

Some Problems on the Measurement of Quantum Observables and Determination of Joint Entropy in Quantum Statistics

V. D. Vainshtein and S. D. Tvorogov

The Institute of Atmospheric Optics, USSR Academy of Sciences, Siberian Branch,
Tomsk, USSR

Received January 20, 1975

Abstract. In the paper the equivalency of a successive measurement of observables a and b on the Hilbert space H and a specially organized measurement of the observable $a \otimes b$ on the Hilbert space $\mathcal{H} = H \otimes H$ is determined. The state, wherein the measurement on \mathcal{H} is performed, is shown to be an operator analog of classical joint density of probability distribution. The functionals such as joint and conventional entropy are constructed; the entropy defect of a quantum ensemble and the information quantity contained in quantum measurement are determined.

A number of problems can be formulated where the necessity arises for extension of statistical interpretation of quantum mechanics, construction of the objects of the probability theory such as joint and conventional densities of distribution, determination of functionals generalizing the classical notions of joint and conventional entropy for a quantum case. In particular, these are the problems of the information quantum theory, quantum fluctuation theory, some problems on nonequilibrium statistics. The specific feature of the problems in the list is that the probabilistic distributions (in a classical sense) obtained on the basis of quantum approach are insufficient for their strict formulation. We need to have the operator analogs of joint and conventional distributions, i.e., corresponding generalization of a density matrix.

Realization of the mentioned generalization requires successive inclusion of the measurement operation and measuring instrument in the theory. In the present paper the equivalency of a successive measurement of observables a and b on the Hilbert space H and a specially organized measurement of the observable $a \otimes b$ on the space $\mathcal{H} = H \otimes H$ is determined following the method developed in [1]. The state wherein the measurement is performed on is the above generalization of a density matrix. The latter enables one to determine strictly the notions of joint and conventional entropy in quantum statistics and also the functionals such as the entropy defect of a quantum ensemble and the information quantity contained in quantum measurement.

The initial conceptions for this are the following [1]. The space of self-adjoint trace class operators on Hilbert space H forms a real Banach space V under trace norm. A state space is a triple $(V, K, 1)$ consisting of a real Banach space V , a positive closed cone K which generates V and a linear functional on V such that $1(v) = \langle 1, v \rangle = \|v\|$ for all $v \in V$; V is endowed with a partial ordering by putting

$w \geq v$ if and only if $(w - v) \in K$. A linear functional φ on V is said to be positive if $\langle \varphi, v \rangle \geq 0$ for all $v \in K$. Every positive linear functional on a state space is necessarily continuous and set of all positive linear functionals forms a closed cone K^* in Banach dual space V^* endowed with the weak* topology. A linear mapping $T: V \rightarrow V$ is said to be positive if $T(K) \subseteq K$. A positive linear mapping on state space is necessarily bounded and the set of all positive linear mappings forms a closed cone $\mathcal{L}^+(V, V)$ in the topological linear space $\mathcal{L}(V, V)$ of all bounded linear mappings of V into itself, endowed with the strong operator topology.

A state v is a non-zero element of K . An observable is a triple (X, \mathcal{A}, a) consisting of a set X with a σ -field \mathcal{A} of subsets of X and a mapping $a: \mathcal{A} \rightarrow V^*$ satisfying

$$(A1) \quad 0 \leq a(E) \leq a(X) \text{ for all } E \in \mathcal{A};$$

$$(A2) \quad a(X) = 1;$$

$$(A3) \quad \text{for each countable family of pairwise disjoint sets in } \mathcal{A}$$

$$a\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} a(E_i),$$

where the right-hand side converges in weak* topology of V^* . The Borel space (X, \mathcal{A}) is called the value space of the observable.

An instrument is a triple $(X, \mathcal{A}, \mathcal{E})$ consisting of a Borel space (X, \mathcal{A}) and a mapping $\mathcal{E}: \mathcal{A} \rightarrow \mathcal{L}^+(V, V)$ satisfying

$$(B1) \quad \langle 1, \mathcal{E}(X)v \rangle = \langle 1, v \rangle \text{ for all } v \in V;$$

$$(B2) \quad \text{for each countable family } \{E_i\} \text{ of pairwise disjoint sets in}$$

$$\mathcal{E}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathcal{E}(E_i)$$

where the right-hand side converges in strong operator topology.

The instrument is said to be discrete if its value space is discrete.

To every instrument \mathcal{E} on the value space (X, \mathcal{A}) there is a unique observable such that (Theorem 1, [1])

$$\langle 1, \mathcal{E}(E)v \rangle = \langle a(E), v \rangle$$

for all $v \in V$ and $E \in \mathcal{A}$. Every observable a with the value space (X, \mathcal{A}) is determined by at least one instrument and the equation

$$\mathcal{E}(E)v = \sum_{i=1}^{\infty} \langle a(E \cap E_i), v \rangle v_i$$

gives the observable a connected with the instrument \mathcal{E} . Here $\{E_i\}$ is the set of pairwise disjoint subsets of X and $\{v_i\}$ is the set of states such that $\langle 1, v_i \rangle = 1$ for every v_i .

The definitions introduced have direct statistical interpretation in terms of statistical ensembles. An ensemble corresponds to a state. (A state with the unit trace norm is a density matrix.) An instrument \mathcal{E} measuring a physical quantity taking values in X , is a family of filters: the filter corresponding to $\mathcal{E}(E)$ transmits to the output only those subsets wherein the physical quantity value lies in E , and in so doing transforms the input state v into the output state $v' = \mathcal{E}(E)v$.

Various filters possess the property of additivity: if E_1 and E_2 are disjoint, the output state $v' = \mathcal{E}(E_1 \cup E_2)v = \mathcal{E}(E_1)v + \mathcal{E}(E_2)v$.

To sum up, for each E the mapping $\mathcal{E} \rightarrow \langle 1, \mathcal{E}(E)v \rangle / \langle 1, v \rangle$ is a probability measure on X , i.e., the probability that the observable a should take value in E while being measured in the state v .

Now we shall proceed to consideration of complex measurements. Below we shall bound the class of observables by the case of discrete spectrum. Let H and $\mathcal{H} = H \otimes H$ be the Hilbert spaces and a be an observable on H with the discrete value space (X, \mathcal{A}) and b be an observable on H with the discrete value space (Y, \mathcal{B}) . It is easy to see the product $a \otimes b = d$ is the observable on $\mathcal{H} = H \otimes H$ with the discrete value space $(Z, C) = (X \otimes Y, \mathcal{A} \otimes \mathcal{B})$.

According to Theorem 10 [1]

$$\begin{aligned} a(E) &= \sum_{x \in E} P_x; & \mathcal{E}(E)v_H &= \sum_{x \in E} P_x v_H P_x; & E \in \mathcal{A}, v_H &\in K_H; \\ b(F) &= \sum_{y \in F} J_y; & \mathcal{F}(F)v_H &= \sum_{y \in F} J_y v_H J_y; & F \in \mathcal{B}, v_H &\in K_H; \\ d(Z) &= \sum_{w \in Z} t_w; & \mathcal{G}(Z)^v &= \sum_{w \in Z} t_w v t_w; & z \in C; v \in K \supset K_H \otimes K_H, \end{aligned}$$

where P_x, J_y are the corresponding orthogonal projectors in H , and $t_w = (P_x \otimes J_y)$ is the orthogonal projector in $\mathcal{H} = H \otimes H$.

In [1] it is shown that successive measurement of the observable a and observable b following a associates with a composition of instruments $\mathcal{Y} = \mathcal{F} \circ \mathcal{E}$ such that, while measuring the observable b following the observable a in the state v_H , the value $\langle 1_H, \mathcal{F}(F)\mathcal{E}(E)v_H \rangle / \langle 1_H, v_H \rangle$ is a measure of joint probability of obtaining the value a on E and b on F .

We shall choose the state v in the form

$$v = \sum_x \langle P_x, v_H \rangle P_x \otimes P_x.$$

It is easy to check, in this case we have

$$\langle 1, \mathcal{G}(E \otimes F)v \rangle / \langle 1, v \rangle = \langle 1_H, \mathcal{F}(F)\mathcal{E}(E)v_H \rangle / \langle 1_H, v_H \rangle.$$

Hence the measurement in \mathcal{H} by the instrument \mathcal{G} in the state v gives the same probability distribution that the measurement by the superposition $\mathcal{F} \circ \mathcal{E}$ in the state v_H in H .

Let us consider the state v in detail. For this consideration we shall use the following terms. We assume that the measurement $\mathcal{E}(X)$ is being made in the state $v_H \in K_H$. As a result, the state v_H will be transformed into the state

$$v_{Ha} = \mathcal{E}(X)v_H = \sum_x \langle P_x, v_H \rangle P_x.$$

Let Q be the class of linear mappings R of the space V into itself

$$R: v \rightarrow R(v), \quad v \in V,$$

and Q^* be the class of linear mappings R^* of the space V^* into itself

$$R^*: B \rightarrow R^*(B) \quad B \in V^*$$

so that $R^*(1) = 1$ and

$$\langle B, R(v) \rangle = \langle R^*(B), v \rangle.$$

The mapping R is obviously positive

$$R(K) \subseteq K.$$

It is easy to verify that if v_H is a state on H then $v_R = \sum_x \langle P_x, v_H \rangle P_x \otimes R(P_x)$ is a state on $\mathcal{H} = H \otimes H$.

Partial trace norms ν_R are the states on H .

$$\langle 1_H \otimes \cdot, \nu_R \rangle = \sum_x \langle P_x, \nu_H \rangle \langle 1_H, P_x \rangle R(P_x) = R(\nu_{Ha}) \in K_H;$$

$$\langle \cdot \otimes 1_H, \nu_R \rangle = \sum_x \langle P_x, \nu_H \rangle P_x \langle 1_H, R(P_x) \rangle = \nu_{Ha} \in K_H.$$

In the state ν_R we shall perform measurement on the observable $a \otimes b$. The probability measure that the observable $a \otimes b$ should take value in $E \otimes F$ is given by

$$\begin{aligned} \langle 1, \mathcal{G}(E \otimes F) \nu_R \rangle / \langle 1, \nu_R \rangle &= \sum_x \langle P_x, \nu_H \rangle \langle a(E), P_x \rangle \langle b(F), R(P_x) \rangle / \langle 1_H, \nu_H \rangle \\ &= \sum_{x \in E} \langle P_x, \nu_H \rangle \langle R^*(b(F)), P_x \rangle / \langle 1_H, \nu_H \rangle. \end{aligned}$$

It is easy to check that $R^*(b(F))$ is the observable on H . Thus, $\langle 1, \mathcal{G}(E \otimes F) \nu_R \rangle = \langle a(E) \otimes b(F), \nu_R \rangle = \langle a(E) \otimes R^*(b(F)), \nu \rangle$.

Let $\mathcal{F}_R(F)$ be an instrument corresponding to the observable $R^*(b(F))$ on H and $\mathcal{G}(E \otimes F_R)$ be an instrument corresponding to the observable $a(E) \otimes R^*(b(F))$ on \mathcal{H} .

Then

$$\begin{aligned} \langle 1, \mathcal{G}(E \otimes F) \nu_R \rangle / \langle 1, \nu_R \rangle &= \langle 1, \mathcal{G}(E \otimes F_R) \nu \rangle / \langle 1, \nu \rangle = \langle 1_H, \mathcal{F}_R(F) \mathcal{E}(E) \nu_H \rangle / \\ &\langle 1_H, \nu_H \rangle. \end{aligned}$$

The obtained equations show that the joint probability of the observable a to take value in E and of the observable b , measured following a , to take value in F is invariant, firstly, with reference to representations $\{\nu_R, \mathcal{G}(E \otimes F), a \otimes b\}$ and $\{\nu, \mathcal{G}(E \otimes F_R), a \otimes R^*(b)\}$ on \mathcal{H} , and, secondly, with reference to description on \mathcal{H} and H . The measurement \mathcal{G} on \mathcal{H} over the state ν_R corresponds to a sequence of measurements on H performed in the following manner: first the observable a in state ν_H is being measured. The state obtained as a result of measurement a is subjected to a mapping R and the observable b is being measured in the state obtained as a result of the mapping R .

Now we shall proceed to the interpretation of the state ν_R . The fact of that the partial trace norms ν_R being the states on H the statistical sense of which is obvious and of that the probability measure $\frac{\langle 1, \mathcal{G}(E \otimes F) \nu_R \rangle}{\langle 1, \nu_R \rangle}$ being interpreted as a joint probability enables one to make a conclusion that the state ν_R jointly describes the statistics of the ensembles ν_{Ha} and $R(\nu_{Ha})$.

Let ν_H be a state on H such that $\langle 1_H, \nu_H \rangle = 1$. Then

$$\langle 1, \nu_R \rangle = 1.$$

We shall determine the entropy of the state ν_R .

$$S(\nu_R) = -\langle \ln \nu_R, \nu_R \rangle.$$

Using the designation $\langle A, \nu \rangle = Sp A \nu$ we shall rewrite $S(\nu_R)$ as

$$\begin{aligned} S(\nu_R) &= -Sp \nu_R \ln \nu_R \\ &= -Sp \nu_{Ha} \ln \nu_{Ha} - \sum_x \langle P_x, \nu_H \rangle Sp R(P_x) \ln R(P_x), \end{aligned}$$

where the last equation is performed on account of orthogonality P_x .

v_R -state entropy interpretation is sufficiently obvious. The meaning of $S(v_{Ha}) = -Spv_{Ha} \ln v_{Ha}$ is apparent without discussion. $S(R|P_x) = -SpR(P_x) \ln R(P_x)$ is an entropy of the ensemble obtained by the mapping R from the state P_x being clear with respect to the observable a , i.e., is a conventional entropy of the image with a given realization of preimage of mapping R , and $S\{R(v_{Ha})|v_{Ha}\} = \sum_x \langle P_x, v_H \rangle \cdot S(R|P_x)$ is the conventional entropy of image of mapping R with reference to its preimage. Therefore, the entropy of v_R -state is a joint entropy $S\{R(v_{Ha}), v_{Ha}\}$ of the states $R(v_{Ha})$ and v_{Ha} on H . This permits to conclude that the state v_R is the operator analog of a joint density of probability distribution. The expression for conventional entropy shows that the state $R(P_x)$ corresponding to the ensemble obtained by the mapping R from a clear ensemble with reference to the observable a is the operator analog of classical conventional density of distribution.

Let the measurement $\mathcal{G}(X \otimes Y)$ of the observable $a \otimes b$ be now performed in the state v_R . As a result of this measurement the state v_R will be transformed into the state

$$v'_R = \mathcal{G}(X \otimes Y) v_R = \sum_x \langle P_x, v_H \rangle \langle J_y, R(P_x) \rangle P_x \otimes J_y.$$

Partial trace norms v'_R are the states on H .

$$\langle \cdot \otimes 1_H, v'_R \rangle = \sum_{xy} \langle P_x, v_H \rangle \langle J_y, R(P_x) \rangle P_x = \sum_x \langle P_x, v_H \rangle P_x = v_{Ha} \in K_H;$$

$$\langle 1_H \otimes \cdot, v'_R \rangle = \sum_{xy} \langle P_x, v_H \rangle \langle J_y, R(P_x) \rangle J_y = \sum_x \langle P_x, v_H \rangle R_b(P_x) = R_b(v_{Ha}) \in K_H.$$

Here the designation $\sum_y \langle J_y, R(P_x) \rangle J_y = R_b(P_x)$ is taken.

The entropy

$$\begin{aligned} S(v'_R) &= -Spv'_R \ln v'_R \\ &= -Spv_{Ha} \ln v_{Ha} - \sum_x \langle P_x, v_H \rangle SpR_b(P_x) \ln R_b(P_x) \\ &= -\sum_x \langle P_x, v_H \rangle \ln \langle P_x, v_H \rangle - \sum_x \langle P_x, v_H \rangle \sum_y \langle J_y, R(P_x) \rangle \ln \langle J_y, R(P_x) \rangle \end{aligned}$$

is the joint entropy of the value ensembles of the observable a and observable b obtained by the instrument \mathcal{G} in the state v_R . Such an interpretation is resulted from probabilistic meaning of $\langle P_x, v_H \rangle$ and $\langle J_y, R(P_x) \rangle$. $S(R_b|P_x) = -SpR_b(P_x) \cdot \ln R_b(P_x)$ is obviously a conventional entropy of the b -observable value ensemble, provided that the observable a took the value x , and $S\{R_b(v_{Ha})|v_{Ha}\} = \sum_x \langle P_x, v_H \rangle \cdot S(R_b|P_x)$ is a conventional entropy of the b -observable value ensemble relative to the a -observable value ensemble.

We can now give a strict definition to entropy defect of a quantum state and information quantity on the observable a , contained in the observable b .

Let us determine the v_{Ha} -state entropy defect in consequence of the mapping R as

$$\begin{aligned} \mathcal{D} &= S(v_{Ha}) + S(R(v_{Ha})) - S\{R(v_{Ha}), v_{Ha}\} \\ &= S(R(v_{Ha})) - S\{R(v_{Ha})|v_{Ha}\}. \end{aligned}$$

The information quantity supplied by b -observable measurement in the state $R(v_{Ha})$, with reference to the state v_{Ha} , is

$$\begin{aligned} I &= S(v_{Ha}) + S(R_b(v_{Ha})) - S\{R_b(v_{Ha}), v_{Ha}\} \\ &= S\{R_b(v_{Ha})\} - S\{R_b(v_{Ha})|v_{Ha}\}. \end{aligned}$$

It is easy to show that in contrast to a classical case where the information quantity is strictly equal to entropy defect in quantum region

$$I \leq \mathcal{D}$$

the equation is possible if and only if $R(P_x)R(P_{x'})=R(P_{x'})R(P_x)$ for all x and x' [3].

The obtained operator analog of joint distribution density and the functionals constructed with the help of it enable us strictly to describe quantum systems of information transmission and to use the theoretical-information method for analysis of some problems of quantum statistics.

References

1. Davies, E. B., Lewis, J. T.: *Commun. math. Phys.* **17**, 239—260 (1970)
2. von Neumann, J.: *Mathematical foundations of quantum mechanics*. Trans. by R. T. Beyer. Princeton: University Press 1955
3. Levitin, L. B.: *Trudy IV Vsesoyuznoi Konfer. po Teorii Inform. i Kodirovaniya* (English translation: *Proceedings of the IV All-Union Conference on Theory of Information and Encoding*) Tashkent 111—115, 1969

Communicated by R. Haag

V. D. Vainshtein
The Institute of Atmospheric Optics
USSR Academy of Sciences
Siberian Branch
Tomsk, USSR