ia^\dagger a\phi}|\gamma\rangle\langle\gamma|e^{ia^\dagger a\phi} \equiv \frac{d\phi}{2\pi} \sum_{nm=0}^{\infty} e^{-i(n-m)\phi + i(\chi_n - \chi_m)}|n\rangle\langle m| , \tag{139}$$

where the phases χ_n depend on the state as follows

$$\chi_n = \begin{cases} \arg(\langle n|\psi\rangle), & \langle n|\psi\rangle \neq 0 \\ 0, & \langle n|\psi\rangle = 0 \end{cases} \tag{140}$$

In practice, it is not too restrictive to consider states with a well defined phase—i.e. with $\chi_n = n\nu$—which are just rotated by an angle ν of real positive states having all $\chi_n = 0$. In this case the optimal POM takes the canonical form

$$d\hat{\Pi}(\phi) = \frac{d\phi}{2\pi} \sum_{nm=0}^{\infty} e^{-i(n-m)\phi}|n\rangle\langle m| . \tag{141}$$

We have find the optimal POM according to the max-likelihood criterion that corresponds to a δ-like cost function. However, the same POM is optimal for any cost-function of the form $C(\phi_*, \phi) = C(\phi - \phi_*)$ where $C(\phi)$ is an even 2π-periodic function on $\mathbf{R}$ satisfying

$$\int_0^\infty C(\phi) \cos k\phi \, \mathrm{d}\phi \leq 0 \,, \quad k = 1, 2 \ldots \tag{142}$$

In fact, any function $C(\phi)$ satisfying Eq. (142) has Fourier series of the form

$$C(\phi) = c_0 - \sum_{k=0}^{\infty} c_k \cos k\phi \,, \qquad c_k \geq 0 \,. \tag{143}$$

Then, consider a general covariant POM

$$\mathrm{d}\hat{\Pi}(\phi) = \frac{\mathrm{d}\phi}{2\pi} \sum_{nm=0}^{\infty} e^{-i(n-m)\phi} \xi_{nm} |n\rangle\langle m| \,, \tag{144}$$

with $\xi_{nm} = \langle n|\hat{\xi}_0|m\rangle$ and $\hat{\xi}_0$ a generic selfadjoint operator. The average cost is given by

$$2\pi\overline{C} = c_0 - \sum_{k=1}^{\infty} c_k \int \frac{\mathrm{d}\phi}{2\pi} \cos k\phi \sum_{nm=0}^{\infty} e^{-i(n-m)\phi} \xi_{nm} \langle m|\psi\rangle\langle\psi|n\rangle \tag{145}$$

$$= c_0 - \sum_{k=1}^{\infty} c_k \frac{1}{2} \sum_{|n-m|=k} \langle\psi|n\rangle \xi_{nm} \langle m|\psi\rangle \,. \tag{146}$$

Positivity of $\mathrm{d}\hat{\Pi}(\phi)$ implies that $|\xi_{nm}| \leq \sqrt{\xi_{nn}\xi_{mm}} = 1$, hence

$$\sum_{|n-m|=k} \langle\psi|n\rangle \xi_{nm} \langle m|\psi\rangle \leq \sum_{|n-m|=k} |\langle\psi|n\rangle| |\langle m|\psi\rangle| \,, \tag{147}$$

and the equality is achieved only if $\xi_{nm} = \gamma_n^* \gamma_m$, with $|\gamma_n| = 1$. It follows that the minimum cost is

$$2\pi\overline{C}_{min} = c_0 - \frac{1}{2} \sum_{k=1}^{\infty} c_k \sum_{nm=0}^{\infty} |\langle\psi|n\rangle| |\langle m|\psi\rangle| \tag{148}$$

and this is attained by the POM (139).[2]

Notice that in the category (142) one also has the following cost functions

$$4 \sin^2 \frac{\phi}{2} = 2 - 2\cos\phi \,, \tag{149}$$

$$\min\{\phi, 2\pi - \phi\} = \frac{\pi}{2} - \frac{4}{\pi} \sum_{k=0}^{\infty} \sum_{k=0}^{\infty} \frac{\cos(2k+1)\phi}{(2k+1)^2} \,, \tag{150}$$

$$|\sin \frac{\phi}{2}| = \frac{2}{\pi} - \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\cos k\phi}{4k^2 - 1} \,, \tag{151}$$

$$-\delta_{2\pi}(\phi) = -\frac{1}{2\pi} - \frac{1}{\pi} \sum_{k=0}^{\infty} \cos k\phi \,. \tag{152}$$

[2]This is another derivation of the optimal POM for the quantum estimation problem of phase shift of pure states, but for more general cost functions. The second derivation is due to Holevo [14], whereas the previous one is an extended version of the Helstrom's proof [1].

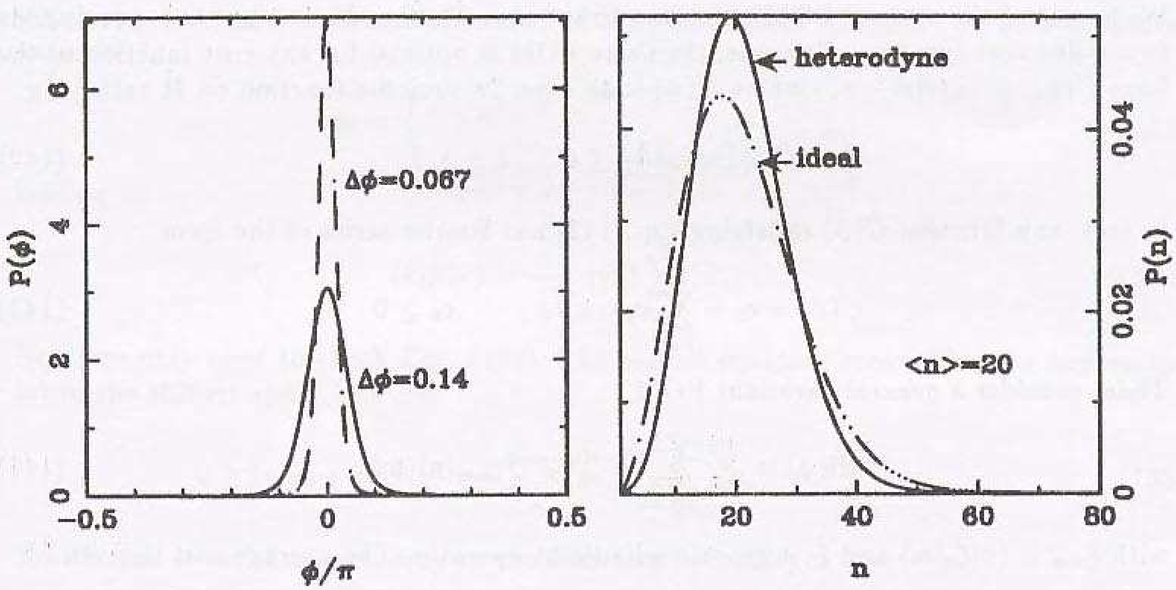

Figure 6. Number and phase probability distribution of optimal phase states with $\langle \bar{n} \rangle = 20$ for ideal and heterodyne phase detection. Optimal states minimize the r.m.s. phase fluctuations (the procedure for deriving such states is explained in Ref. [15]).

Minimizing the cost is equivalent to minimize the corresponding periodicized fluctuations (r.m.s fluctuations in Eq. (149)), or to maximize the likelihood in Eq. (152). Notice that the function $\min\{\phi^2, (\phi - 2\pi)^2\}$ does not belong to category (142). As the optimal POM is rather insensitive to the choice of the cost function, it can deserve the name "ideal" POM for the phase.

In Fig. 6 I report a numerical comparison between the ideal and the marginal heterodyne phase detection. For both cases a state with 20 photons is considered that minimizes the r.m.s. phase deviation $\Delta\phi = \sqrt{\langle \Delta\phi^2 \rangle}$ of the corresponding detection probability. It is evident that the ideal POM leads to probability distribution sharper than the heterodyne POM. Correspondingly, the number probability of optimal states for ideal detection are slightly broader than the number probability of optimal states for heterodyne detection. A more detailed analysis on marginal phase detection can be found in Ref. [15].

6. Beyond the POM: state reduction and "instrument"

Insofar we have considered only measurements that completely destroy the quantum mechanical description of the system after the interaction with a macroscopic detector: this is the case, for example, of photodetection, where radiation is completely absorbed. We are now interested in a different kind of measurements, which do not destroy the quantum mechanical description of the system, so that in principle a second measurement on the system can be performed after the first one. We call this type of measurements "measurements of the first kind", generalizing a term introduced by W. Pauli [16]. Henceforth, the customary measurement—i.e. those that destroy the state of the system—will be referred to as "measurements of the second kind".[3] More precisely, the definition of measurements

[3]This nomenclature has been used by M. Ozawa in Ref. [17]

of the first and second kind are as follows. For the second kind measurements the quantum mechanical description is provided just by the Born's rule: hence, these measurements are in one-to-one correspondence with POM's. For the first kind measurements, on the other hand, the description provides also the "state reduction" $\hat{\rho} \to \hat{\rho}_\Delta$, which gives the state $\hat{\rho}_\Delta$ immediately after the measurement, for a given experimental event Δ and for state $\hat{\rho}$ immediately before the measurement.[4] The "state reduction" is needed in order to evaluate the statistics of repeated measurements. In the following, we will refer to as "statistics of the measurements" including both the Born's rule and the state reduction.[5] The physical design and the preparation of the measuring apparatus determines the whole statistics of the measurement. In the following we will analyze the mathematical notion of "instrument", which synthetically describes the statistics of a measurement of the first kind.

6.1. INDIRECT MEASUREMENTS

The first kind measurement can be defined as a special type of "indirect" measurement of the second kind. An indirect measurement is a measurement that, instead of being performed directly on the system of interest, is carried out on a different system, which may include also the original system itself. Observables are measured that support informations on the desired quantities (for example, they have the same expectation values), but are different from them, and hence have different statistical distributions. We have already considered this kind of measurement in these lectures, when we analyzed joint measurements of conjugated quadratures of the field by heterodyne detection. The general scheme of this kind of measurements is sketched in Fig. 7. There is a system S and a probe P that interact (but not necessarily) each other; a measurement of the second kind is performed on compatible observables corresponding to operators acting on the whole Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_P$. In this case the state of the system itself is destroyed, and overall the measurement is of the second kind: the probe P is needed only in order to make observables compatible. For example, if $S + P$ have orthogonal projection-valued observables $\mathrm{d}\hat{E}_{SP}(\mathbf{x})$ and preparation $\hat{\rho} \otimes \hat{\rho}_P$ (I drop the subindex-S from the system density matrix) and the measurement is performed after the interaction $\hat{U}$, the Born's rule is given by

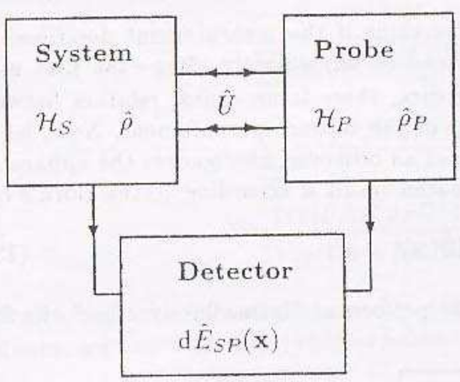

Figure 7. General scheme of indirect measurement of the second kind.

$$\mathrm{d}P(\mathbf{x}) = \mathrm{Tr}_{S+P}[\hat{U}\hat{\rho} \otimes \hat{\rho}_P \hat{U}^\dagger \mathrm{d}\hat{E}_{SP}(\mathbf{x})] = \mathrm{Tr}_S[\hat{\rho}\mathrm{d}\hat{\Pi}(\mathbf{x})] , \tag{153}$$

[4]In this way the nomenclature "first" and "second kind" can be put into correspondence to the "levels of description" of Holevo [18]. Here level I of description is the pure probabilistic one provided by the POM's, level II is the description of state reduction; finally, level III is the complete unitary description of the microscopic apparatus.

[5]Again this nomenclature is due to Ozawa [17].

corresponding to the POM

$$\mathrm{d}\hat{\Pi}(\mathbf{x}) = \mathrm{Tr}_P[\hat{\rho}_P \hat{U}^\dagger \mathrm{d}\hat{E}_{SP}(\mathbf{x})\hat{U}] \,. \tag{154}$$

There is a simple way to change the above scheme in order to make it suited to first kind measurements: just make a second kind measurement only on a probe observable $\mathrm{d}\hat{E}_P(\mathbf{x})$ (it is clear that now the interaction $\hat{U}$ is strictly needed). The resulting measurement scheme is sketched in Fig. 8. The Born's rule is

$$\mathrm{d}P(\mathbf{x}) = \mathrm{Tr}_{SP}[\hat{U}\hat{\rho}\otimes\hat{\rho}_P\hat{U}^\dagger \hat{1}_S \otimes \mathrm{d}\hat{E}_P(\mathbf{x})] = \mathrm{Tr}_S\{\hat{\rho}\mathrm{Tr}_P[\hat{\rho}_P\hat{U}^\dagger\hat{1}_S\otimes \mathrm{d}\hat{E}_P(\mathbf{x})\hat{U}]\} \,, \tag{155}$$

corresponding to the POM

$$\mathrm{d}\hat{\Pi}(\mathbf{x}) = \mathrm{Tr}_P[\hat{\rho}_P\hat{U}^\dagger\hat{1}_S\otimes \mathrm{d}\hat{E}_P(\mathbf{x})\hat{U}] \,. \tag{156}$$

Now, in order to determine the state reduction, one assumes that immediately after this measurement the system S is subjected to another measurement of an arbitrary observable of S, say with spectral resolution $\mathrm{d}\hat{E}_S(\mathbf{y})$.[6] The joint probability for the two combined measurements is

$$\mathrm{d}P(\mathbf{x},\mathbf{y}) = \mathrm{Tr}_{SP}[\hat{U}\hat{\rho}\otimes\hat{\rho}_P\hat{U}^\dagger \mathrm{d}\hat{E}_S(\mathbf{y})\otimes \mathrm{d}\hat{E}_P(\mathbf{x})] \,. \tag{157}$$

It is clear that the result would be exactly the same if the measurement described by $\mathrm{d}\hat{E}_S(\mathbf{y})$ is performed immediately before—instead of immediately after—the first measurement (however after the interaction $\hat{U}$). Hence, there is no causal relation between the "reading of the result $\mathbf{x}$" and the result $\mathbf{y}$ of the second measurement. Now, let us consider the experiment from the point of view of an observer who ignores the apparatus. He asserts that the first measurement has produced result $\mathbf{x}$ according to the Born's rule

$$\mathrm{d}P(\mathbf{x}) = \mathrm{Tr}[\hat{\rho}\,\mathrm{d}\hat{\Pi}(\mathbf{x})] \,. \tag{158}$$

Then, he considers the second measurement as performed "immediately after" the first

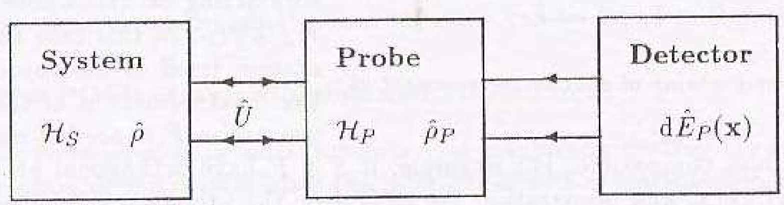

Figure 8. Scheme of an indirect measurement of the first kind.

one, described by the Born's rule with POM $\mathrm{d}\hat{E}_S(\mathbf{y})$, but now the measurement is performed on a different state, say $\hat{\rho}_\mathbf{x}$, that depends on the result $\mathbf{x}$ of the first measurement.

[6] Here I consider that the second measurement is of the second kind. The argument can be easily extended to the case that the second measurement is itself of the first kind: however, for the present purpose, this would create a logical loop.

In formulas, the conditional probability $dP(\mathbf{x}|\mathbf{y})$ of obtaining $\mathbf{y}$ given the result of the first measurement was $\mathbf{x}$, is given by

$$dP(\mathbf{x}|\mathbf{y}) = \mathrm{Tr}_S[\hat{\rho}_{\mathbf{x}} d\hat{E}_S(\mathbf{y})] , \tag{159}$$

and hence the joint probability of obtaining $\mathbf{x}$ and $\mathbf{y}$ can be written as follows

$$dP(\mathbf{x},\mathbf{y}) = dP(\mathbf{x}|\mathbf{y})\, dP(\mathbf{x}) = \mathrm{Tr}_S[\hat{\rho}_{\mathbf{x}}\, d\hat{E}_S(\mathbf{y})]\mathrm{Tr}_S[\hat{\rho}\, d\hat{\Pi}(\mathbf{x})] . \tag{160}$$

On the other hand, the probability (160) must be equal to the probability (157): in this way the following identity is obtained

$$\mathrm{Tr}_S[\hat{\rho}_{\mathbf{x}} d\hat{E}_S(\mathbf{y})] = \frac{\mathrm{Tr}_{SP}[\hat{U}\hat{\rho}\otimes\hat{\rho}_P\hat{U}^\dagger d\hat{E}_S(\mathbf{y})\otimes d\hat{E}_P(\mathbf{x})]}{\mathrm{Tr}_{SP}[\hat{U}\hat{\rho}\otimes\hat{\rho}_P\hat{U}^\dagger \hat{1}_S\otimes d\hat{E}_P(\mathbf{x})]} . \tag{161}$$

The arbitrariness of the choice of the second measurement yields the following relation for any basis $\{|n\rangle\}$

$$\langle n|\hat{\rho}_{\mathbf{x}}|m\rangle = \frac{\mathrm{Tr}_{SP}[\hat{U}\hat{\rho}\otimes\hat{\rho}_P\hat{U}^\dagger |m\rangle\langle n|\otimes d\hat{E}_P(\mathbf{x})]}{\mathrm{Tr}_{SP}[\hat{U}\hat{\rho}\otimes\hat{\rho}_P\hat{U}^\dagger \hat{1}_S\otimes d\hat{E}_P(\mathbf{x})]} , \tag{162}$$

namely

$$\begin{aligned}
\hat{\rho}_{\mathbf{x}} &= \sum_{nm} |n\rangle\langle m| \frac{\mathrm{Tr}_{SP}[\hat{U}\hat{\rho}\otimes\hat{\rho}_P\hat{U}^\dagger |m\rangle\langle n|\otimes d\hat{E}_P(\mathbf{x})]}{[\hat{U}\hat{\rho}\otimes\hat{\rho}_P\hat{U}^\dagger \hat{1}_S\otimes d\hat{E}_P(\mathbf{x})]} \\
&= \sum_{nm} |n\rangle\langle m| \frac{\mathrm{Tr}_S\{\mathrm{Tr}_P[\hat{U}\hat{\rho}\otimes\hat{\rho}_P\hat{U}^\dagger \hat{1}_s\otimes d\hat{E}_P(\mathbf{x})]|m\rangle\langle n|\}}{\mathrm{Tr}_{SP}[\hat{U}\hat{\rho}\otimes\hat{\rho}_P\hat{U}^\dagger \hat{1}_S\otimes d\hat{E}_P(\mathbf{x})]} \\
&= \frac{\mathrm{Tr}_P[\hat{U}\hat{\rho}\otimes\hat{\rho}_P\hat{U}^\dagger \hat{1}_s\otimes d\hat{E}_P(\mathbf{x})]}{\mathrm{Tr}_{SP}[\hat{U}\hat{\rho}\otimes\hat{\rho}_P\hat{U}^\dagger \hat{1}_S\otimes d\hat{E}_P(\mathbf{x})]} .
\end{aligned} \tag{163}$$

Notice that the denominator in Eq. (163) is just the trace of the numerator over $\mathcal{H}_S$. Hence, we can write Eq. (163) as follows

$$\hat{\rho}_{\mathbf{x}} = \frac{dI(\mathbf{x})\hat{\rho}}{\mathrm{Tr}_S[dI(\mathbf{x})\hat{\rho}]} , \tag{164}$$

where the mapping $dI(\mathbf{x})$ is defined as

$$dI(\mathbf{x})\hat{\rho} = \mathrm{Tr}_P[\hat{U}\hat{\rho}\otimes\hat{\rho}_P\hat{U}^\dagger \hat{1}_S\otimes d\hat{E}_P(\mathbf{x})] , \tag{165}$$

and is called "instrument".[7]

Thus we have seen that the description of a first kind measurement in terms of Born's rule and state reduction $\hat{\rho} \to \hat{\rho}_{\mathbf{x}}$ pertains an observer who ignores the microscopic description of the apparatus and focuses attention on the system S only. This also makes clear that the state reduction is not a causal evolution: it is not the consequence of the first

[7]For many authors [19] the names "instrument" and "POM" are substituted by "operations" and "effects", respectively.

observation and of "knowing the result", but just the statistical correlation between the results of the two measurements due to the interaction of the system with the probe.[8]

Now, let us consider the properties of the map $dI(\mathbf{x})$ more abstractly. The result of the measurement is not just a point $\mathbf{x}$: more generally, it is a Borel set, practically an interval $\Delta \in \mathbf{R}$ which the readout of the measurement is known to belong to. The above derivation of the state reduction can be generalized to the following rule (hereafter we drop the subindex S everywhere)

$$\hat{\rho} \to \hat{\rho}_\Delta = \frac{I(\Delta)\hat{\rho}}{\mathrm{Tr}[I(\Delta)\hat{\rho}]}\,, \qquad I(\Delta) = \int_\Delta dI(\mathbf{x})\,. \tag{166}$$

6.2. REALIZABLE INSTRUMENTS

Mathematically, the map $I(\Delta)$ is a linear transformation of trace class operators with the following rules

$$\begin{gathered} 0 \le \mathrm{Tr}[I(\Delta)\hat{\rho}] \le 1\,, \quad \mathrm{Tr}[I(\mathbf{R})\hat{\rho}] = 1 \\ \mathrm{Tr}[I(\cup_n \Delta_n)\hat{\rho}] = \sum_n \mathrm{Tr}[I(\Delta)\hat{\rho}]\,, \qquad \{\Delta_n\} \text{ countable disjoint}\,. \end{gathered} \tag{167}$$

Notice that $I(\mathbf{R})\hat{\rho} \neq \hat{\rho}$, in general.[9] It is easy to check that the map defined in Eq. (165) along with Eq. (166) satisfies the above axioms.[10] On the other hand, an abstract map I satisfying Eqs. (167) fully describes the statistics of a measurement of the first kind. It gives both the state reduction and the Born's rule as follows

$$\hat{\rho} \to \hat{\rho}_\Delta = \frac{I(\Delta)\hat{\rho}}{\mathrm{Tr}[I(\Delta)\hat{\rho}]}\,, \qquad P(\Delta) = \int_\Delta dP(\mathbf{x}) = \mathrm{Tr}[I(\Delta)\hat{\rho}]\,. \tag{168}$$

Now we address the problem if an instrument that satisfies axioms (167) can be physically realized in terms of an indirect measurement involving an interaction with some probe P, and for a suitable preparation of the probe. In other words, we want to know if any "mathematically given" instrument dI is "physically realizable" with a unitary interaction between S and P, as in Eq. (165). To this purpose, first notice that Eq. (165) leads to

[8] In my knowledge, this point was first clarified by Ozawa [20].

[9] The map $I(\mathbf{R})$ describes a "measurement without reading": this is the evolution of an "open system" S in interaction with an "environment" P.

[10] In particular, let us check positivity of the map, namely

$$\langle v|dI(\mathbf{x})\hat{\rho}|v\rangle \ge 0 \quad \forall v \in \mathcal{H}\,, \forall \hat{\rho} \text{ traceclass}\,.$$

One has

$$\langle v|dI(\mathbf{x})\hat{\rho}|v\rangle = \mathrm{Tr}[\hat{U}\hat{\rho} \otimes \hat{\rho}_P \hat{U}^\dagger |v\rangle\langle v| \otimes d\hat{E}_P(\mathbf{x})]\,.$$

Both density matrices can be written as convex linear combination of pure states. Hence, it is sufficient to prove positivity for pure states only, say $\hat{\rho} \equiv |\psi\rangle\langle\psi|$ and $\hat{\rho}_P \equiv |\varphi\rangle\langle\varphi|$. One has

$$\mathrm{Tr}[\hat{U}|\psi\rangle\langle\psi| \otimes |\varphi\rangle\langle\varphi|\hat{U}^\dagger|v\rangle\langle v| \otimes d\hat{E}(\mathbf{x})] = d\mathbf{x}\left\|(|\psi\rangle\langle\psi| \otimes |\varphi\rangle\langle\varphi|)\hat{U}^\dagger|v\rangle|\mathbf{x}\rangle\right\|^2 \ge 0\,.$$

a property for $\mathrm{d}I$ which is stronger than positivity: this is "complete positivity". We say that an instrument is completely positive if it satisfies the following requirement: for any finite sequence of vectors $|v_k\rangle$ and $|w_k\rangle$, $k = 1, \ldots, n$ one has

$$\sum_{k,l=1}^{n} \langle v_k|[\mathrm{d}I(\mathbf{x})|w_k\rangle\langle w_l|]|v_l\rangle \geq 0 \ . \tag{169}$$

From Eq. (165) with $\hat{\rho}_P = |\varphi\rangle\langle\varphi|$ pure state we have

$$\begin{aligned}\sum_{k,l=1}^{n} \langle v_k|[\mathrm{d}I(\mathbf{x})|w_k\rangle\langle w_l|]|v_l\rangle &= \sum_{k,l=1}^{n} \langle w_l|\langle\varphi|\hat{U}^\dagger[|v_l\rangle\langle v_k| \otimes \mathrm{d}\hat{E}_P(\mathbf{x})]\hat{U}|\varphi\rangle|w_k\rangle \\ = \sum_{k,l=1}^{n} \langle w_l|\langle\varphi|\hat{U}^\dagger|\mathbf{x}\rangle|v_l\rangle\langle v_k|\langle\mathbf{x}|\hat{U}|\varphi\rangle|w_k\rangle \mathrm{d}\mathbf{x} &= \left|\sum_{k=1}^{n}\langle w_l|\langle\varphi|\hat{U}^\dagger|\mathbf{x}\rangle|v_l\rangle\right|^2 \mathrm{d}\mathbf{x} \geq 0 \ . \end{aligned} \tag{170}$$

Hence, a realizable instrument is completely positive. Remarkably, Ozawa [20] has proven also the converse assertion, more precisely: Every completely positive instrument $\mathrm{d}I(\mathbf{x})$ is realizable, i.e. there is an extension $\mathcal{H} \otimes \mathcal{H}_P$ of the Hilbert space, a pure state preparation $|\varphi\rangle \in \mathcal{H}_P$ of the probe, a unitary operator $\hat{U}$ acting on $\mathcal{H} \otimes \mathcal{H}_P$, and a selfadjoint operator on $\mathcal{H}_P$ with spectral resolution $\mathrm{d}\hat{E}(\mathbf{x})$, such that

$$\mathrm{d}I(\mathbf{x})\hat{\rho} = \mathrm{Tr}_P[\hat{U}\hat{\rho} \otimes |\varphi\rangle\langle\varphi|\hat{U}^\dagger \hat{1} \otimes \mathrm{d}\hat{E}(\mathbf{x})] \ . \tag{171}$$

For the proof of the theorem the reader is referred to the original work of Ozawa [20].

We have seen that every instrument $\mathrm{d}I(\mathbf{x})$ is associated to a POM $\mathrm{d}\hat{\Pi}(\mathbf{x})$ by trace-duality as follows

$$\mathrm{Tr}[\mathrm{d}\hat{\Pi}(\mathbf{x})\hat{\rho}] = \mathrm{Tr}[\mathrm{d}I(\mathbf{x})\hat{\rho}] \ . \tag{172}$$

It is also true that for every POM $\mathrm{d}\hat{\Pi}(\mathbf{x})$ there is always at least an instrument $\mathrm{d}I(\mathbf{x})$ that satisfies Eq. (172). As a consequence, every POM can be achieved by a measurement of the first kind, and thus the Ozawa's theorem generalizes the Naimark's theorem.

I emphasize that an instrument $\mathrm{d}I(\mathbf{x})$ unambiguously determines a POM $\mathrm{d}\hat{\Pi}(\mathbf{x})$, whereas a POM $\mathrm{d}\hat{\Pi}(\mathbf{x})$ can be generally obtained from many different instruments: in other words, one has the same Born's rule with different state reductions. The POM does not contain sufficient details on the apparatus to describe the back action on the system, whereas the instrument provides a complete description of the statistics of the measurement. On the other hand, there can be still many different apparatus—i.e. different probes, probe preparations, and system-probe interactions—that are described by the same instrument $\mathrm{d}I(\mathbf{x})$.[11]

Now I want to make the relation between instrument and POM more explicit. Let us consider, for simplicity, the case of continuous spectrum and pure state preparation $|\varphi\rangle$ for the probe. One has

$$\begin{aligned}\int_\Delta \mathrm{d}I(\mathbf{x})\hat{\rho} &= \mathrm{Tr}_P\left[\hat{U}(\hat{\rho} \otimes |\varphi\rangle\langle\varphi|)\hat{U}^\dagger\left(\hat{1} \otimes \int_\Delta \mathrm{d}\mathbf{x}|\mathbf{x}\rangle\langle\mathbf{x}|\right)\right] \\ &= \int_\Delta \mathrm{d}\mathbf{x} \int \mathrm{d}\mathbf{x}' \langle\mathbf{x}'|\hat{U}|\varphi\rangle\, \hat{\rho}\, \langle\varphi|\hat{U}^\dagger|\mathbf{x}\rangle\delta(\mathbf{x}-\mathbf{x}') \\ &= \int_\Delta \mathrm{d}\mathbf{x}\, \langle\mathbf{x}|\hat{U}|\varphi\rangle\, \hat{\rho}\, \langle\varphi|\hat{U}^\dagger|\mathbf{x}\rangle = \int_\Delta \mathrm{d}\mathbf{x}\, \hat{\Omega}(\mathbf{x})\, \hat{\rho}\, \hat{\Omega}^\dagger(\mathbf{x}) \ ,\end{aligned} \tag{173}$$

[11] This is the level III of description of Holevo [18].

where $\hat{\Omega}(\mathrm{x})$ is the (non-unitary) operator acting on the Hilbert space $\mathcal{H}$ of the system only

$$\hat{\Omega}(\mathrm{x}) = \langle \mathrm{x}|\hat{U}|\varphi\rangle \tag{174}$$

which satisfies the completeness relation[12]

$$\int \mathrm{d}\mathrm{x}\, \hat{\Omega}^\dagger(\mathrm{x})\, \hat{\Omega}(\mathrm{x}) = \hat{1}\,. \tag{175}$$

The POM associated with the instrument can be obtained upon substituting Eq. (174) into (172), and using invariance of trace under cyclic permutations. One has

$$\mathrm{d}\hat{\Pi}(\mathrm{x}) = \mathrm{d}\mathrm{x}\, \hat{\Omega}^\dagger(\mathrm{x})\, \hat{\Omega}(\mathrm{x})\,, \tag{176}$$

namely

$$\hat{\Pi}(\Delta) = \int_\Delta \mathrm{d}\mathrm{x}\, \hat{\Omega}^\dagger(\mathrm{x})\hat{\Omega}(\mathrm{x})\,. \tag{177}$$

The generalization of Eq. (177) to the case of discrete spectrum is straightforward. Now we can immediately see that a way to change the instrument without changing the POM is the following "local" unitary transformation of operators $\hat{\Omega}(\mathrm{x})$

$$\hat{\Omega}(\mathrm{x}) \Longrightarrow \hat{V}(\mathrm{x})\hat{\Omega}(\mathrm{x})\,, \qquad \hat{V}^\dagger(\mathrm{x})\hat{V}(\mathrm{x}) = \hat{1}\,. \tag{178}$$

The transformation (178) does not affect the complete positivity of the instrument $\mathrm{d}I(\mathrm{x})$, however it gives a different state reduction (or "back-action") with the same POM.

6.2.1. *Example 1: the standard von Neumann model*

As a first example, we consider the von Neumann model [21] for a first kind measurement. Originally the model was conceived for the measurement of the position of a particle: here I translate it in the language of quantum optics, and I will describe an indirect unsharp measurement of a field quadrature. The interaction Hamiltonian is given by

$$\hat{H} = 2\hbar\kappa\hat{X}\hat{Y}_P\,, \tag{179}$$

where $\hat{X} = \frac{1}{2}(a^\dagger + a)$ and $\hat{Y}_P = \frac{i}{2}(a_P^\dagger - a_P)$ are quadratures of the field modes pertaining the system and the probe, respectively. The quadratures $\hat{X}$ and $\hat{Y}_P$ are the optical equivalent of position and momentum $\hat{q}_1$ and $\hat{p}_2$ of two different interacting particles, as it was considered in the original von Neumann model. Notice that the present optical model is given only for the sake of exemplification, because is would be difficult to achieve the Hamiltonian (179) optically (but also mechanically!) In the impulsive case (i.e. for strong coupling and short interaction time $\tau = \kappa^{-1} \to 0$) the operator $\hat{\Omega}(\mathrm{x})$ in Eq. (174) is given by

$$\hat{\Omega}(x) = \langle x|\exp(-2i\hat{X}\hat{Y}_P)|\varphi\rangle = \int \mathrm{d}x'\langle x|x'\rangle \exp\left[-\hat{X}\frac{\mathrm{d}}{\mathrm{d}x'}\right]\varphi(x') = \varphi(x - \hat{X})\,, \tag{180}$$

[12]I remind that in our notation the domain of the integral, when not specified, is the spectrum of the considered observable. Also one should keep in mind that vectors $|\mathrm{x}\rangle$ and $|\varphi\rangle$ in Eq. (174) belongs to the Hilbert space $\mathcal{H}_P$, so that the matrix element of $\hat{U}$ between them is an operator acting on $\mathcal{H}$ only ($\hat{U}$ is an operator acting on $\mathcal{H} \otimes \mathcal{H}_P$).

and hence the instrument is

$$\mathrm{d}I(x)|\psi\rangle\langle\psi| = \mathrm{d}x\,\varphi(x-\hat{X})|\psi\rangle\langle\psi|\,\varphi(x-\hat{X})\,. \tag{181}$$

To obtain the von Neumann state reduction—i.e. the projection over eigenvectors $|x\rangle$ of $\hat{X}$—let us consider the squeezed-vacuum preparation for P

$$\varphi(x) = \frac{1}{(2\pi\epsilon^2)^{1/4}}\exp\left(-\frac{1}{4}\frac{x^2}{\epsilon^2}\right)\,. \tag{182}$$

In the limit of vanishingly small $\epsilon \to 0$ (ϵ plays the role of the measurement precision) the reduced state will localize on a narrower and narrower Gaussian centered around the value x. Formally, we write the limit as follows

$$\lim_{\epsilon\to 0}\frac{\mathrm{d}I(x)\hat{\rho}}{\mathrm{Tr}[\mathrm{d}I(x)\hat{\rho}]} = |x\rangle\langle x|\,. \tag{183}$$

6.2.2. *Example 2: the Arthurs-Kelly model for joint measurements*

The previous example can be easily generalized to the case of a joint measurement of $\hat{X}$ and $\hat{Y}$. In this case we need two different probe modes that commute each other. Such a measurement model was considered for the first time by Arthurs and Kelly [22]: here I extend their analysis in order to derive also the state reduction of the model.

The impulsive interaction Hamiltonian can be chosen as follows

$$\hat{H} = \hbar(\kappa_1\hat{X}\hat{Y}_1 - \kappa_2\hat{Y}\hat{X}_2)\,, \tag{184}$$

where the choice of signs and constants is for later convenience. For simplicity of notation we set the interaction time $\tau = 1$. Let us analyze this model in the Heisenberg picture. One has

$$\hat{X}_1' \;=\; \hat{U}^\dagger\hat{X}_1\hat{U} = \hat{X}_1 + \frac{1}{2}\kappa_1\hat{X} - \frac{1}{8}\kappa_1\kappa_2\hat{X}_2\,, \tag{185}$$

$$\hat{Y}_2' \;=\; \hat{U}^\dagger\hat{Y}_2\hat{U} = \hat{Y}_2 + \frac{1}{2}\kappa_2\hat{Y} - \frac{1}{8}\kappa_1\kappa_2\hat{Y}_1\,. \tag{186}$$

It is convenient to require that the indirect measurement of $\hat{X}$ and $\hat{Y}$ be "unbiased", namely that the time evolved expectation values of $\hat{X}_1$ and $\hat{Y}_2$ are equal to those of $\hat{X}$ and $\hat{Y}$ that we want to measure at $t=0$, i.e. $\langle\hat{X}_1'\rangle = \langle\hat{X}\rangle$ and $\langle\hat{Y}_2'\rangle = \langle\hat{Y}\rangle$. This can be accomplished by choosing $\kappa_1 = \kappa_2 = 2$, and putting the probe modes into the vacuum state. However, considering that the vacuum fluctuations for each quadrature is equal to $\frac{1}{4}$, one has $\langle\Delta\hat{X}_1'^2\rangle = \langle\Delta\hat{Y}_2'^2\rangle = \frac{9}{16}$, which is $\frac{1}{16}$ larger that the minimum noise for a joint measurement. This suggests the further unitary transformation

$$\hat{X}_1 \to \frac{1}{\sqrt{2}}\hat{X}_1\,,\quad \hat{X}_2 \to \sqrt{2}\hat{X}_2\,,$$
$$\hat{Y}_1 \to \sqrt{2}\hat{Y}_1\,,\quad \hat{Y}_2 \to \frac{1}{\sqrt{2}}\hat{Y}_2\,, \tag{187}$$

which minimizes the noise, and can be achieved by the unitary operator

$$\hat{U}_{sq} = e^{i\log 2(\hat{X}_1\hat{Y}_1 - \hat{X}_2\hat{Y}_2)}\,. \tag{188}$$

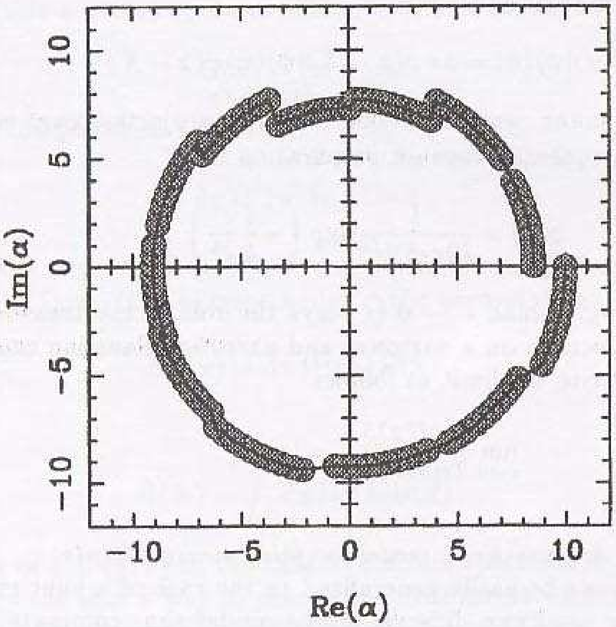

Figure 9. Brownian-motion effect on the free evolution due to the state reduction of the joint first-kind measurement in Eq. (193). This kind of measurement could account for the trajectory description of measurements from cloud or bubble chamber tracks.

The operator (188) squeezes the vacuum state of the probe '1' by a factor two in the $\hat{X}$ direction, and correspondingly unsqueezes the state of the probe '2' by the same factor. The squeezing operator $\hat{U}_{sq}$ acts before the interaction Hamiltonian $\hat{H}$. Hence, in summary, the model corresponds to the Hamiltonian (184) with probe preparation given by

$$\hat{\rho}_P = \hat{U}_{sq}|0,0\rangle\langle 0,0|\hat{U}^{\dagger}_{sq}\ , \tag{189}$$

and with $\hat{U}_{sq}$ defined in Eq. (188). In order to obtain the instrument corresponding to the present measurement scheme, we evaluate the operators $\hat{\Omega}(\mathbf{x})$ as follows

$$\begin{aligned}\hat{\Omega}(x,y) &= {}_1\langle x|\,{}_2\langle y|\exp[-2i(\hat{X}\hat{Y}_1 - \hat{Y}\hat{X}_2)]\hat{U}_{sq}|0,0\rangle \\ &= {}_1\langle x|\,{}_2\langle y|\hat{D}^{\dagger}(\hat{X}_2 + i\hat{Y}_1)\hat{U}_{sq}|0,0\rangle\ ,\end{aligned} \tag{190}$$

where $\hat{D}(x+iy) \equiv \exp[-2i(x\hat{Y} - y\hat{X})]$ denotes the displacement operator acting on $\mathcal{H}$ [$\hat{Y}_1$ and $\hat{X}_2$ can be treated as c-numbers, because they commute each other and with any system operator]. Using the resolutions of the identity in terms of eigenstates of $\hat{X}_1$ and $\hat{Y}_2$ we obtain

$$\begin{aligned}\hat{\Omega}(x,y) &= \int \mathrm{d}x'\mathrm{d}y'\,{}_1\langle x|y'\rangle_1\,{}_2\langle y|x'\rangle_2\,{}_1\langle y'|\,{}_2\langle x'|\hat{D}^{\dagger}(\hat{X}_2 + i\hat{Y}_1)\hat{U}_{sq}|0,0\rangle \\ &= \int \mathrm{d}x'\mathrm{d}y'\frac{1}{\pi}e^{2i(xy'-yx')}\hat{D}^{\dagger}(x'+iy')\,{}_1\langle y'|\,{}_2\langle x'|\hat{U}_{sq}|0,0\rangle \\ &= \int \frac{\mathrm{d}x'\mathrm{d}y'}{\pi}e^{2i(xy'-yx')}\hat{D}^{\dagger}(x'+iy')\frac{1}{\sqrt{\pi}}e^{-\frac{1}{2}(x'^2+y'^2)}\ ,\end{aligned} \tag{191}$$

and introducing the complex variables $\alpha = x + iy$ and $\lambda = x' + iy'$ we recognize in Eq. (191) the coherent state projector

$$\hat{\Omega}(x,y) \equiv \hat{\Omega}(\alpha,\overline{\alpha}) = \frac{1}{\sqrt{\pi}} \int \frac{d^2\lambda}{\pi} e^{\lambda\overline{\alpha}-\overline{\lambda}\alpha} : \hat{D}^\dagger(\lambda) :_A \equiv \frac{1}{\sqrt{\pi}} |\alpha\rangle\langle\alpha| . \quad (192)$$

Hence, the state reduction is given by

$$\hat{\rho}_{(\alpha,\overline{\alpha})} = \frac{|\alpha\rangle\langle\alpha|\hat{\rho}|\alpha\rangle\langle\alpha|}{\mathrm{Tr}[|\alpha\rangle\langle\alpha|\hat{\rho}]} \equiv |\alpha\rangle\langle\alpha| . \quad (193)$$

In Fig. 9 the effect of the state reduction (193) is illustrated for a freely evolving field mode. The instrument in Eq. (193) reduces the state to a coherent state $|\alpha\rangle\langle\alpha|$ that depends only on the outcome α of the measurement, whatever the starting state $\hat{\rho}$ is. Such kind of measurement—where the reduced state is independent on the input state—is referred to as *Gordon-Louisell*[13] measurement [23]. In general, a Gordon-Louisell measurement has an $\hat{\Omega}$ operator of the form $\hat{\Omega}(\mathrm{x}) = |\psi_{\mathrm{x}}\rangle\langle\varphi_{\mathrm{x}}|$, where $|\psi_{\mathrm{x}}\rangle$ denotes a normalized state vector that depends on the reading x, and $\{|\varphi_{\mathrm{x}}\rangle\}$ is a complete (generally not normalizable) set of vectors in the Hilbert space.

References

1. C. W. Helstrom, *Quantum Detection and Estimation Theory*, Academic Press, (New York, 1976). The reader is also addressed to the original papers: C. W. Helstrom, Found. Phys. 4, 453 (1974); Int. J. Theor. Phys. **11**, 357 (1974).
2. M. A. Naimark, Iza. Akad. Nauk USSR, Ser. Mat. **4** 277 (1940).
3. N. I. Akhierzer and I. M. Glezman, *Theory of Linear Operators in the Hilbert Space*, (two volumes bound as one), Dover Publ. Inc. N.Y 1993, pag. 121 of Vol. 2.
4. A. Peres, *Quantum Theory: Concepts and Methods*, vol. 57 of *Fundamental Theories of Physics*, Kluwer Academic Publisers (Dordrecht, Boston, London 1993)
5. G. M. D'Ariano, *Measuring Quantum States*, in this volume.
6. L. Mandel, Proc. Phys. Soc. **72** 1037 (1958); *ibid.* **74** 233 (1959)
7. P. L. Kelley, and W. H. Kleiner, Phys. Rev. A **30** 844 (1964)
8. H.Carmichael, *An Open System Approach to Quantum Optics*, Cap. 1 e 2, Springer-Verlag, Heidelberg (1993).
9. G. M. D'Ariano, Int. J. Mod. Phys. B **6** 1291 (1992)
10. G. M. D'Ariano, Nuovo Cimento **107**B 643 (1992).
11. H. P. Yuen and J. H. Shapiro, IEEE Trans. Inf. Theory, **24** 657 (1978); **25** 179 (1979); **26** 78 (1980).
12. H. P. Yuen, Phys. Lett. **91A**, 101 (1982).
13. E. Arthurs and M. S. Goodman, Phys. Rev. Lett. **60** 2447 (1988)
14. A. S. Holevo. *Probabilistic and statistical aspects of quantum theory*, North-Holland, (Amsterdam, 1982).
15. G. M. D'Ariano and M. G. A. Paris, Phys. Rev. A **49** 3022 (1994).
16. See R. Omnès, *The interpretation of Quantum Mechanics*, Princeton Series in Physics, Princeton University Press, (Princeton, 1994), pag. 84. The original reference is: W. Pauli, Handbuch der Physik, **24** 1 (1933)
17. M. Ozawa, in *Squeezed States and Nonclassical Light*, ed. by P. Tombesi and E. R. Pike, pag. 263, Plenum, (New York, 1989).
18. A. S. Holevo, in *Quantum Aspects of Optical Communications*, (Paris 1990), Lecture Notes in Physics **378**, eds. C. Bendjaballah, O. Hirota and S. Reynaud, Springer, Berlin-New York, (1991) pag. 127

[13] This nomenclature was introduced by Yuen

19. G. Ludvig, *Foundations of Quantum Mechanics I and II*, Springer, (Berlin, 1983). K. Kraus, *States, Effects, and Operations: fundamental notions of Quantum theory*, Springer, (Berlin, 1983). E. B. Davies, *Quantum Theory of Open Systems*, Academic, (London, 1976).
20. M. Ozawa, J. Math. Phys. **25** (1984).
21. J. von Neumann. *Mathematical Foundations of Quantum Mechanics*, Princeton UP, (Princeton, NJ, 1955).
22. E. Arthurs and J. L. Kelly, Bell. Syst. Tech. J., **44** 725-729 (1965).
23. J. P. Gordon and W. H. Louisell, in *Physics of Quantum Electronics*, pp. 833-840, McGraw-Hill, (New York, 1966).