Strong correspondence principle for joint measurement of conjugate observables

A. Di Lorenzo

Instituto de Física, Universidade Federal de Uberlândia, 38400-902 Uberlândia, Minas Gerais, Brazil

(Received 2 September 2010; published 18 April 2011)

It is demonstrated that the statistics for a joint measurement of two conjugate variables in quantum mechanics are expressed through an equation identical to the classical one, provided that joint classical probabilities are replaced by Wigner functions and that the interaction between the system and the detectors is accounted for. This constitutes an extension of Ehrenfest's correspondence principle and is thereby dubbed the strong correspondence principle. Furthermore, it is proved that the detectors provide an additive term to all the cumulants and that if they are prepared in a Gaussian state they contribute only to the first and second cumulants.

DOI: 10.1103/PhysRevA.83.042104 PACS number(s): 03.65.Ta

I. INTRODUCTION

Simultaneous knowledge of two conjugate observables is a discriminating feature between classical and quantum mechanics: In the former it is possible with arbitrary precision, at least in principle, while in the latter it is limited by the uncertainty relation. In particular, the question of whether it is possible to attribute a joint reality to position and momentum, independently of measurement, gave rise to the Einstein-Bohr debate and would later culminate in the Einstein-Podolsky-Rosen argument [1], from which the concept of entanglement arose. The measurement proposed in Ref. [1] is indeed a joint measurement of position and momentum, but it applies only to the particular entangled state considered therein. In a pioneering work, Arthurs and Kelly [2] demonstrated that a scheme exists for joint measurement of the position and momentum of a particle prepared in an arbitrary state, albeit at the cost of sacrificing the precision of both measurements. In Ref. [2] specific hypotheses were made about the initial preparation of the detectors. Later on it was shown that under these assumptions there is a connection between the joint probability and the Husimi Q function [3]; however, the hypotheses were relaxed in order to obtain the most general expression of the statistics [4,5]. Recently, the Arthurs-Kelly scheme was extended to arbitrary time-dependent coupling [6].

A gedanken experiment for the joint determination of Q and P based on time-of-flight measurements was also proposed [7]. However, the momentum would be inferred from Newtonian mechanics and not actually measured. The joint measurability of position and momentum was also investigated with no explicit reference to a detector, but considering only positive-operator measures [8,9]. In this context the connection of the joint measurement to the Neumark (or Naimark) embedding was demonstrated [10]; that is, by having a system interact with two detectors, the noncommuting operators $\hat{Q}$ and $\hat{P}$ in a Hilbert space $\mathcal{H}$ are associated with two commuting operators $\hat{I}_Q$ and $\hat{I}_P$ in an extended Hilbert space. A rigorous formalization and characterization of the joint measurement can be found in Ref. [11].

Remarkably, joint measurements of conjugate variables have been done in optics. The conjugate pair was realized through the amplitude and phase of light. Multiport techniques [12] or homodyne detection [13] were used to perform the measurement.

Furthermore, it was demonstrated [14] that in the joint measurement the product of the uncertainty of the conjugate variables is not less than $\hbar$. This is twice as much as the limit established by the uncertainty principle, which applies to separate measurements of position and momentum, not to joint ones. This result sparked much discussion about the interpretation of the uncertainty relations, starting from the 1990s [15] up to recent years [16]. On another front, the conditional state of the system after the measurement, already considered in Ref. [2], was studied from a more general viewpoint in Ref. [17] by assuming that the statistics are fully determined by the observed values and the associated spreads. From this hypothesis a Gaussian function was constructed according to the principle of maximum entropy [18]. This assumption is generally incorrect since the statistics of the outcomes are not always Gaussian and higher-order cumulants are needed to characterize it.

In the present work, it shall be proved that the formula giving the probability density for the output of the detectors is identical to the one obtained in the classical case provided that joint probabilities for the system observables and for the detectors observables are replaced by the Wigner quasiprobability functions. This result goes beyond Ehrenfest's theorem, which applies only to average values. A remarkably simple expression for the characteristic function is presented, thereby allowing us to draw general conclusions about the influence of the detectors on cumulants of any order. Furthermore, it is demonstrated that, by considering the conditional state of the system given the outcomes, the formula that expresses such a state in terms of Wigner functions is identical to the classical equation as well. Finally, a more general expression for the conditional state of the system is derived and it is found to differ from the one surmised in Ref. [17].

II. THE PROBLEM

Let us consider detection of two conjugate variables, denoted $\hat{Q}$ and $\hat{K}$, with units such that $[\hat{Q},\hat{K}] = i$. The interaction between the detectors and the system is written

$$H_{\rm int} = -\delta(t-t_0) \sum_{A=Q,K} \lambda_A \hat{\phi}_A \hat{A}, \tag{1}$$

with $\hat{\phi}_A$ an operator in the Ath-detector Hilbert space, $\hat{A}$ the observables of the system, λ_A coupling constants, and t_0

the time of the measurement. This coupling corresponds to the standard von Neumann detection scheme [19], duplicated to allow joint measurement of noncommuting observables, and as such it is known as the Arthurs-Kelly scheme [2].

A. Measurement of a single observable

For a single coupling (e.g., $\lambda_K = 0$), in order to have an ideal measurement, the following requirements must be met: (i) Initially, the density matrix of the detector and the system is $\rho = \rho_S \otimes \rho_A$ and (ii) the detector is prepared in a sharp state $\rho_A(I_A, I'_A) = |I_A = 0\rangle\langle I_A = 0|$, which peaks around $I_A = 0$. Indeed, under assumptions (i) and (ii), after the interaction with the system given in Eq. (1), the probability distribution for the variable I_A is

$$\Pi(I_A|\hat{A}) = \int d\mu(a)\delta(I_A - \lambda_A a)\langle a|\hat{\rho}_S(t_0^-)|a\rangle, \tag{2}$$

with a and $|a\rangle$ respectively, the eigenvalues and eigenstates, of $\hat{A}$, and $\mu(a)$ a measure. Equation (2) implies that in any single measurement I_A will be found to have one of the values $\lambda_A a$. Hypothesis (ii) can be replaced by a more realistic condition, requiring that (ii′) initially the detector is prepared in a state $\rho_A(I_A, I'_A)$, with $\Pi(I_A) \equiv \rho_A(I_A, I_A)$, which peaks around $I_A = 0$. If the range over which the off-diagonal elements of $\rho_A(I_A, I'_A)$ vanish exceeds the scale $\lambda_A \delta_a$, with δ_a the minimum distance between the eigenvalues of $\hat{A}$, then the measurement is a weak one [20,21]. Under assumption (ii′), Eq. (2) becomes the convolution

$$\Pi(I_A|\hat{A}) = \int d\mu(a)\Pi_A(I_A - \lambda_A a)\langle a|\rho_S(t_0^-)|a\rangle. \tag{3}$$

In terms of characteristic functions,

$$Z_A(\chi_A; t_0^+) = Z_S(\chi_A; t_0^-)Z_A(\chi_A; t_0^-). \tag{4}$$

The procedure above can be generalized to a nondemolition measurement [22]. One can consider a finite duration measurement, i.e., replace the δ function with a regular function $g(t)$, vanishing outside a finite-time window, under the following additional assumptions: (iii) The Hamiltonian of the detector depends only on $\hat{I}_A$, the variables conjugated to $\hat{\phi}_A$, so that $[\hat{\phi}_A, \hat{I}_B] = \delta_{AB} i\hbar$, and (iv) the observable $\hat{A}$ is conserved during the free evolution of the system (at least approximately). Since here we are interested in measuring noncommuting observables, the only Hamiltonian conserving $\hat{Q}$ and $\hat{K}$ simultaneously would be the trivial constant Hamiltonian; thus we keep the instantaneous interaction and make use of assumptions (i) and (ii′) throughout the rest of this paper. (In the case of a joint measurement of spin components, any spin-independent Hamiltonian conserves all spin components simultaneously and then one could consider a finite duration measurement.)

B. Measurement of two observables

Let us now consider a simultaneous interaction with both detectors, i.e., $\lambda_A \neq 0$, with $A = Q$ and K. In the following the operators are rescaled to eliminate the coupling constants and Planck's constant: $\lambda_A \hat{\phi}_A/\hbar \to \hat{\phi}_A$ and $\hat{I}_A/\lambda_A \to \hat{I}_A$. This way I_Q and Φ_K have the same dimensions as Q, while I_K and Φ_Q have the same dimensions as K. We shall also indicate for brevity (but not always) with I, without indexes, the pair $\{I_K, I_Q\}$, and define analogously $\Phi = \{\Phi_K, \Phi_Q\}$ and $\chi = \{\chi_K, \chi_Q\}$. The notation $\chi \cdot I$ is taken to mean $\sum_A \chi_A I_A$, etc. The formalism of quantum mechanics allows us to derive the probability density for I_Q and I_K. Indeed, by applying Born's rule to the time evolution of the total density matrix,

$$\Pi(I) = \mathrm{Tr}_S\langle I|e^{i\hat{\phi}\cdot\hat{A}}\rho^- e^{-i\hat{\phi}\cdot\hat{A}}|I\rangle, \tag{5}$$

with Tr_S the trace over the system degrees of freedom and $\rho^- \equiv \rho_S \otimes \rho_{\mathrm{det}}$ the density matrix of the system and the detectors evaluated at time $t_0^- = t_0 - \varepsilon$. After introducing twice the identity over the detector Hilbert space in terms of the eigenstates of $\hat{\phi}_A$, the probability density is

$$\Pi(I) = \int \left(\prod_A \frac{d\chi_A}{2\pi}\right) \exp[-i\chi \cdot I] Z(\chi), \tag{6}$$

where the generating function is

$$Z(\chi) = \int \left(\prod_A \frac{d\chi_A}{2\pi}\right) N(\Phi, \chi)\rho_{\mathrm{det}}(\Phi + \chi/2, \Phi - \chi/2), \tag{7}$$

with the kernel

$$N(\Phi, \chi) = \mathrm{Tr}_S\{\hat{V}_+ \rho_S \hat{V}_-^\dagger\},$$
$$\hat{V}_\pm \equiv \exp\left[i\sum_A \left(\Phi_A \pm \frac{\chi_A}{2}\right)\hat{A}\right]. \tag{8}$$

Since the operators $\hat{Q}$ and $\hat{K}$ do not commute with each other, the kernel in Eq. (7) generally will not be a function of only χ. Equations (5)–(8) actually apply to an arbitrary number of joint measurements. In the following, we consider the case of only two variables.

For general $\hat{Q}$ and $\hat{K}$ with a nonconstant commutator, it is not possible to obtain analytical results from Eqs. (7) and (8). For instance, if $\hat{Q}$ and $\hat{K}$ represented spin components, the integrals in Eq. (7) can only be done numerically.

The problem can be handled analytically when $[\hat{Q}, \hat{K}] = i$. In this case, Eq. (8) can be rewritten, after applying the Baker-Campbell-Hausdorff formula,

$$N = e^{i(\Phi_K\chi_Q + \chi_K\Phi_Q)/2}\,\mathrm{Tr}_S(\hat{V}_{Q+}\hat{V}_{K+}\rho_S\hat{V}_{K-}^\dagger\hat{V}_{Q-}^\dagger), \tag{9}$$

where the functional dependence was omitted for brevity and we define $\hat{V}_{A\pm} \equiv \exp[i(\Phi_A \pm \chi_A/2)\hat{A}]$.

The position eigenstates can be used to express the trace, giving

$$N(\Phi, \chi) = e^{i(\chi_K\Phi_Q - \Phi_K\chi_Q)/2} Z_S^W(\chi), \tag{10}$$

with

$$Z_S^W(q,k) \equiv \widetilde{\Pi}_S^W(q, -k) = \int dQ e^{ikQ}\rho_S(Q + q/2, Q - q/2) \tag{11}$$

the Fourier transform of the Wigner quasiprobability $\Pi_S^W(K, Q)$. The sign of the second variable has been changed so that $Z_S^W(0, \chi_Q)$ is the generating function for the probability

$\Pi_S(Q) = \langle Q|\rho_S|Q\rangle$ and $Z_S^W(\chi_K,0)$ is the generating function for the probability $\breve{\Pi}_S(K) = \langle K|\rho_S|K\rangle$. If the Wigner quasiprobability were positive definite, $Z_S^W(\chi_K,\chi_Q)$ would be the corresponding characteristic function. Since this is generally not the case, Z_S^W will be called a quasicharacteristic function.

III. RESULTS

It follows readily from Eqs. (7) and (10) that the characteristic function is

$$Z(\chi_K,\chi_Q) = Z_S^W(\chi_K,\chi_Q)Z_{\rm det}^W\left(\chi_K, -\frac{\chi_Q}{2};\chi_Q,\frac{\chi_K}{2}\right), \quad (12)$$

where

$$Z_{\rm det}^W(\chi_K,j_K;\chi_Q,j_Q) \equiv \widetilde{\Pi}_{\rm det}^W(\chi_K,-j_K;\chi_Q,-j_Q) = \int d^2\Phi\, e^{i\Phi\cdot j}\rho_{\rm det}\left(\Phi+\frac{\chi}{2},\Phi-\frac{\chi}{2}\right) \quad (13)$$

is the quasicharacteristic function of the detectors. Equation (12) should be contrasted with the case when first a measurement of K is made and shortly thereafter Q is observed

$$Z_<(\chi_K,\chi_Q) = Z_S^W(\chi_K,\chi_Q)Z_{\rm det}^W(\chi_K,-\chi_Q;\chi_Q,0), \quad (14)$$

or vice versa

$$Z_>(\chi_K,\chi_Q) = Z_S^W(\chi_K,\chi_Q)Z_{\rm det}^W(\chi_K,0;\chi_Q,\chi_K). \quad (15)$$

An important conclusion that can be drawn from Eq. (12) is that the contribution of the detectors to all the cumulants (defined as logarithmic derivatives of Z calculated at $\chi_Q = 0$ and $\chi_K = 0$) is simply an additive term. In particular, if the initial density matrix of the detectors is Gaussian, so is $Z_{\rm det}$ and thus the contribution of the detectors to cumulants higher than the second one vanishes.

The probability density is obtained by Fourier transforming Eq. (12); it consists of a convolution of the Wigner quasiprobability densities:

$$\Pi(I) = \int d^2I'd^2\Phi'\Pi_{\rm det}^W(I'_K,\Phi'_K;I'_Q,\Phi'_Q) \times \Pi_S^W\left(I_K - I'_K - \frac{\Phi'_Q}{2}, I_Q - I'_Q + \frac{\Phi'_K}{2}\right). \quad (16)$$

A priori, it is not obvious that the convolution of the two arbitrary quasiprobability functions presented in Eq. (16) is positive definite. However, $\Pi(I_K,I_Q)$ is positive by construction, thus an interesting mathematical corollary follows from the derivation presented above: Given any two quasiprobability functions, their convolution as defined in Eq. (16) gives a proper probability distribution. In particular, one can consider the detectors to be initially independent of one another so that

$$\Pi(I) = \int d^2I'd^2\Phi'\Pi_K^W(I'_K,\Phi'_K)\Pi_Q^W(I'_Q,\Phi'_Q) \times \Pi_S^W\left(I_K - I'_K - \frac{\Phi'_Q}{2}, I_Q - I'_Q + \frac{\Phi'_K}{2}\right) \quad (17)$$

is positive definite for any three Wigner functions.

IV. COMPARISON WITH CLASSICAL MECHANICS

Equation (16) would have a simple interpretation if the Wigner quasiprobabilities were positive definite: Before the interaction the observables of the detector A possess the values I'_A and Φ'_A with probability $\Pi_{\rm det}^W(I'_K,\Phi'_K;I'_Q,\Phi'_Q)$ and the K and Q variables of the system have the values K' and Q' with probability $\Pi_S^W(K',Q')$. Due to the interaction, the value of I_K is shifted deterministically by $K' + \Phi'_Q/2$ and that of I_Q is shifted by $Q' - \Phi'_K/2$. It is interesting to note that this is indeed the result one would obtain in the classical case if the interaction term is given by Eq. (1). Solving the classical Hamiltonian equations yields the values of I_K and I_Q immediately after the interaction, which are (primed quantities are calculated at $t_0^- = t_0 - \varepsilon$ and unprimed ones at $t_0^+ = t_0 + \varepsilon$)

$$I_K = K' + I'_K + \Phi'_Q/2, \qquad K = K' + \Phi'_Q, \quad (18a)$$

$$I_Q = Q' + I'_Q - \Phi'_K/2, \qquad Q = Q' - \Phi'_K, \quad (18b)$$

from which Eq. (16) readily follows.

From Eq. (12) we can derive the average,

$$\langle I_K\rangle = \langle K\rangle_S + \langle I_K\rangle_{\rm det} + \langle\Phi_Q\rangle_{\rm det}/2, \quad (19a)$$

$$\langle I_Q\rangle = \langle Q\rangle_S + \langle I_Q\rangle_{\rm det} - \langle\Phi_K\rangle_{\rm det}/2, \quad (19b)$$

and the spread of the measurements

$$\langle\Delta I_K^2\rangle = \langle\Delta\hat{K}^2\rangle_S + \langle\Delta\hat{I}_K^2\rangle_{\rm det} + \tfrac{1}{4}\langle\Delta\hat{\Phi}_Q^2\rangle_{\rm det} + \langle\Delta\hat{I}_K\Delta\hat{\Phi}_Q\rangle_{\rm det},$$

$$\langle\Delta I_Q^2\rangle = \langle\Delta\hat{Q}^2\rangle_S + \langle\Delta\hat{I}_Q^2\rangle_{\rm det} + \tfrac{1}{4}\langle\Delta\hat{\Phi}_K^2\rangle_{\rm det} + \langle\Delta\hat{I}_Q\Delta\hat{\Phi}_K\rangle_{\rm det},$$

where the indexed brackets indicate averaging over the density matrices of the system and the detectors before the interaction, while the unindexed ones indicate averaging over the probability $\Pi(I_K,I_Q)$ given in Eq. (16). Equations (18) are identical in form to Eqs. (19). This is a consequence of Ehrenfest's theorem, which implies that, for quadratic Hamiltonians, the equations of motion for the average values of an observable are identical to the corresponding classical equations. The formal identity of Eq. (16) in the classical and quantum cases, however, is a result that goes well beyond Ehrenfest's theorem. Herein this result is called the strong correspondence principle. It amounts to the following prescription: (i) Solve the classical equations of motion for the interaction between the detectors and the system, (ii) assume an initial joint probability distribution $\Pi(I_K,\Phi_K;I_Q,\Phi_Q)$ for the detectors and $\Pi(K,Q)$ for the system, (iii) find the joint probability of observing outcomes I_K and I_Q in terms of the initial probabilities, and (iv) substitute the classical probability distributions with the Wigner quasiprobability ones.

V. POSTDETECTION STATE

Finally, we consider the state of the system after the detection, conditioned on the fact that the readouts of the detectors were I_K and I_Q. Rather than working with the density matrix, here we study, equivalently, the conditional quasicharacteristic function and quasiprobability distribution of the system. The unnormalized conditional quasicharateristic function $\Pi(I_K,I_Q)Z_S^W(q,k|I_K,I_Q)$ can be found from Eq. (7) by replacing the kernel in Eq. (8) with

$$N = \int dQ\, e^{ikQ} \left\langle Q + \frac{q}{2} \right| \hat{V}_+ \rho_S \hat{V}_-^\dagger \left| Q - \frac{q}{2} \right\rangle. \quad (20)$$

This readily gives

$$Z(q,k|I) = \int d\chi_K d\chi_Q \frac{\exp(-i\chi\cdot I)}{\Pi(I)} \times Z_S^W(q+\chi_K,k+\chi_Q) \times Z_{\rm det}^W\left(\chi_K, -k - \frac{\chi_Q}{2}; \chi_Q, q + \frac{\chi_K}{2}\right)$$

and the corresponding quasiprobability distribution is

$$\Pi_S^W(K,Q|I) = \frac{1}{\Pi(I)} \int d^2\Phi\, \Pi_S^W(K-\Phi_Q,Q+\Phi_K)\Pi_{\rm det}^W\left(I_K - K + \frac{\Phi_Q}{2},\Phi_K; I_Q - Q - \frac{\Phi_K}{2},\Phi_Q\right) \quad (21)$$

or

$$\Pi_S^W(K,Q|I) = \frac{1}{\Pi(I)} \int dK'dQ'\, \Pi_S^W(K',Q')\Pi_{\rm det}^W\left(I_K - \frac{K+K'}{2},Q'-Q; I_Q - \frac{Q+Q'}{2},K-K'\right). \quad (22)$$

Equation (21) has a simple classical interpretation: According to Bayes's theorem, the conditional probability of finding the system with values K and Q, given that the detectors gave the output I_K and I_Q, satisfies

$$\Pi(I)\Pi_S(K,Q|I) = \Pi(K,Q,I). \quad (23)$$

The classical joint probability $\Pi(K,Q,I)$ can be derived from the classical equations of motion [Eqs. (18)] through the following reasoning: For given Φ_K and Φ_Q, the values of the system before the interaction must be $K-\Phi_Q$ and $Q+\Phi_K$; this happens with probability $\Pi_S^W(K-\Phi_Q,Q+\Phi_K)$. The values of I_K and I_Q before the interaction must have been $I'_K = I_K - K + \Phi_Q/2$ and $I'_Q = I_Q - Q - \Phi_K/2$ with arbitrary Φ_K and Φ_Q; this happens with probability $\Pi_{\rm det}(I_K - K + \Phi_Q/2,\Phi_K; I_Q - Q - \Phi_K/2,\Phi_Q)$. By integrating over all possible values of Φ_K and Φ_Q we obtain Eq. (21). Thus the strong correspondence principle has a further application: One could derive the joint conditional probability through classical reasoning and then replace in the formulas the Wigner quasiprobabilities distributions of the detectors and the system for the positive definite classical probabilities. It should be noted that, for general preparation of the detectors, the conditional state of the system depends on its initial state and it is not Gaussian, contrary to what was concluded in Ref. [17] by applying the principle of maximum entropy.

VI. CONCLUSIONS

A rich correspondence between classical and quantum mechanics has been demonstrated herein: Not only do the average values of an observable obey the classical equations of motion, as established by Ehrenfest's theorem, but the full joint probability of the outcomes has the same expression in the classical case as in the quantum case provided that the classical joint probabilities are replaced by the Wigner quasiprobabilities. Due to the uncertainty relations, the Wigner quasiprobabilities come in such combinations that they give rise to a positive probability distribution. It was also demonstrated that the characteristic function of the joint outcomes has a remarkably simple expression in terms of the quasicharacteristic function [see Eq. (12)]. From this we can conclude that detectors contribute an additive term to the cumulants of all orders. The strong correspondence between the classical and quantum cases was shown to hold also for the determination of the conditional state of the system after the measurement.

ACKNOWLEDGMENTS

The author acknowledges useful comments on the manuscript by Rosario Fazio and Saverio Russo. This work was supported by Fundação de Amparo à Pesquisa do Estado de Minas Gerais through Process No. APQ-02804-10.

[1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[2] E. Arthurs and J. L. Kelly, Bell Syst. Tech. J. **44**, 725 (1965).
[3] K. Husimi, Proc. Phys. Math. Soc. Jpn. **22**, 264 (1940).
[4] S. L. Braunstein, C. M. Caves, and G. J. Milburn, Phys. Rev. A **43**, 1153 (1991).
[5] S. Stenholm, Ann. Phys. (NY) **218**, 233 (1992).
[6] M. Bußhardt and M. Freyberger, Phys. Rev. A **82**, 042117 (2010).
[7] J. Park and H. Margenau, Int. J. Theor. Phys. **1**, 211 (1968).
[8] W. M. de Muynck and J. M. V. A. Koelman, Phys. Lett. A **98**, 1 (1983); P. Kruszynski and W. M. de Muynck, J. Math. Phys. **28**, 1761 (1987); O. Abu-zeid, Phys. Lett. A **125**, 162 (1987).
[9] P. Busch and P. J. Lahti, Phys. Rev. D **29**, 1634 (1984).
[10] M. Naimark, Izv. Akad. Nauk SSSR, Ser. Mat. **4**, 277 (1940); C. R. Acad. Sci. URSS **41**, 359 (1943); R. Y. Levine and R. R. Tucci, Found. Phys. **19**, 161 (1989).

[11] C. Carmeli *et al.*, J. Phys. A **38**, 5253 (2005).
[12] N. G. Walker and J. E. Carroll, Electron. Lett. **20**, 981 (1984); Opt. Quantum Electron. **18**, 355 (1986); N. G. Walker, J. Mod. Opt. **34**, 15 (1987).
[13] J. W. Noh, A. Fougères, and L. Mandel, Phys. Rev. Lett. **67**, 1426 (1991); Phys. Rev. A **45**, 424 (1992); **46**, 2840 (1992).
[14] E. Arthurs and M. S. Goodman, Phys. Rev. Lett. **60**, 2447 (1988).
[15] M. G. Raymer, Am. J. Phys. **62**, 986 (1994); U. Leonhardt, B. Böhmer, and H. Paul, Opt. Comm. **119**, 296 (1995); D. M. Appleby, J. Phys. A **31**, 6419 (1998); Int. J. Theor. Phys. **37**, 1491 (1998); **38**, 807 (1999).
[16] G. M. D'Ariano, Fortschr. Phys. **51**, 318 (2003); M. Ozawa, Phys. Rev. A **67**, 042105 (2003); M. J. W. Hall, *ibid.* **69**, 052113 (2004); R. F. Werner, in *Quantum Information, Statistics, Probability*, edited by O. Hirota (Rinton Pr. Inc., 2004), reprinted in Quant. Inf. Comput. **4**, 546 (2004); E. Andersson, S. M. Barnett, and A. Aspect, Phys. Rev. A **72**, 042104 (2005); Yu. I. Vorontsov, Phys. Usp. **48**, 999 (2005); P. Busch, T. Heinonen, and P. Lahti, Phys. Rep. **452**, 155 (2007).
[17] C. Y. She and H. Heffner, Phys. Rev. **152**, 1103 (1966).
[18] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).
[19] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932) [*Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955)].
[20] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. **60**, 1351 (1988).
[21] A. Di Lorenzo and J. C. Egues, Phys. Rev. A **77**, 042108 (2008). Interestingly, if $\hat{A}$ has a continuous spectrum, $\delta_a = 0$ and the distinction between strong and weak measurements is a sharp one; otherwise, there is a smooth crossover between the two.
[22] V. B. Braginsky and F. Ya. Khalili, *Quantum Measurement* (Cambridge University Press, Cambridge, 1992).