Optimum phase-shift estimation and the quantum description of the phase difference

A. Luis[*] and J. Peřina[†]

Laboratory of Quantum Optics, Palacký University, 17. listopadu 50, 77207 Olomouc, Czech Republic

(Received 7 June 1996)

The problem of a correct quantum description of the phase difference is examined from the perspective of parameter estimation theory. It is shown that an optimum phase-shift measurement defines a phase difference operator which coincides with other approaches to the same problem. We also study the fundamental limit to the accuracy of a phase difference shift detection. We show that this limit can be reached by a measurement having countable outcomes despite the fact that a phase shift can take any value. We show that this is the case of the phase difference operator defined by an optimum phase-shift measurement. [S1050-2947(96)09911-8]

PACS number(s): 42.50.Dv

I. INTRODUCTION

The description of the phase in quantum terms has been influenced by the difficulty of ascribing an operator to it in the usual quantum sense. This occurs because a polar decomposition of the one-mode-field complex amplitude operator does not give a unitary operator exponential of the phase [1]. Given the relevance of this variable in classical terms there have been several attempts to provide a suitable translation into the framework of Hilbert space formalism [2]. Although this is still a controversial subject, recently significant progress has been made in unifying all different approaches to the problem [3].

One of them considers the phase as a parameter instead of as a dynamical variable. From this point of view the phase problem can be reformulated as the proper detection of a phase change. This leads to a study of measurement strategies allowing a determination as precise as possible of a variation of this parameter. The evaluation and comparison of different schemes falls then within the quantum estimation theory which provides the tools for the study of their performance [4]. The conclusions of such an analysis are important not only for possible applications but also for every quantum description of the phase. Although in this context the phase is regarded just as a parameter, we can expect that an optimum measurement of a phase shift should provide, more or less directly, a suitable description of the phase as a dynamical variable. After specifying conditions to be fulfilled by an optimum detection, this analysis can lead to the same result as that provided by other approaches quite distinct in principle [4,5]. Such a description is the positive operator measure given by the nonorthogonal Susskind-Glogower phase states.

The main part of this work is devoted to the phase of a one-mode field or absolute phase. Nevertheless, we can regard the phase difference as a more fundamental variable. This is because any observation of the phase must be relative to a reference phase. The relative phase is essentially the relevant variable in most of the practical cases of a phase-shift detection as it occurs when interferometric schemes are used. It could be thought that this reasoning cannot lead to any new conclusion and the phase difference would inherit the same difficulties and the same solutions of the absolute phase problem. However, the actual situation appears to be different.

For instance, we can point out the possibility of defining a unitary operator exponential of the phase difference by means of a polar decomposition of the product of two-mode-field complex amplitudes [6]. The solution for the exponential of the phase difference can be unitary, contrary to the result for the same problem in the one-mode case. We can also describe naturally the phase difference by using previous approaches for the absolute phases [7]. These two procedures lead in principle to different results, although a closer examination reveals some points of coincidence which are not explicit from the beginning. In any case, this kind of reasoning opens the possibility of applying, directly to the phase difference, the same tools used in the absolute phase problem with the hope of finding new conclusions.

Here we will examine the quantum description of the phase difference from the perspective of the estimation theory. We will consider that a phase shift is in fact observed as a phase difference shift. We expect that a measurement providing an estimation with desirable properties must serve as a suitable quantum description for the phase difference as a two-mode-field variable. Following the usual procedure of ascribing an operator to each variable, we will impose that such optimum estimation relies on a measurement described by a projection measure, arbitrary in principle. Our aim is to find out which projection measure is defined by imposing desirable properties to the estimation problem. Afterwards, it will be compared with the phase difference operator arising from a two-mode polar decomposition and also with the positive operator measure emerging from the description of the absolute phases in terms of the Susskind-Glogower states.

An interesting question arises from the comparison of these two last possibilities. On one hand, the positive operator measure derived from the absolute phase results in a continuous range of variation for the phase difference (although this behavior seems to be more apparent than effective [7]).

---

[*]Permanent address: Departamento de Optica, Facultad de Ciencias Físicas, Universidad Complutense, Madrid, Spain.

[†]Also at Joint Laboratory of Optics, Palacký University and Physical Institute of Czech Academy of Sciences, Olomouc, Czech Republic.

On the other hand, the phase difference operator results in a discrete character for this variable. The number of allowed values grows with the total photon number so that discreteness is only relevant in the limit of small photon numbers. In any case, a discrete character seems to be in apparent contradiction with its behavior as a parameter, because a phase difference shift can take any value. Accordingly, the positive operator measure would be preferable even if it does not give an operator for this variable.

Here again the quantum estimation theory provides a valuable tool for a proper discrimination between these two possibilities. In particular, the existence of a fundamental limit to the minimum detectable phase change depending on the total number of photons involved in the measurement is relevant [8–11]. This limit must play a significant role in the examination of whether a discrete character for the phase difference as a variable is in contradiction or not with its continuous behavior when considered as a shift parameter. In Sec. II we define the phase difference operator using an optimum phase-shift estimation and then we study the limit to the phase resolution in the detection of a phase difference shift in Sec. III.

II. PHASE DIFFERENCE OPERATOR DEFINED BY AN OPTIMUM PHASE-SHIFT ESTIMATION

In a phase estimation problem a field state $|\psi\rangle$, pure for simplicity, undergoes an unknown, nonrandom, phase shift θ. By making an appropriate quantum measurement and knowing the input state $|\psi\rangle$, the purpose is to estimate the phase shift θ.

As it has been discussed in the Introduction, we are interested in the phase difference between two modes rather than in the absolute phase and therefore we will consider that $|\psi\rangle$ is a two-mode field with complex amplitude operators a_1,a_2. Accordingly, the measurement is performed over these two modes. Although the estimation problem views the phase as a parameter instead of as a dynamical variable, it seems that the better the estimation is, the closer should be the measurement performed to a true measurement of the phase, or, in our case, of the phase difference. This means that after the specification of some desirable properties valid for every field state $|\psi\rangle$, an optimum phase estimation should lead to a quantum description of the phase difference as a variable. Such description will be given by the measurement that should be carried out in order to fulfill the said properties.

The first requirement that we will consider is that the measurement is described by an operator A or, equivalently, by a projection measure. We exclude from the beginning the possibility of a measurement relying on a positive operator measure. Such kinds of measurements are in fact projection measures defined over a wider set of field modes (or other auxiliary quantum degrees of freedom) when the state in the other modes is fixed and known in advance [12,11]. Since we are looking for a quantum description of a two-mode-field property, we dismiss this possibility which involves implicitly other degrees of freedom.

As a further requirement on A we will impose that it commutes with the total photon number $[A,a_1^\dagger a_1 + a_2^\dagger a_2]=0$. Here again this requirement is moti-

vated by our purpose of arriving at a suitable description of the phase difference. This commutation relation is the quantum translation of the corresponding classical Poisson bracket. The total photon number is the infinitesimal generator of translations of the phase sum and this commutation ensures that the phase difference is not modified by an equal shift of the two phases. This commutation or compatibility with the total photon number is, in fact, verified by a very broad class of phase approaches when we derive their associated description of the phase difference [7].

This commutation has another advantage since it means that the measurement can be carried out with no active elements involved and then all the energy available is that supplied by the input field state $|\psi\rangle$. This can be relevant in the examination of the ultimate phase resolution achievable since it depends on the total energy involved. Otherwise it may happen that the practical realization of the measurement could demand the use of other sources of energy not taken into account explicitly, as it is the case of a quadrature measurement, for example, where an intense local oscillator is needed.

After that, we will consider a measurement described by an operator A commuting with the total photon number. Their common eigenvectors $|N,a^{(N,k)}\rangle$ can be written in the number basis as

$$|N,a^{(N,k)}\rangle = \sum_{n=0}^N a_n^{(N,k)}|n,N-n\rangle, \qquad (2.1)$$

where $|n,N-n\rangle$ are number states with n photons in mode a_1 and $N-n$ photons in mode a_2, and $k=0,1,\ldots,N$ labels the $N+1$ eigenvectors of A within the subspace $\mathcal{H}_N$ with total photon number N. These vectors satisfy the resolution of the identity

$$\sum_{k=0}^N |N,a^{(N,k)}\rangle\langle N,a^{(N,k)}| = I_N, \qquad (2.2)$$

where I_N is the identity on $\mathcal{H}_N$ and the orthogonality relations

$$\langle N',a^{(N',k')}|N,a^{(N,k)}\rangle = \delta_{N,N'}\delta_{k,k'} \qquad (2.3)$$

hold. Otherwise, the vectors $|N,a^{(N,k)}\rangle$ and A itself are arbitrary.

The basic ingredient of a phase estimation problem is the conditional probability distribution function

$$P(N,k;\theta) = |\langle N,a^{(N,k)}|e^{i\theta a_1^\dagger a_1}|\psi\rangle|^2 \qquad (2.4)$$

of getting the outcome N,k when the true phase shift is θ. If we express the input field $|\psi\rangle$ in the number basis as

$$|\psi\rangle = \sum_{N=0}^\infty \sum_{n=0}^N \psi_n^{(N)}|n,N-n\rangle, \qquad (2.5)$$

we have

$$P(N,k;\theta) = \sum_{n,n'=0}^N a_n^{(N,k)*} a_{n'}^{(N,k)} \psi_n^{(N)} \psi_{n'}^{(N)*} e^{i(n-n')\theta}. \qquad (2.6)$$

This function contains all the relevant information concerning the estimation. It provides the estimate $\widetilde{\theta}^{(N,k)}$ of the phase shift θ as a function of the particular outcome N,k and also gives its accuracy. Since this information can be extracted in many different ways, we will consider general conditions on $P(N,k;\theta)$ independent of any procedure of data analysis.

As a desirable and natural property for $P(N,k;\theta)$ we will impose that its form as a function of θ does not depend on the particular outcome k which will appear simply as a parameter shifting the whole function. Choosing $k=0$ for example as a reference, we assume that for every N,k there are two quantities, $\mu(N,k)$ and $\phi^{(N,k)}$, such that the shifting property

$$P(N,k;\theta) = \mu(N,k)P(N,0;\theta - \phi^{(N,k)}) \quad (2.7)$$

is satisfied. The function $P(N,0;\theta)$ does not depend on k; $\mu(N,k)$ is a factor of proportionality and $\phi^{(N,k)}$ are phase parameters shifting the probability distribution as a whole, with $\mu(N,0)=1$ and $\phi^{(N,0)}=0$. Among other possible consequences of this shifting property we have that the uncertainty in the phase-shift estimation will be independent of the particular outcome k for a given value of N.

Next we proceed to the determination of the most general measurement satisfying these properties. Since the input vector $|\psi\rangle$ is arbitrary, the shifting condition (2.7) reads

$$a_n^{(N,k)*}a_{n'}^{(N,k)} = \mu(N,k)a_n^{(N,0)*}a_{n'}^{(N,0)}e^{i(n'-n)\phi^{(N,k)}}. \quad (2.8)$$

Taking $n=n'$ and using the normalization of $|N,a^{(N,k)}\rangle$ we get $\mu(N,k)=1$. Furthermore, if we sum Eq. (2.8) over k when $n=n'$ and if we use the resolution of the identity (2.2), we get

$$a_n^{(N,0)} = \frac{1}{\sqrt{N+1}}e^{i\delta_n^{(N)}}, \quad (2.9)$$

where $\delta_n^{(N)}$ are arbitrary phases. Then we have from Eq. (2.8)

$$a_n^{(N,k)} = \frac{1}{\sqrt{N+1}}e^{i\delta_n^{(N)}}e^{in\phi^{(N,k)}}. \quad (2.10)$$

The allowed values for $\phi^{(N,k)}$ can be determined from the orthogonality condition (2.3),

$$\langle N,a^{(N,k')}|N,a^{(N,k)}\rangle = \frac{1}{N+1}\sum_{n=0}^{N}e^{in(\phi^{(N,k)}-\phi^{(N,k')})} = \delta_{k,k'}, \quad (2.11)$$

which gives $\phi^{(N,k)} = 2\pi m(k)/(N+1)$, where $m(k)$ are integers satisfying $m(k) \neq m(k')$ (modulus $N+1$) if $k \neq k'$, and $m(0)=0$. If we rearrange the index k, we can write

$$\phi^{(N,k)} = \frac{2\pi}{N+1}k. \quad (2.12)$$

This is everything we can derive from the conditions imposed on the measurement. We can rename the vectors $|N,a^{(N,k)}\rangle$ as $|N,\phi^{(N,k)}\rangle$ and we have that the most general measurement having the desired properties is described by the countable set of orthogonal and normalized vectors

$$|N,\phi^{(N,k)}\rangle = \frac{1}{\sqrt{N+1}}\sum_{n=0}^{N}e^{in\phi^{(N,k)}}e^{i\delta_n^{(N)}}|n,N-n\rangle$$

$$= e^{i\phi^{(N,k)}a_1^{\dagger}a_1}|N,\phi^{(N,0)}\rangle, \quad (2.13)$$

where the $\phi^{(N,k)}$ phases are given by Eq. (2.12).

It is interesting to examine whether this solution has further properties or whether it is related to other approaches to the quantum description of the phase difference. If we take $\delta_n^{(N)} = n\delta^{(N)}$, we can build the unitary operator

$$E_{12} = \sum_{N=0}^{\infty}\sum_{k=0}^{N}|N,\phi^{(N,k)}\rangle e^{i[(2\pi/N+1)k+\delta^{(N)}]}\langle N,\phi^{(N,k)}|, \quad (2.14)$$

which satisfies the polar decomposition [6]

$$a_1 a_2^{\dagger} = E_{12}\sqrt{a_1^{\dagger}a_1(a_2^{\dagger}a_2+1)} = \sqrt{a_2^{\dagger}a_2(a_1^{\dagger}a_1+1)}E_{12} \quad (2.15)$$

defining the quantum translation for the exponential of the phase difference. Therefore, E_{12} can be properly considered as the unitary operator exponential of the phase difference and its eigenvectors $|N,\phi^{(N,k)}\rangle$ as phase difference states. Other possible solutions in Eq. (2.13) are simply related with this one by the unitary transformation

$$U = \sum_{N=0}^{\infty}\sum_{n=0}^{N}|n,N-n\rangle e^{i(\delta_n^{(N)}-n\delta^{(N)})}\langle n,N-n|, \quad (2.16)$$

and from now on we will consider for simplicity that the relation $\delta_n^{(N)} = n\delta^{(N)}$ is satisfied. It is worth noting that when one of the modes is in a very intense coherent state, the projection measure defined by the vectors (2.13) leads to the one-mode positive operator measure defined in terms of the Susskind-Glogower phase states [6].

We can see that the very general specification of good properties for a phase estimation problem leads essentially to a well behaved phase difference operator. This situation can be compared with the same case for the absolute phase. There a similar procedure leads to a positive operator measure instead of an operator since no unitary solution exists for the corresponding polar decomposition of a one-mode complex amplitude operator. In our case it leads to a projection measure since an operator description for the phase difference is possible. Although the method for implementing a direct measurement of this operator is not known yet, it is possible to obtain the quantities $|\langle N,\phi^{(N,k)}|\psi\rangle|^2$ from a measurement performed by means of an eight-port homodyne detector [13].

On the other hand, the positive operator measure for the absolute phases can be used to get a quantum description of the phase difference by means of very general methods [7]. The result is also a positive operator measure $\Delta(\phi)$,

$$\Delta(\phi) = \sum_{N=0}^{\infty}\Delta(N,\phi) = \sum_{N=0}^{\infty}\frac{N+1}{2\pi}|N,\phi\rangle\langle N,\phi|, \quad (2.17)$$

where ϕ can take any value in a 2π interval and $|N,\phi\rangle$ are the counterparts of Eq. (2.13) for arbitrary values of ϕ,

$$|N,\phi\rangle = \frac{1}{\sqrt{N+1}}\sum_{n=0}^{N} e^{in\phi}|n,N-n\rangle. \qquad (2.18)$$

This positive operator measure has been discussed recently as providing an optimum measurement for a phase-shift estimation [14]. The family $\Delta(N,\phi)$ commutes with the total photon number and also verifies the same shifting property (2.7) replacing $\phi^{(N,k)}$ by ϕ. This different range of variation distinguishes it from the projection measure defined by $|N,\phi^{(N,k)}\rangle$. As a consequence of this, the positive operator measure (2.17) cannot define a unitary operator exponential of the phase difference. On the other hand, the vectors $|N,\phi\rangle$ are not linearly independent because the subspaces $\mathcal{H}_N$ with total photon number N are finite dimensional. Then, the information they convey is redundant and a finite number of them is enough within each subspace $\mathcal{H}_N$.

In addition to these remarks, it is interesting to study whether this description has better properties than the one based on the countable set of vectors (2.13) especially if we take into account that a phase shift can take any value. To this end, the phase estimation problem offers a relevant test. We can examine which is the best possible resolution achievable and whether it can be reached with a measurement with countable outcomes. If there is a fundamental limit to the accuracy, we can expect that the detailed structure of the measurement below it will have no practical consequences. In such a case there would be no contradiction between the continuous character of the phase difference as a parameter and a discrete description as a variable. These points are studied in the next section.

III. LIMIT TO THE ACCURACY OF A PHASE DIFFERENCE SHIFT ESTIMATION

The purpose of this section is the precise determination of the ultimate limit to the accuracy in the detection of a phase difference shift and the measurement that should be performed in order to achieve it. We will also examine whether it can be matched by a measurement described by the vectors $|N,\phi^{(N,k)}\rangle$ found in the preceding section as providing a phase estimation having good properties.

As before, we regard the change of phase to be detected as a phase difference shift and therefore a two-mode description of the field state and the measurement is considered. The only prior requirement we will impose is that the measured observable commutes with the total photon number in accordance with the discussion in the preceding section. As to its precise form, the measurement is in principle arbitrary and we will describe it by a positive operator measure $\Delta(N,k)$ with the properties

$$[\Delta(N,k),a_1^\dagger a_1 + a_2^\dagger a_2]=0, \quad \sum_{k=0}^{M}\Delta(N,k)=I_N. \quad (3.1)$$

For simplicity we have labeled these operators by a discrete k taking an arbitrary number of values $M+1$, but it must be understood that it can be replaced throughout by a continu-

ous ϕ with the replacement of $\Sigma_{k=0}^{M}$ by $\int d\phi$. It can be checked at any step that such a replacement has no consequences on the following.

Any performance analysis of a detection scheme is based on the conditional probability distribution of getting the outcome N,k when the true value of the phase shift is θ,

$$P(N,k;\theta)=\langle\psi|e^{-i\theta a_1^\dagger a_1}\Delta(N,k)e^{i\theta a_1^\dagger a_1}|\psi\rangle, \qquad (3.2)$$

where the input field $|\psi\rangle$ is here again pure for simplicity. This function provides the estimate $\widetilde{\theta}^{(N,k)}$ of θ after the outcome N,k has been obtained and it also enters in the measure of its performance. For this purpose it is necessary to specify the importance of the deviations of the estimate $\widetilde{\theta}^{(N,k)}$ from the true value θ. This can be done by introducing a cost function $C(\widetilde{\theta}^{(N,k)},\theta)$. Assuming a uniform prior distribution for θ (maximum initial ignorance) the average cost $\bar{C}$,

$$\bar{C}=\int d\theta \sum_{N=0}^{\infty}\sum_{k=0}^{M} C(\widetilde{\theta}^{(N,k)},\theta)P(N,k;\theta), \qquad (3.3)$$

represents the performance measure and the minimum of $\bar{C}$ gives the best accuracy achievable according to the criterion defined by $C(\widetilde{\theta}^{(N,k)},\theta)$. Here we choose as the cost function

$$C(\widetilde{\theta}^{(N,k)},\theta)=|e^{i\widetilde{\theta}^{(N,k)}}-e^{i\theta}|^2=4\sin^2[(\widetilde{\theta}^{(N,k)}-\theta)/2]$$
$$=2[1-\cos(\widetilde{\theta}^{(N,k)}-\theta)], \qquad (3.4)$$

because of its good properties [15].

From the last two equations we can see that the average cost depends on the quantities

$$\int d\theta e^{i\theta}P(N,k;\theta). \qquad (3.5)$$

These quantities are interesting on their own since they provide a measure of the width of the conditional probability distribution $P(N,k;\theta)$ evaluating the accuracy of the phase estimate $\widetilde{\theta}^{(N,k)}$ after the outcome N,k. Specifically, we can define a normalized probability distribution $\mathcal{P}(N,k;\theta)$,

$$\mathcal{P}(N,k;\theta)=\frac{1}{\int d\theta' P(N,k;\theta')}P(N,k;\theta), \qquad (3.6)$$

which can be considered as a probability distribution for the inferred phase shift when we do not have any prior knowledge about its true value. The dispersion [8,9,16,17]

$$D^2(N,k)=1-|\langle e^{i\theta}\rangle^{(N,k)}|^2=1-\left|\int d\theta e^{i\theta}\mathcal{P}(N,k;\theta)\right|^2 \qquad (3.7)$$

gives a measure of the accuracy of the phase-shift measurement provided by the outcome N,k. We will begin looking for the minimum value possible for this dispersion. In addition to its own interest, it will be necessary for the minimization of the average cost.

The normalized distribution $\mathcal{P}(N,k;\theta)$ is given by

$$\mathcal{P}(N,k;\theta) = \mathcal{N}(N,k) \sum_{n,n'=0}^{N} \psi_n^{(N)*} \Delta_{n,n'}(N,k) \psi_{n'}^{(N)} e^{i(n'-n)\theta},$$

$$(3.8)$$

where

$$\mathcal{N}(N,k) = \frac{1}{2\pi \sum_{n=0}^{N} |\psi_n^{(N)}|^2 \Delta_{n,n}(N,k)}, \quad (3.9)$$

and

$$\Delta_{n,n'}(N,k) = \langle n, N-n | \Delta(N,k) | n', N-n' \rangle. \quad (3.10)$$

This gives

$$\langle e^{i\theta} \rangle^{(N,k)} = 2\pi \mathcal{N}(N,k) \sum_{n=1}^{N} \psi_n^{(N)*} \Delta_{n,n-1}(N,k) \psi_{n-1}^{(N)}. \quad (3.11)$$

The minimum of $D^2(N,k)$ or, equivalently, the maximum of $|\langle e^{i\theta} \rangle^{(N,k)}|$ can be found by the method of Lagrange multipliers. Nevertheless, there is another procedure which provides a better understanding of the result. We can note that Eq. (3.11) can be written formally as the following mean value:

$$\langle e^{i\theta} \rangle^{(N,k)} = \mathrm{tr}[\rho^{(N,k)} E_1 E_2^{\dagger}], \quad (3.12)$$

where $\rho^{(N,k)}$ is the normalized density matrix in the $\mathcal{H}_N$ subspace,

$$\rho^{(N,k)} = 2\pi \mathcal{N}(N,k) \sum_{n,n'=0}^{N} |n, N-n\rangle \psi_n^{(N)*} \Delta_{n,n'}(N,k) \psi_{n'}^{(N)}$$

$$\times \langle n', N-n' |, \quad (3.13)$$

and E_1, E_2 are the Susskind-Glogower operators

$$E|n\rangle = |n-1\rangle, \quad E|0\rangle = 0. \quad (3.14)$$

Since $\Delta(N,k)$ is a positive operator measure it can be verified that Eq. (3.13) defines a legitimate density matrix.

We are interested in the modulus of $\langle e^{i\theta} \rangle^{(N,k)}$ and so it will be convenient to extract its argument

$$\delta^{(N,k)} = \arg\left[\int d\theta e^{i\theta} P(N,k;\theta) \right] = \arg\{\mathrm{tr}[\rho^{(N,k)} E_1 E_2^{\dagger}]\}, \quad (3.15)$$

and then

$$\langle e^{i\theta} \rangle^{(N,k)} = e^{i\delta^{(N,k)}} \mathrm{tr}[\widetilde{\rho}^{(N,k)} E_1 E_2^{\dagger}], \quad (3.16)$$

where

$$\widetilde{\rho}^{(N,k)} = e^{-i\delta^{(N,k)} a_1^{\dagger} a_1} \rho^{(N,k)} e^{i\delta^{(N,k)} a_1^{\dagger} a_1}. \quad (3.17)$$

Now the trace in Eq. (3.16) is a positive real number and we have

$$\langle e^{i\theta} \rangle^{(N,k)} = e^{i\delta^{(N,k)}} \mathrm{tr}[\widetilde{\rho}^{(N,k)} C], \quad (3.18)$$

where

$$C = \tfrac{1}{2}(E_1 E_2^{\dagger} + E_1^{\dagger} E_2) \quad (3.19)$$

is the operator cosine of the phase difference introduced by Carruthers and Nieto [1], which has the following eigenvectors and eigenvalues:

$$C|N, \cos\theta_{N,r}\rangle = \cos\theta_{N,r} |N, \cos\theta_{N,r}\rangle,$$

$$|N, \cos\theta_{N,r}\rangle = \sqrt{\frac{2}{N+2}} \sum_{n=0}^{N} \sin[(n+1)\theta_{N,r}] |n, N-n\rangle,$$

$$(3.20)$$

with

$$\theta_{N,r} = \frac{\pi}{N+2} r, \quad r = 1,2,\ldots,N+1. \quad (3.21)$$

Since $|\langle e^{i\theta} \rangle^{(N,k)}|$ is the mean value of C in the state described by $\widetilde{\rho}^{(N,k)}$, its maximum value (minimum dispersion) is obtained when $\widetilde{\rho}^{(N,k)}$ is formed by the eigenvector with the maximum eigenvalue $\cos\theta_{N,1}$, $\widetilde{\rho}^{(N,k)} = |N, \cos\theta_{N,1}\rangle\langle N, \cos\theta_{N,1}|$. In such a case we have

$$\langle e^{i\theta} \rangle^{(N,k)} = e^{i\delta^{(N,k)}} \cos\frac{\pi}{N+2}, \quad (3.22)$$

and the minimum dispersion is

$$D_{\min}^2(N) = 1 - \cos^2\frac{\pi}{N+2} = \sin^2\frac{\pi}{N+2}, \quad (3.23)$$

which is reached whenever it is satisfied that

$$\rho^{(N,k)} = e^{i\delta^{(N,k)} a_1^{\dagger} a_1} |N, \cos\theta_{N,1}\rangle\langle N, \cos\theta_{N,1}| e^{-i\delta^{(N,k)} a_1^{\dagger} a_1}, \quad (3.24)$$

for any value of $\delta^{(N,k)}$. When $N \gg 1$ the limit of the resolution has the expected dependence $D_{\min}(N) \propto \pi/N$. The minimum detectable phase shift is of the order of the spacing of the phase difference values $\phi^{(N,k)} = 2\pi k/(N+1)$ derived in Eq. (2.12) in the preceding section.

We have found that in order to achieve the minimum dispersion, $\rho^{(N,k)}$ must be formed by a pure vector and then $\Delta(N,k)$ must be also formed by a pure vector,

$$\Delta(N,k) \propto |N, b^{(N,k)}\rangle\langle N, b^{(N,k)}|. \quad (3.25)$$

Given the definition (3.13) of $\rho^{(N,k)}$ and the form (3.24) that it must take in order to reach the minimum dispersion, the coefficients in the number basis of $|\psi\rangle$ and $|N, b^{(N,k)}\rangle$ must satisfy the relations

$$\psi_n^{(N)*} b_n^{(N,k)} \propto \sin\left[(n+1)\frac{\pi}{N+2}\right] e^{in\delta^{(N,k)}}. \quad (3.26)$$

The limit (3.23) can be reached in principle by any detection scheme for those outcomes having $b_n^{(N,k)} \neq 0$ for all n if the input state $|\psi\rangle$ is prepared according to Eq. (3.26). In general, for different N,k these conditions will be incompatible and the minimum will be reached only provided that just a particular result is obtained. Nevertheless, if the measurement is that described by the vectors (2.13) or the positive

operator measure formed by the vectors (2.18), it can be verified that the conditions (3.26) can be satisfied for every N,k if we identify $\delta^{(N,k)}$ with $\phi^{(N,k)}$ or ϕ, respectively. In such a case the minimum dispersion is reached irrespectively of the particular outcome obtained by means of a suitable preparation of $|\psi\rangle$ according to Eq. (3.26).

Next we consider the minimization of the average cost. For this purpose we have to evaluate

$$
\langle e^{i(\theta-\widetilde{\theta}^{(N,k)})}\rangle = \sum_{N=0}^{\infty}\sum_{k=0}^{M}\int d\theta\, e^{i(\theta-\widetilde{\theta}^{(N,k)})}P(N,k;\theta)
$$
$$
= \sum_{N=0}^{\infty}\sum_{k=0}^{M}\frac{1}{\mathcal{N}(N,k)}e^{-i\widetilde{\theta}^{(N,k)}}\int d\theta\, e^{i\theta}\mathcal{P}(N,k;\theta).
$$
(3.27)

The maximum possible value of the real part of $\langle e^{i(\theta-\widetilde{\theta}^{(N,k)})}\rangle$ will give us the minimum average cost. In the first place we can focus on the θ integration which just gives Eq. (3.11). It must take its maximum modulus for every N,k. We have found it to be Eq. (3.22) when Eq. (3.26) is satisfied. In such a case we have

$$
\langle e^{i(\theta-\widetilde{\theta}^{(N,k)})}\rangle = \sum_{N=0}^{\infty}\sum_{k=0}^{M}\frac{1}{\mathcal{N}(N,k)}e^{i(\delta^{(N,k)}-\widetilde{\theta}^{(N,k)})}\cos\frac{\pi}{N+2}.
$$
(3.28)

To obtain its maximum real part we can assume that the estimate $\widetilde{\theta}^{(N,k)}$ is naturally defined to be $\delta^{(N,k)}$ [17]

$$
\widetilde{\theta}^{(N,k)}=\delta^{(N,k)}=\arg\left\{\int d\theta\, e^{i\theta}P(N,k;\theta)\right\}, \quad (3.29)
$$

and we have that the minimum cost is associated with

$$
\langle e^{i(\theta-\widetilde{\theta}^{(N,k)})}\rangle = 2\pi\sum_{N=0}^{\infty}\cos\frac{\pi}{N+2}\sum_{n=0}^{N}|\psi_n^{(N)}|^2\sum_{k=0}^{M}\Delta_{n,n}(N,k).
$$
(3.30)

Using the resolution of the identity (3.1), the sum over k gives unity and we have

$$
\langle e^{i(\theta-\widetilde{\theta}^{(N,k)})}\rangle = 2\pi\sum_{N=0}^{\infty}\cos\frac{\pi}{N+2}\sum_{n=0}^{N}|\psi_n^{(N)}|^2
$$
$$
= 2\pi\sum_{N=0}^{\infty}P_{\psi}(N)\cos\frac{\pi}{N+2}
$$
$$
= 2\pi\langle\psi|\cos\frac{\pi}{\hat{N}+2}|\psi\rangle, \quad (3.31)
$$

where $P_{\psi}(N)$ is the probability of having a total photon number N in the state $|\psi\rangle$ and $\hat{N}=a_1^{\dagger}a_1+a_2^{\dagger}a_2$ is the total photon number operator. From the resolution of the identity (3.1) it follows that

$$
\int d\theta\sum_{N=0}^{\infty}\sum_{k=0}^{M}P(N,k;\theta)=2\pi, \quad (3.32)
$$

and we obtain that the minimum average cost is finally

$$
\overline{\mathcal{C}}_{\min}=4\pi\left(1-\langle\psi|\cos\frac{\pi}{\hat{N}+2}|\psi\rangle\right)=8\pi\langle\psi|\sin^2\frac{\pi}{2\hat{N}+4}|\psi\rangle.
$$
(3.33)

We recall that this minimum can be achieved provided that one requirement is fulfilled. The dispersion (3.7) must take its minimum value (3.23) for every N,k. To reach this minimum value the input vector and the measurement performed must satisfy the relation (3.26) for every N,k. It can be checked that it does not matter whether the measurement has discrete or continuous outcomes.

As we noted after Eq. (3.26), this requirement is satisfied when the measurement is described by the projection measure defined by the vectors (2.13) and also when it corresponds to the positive operator measure (2.17). According to Eqs. (3.26), (2.13), and (2.18) the input vector must be of the form

$$
|\psi\rangle = \sum_{N=0}^{\infty}c_N e^{i\delta'^{(N)}a_1^{\dagger}a_1}|N,\cos\theta_{N,1}\rangle, \quad (3.34)
$$

being an arbitrary superposition of rotated eigenvectors of the cosine operator (3.19) with the eigenvalue $\cos\theta_{N,1}$. In such a case, from Eqs. (3.29), (3.15), and (3.26), the estimates $\widetilde{\theta}^{(N,k)}$ are

$$
\widetilde{\theta}^{(N,k)}=\delta^{(N,k)}=\frac{2\pi}{N+1}k+\delta^{(N)}-\delta'^{(N)}, \quad (3.35)
$$

if the measurement is described by the vectors (2.13), while they are

$$
\widetilde{\theta}^{(N,\phi)}=\delta^{(N,\phi)}=\phi-\delta'^{(N)}, \quad (3.36)
$$

if the measurement corresponds to the vectors (2.18). In both the cases the estimates are given by the phase difference variables $\phi^{(N,k)}$ and ϕ, respectively.

This answers the question posed at the end of the preceding section. We have found that the measurement with countable outcomes described by the vectors (2.13) can reach the best possible accuracy of an arbitrary phase difference shift.

The minimum average cost (3.33) still depends on the total photon number distribution of the input vector (3.34). We can optimize it constraining the total mean photon number to be some given integer N (or its nearest integer). The minimum is then obtained when $|\psi\rangle = e^{i\delta'^{(N)}a_1^{\dagger}a_1}|N,\cos\theta_{N,1}\rangle$ and $\overline{\mathcal{C}}_{\min}$ takes the value $\overline{\mathcal{C}}_{\min}=8\pi\sin^2[\pi/(2N+4)]$ which scales as $1/N^2$ for large N.

We can note that there is some lack of symmetry between the phase difference states (2.13) and (2.18), describing optimal detection schemes and the input state (3.34) needed to reach the best resolution available. This is merely an effect of the cost function used here which is based on the dispersion. We have made such a choice because of its good properties. Nevertheless, the same analysis could be carried out with a different cost function. The choice $\mathcal{C}(\widetilde{\theta}^{(N,k)},\theta)=-\delta(\widetilde{\theta}^{(N,k)}-\theta)$, for instance, corresponds to the maximum likelihood and reciprocal peak criteria. It is not difficult to show that this leads to the same conclusions in

relation with the measurement providing the minimum cost which is obtained then when the input state is precisely a phase difference state (2.13) or (2.18).

Finally we can point out that, since the subspaces $\mathcal{H}_N$ have finite dimension $N+1$, the conclusions of this and the preceding sections for the phase difference can be translated straightforwardly to the azimuthal angle of an angular momentum j if we restrict the analysis to a fixed value of N with $N=2j$.

IV. CONCLUSIONS

In this paper we have demonstrated that an optimum measurement of a phase shift provides a suitable quantum description of the phase as a variable. To examine the optimum character of a measurement we have analyzed it from the point of view of a quantum estimation problem. We have applied this reasoning to the phase difference as being more meaningful than the absolute phase. After imposing some very general properties we have shown that this procedure leads to an operator description of this variable in the usual quantum sense, contrary to what happens for the absolute phase. Such an operator defined by an optimum measurement coincides with the solution of a polar decomposition for a two-mode field.

A particular feature of this operator is that its eigenvalues form a countable set which seems to be in contradiction with the fact that a phase shift can take any value. This contradiction is solved by the existence of an ultimate limit to the accuracy achievable in the measurement of a phase difference shift. We have shown that this limit can be reached by a measurement scheme having a countable set of outcomes and, in particular, by the phase difference operator previously found. Due to this limit, a discrete description of the phase difference as a variable is consistent with the fact that it can take any value as a parameter since a measurement with continuously distributed outcomes cannot provide a better estimation.

[1] P. Carruthers and M. M. Nieto, Rev. Mod. Phys. **40**, 411 (1968).

[2] J. Bergou and B. G. Englert, Ann. Phys. (N.Y.) **209**, 479 (1991); Special issue on quantum phase, Phys. Scr. **T48** (1993); A. Lukš and V. Peřinová, Quantum Opt. **6**, 125 (1994); R. Lynch, Phys. Rep. **256**, 367 (1995).

[3] U. Leonhardt, J. A. Vaccaro, B. Böhmer, and H. Paul, Phys. Rev. A **51**, 84 (1995); M. J. W. Hall, Quantum Opt. **3**, 7 (1991).

[4] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).

[5] J. H. Shapiro and S. R. Shepard, Phys. Rev. A **43**, 3795 (1991); M. J. W. Hall and I. G. Fuss, Quantum Opt. **3**, 147 (1991).

[6] A. Luis and L. L. Sánchez-Soto, Phys. Rev A **48**, 4702 (1993); L. L. Sánchez-Soto and A. Luis, Opt. Commun. **105**, 84 (1994).

[7] A. Luis and L. L. Sánchez-Soto, Phys. Rev A **53**, 495 (1996); Acta Phys. Slov. **45**, 387 (1995).

[8] A. Bandilla, H. Paul, and H. H. Ritze, Quantum Opt. **3**, 267 (1991).

[9] Z. Hradil, Quantum Opt. **4**, 93 (1992).

[10] A. S. Lane, S. L. Braunstein, and C. M. Caves, Phys. Rev. A **47**, 1667 (1993).

[11] J. Peřina, Z. Hradil, and B. Jurčo, *Quantum Optics and Fundamentals of Physics* (Kluwer, Dordrecht, 1994).

[12] H. P. Yuen, Phys. Lett. **91A**, 101 (1982).

[13] A. Luis and J. Peřina, Quantum Semiclass. Opt. **8**, 873 (1996).

[14] B. C. Sanders and G. J. Milburn, Phys. Rev. Lett. **75**, 2944 (1995); Z. Hradil, R. Myška, J. Peřina, M. Zawisky, Y. Hasegawa, and H. Rauch, *ibid.* **76**, 4295 (1996).

[15] Z. Hradil, Phys. Rev. A **46**, R2217 (1992).

[16] J. M. Lévy-Leblond, Ann. Phys. (N.Y.) **101**, 319 (1996).

[17] T. Opatrný, J. Phys. A **27**, 7201 (1994); A. Lukš and V. Peřinová, Czech. J. Phys. **41**, 1205 (1991).