ОПТИКА И СПЕКТРОСКОПИЯ, 2001, том 91, № 3, с. 402–405

КВАНТОВАЯ ИНФОРМАЦИЯ И КВАНТОВЫЕ ИЗМЕРЕНИЯ

УДК 535.14

QUANTUM THEORY OF INCOMPATIBLE OBSERVATIONS[1]

© 2001 г. Z. Hradil

Department of Optics, Palacký University, Olomouc 772 07, Czech Republic

E-mail: hradil@optnw.upol.cz

Received November 17, 2000

Quantum theory allows to measure incompatible observables sequentially in the course of repeated measurements. To get information about the observed system, all the observations must be synthetized. This is the main idea of quantum tomographical methods. As shown in this contribution, the maximum likelihood principle provides the best measure for relating the experimental data with predictions of quantum theory. Synthesis of incompatible observations appears to be a novel quantum measurement described by a positive operator-valued measure. Besides this the procedure finds the optimal state of the system, which fitts such a measurement in optimal way.

INTRODUCTION

Quantum theory describes events on the most fundamental level currently available. The synthesis of information from mutually incompatible quantum measurements plays the key role in testing of the theory. Bell inequalities with a pair of spin-1/2 particles may be considered as an example. Correlations for two settings of polarizators are measured on each component of the pair. These observations are incompatible, since they cannot be obtained in the same simultaneous measurement. Information must be collected in subsequent measurements, when the experiment is repeated with different setting of polarizators. The purpose of this contribution is to develop general quantum theory of such observations. As will be shown, there is a unique relationship between quantum theory and the mathematical statistics: Quantum theory prefers the relative entropy (maximum likelihood principle) as the proper measure for evaluation of the distance between measured data and probabilities defined by quantum theory. For an experimentalist working in quantum physics, it means that data should be fitted to the theory preferably using the maximum likelihood estimation.

For the sake of simplicity and brevity we assume a discrete spectrum of the observed variable. This corresponds to the case of sharp and precise quantum measurements. Notice, however, that these ideal assumptions are not detrimental. The more realistic case of observables with a continuous spectrum and finite experimental precision can be incorporated into this framework by replacing the corresponding projectors by a probability-valued operator measure (POVM) [1, 2]. Our main result is independent of a particular implementation of the quantum measurement and works in the very general case as well. In the following, we shall use the Dirac notation.

Let us start the exposition with the simplest model. Quantum mechanics of spin-1/2 particles often serves as an illustrative example of key quantum physical concepts in standard textbooks of theoretical physics [3]. The importance of spin-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-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. Let us review briefly the basic properties of spin-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

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

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

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

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

[1] Статья представлена автором на английском языке.

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

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

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

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

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 $|\mathbf{r}|^2 \le 1$. 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 [4]. 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 $\sqrt{N(1-(\mathbf{r}\cdot\mathbf{a}^j)^2)}/2$ depending on the deviations between projections and the true (but unknown!) direction of the spin $\mathbf{r}$. Therefore the data from 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 takes into account data fluctuations. This indicates the inevitable nonlinearity of such a kind of algorithm. As will be demonstrated in the following section, the synthesis of incompatible measurements may be considered as a novel concept of measuring quantum states.

GENERAL THEORY

Let us review briefly the standard theory treated in the textbooks [3]. Any observation is represented by a hermitian operator $\hat{A}$, whose spectrum determines the possible results of the measurement

$$\hat{A}|a\rangle = a|a\rangle. \tag{9}$$

Eigenstates are orthogonal $\langle a|a'\rangle = \delta_{aa'}$ and the corresponding projectors provide the closure relation

$$\sum_a |a\rangle\langle a| = \hat{1}. \tag{10}$$

Projectors predict the probability for detecting a particular value of the q-variable a represented by the operator $\hat{A}$ as $p_a = \langle a|\rho|a\rangle$, provided that the system has been prepared in a quantum state ρ. This mathematical picture corresponds to the experimental reality in the following sense: When the measurement represented by the operator $\hat{A}$ is repeated N times on identical copies of the system, the number a particular output a is collected N_a times. The relative frequencies $f_a = \frac{N_a}{N}$ will sample the true probability as $f_a \longrightarrow p_a$ fluctuating around them. The exact values are reproduced only in the asymptotical limit $N \longrightarrow \infty$. Experimentalist's knowledge may be expressed in the form of a diagonal density matrix

$$\hat{\rho}_{\mathrm{est}} = \sum_a f_a|a\rangle\langle a|, \tag{11}$$

provided that error bars of the order $1/\sqrt{N}$ are associated with the sampled relative frequencies. This should be understood as mere rewriting of the experimental data $\{N, N_a\}$. Similar knowledge may be obtained by observations, which can be parameterized by operators diagonal in the $|a\rangle$ basis, i.e. by operators commuting with operator $\hat{A}$. But the possible measurement of *noncommuting* operators yields new information, which cannot be derived from the measurement of $\hat{A}$.

Consider now the sequential synthesis of various noncommuting observables. In this case, several operators $\hat{A}_j$, $j = 1, 2, \ldots$ will be measured by probing of the system N times together. Now, one expects to gain more than just the knowledge of the diagonal elements of the density matrix in some a priori given basis. This sequential measurement of noncommuting observables should be distinguished from the similar problem of approximate simultaneous measurement of noncommuting observables[2]. As in the case of the measurement of a hermitian operators, the result of sequential measurements of noncommuting operators may be represented by a series of projectors $|y_i\rangle\langle y_i|$. This should be accompanied by relative frequencies f_i indicating how many times a particular output i has been registered, $\sum_i f_i = 1$. Various states need not be orthogonal $\langle y_i|y_j\rangle \neq \delta_{ij}$, in contrast to the previous case of a hermitian operator. However, this substantial difference has its deep consequences. The result of the measurement cannot be meaningfully represented in the same manner as previously. For example, direct linking of probabilities with relative frequencies used in standard reconstructions [5, 6] $\rho_{ii} = f_i$, $\rho_{ii} = \langle y_i|\hat{\rho}|y_i\rangle$, may appear as inconsistent, since the system of linear equations is overdetermined, in general.

To develop the novel approach, let us assume the existence of a quantum measure $F(\rho_{ii}|f_i)$ parameterizing the distance between measured data and probabilities. Then we will search for the state(s) located in the closest neighborhood of the data. A general state may be parameterized in its diagonal basis as

$$\hat{\rho} = \sum_i r_i|\varphi_i\rangle\langle\varphi_i|. \tag{12}$$

The equation for the extremal states may be found analogously to the treatment developed in [4, 7]. Particularly, the formal necessary condition for extremal solution reads

$$\frac{\delta F(\rho_{ii}|f_i)}{\delta\hat{\rho}} = 0. \tag{13}$$

Since the density matrix is parametrized according to the relation (6) with the help of independent (orthogonal) states $|\varphi_k\rangle$, the variation may be done along these rays yielding the system of coupled equations $\frac{\delta F(\rho_{ii}|f_i)}{\delta\langle\varphi_k|} = 0$ for any allowed k. Using the relation

$$\frac{\delta F(\rho_{ii}|f_i)}{\delta\langle\varphi_k|} = \sum_i \frac{\partial F(\rho_{ii}|f_i)}{\delta\rho_{ii}}|y_i\rangle\langle y_i|\varphi_k\rangle,$$

the system of equations may be rewritten as the equation for the density matrix

$$\sum_i \frac{\partial F}{\partial\rho_{ii}}|y_i\rangle\langle y_i|\hat{\rho} = \lambda\hat{\rho}, \tag{14}$$

where λ is a Lagrange multiplier. The normalization condition $\mathrm{Tr}\,\hat{\rho} = 1$ sets its value to

$$\lambda = \sum_i \frac{\partial F}{\partial\rho_{ii}}\rho_{ii}.$$

Any composed function $G(F(\rho_{ii}|f_i))$ fulfills the same extremal equation (14) with the Lagrange multiplier rescaled as $\lambda\frac{dG}{dF}$. Without loss of generality it is therefore enough to consider the normalization condition $\lambda = 1$.

The extremal equation (14) has the form of a decomposition of the identity operator on the subspace, where the density matrix is defined by

$$\sum_i \frac{\partial F}{\partial\rho_{ii}}|y_i\rangle\langle y_i| = \hat{1}_\rho. \tag{15}$$

This resembles the definition of POVM characterizing a generalized measurement [1, 2]. To link the above extremalization with quantum theory, let us postulate the natural condition for the quantum expectation value

$$\mathrm{Tr}\left(\frac{\partial F}{\partial\rho_{ii}}|y_i\rangle\langle y_i|\hat{\rho}\right) = f_i. \tag{16}$$

This assumption seems to be reasonable. The synthesis of sequential incompatible observations may be regarded as a new measurement scheme, namely the measurement of the quantum state.

The quantum measure F then fulfills the differential equation

$$\frac{\partial F}{\partial\rho_{ii}}\rho_{ii} = f_i \tag{17}$$

and singles out the solution in the form

$$F(\rho_{ii}|f_i) = \sum_i f_i \ln\rho_{ii}. \tag{18}$$

[2] The approximate simultaneous measurement of noncommuting operators $[\hat{A}, \hat{B}] \neq 0$ can always be represented by measurement of commuting operators $\hat{\mathcal{A}}$, $\hat{\mathcal{B}}$ defined on the extended Hilbert space $\mathcal{H} = H_s \otimes H_a$, where H_s, H_a is the space of original system and space of auxiliary field (ancilla), respectively.

This is nothing else than the log likelihood or Kullback–Leibler relative information [8][3]. Formal requirements of quantum theory, namely the interpretation of the extremal equation as a POVM, result in the concept of maximum likelihood in mathematical statistics. The analogy between the standard quantum measurement associated with a single hermitian operator and a series of sequential measurements associated with many noncommuting operators is apparent now. The former determines the diagonal elements in the basis of orthonormal eigenvectors, whereas the latter estimates not only the diagonal elements, but the diagonalizing basis itself. This is the difference between measurement of the quantum observable $\hat{A}$ and measurement of the quantum state. In this sense maximum likelihood estimation may be considered as a new quantum measurement. The observed quantum state is given by the solution of the nonlinear operator equation

$$\hat{R}(\hat{\rho})\hat{\rho} = \hat{\rho}, \tag{19}$$

where

$$\hat{R}(\hat{\rho}) = \sum_i \frac{f_i}{\rho_{ii}} |y_i\rangle\langle y_i|.$$

Extremal equation is, in fact, the completeness relation of a POVM, expectation values of which are the measured data $\{f_i\}$ [9]. The equation of this type is well known in mathematical statistics and its solution is given by the so called EM (expectation–maximization) algorithm [10]. In quantum domain, the EM algorithm must be completed by a unitary transformation changing the diagonalizing basis of extremal density matrix [11].

Maximum likelihood has been used recently for solution of various problems in quantum theory. Ideal phase concepts have been considered from the viewpoint of maximum likelihood in Refs. [12]. Special cases of the solution (19) have been discussed for the operational phase concepts [13], determination of diagonal elements of the density matrix [14]. Reconstruction of the 1/2 spin state using as an introductory example has been considered in the Ref. [15]. A numerical technique for maximum likelihood estimation of density matrices has been suggested in Ref. [16].

SUMMARY

The quantum interpretation offers a new viewpoint on the maximum likelihood estimation. This method is customarily considered as just one of many estimation methods, unfortunately one of the most complicated ones. It is often considered as rather subjective, since likelihood quantifies the degree of belief in a certain hypothesis. Any physicist, an experimentalist above all, would perhaps use as his first choice another fitting procedure, for example the least–squares method. However, such fitting will not reveal the structure of quantum measurement. Only the maximum likelihood estimation interprets the measured data as expectation values of some new POVM. In this sense the maximum likelihood seems to be unique and exceptional.

This work was supported by TMR Network ERB FMRXCT 96–0057 “Perfect Crystal Neutron Optics” of the European Union and by the grants of Czech Ministry of Education VS 96028 and CEZ J14/98.

[3] Notice the asymmetry between the arguments f and p in definition of Kullback–Leibler relative information $K(f/p) = \sum_i f_i \ln(f_i/p_i)$. In the paper of B.R. Frieden, in *Maximum Entropy and Bayesian Methods in Inverse Problems*, edited by C.R. Smith, W.T. Grandy Jr. (Reidel, Dordrecht 1985), p. 133, the term Kullback–Leibler norm is used for opposite ordering of data and probabilities. The case discussed here is called generalized Burg principle.

REFERENCES

1. C. W. Helstrom, Quantum Detection and Estimation Theory (Academic Press, New York, 1976).
2. A. S. Holevo, Probabilistic and Statistical Aspects of Quantum Theory (North–Holland Publ., Amsterdam; 1982).
3. J. J. Sakurai, *Modern Quantum Mechanics* (Addison–Wesley, 1994).
4. Z. Hradil, J. Summhammer, and H. Rauch, Phys. Lett. **A 261**, 20 (1999).
5. U. Leonhardt, *Measuring of the Quantum State of Light* (Cambridge Press, 1997).
6. D.-G. Welsch, W. Vogel, and T. Opatrny, Progress in Optics, Ed. by E. Wolf, **39** (1999).
7. Z. Hradil, *Phys.* Rev. **A 55**, R1561 (1997).
8. S. Kullback and R.A. Leibler, Ann. of Math. Stat. **22**, 79 (1951).
9. Z. Hradil, J. Summhammer, Quantum theory of incompatible observations, submitted to J. Phys. A.
10. A. P. Dempster, N. M. Laird, and D. B. Rubin, J.R. Stat. Soc. B **39**, 1 (1977); Y. Vardi and D. Lee, J.R. Stat. Soc. B **55**, 569 (1993).
11. Z. Hradil, J. Rehacek, and M. Jezek in preparation (to be published).
12. J. H. Shapiro, S. R. Shepard, and N. C. Wong, Phys. Rev. Lett. 62, 2377 (1989); A.S. Lane, S.L. Braunstein, and C.M. Caves, Phys. Rev. A **47**, 1667 (1993).
13. J. Rehacek Z. Hradil, M. Zawisky, S. Pascazio, H. Rauch, and J. Perina, Phys. Rev. A **60**, 473 (1999).
14. Z. Hradil and R. Myska, Acta phys. slov. **48**, 199 (1998); D. Mogilevtsev, Z. Hradil, and J. Peina, J. Mod. Opt. **44**. 2261 (1997); D. Mogilevtsev, Z. Hradil, and J. Peina, Quantum Semiclass. Opt. **10**, 345. (1998).
15. Z. Hradil, J. Summhammer, G. Badurek, and H. Rauch, Spin state reconstruction, quant-ph/9911068, to be published in Phys. Rev. A.
16. K. Banaszek, G.M. D’Ariano, M.G.A. Paris, and M. F. Sacchi, Phys. Rev. A **61**, 010304(R) (1999).