Reconstruction of the spin state

Z. Hradil,* J. Summhammer, G. Badurek, and H. Rauch
Atominstitut der Österreichischen Universitäten, Stadionallee 2, A-1020 Wien, Austria

(Received 13 January 2000; published 9 June 2000)

An ensemble of spin-$\frac{1}{2}$ particles is observed repeatedly using Stern-Gerlach devices with varying orientations. Synthesis of such noncommuting observables is analyzed using the maximum likelihood estimation as an example of quantum-state reconstruction. Repeated incompatible observations represents a new generalized measurement. This idealized scheme will serve for analysis of future experiments in neutron and quantum optics.

PACS number(s): 03.65.−w

I. INTRODUCTION

Quantum mechanics of spin-$\frac{1}{2}$ particles often serves as an illustrative example of key quantum physical concepts in standard textbooks of theoretical physics [1]. The importance of spin-$\frac{1}{2}$ states is enhanced by the fact that they represent the smallest possible amount of quantum information—quantum bits (q bits). Aside from theoretically valuable ''Gedanken'' experiments, spin-$\frac{1}{2}$ particles such as electrons, neutrons, or the circular polarization states of light quanta have allowed the realization of a variety of fundamental experiments in matter wave and quantum optics. They play a crucial role in many sophisticated schemes involving entanglement, Bell state analysis, or teleportation. Spin coherence and the possibility to reconstitute the beam of spin-$\frac{1}{2}$ particles after the Stern-Gerlach (SG) detection have been considered as the ''Humpty-Dumpty'' problem [2]. Several approaches for measurement and estimation of spin states have been considered recently [3–6]. In this Brief Report, the maximum likelihood (MaxLik) estimation of a spin-$\frac{1}{2}$ quantum state will be formulated as an illustrating example of a more general treatment [7,8]. The formalism presented here reveals the tight relationship between quantum theory and statistics. The synthesis of many independent and nonequivalent idealized detection schemes of the SG type will be interpreted as a kind of generalized measurement of quantum states. It will be an obviously useful tool for spin-state analysis in neutron depolarization experiments, neutron as well as photon interferometry, and for quantum state reconstruction, to name just a few typical examples.

Let us begin with a brief review of the basic properties of spin-$\frac{1}{2}$ quantum systems. A pure state (projector) shall be represented by the expression

$$|\mathbf{a}\rangle\langle\mathbf{a}| = \frac{1}{2}(1 + a_i\sigma_i), \tag{1}$$

where $\mathbf{a} = (a_1, a_2, a_3)$ is the three-dimensional normalized state vector σ_i, $i = 1,2,3$ represent the Pauli matrices, and the summation convention for repeated indices is used. Since

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k,$$

the scalar product of two projectors is given as

$$|\langle a|b\rangle|^2 = \frac{1}{2}(1 + a_i b_i).$$

A mixed state, which is described by a density matrix, can be parametrized by

$$\hat{\rho} = p_+|\mathbf{a}\rangle\langle\mathbf{a}| + p_-|-\mathbf{a}\rangle\langle-\mathbf{a}| \tag{2}$$

$$= \frac{1}{2} + \frac{1}{2}\sigma_i a_i(p_+ - p_-), \tag{3}$$

where $p_+ + p_- = 1$ and the states $|\pm\mathbf{a}\rangle$ denote a general orthogonal basis. Alternatively the spin state is completely determined if the associated polarization vector

$$r_i = \langle\sigma_i\rangle = a_i(p_+ - p_-) \tag{4}$$

is known, where, as usual, the brackets $\langle\rangle$ denote an expectation value. The degree of polarization is defined by

$$|\mathbf{r}|^2 \leqslant 1,$$

with $|\mathbf{r}|^2 = 0$ for completely unpolarized (mixed) state and $|\mathbf{r}|^2 = 1$ for fully polarized (pure) states.

The polarization or spin may be measured by projecting the state into the given directions $\pm\mathbf{a}$ of a SG apparatus. Closure relation and operator representation of such a device can be written as

$$|\mathbf{a}\rangle\langle\mathbf{a}| + |-\mathbf{a}\rangle\langle-\mathbf{a}| = \hat{1}, \tag{5}$$

$$\hat{A} = \frac{1}{2}[|\mathbf{a}\rangle\langle\mathbf{a}| - |-\mathbf{a}\rangle\langle-\mathbf{a}|]. \tag{6}$$

Assuming for the sake of simplicity always the same total number of particles N, the number of particles with either spin ''up'' or ''down'' yields estimates of the projections of the polarization vector according to the relations

$$n_\pm = Np(\pm\mathbf{a}) = \frac{1}{2}N(1 \pm \mathbf{ra}). \tag{7}$$

*Permanent address: Department of Optics, Palacký University, 17. listopadu 50, 772 07 Olomouc, Czech Republic.

Since this may be done for three orthogonal directions in space $\mathbf{x}_i$ $(i=1,2,3)$ the polarization vector may be found by eliminating the total number of particles N

$$r_i = \frac{n_{i+} - n_{i-}}{n_{i+} + n_{i-}}. \tag{8}$$

By this procedure each polarization component is determined separately. It represents a correct solution, provided that the resulting polarization lies upon or inside the Poincaré sphere $|\mathbf{r}|^2 \leqslant 1$. However, the ‘‘states’’ outside the Poincaré sphere violate the positive semidefiniteness of quantum states and thus leads to an improper quantum physical description of noise [8]. Similar problems appear in the case when more than three projections are used. Some results of SG projections might appear as incompatible among themselves due to the fluctuations and noises involved. Various SG measurements are not equivalent, since they are observing different ‘‘faces’’ of the spin system. Such measurements, even when done with an equal number of particles, determine different projection with different errors. Detected data $n_{i,\pm}$ collected from SG observations in M directions $\pm\mathbf{a}^i$, $i=1,2,\ldots,M$ sample a variety of binomial distributions. Significantly, the detected data $n_{i,\pm}$ fluctuate with the root-mean-square errors given by $\sqrt{N[1-(\mathbf{r}\mathbf{a}^j)^2]/2}$, depending on the deviations between projections and the true (but unknown) direction of the spin $\mathbf{r}$. Therefore the various projections cannot be trusted with the same degree of credibility, since they are affected by different errors. The incompatibility of various SG measurements becomes manifest in quantum theory as the corresponding operators (1) do not commute for different orientations $\mathbf{a}^j$. Such data cannot be obtained in the course of the same measurement, but may be collected by repeated experiments. Thus an optimal procedure must predict an unknown state and simultaneously take into account data fluctuations. This indicates the inevitable nonlinearity of such a kind of algorithm. MaxLik estimation does this job and fits the data to a quantum state. As will be demonstrated in the following section, the synthesis of incompatible measurements may be considered as a unique concept of measuring quantum states.

II. SPIN ESTIMATION

It is assumed that all spin-$\frac{1}{2}$ particles supplied by the source are prepared in the same mixed state, and an ideal lossless SG analysis is performed repeatedly on a system of N such particles with varying orientation of the SG device. Provided that the detection has been done with M different orientations, $N*M$ particles have been used altogether and the unknown quantum state should be determined. The results of the measurement may be characterized by the settings $\pm\mathbf{a}^j$ of the SG apparatus and by the relative frequencies of the outcomes $1/2(1\pm X_j) = n_{j,\pm}/N$. Then the problem is to find the state(s) that fit(s) the data in an optimal way. One might be tempted to sample and invert the probability, predicted by quantum theory, as it is done in the case of Eq. (8). Because each SG detection is represented by a complete measurement, the sum of the relations (5) over all orientations of the SG device reads

$$\frac{1}{M}\sum_j^M |\mathbf{a}^j\rangle\langle\mathbf{a}^j| + |-\mathbf{a}^j\rangle\langle-\mathbf{a}^j| = \hat{1}, \tag{9}$$

where $\hat{1}$ denotes the unit matrix. In general, however, the expected relations

$$\mathrm{Tr}\left\{\hat{\rho}\frac{1}{M}|\pm\mathbf{a}^j\rangle\langle\pm\mathbf{a}^j|\right\} = \frac{1}{2M}(1\pm X_j) \tag{10}$$

cannot be fulfilled because the system is overdetermined and the data are fluctuating. Hence, the probabilities cannot be mapped in a straightforward manner just according to the relative frequencies of outcomes.

The MaxLik principle serves as a tool to overcome this problem. It allows one to find the most probable state consistent with the data. As the measure of probability the likelihood functional may be constructed, which corresponds to the product of all probabilities for all detected data,

$$\mathcal{L}(\hat{\rho}) = \prod_j \left(\langle\mathbf{a}^j|\hat{\rho}|\mathbf{a}^j\rangle\right)^{N(1+X_j)/2}\left(\langle-\mathbf{a}^j|\hat{\rho}|-\mathbf{a}^j\rangle\right)^{N(1-X_j)/2}. \tag{11}$$

The extremal states of this likelihood functional satisfy the nonlinear operator equation [8],

$$\frac{1}{2M}\sum_j\left[(1+X_j)\frac{|\mathbf{a}^j\rangle\langle\mathbf{a}^j|}{\langle\mathbf{a}^j|\hat{\rho}|\mathbf{a}^j\rangle} + (1-X_j)\frac{|-\mathbf{a}^j\rangle\langle-\mathbf{a}^j|}{\langle-\mathbf{a}^j|\hat{\rho}|-\mathbf{a}^j\rangle}\right]\hat{\rho} = \hat{\rho}. \tag{12}$$

The quantum state shall be represented by its respective polarization. Using the relation (1), multiplying both sides by σ_k and performing the trace, one obtains

$$R(\mathbf{r})\mathbf{r} + \mathbf{K}(\mathbf{r}) + i\mathbf{K}(\mathbf{r})\times\mathbf{r} = \mathbf{r}, \tag{13}$$

where the functions $R(\mathbf{r})$ and $K(\mathbf{r})$ are defined as

$$R(\mathbf{r}) = \frac{1}{2M}\sum_j\left(\frac{1+X_j}{1+\mathbf{a}^j\cdot\mathbf{r}} + \frac{1-X_j}{1-\mathbf{a}^j\cdot\mathbf{r}}\right),$$

$$\mathbf{K}(\mathbf{r}) = \frac{1}{2M}\sum_j\left(\frac{1+X_j}{1+\mathbf{a}^j\cdot\mathbf{r}} - \frac{1-X_j}{1-\mathbf{a}^j\cdot\mathbf{r}}\right)\mathbf{a}^j.$$

Since real and imaginary parts are independent from each other; the real part of Eq. (13) is sufficient to derive the necessary conditions. Hence the final equation for the polarization vector reads

$$R(\mathbf{r})\mathbf{r} + \mathbf{K}(\mathbf{r}) = \mathbf{r}, \tag{14}$$

which can be solved conveniently by iteration. Starting from the center $\mathbf{r}=0$ of the Poincaré sphere, the left side of Eq. (14) yields the first correction, which in turn may be used as input for a subsequent iteration step. This procedure provides a rapidly converging algorithm for MaxLik fitting of an unknown quantum state inside the Poincaré sphere.

An equivalent result may be derived by parametrizing the likelihood function directly in terms of polarization. The rel-

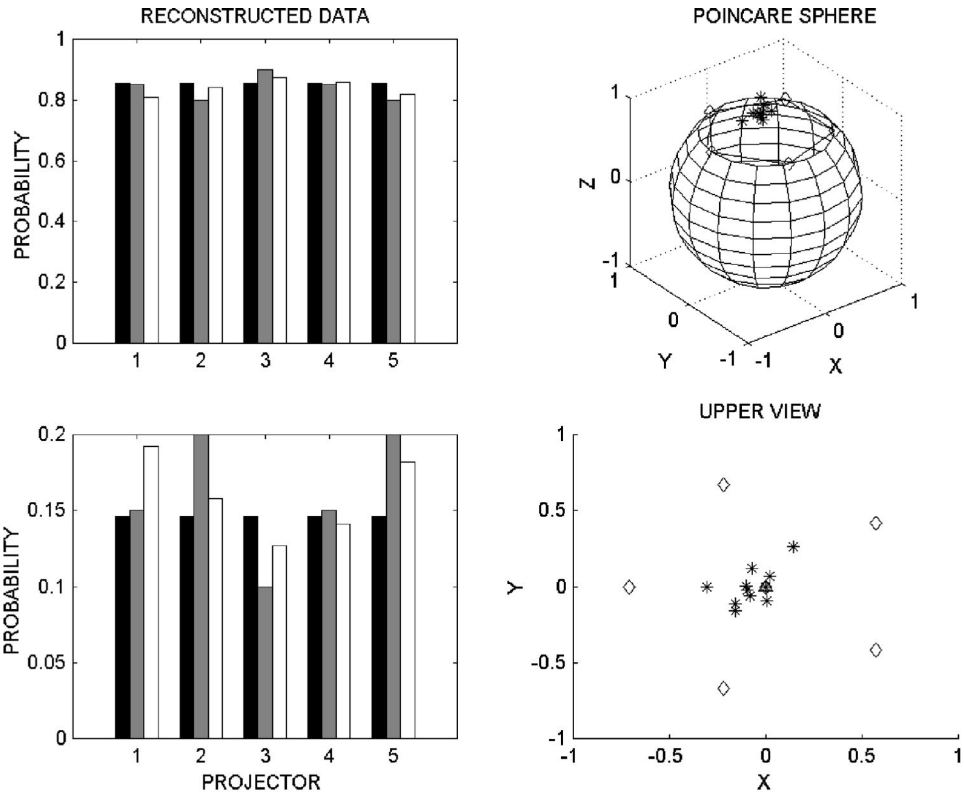

FIG. 1. Results and interpretation of the spin-state reconstruction for numerical simulation of the Stern-Gerlach detection with five different orientations of the apparatus (see the text).

evant part of the likelihood function corresponding to the observation of particular data then can be written as

$$\mathcal{L}(\mathbf{r})=\prod_j (1+\mathbf{r}\cdot\mathbf{a}^j)^{N(1+X_j)/2}(1-\mathbf{r}\cdot\mathbf{a}^j)^{N(1-X_j)/2}. \quad (15)$$

The vector $\mathbf{r}$ parametrizes an arbitrary unknown polarization inside the Poincaré sphere and the products runs over all M directions. The standard statistical approach using MaxLik $(\partial/\partial\mathbf{r})\ln\mathcal{L}$ leads to a vector equation for the extreme value of the polarization [4]

$$\sum_j \frac{X_j-\mathbf{a}^j\cdot\mathbf{r}}{1-(\mathbf{a}^j\cdot\mathbf{r})^2}\mathbf{a}^j=0. \quad (16)$$

Equation (14) is equivalent to Eq. (16). Indeed, Eq. (16) is nothing else than $\mathbf{K}(\mathbf{r})=0$, implying the relation $R(\mathbf{r})=1$. Vice versa, Eq. (14) could be rewritten in the form of Eq. (16) as well.

In Fig. 1 the results of numerical simulations are shown. Stern-Gerlach detection is simulated here for projection of an ‘‘unknown’’ state (north pole on the Poincaré sphere) in five different directions. Each ‘‘measurement’’ is done with 20 impinging particles registered either with spin up (upper left panel) or with spin down (lower left panel). Both left panels show typical values for a single experiment. For each projector three bars are plotted: the first bars (black) show the true value of the probability whereas the second bars (gray) exhibit the statistical fluctuations of the ‘‘counted’’ events around the respective true value. Finally the hollow bars represent the results of the reconstruction—the statistics of the reconstructed state corresponding to the given projector. Notice here that upper and lower panels are complementary and the sum of the respective true probabilities is always exactly 1. The right panels visually present the results obtained by repeating the experiment 10 times. Diamond symbols denote the positions of five projectors on the Poincaré sphere. Orthogonal projectors in the opposite directions are not depicted. Stars indicate the position of the reconstructed states. The true state corresponds to the north pole. Viewing the sphere from the top yields the lower right panel.

Quantum physical formulation inherently requires a nontrivial interpretation which can hardly be recognized from the equation for the polarization vector (16). Since the above scheme determines a quantum state, a generalized measurement concept described by a probability operator measure (POM) [9] must exist, the result of which is the quantum state. Indeed, such a probability operator measure can be found by proper renormalization of the original SG analysis [8]. Let us define the POM as renormalized SG projectors

$$|\pm\mathbf{a}^j\rangle\langle\pm\mathbf{a}^j|_R=\frac{1\pm X_j}{2M\langle\pm\mathbf{a}^j|\hat{\rho}_e|\pm\mathbf{a}^j\rangle}|\pm\mathbf{a}^j\rangle\langle\pm\mathbf{a}^j|, \quad (17)$$

for each index j. The closure relation then reads

$$\sum_j^M |\mathbf{a}^j\rangle\langle\mathbf{a}^j|_R+|-\mathbf{a}^j\rangle\langle-\mathbf{a}^j|_R=\hat{1}_\rho, \quad (18)$$

and the renormalized POM fulfills the conditions

$$\mathrm{Tr}\{\hat{\rho}_e|\pm\mathbf{a}^j\rangle\langle\pm\mathbf{a}^j|_R\}=\frac{1}{2M}(1\pm X_j). \quad (19)$$

Here $\hat{\rho}_e$ denotes the extremal state—a solution of Eq. (12). Relation (18) indeed coincides with the equation for extremal states (12), whereas the condition for expectation values (19) is fulfilled as an identity. The reconstruction is done in that subspace where the renormalized POM reproduces the identity operator. Specifically this means that the identity operator on the right-hand side of Eq. (18) is spanned by the one-dimensional subspace only (i.e., by a single ray), provided that the extremal state $\hat{\rho}_e$ is a pure state. For a general extremal density matrix the reconstruction is accomplished in the whole two-dimensional Hilbert space. The distinction between relations (9),(10) and (18),(19) characterizes the subtleties of quantum state reconstruction. The MaxLik solution may be also interpreted in the language of probabilities. The detected data $n_{i,\pm}$ sample different binomial probability distributions for $i=1,\ldots,M$. MaxLik estimation finds a common multinomial distribution and thus allows sampling of the data with seemingly the highest likelihood.

The method developed here may be compared with the existing approaches. Jaynes's maximum entropy principle (MaxEnt) [10] has been applied as well to the estimation of spin-$\frac{1}{2}$ states in Refs. [3,6]. In general, however, these two methods are not equivalent. The MaxLik method seeks for the most likely solution consistent with the data, whereas with the MaxEnt method one searches for the worst solution still consistent with the data. This difference may be attributed to the different prior information in the maximum probability principle [11]. But this is not the only difference. External conditions of both approaches differ substantially. MaxLik has been applied to measurements with many projectors. However, the same conditions cannot be applied for the MaxEnt approach because there only three free parameters are necessary for the determination of an unknown spin state. Therefore, the conditions defined by Eq. (10) cannot be fulfilled in general. Obviously the MaxEnt concept is not applicable if more than three independent conditions are imposed upon the density matrix of a spin-$\frac{1}{2}$ system.

In the papers of Ref. [12] an optimal strategy for measuring an unknown two-state system is investigated. As a result, an optimal coherent measurement may be predicted. On the other hand, we aimed to optimize not the measurement itself but its mathematical treatment. This seems to be reasonable from the experimentalist's point of view because it is questionable how to do a general measurement described by a POM. In the Ref. [13] the information content of the large ensemble of identically prepared quantum systems is investigated. It is proved that for the spin-$\frac{1}{2}$ ensembles the Cramér-Rao bound can always be attained and optimal measurements may be well approximated by adaptive separable ones. This corresponds well to the results presented in this Brief Report. Our analysis is not restricted to the asymptotic domain and emphasizes the relationship between quantum theory and mathematical statistics. As we have demonstrated, for a given measurement the MaxLik approach provides an optimal treatment in the sense that it reproduces a generalized quantum measurement.

III. SUMMARY

The synthesis of incompatible observations has been evaluated using the concept of MaxLik estimation. It defines a generalized measurement of a quantum state. The MaxLik procedure provides a quick recipe for an experimentalist. No *a priori* knowledge about the spin state is needed. The iterative algorithm based on the solution of Eq. (14) is capable of finding the polarization of the most probable state, provided that many detections with various settings of SG apparatus have been done. In the near future the formalism developed here will be applied to the investigation of various problems, such as the spin-state estimation in neutron depolarization experiments, the estimation of quantum states inside split beam neutron interferometers, or the analysis of entangled states.

ACKNOWLEDGMENTS

This work was supported by TMR Network No. ERB FMRXCT 96-0057 ''Perfect Crystal Neutron Optics'' of the European Union, by the East-West program of the Austrian Academy of Science, and by the Czech Ministry of Education Grants No. VS 96028 and No. CEZ J14/98.

[1] J. J. Sakurai, *Modern Quantum Mechanics* (Addison-Wesley, Reading, MA, 1994).
[2] B.-G. Englert, J. Schwinger, and M. O. Scully, Found. Phys. **18**, 1045 (1988); J. Schwinger, M. O. Scully, and B.-G. Englert, Z. Phys. D: At., Mol. Clusters **10**, 135 (1988); M. O. Scully, B.-G. Englert, and J. Schwinger, Phys. Rev. A **40**, 1775 (1989).
[3] P. R. Dukes and E. G. Larson, in *Maximum Entropy and Bayesian Methods*, edited by W. T. Grandy and L. H. Schick (Kluwer Academic, Dordrecht, 1991), p. 161.
[4] K. R. W. Jones, Phys. Rev. A **50**, 3682 (1994).
[5] S. Weigert, Phys. Rev. A **45**, 7688 (1992); **45**, 7688 (1992); **53**, 2078 (1996).
[6] V. Bužek, G. Adam, and G. Drobný, Phys. Rev. A **54**, 804 (1996); Ann. Phys. (NY) **245**, 37 (1996).
[7] Z. Hradil, Phys. Rev. A **55**, R1561 (1997).
[8] Z. Hradil, J. Summhammer, and H. Rauch, Phys. Lett. A **261**, 20 (1999).
[9] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[10] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).
[11] B. R. Frieden, in *Maximum Entropy and Bayesian Methods in Inverse Problems*, edited by C. R. Smith and W. T. Grandy, Jr. (Reidel, Dordrecht, 1985), p. 133.
[12] S. Massar and S. Popescu, Phys. Rev. Lett. **74**, 1259 (1995); J. I. Latorre, P. Pascual, and R. Tarrach, Phys. Rev. Lett. **81**, 1351 (1998); G. Vidal, J. I. Latorre, P. Pascual, and R. Tarrach, Phys. Rev. A **60**, 126 (1999).
[13] O. E. Barndorff-Nielsen and R. D. Gill, e-print quant-ph/9808009; R. D. Gill and S. Massar, Phys. Rev. A **61**, 42312 (2000).