International Journal of Theoretical Physics, Vol. 32, No. 12, 1993

Quantum Measurements and Information Theory

K.-E. Hellwig[1]

Received July 22, 1993

1. INTRODUCTION

Max Born (1926) proposed the statistical interpretation of quantum mechanics and there is no doubt that it well describes the statistics of pointer readings. Is this knowledge sufficient to understand the physical reality of microparticles, i.e., can a physical theory be complete if nothing behind the statistics of pointer readings is assumed? Such questions arose shortly after Born's proposal in the Bohr–Einstein debate, which culminated in the famous Einstein–Podolsky–Rosen paradox (Einstein *et al.*, 1935). Later investigations led to Bell (1964)-type inequalities. The experimental tests by A. Aspect and co-workers gave deep insight into the physical situation: If somebody believes in the physical reality of accidental properties of a single microparticle described by some kind of hidden variables, then he or she has to assume the existence of action at a distance.

The occurrence of pointer positions can be described in terms of quantum mechanics and quantum statistics as a physical process. The description of this process can be understood as an analysis of how we get knowledge about microsystems by their interaction with a measuring apparatus. The experiments to be analyzed have a very general scheme: There is one part of the experimental setup that isolates a microparticle from its surroundings in a particular manner and another one that responds because of interaction with it by a certain pointer position. It is suggestive to assume the microparticle to carry some message from the first part of the setup, the source of information, to the second, the receiver. This message, if it is any, can only be transmitted if it is coded into some property of the microparticle. If the language of the statistical theory of

[1]Institut für Theoretische Physik, Technische Universität Berlin, D-10623 Berlin, Germany.

0020-7748/93/1200-2401$07.00/0

information can be used to describe this setup, one probably obtains new aspects about the nature of quantum systems.

In the following, I first give a short introduction to the theory of quantum measurements and then I describe the situation in terms of the statistical theory of information.

2. THE THEORY OF QUANTUM MEASUREMENTS

As already mentioned, quantum experimental setups usually are divided into two parts. The first part is the means by which the quantum object is prepared under the control of relevant macroscopic parameters such that it belongs to a well-defined statistical ensemble described by a density operator $W_{\mathcal{O}}$ acting in the Hilbert space $\mathcal{H}_{\mathcal{O}}$. The second part is called the apparatus and consists of a macroscopic system in a thermodynamically metastable state which can become unstable because of interaction with the object. As a consequence of the interaction, one of several different equilibrium states will arise and indicate the final result of a measurement (Weidlich, 1967), which we call "*pointer position.*" As a many-particle quantum system, the metastable state of the apparatus also will be described by a density operator, say $W_{\mathcal{A}}$ acting in a Hilbert space $\mathcal{H}_{\mathcal{A}}$. The preparations of the object and the apparatus are supposed to be independent of each other such that the initial state of the coupled system will be the uncorrelated one, i.e., it is described by the density operator $(W_{\mathcal{O}} \otimes W_{\mathcal{A}})$. This is a highly simplified scheme and there are many features which I cannot describe in full detail or even mention here.

As long as the irreversible dynamics towards the final equilibrium of the apparatus is not involved, the dynamics of the interaction process is described by a unitary operator. Although it is very doubtful whether the final state is determined by the reversible part of the process and no stochastics enters thereafter with respect to the resulting final equilibrium we will confine ourselves to consider only the unitary dynamics which may be given by the unitarian S acting in $\mathcal{H}_{\mathcal{O}} \otimes \mathcal{H}_{\mathcal{A}}$. It should be mentioned that the influence of the irreversible part of the process on the final result has not been completely investigated up to now. We will consider here $S(W_{\mathcal{O}} \otimes W_{\mathcal{A}})S^{+}$ as the final state which determines the measurement result.

2.1. The Historical Scheme

For the convenience of the reader to enter into this material, I will briefly recall the historical scheme, although it works only for such observables of the object that have discrete spectra. Let the self-adjoint operator acting in $\mathscr{H}_{\mathscr{O}}$ of the observable to be measured be given by

$$B = \sum_{i=1}^{\infty} b_i |\psi_i\rangle\langle\psi_i|$$

The self-adjoint operator acting in $\mathscr{H}_{\mathscr{A}}$ and corresponding to the macroscopic equilibrium quantity finally to be observed shall be given by

$$A = \sum_{j=1}^{\infty} a_j |\Psi_j\rangle\langle\Psi_j|$$

The eigenvalues of this operator can be identified with the final pointer positions. In a realistic situation such as a macroscopic equilibrium the observable A should be highly degenerate. Here we will assume it in the contrary to be nondegenerate because this simplifies our reasoning without a big loss of generality. For the same reason we assume $W_{\mathscr{A}}$ to represent a pure state, say $\Psi_1 \in \mathscr{H}_{\mathscr{A}}$. Finally, let also $W_{\mathscr{O}}$ represent a pure state, say $\phi \in \mathscr{H}_{\mathscr{O}}$. Hence the final state of the coupled system will be $S(\phi \otimes \Psi_1)$. By a suitable choice of S and A, which formally is always possible, one gets a complete correlation of the pointer positions with the eigenvalues of the operator A in the final state. This situation is given if for each $i \in \mathbb{N}$

$$S(\psi_i \otimes \Psi_1) = \psi_i \otimes \Psi_i$$

holds true. By the linearity of S we have immediately

$$S(\phi \otimes \Psi_1) = \sum_{i=1}^{\infty} \langle\psi_i|\phi\rangle(\psi_i \otimes \Psi_i)$$

such that the probability to read the pointer position a_i from the apparatus is just given by $|\langle\psi_i|\phi\rangle|^2$, as it should be.

There are some serious difficulties with this approach. The final state after the pointer position a_j has been fixed should be a mixture rather than a pure state. The simplest explanation is given by Jauch (1964). Since A is an observable of a many-particle system that directly can be (classically) observed, it belongs to a commutative subset of all observables of this system. On the spectrum of any observable of this subset the pure state and the corresponding mixture give the same probability distributions. Hence the problem disappears. Problems arise, however, if the entropy of states is taken into consideration, since the pure state has a lower entropy than the mixture. The first idea to solve them is due to von Neumann and states that this transition to a mixture is caused when the first person gets knowledge about the value by reading it from the apparatus. It has been assumed that a chain of succeeding processes of this kind happen until a human consciousness irreversibly decides about the final result (see Jammer, 1974, Chapter 11.2 for review). Clearly, there arises a problem of objectivity since different persons watching the apparatus may decide differently. Although it seems much more natural that a final pointer position arises without interaction with a human brain, the first idea is still alive. The Everett–Wheeler interpretation of many worlds (DeWitt and Graham, 1973) solves the objectivity problem at least formally. The realistic point of view is to assume that the final pointer position is the result of an irreversible process inside the apparatus as mentioned at the beginning.

Other difficulties are the following: This scheme only works for discrete spectra. Moreover, suitable choices for S and A only exist if B commutes with each universally conserved quantity (Wigner, 1952; Wigner and Yanase, 1963). Since each component of angular momentum of a closed system is a universally conserved quantity and spin-projections for different directions do not commute, the historical approach is too narrow to explain measurements of one of them.

2.2. The Modern Approach

A much more general and simple approach arose at the beginning of the sixties (Kraus, 1983). In contrast to the historical approach where the operator B is assumed to be given and the suitable choice of S and A has to be determined, the modern approach assumes S and A to be given and it is asked what will be measured. This formulation of the problem surrounds the obstructions mentioned at the end of Section 2.1 and includes the problem of approximate measurements as well. Let again A denote the self-adjoint operator that represents the observable finally to be observed on the apparatus. The present approach does not require the spectrum to be discrete. With the spectral resolution E in $\mathscr{H}_{\mathscr{A}}$ the operator A can be written in the form

$$A = \int a dE(a)$$

Let $W_{\mathscr{A}}$ denote the density operator of the metastable equilibrium of the apparatus. Let b be a Borel set on the real line such that

$$\mathrm{tr}(W_{\mathscr{A}} E(b)) = 1$$

b represents the set of pointer positions for which the experimenter would state "zero." Let $W_{\mathscr{O}}$ denote the density operator of the ensemble of objects which are prepared by the first part of the experimental setup acting in $\mathscr{H}_{\mathscr{O}}$. Then by the rules of quantum mechanics the probability to find the pointer finally in the Borel set a is given by

$$p(a, W_{\mathcal{O}}) = \mathrm{tr}((W_{\mathcal{O}} \otimes W_{\mathcal{A}})S^{+}(1 \otimes E(a)S))$$

Since this expression is trace norm continuous in $W_{\mathcal{O}}$ and bounded in between 0 and 1, there is exactly one self-adjoint operator $F(a)$ acting in $\mathscr{H}_{\mathcal{O}}$ and fulfilling $0 \le F(a) \le 1$ such that for each trace class operator $W_{\mathcal{O}}$ acting in $\mathscr{H}_{\mathcal{O}}$ we have

$$p(a, W_{\mathcal{O}}) = \mathrm{tr}(W_{\mathcal{O}} F(a))$$

Since $p(\cdot, W_{\mathcal{O}})$ is a probability measure, $F(\cdot)$ is a positive operator valued measure and $F(\mathbb{R}) = 1$.

Given the self-adjoint operator A acting on $\mathscr{H}_{\mathcal{A}}$ and the unitarian S acting in $\mathscr{H}_{\mathcal{O}} \otimes \mathscr{H}_{\mathcal{A}}$, the observable measured on the object has values in the spectrum of A and is represented by a positive operator valued measure F on the real line with $F(\mathbb{R}) = 1$, which may be projection valued or not. If it is projection valued, then

$$B := \int aF(da)$$

is a self-adjoint operator acting in $\mathscr{H}_{\mathcal{O}}$ with the usual interpretation in quantum mechanics. If it is not, then one may think that it approximates an observable in the usual sense in that it accounts for systematic errors. Since there is no primary principle to determine which self-adjoint operator in some ultraweak neighborhood of B is approximated, one usually generalizes the concept of observables in quantum mechanics to all positive operator valued measures on the Borel sets of the real line $\mathbb{R}$ with $F(\mathbb{R}) = 1$. Among them the projection valued measures are called decision observables by Ludwig (1985).

From the probabilistic point of view and starting from what is given in nature, realizable metastable systems and their interactions with quantum objects, this scheme seems very natural. Moreover, no problems with continuous spectra arise.

2.3. Informational Completeness

Let

$$F: \quad B(\mathbb{R}) \to [0, 1] \subset \mathscr{B}(\mathscr{H}_{\mathcal{O}})$$

where $B(\mathbb{R})$ denotes the algebra of Borel sets on the real line and $\mathscr{B}(\mathscr{H}_{\mathcal{O}})$ the set of bounded operators on $\mathscr{H}_{\mathcal{O}}$, be a positive operator valued measure fulfilling $F(\mathbb{R}) = 1$. Then

$$W_{\mathcal{O}} \mapsto \mathrm{tr}(W_{\mathcal{O}}, F(\cdot))$$

is an affine mapping from the density operators on $\mathscr{H}_{\mathcal{O}}$ into the probability measures on $B(\mathbb{R})$. F is called "*informationally complete*" in case this map is injective. It is well known that there is no informational complete projection valued measure. Moreover, the probability distributions defined by the spectral measures of the position and momentum operators together do not determine a density operator uniquely. But there are a lot of informationally complete positive operator valued measures.

Well-known examples are given by the so-called "*joint position-momentum observables*," which can be generated in the following way: Let $\varphi \in L^2(\mathbb{R})$, $\|\varphi\| = 1$, $\langle\varphi|x\varphi\rangle = 0$, and

$$\left\langle \varphi \left| \frac{1}{i}\frac{\partial}{\partial x}\varphi \right.\right\rangle = 0$$

Consequently, for the Galilean shifts

$$\psi_{p,q}(x) := e^{ipx}\varphi(x-q) \qquad (x, p, q \in \mathbb{R})$$

there hold

$$\langle\psi_{p,q}|x\psi_{p,q}\rangle = q \qquad \text{and} \qquad \left\langle \psi_{p,q} \left| \frac{1}{i}\frac{\partial}{\partial x}\psi_{p,q} \right.\right\rangle = p$$

Now for any density operator $W_{\mathcal{O}}$ of the object the function $\langle\psi_{p,q}|W_{\mathcal{O}}\psi_{p,q}\rangle$ is integrable on the (p, q) plane. Moreover, on the Borel sets of this plane a probability measure is given by

$$a \mapsto \frac{1}{2\pi}\int_a \langle\psi_{p,q}|W_{\mathcal{O}}\psi_{p,q}\rangle\, dp\, dq, \qquad a \in B(\mathbb{R})$$

This probability measure can also be written as

$$\operatorname{tr}(W_{\mathcal{O}}F(a)), \qquad F(a) := \frac{1}{2\pi}\int_a |\psi_{p,q}\rangle\langle\psi_{p,q}|\, dp\, dq$$

the latter defining a positive operator valued measure fulfilling $F(\mathbb{R}) = 1$. It has been shown by Ali and Prugovečki (1977*a*,*b*) that this measure is informationally complete whenever

$$\langle\varphi|\psi_{p,q}\rangle \neq 0 \qquad \text{a.e. on } \mathbb{R}^2$$

holds true.

Finally, using informationally complete joint position-momentum observables and the related "*phase space representations*," Singer and Stulpe (1992) have shown how far classical probability can approximate quantum probability. Let $\rho_{W_{\mathcal{O}}}$ be the probability density on the (p, q) plane defined by

$$\int_a \rho_W\, dp\, dq = \operatorname{tr}(W_{\mathcal{O}}, F(a))$$

or, equivalently, by

$$\rho_{W_{\mathcal{O}}} := \frac{1}{2\pi} \langle \psi_{p,q} | W_{\mathcal{O}} \psi_{p,q} \rangle$$

The mapping $W_{\mathcal{O}} \mapsto \rho_{W_{\mathcal{O}}} \in \mathbb{R}^2$ is affine and injective, the latter being a consequence of the informational completeness of F. Now consider a finite set of density operators $\{W_i\}_{i=1,2,3,\ldots,n}$, $n \in \mathbb{N}$, and let $\epsilon > 0$. Then there is a mapping

$$\mathcal{B}(\mathcal{H}_{\mathcal{O}}) \ni A \mapsto f \in L^\infty(\mathbb{R}^2)$$

such that for each $A \in \mathcal{B}(\mathcal{H}_{\mathcal{O}})$

$$\left| \operatorname{tr}(W_i A) - \int \rho_{W_i} f \, dp \, dq \right| < \epsilon$$

This means: Given a finite set of density operators, the quantum mechanical expectation values can uniformly be approximated up to arbitrary accuracy by the corresponding classical expressions.

3. QUANTUM MEASUREMENTS AND INFORMATION

As mentioned in the Introduction, the statistical theory of information is based on the probability distribution for the events to be observed, and one may ask whether it can help to get deeper insight into the nature of microsystems than one gets if only the statistical interpretation of quantum theory is taken into consideration. The hope is that new knowledge may arise because new aspects of interpretation enter with the statistical theory of information.

3.1. Concepts of Classical Statistical Information Theory

An informational setup consists in at least three things: A "*source*" of information, a transmitting "*channel*," and a "*receiver*." Let the source be able to send one of n "*letters*" D_k, $k = 1, 2, 3, \ldots, n$, and let the receiver be equipped with m different lamps L_l, $l = 1, 2, 3, \ldots, m$. Now let $p_l(D_k)$ be the probability that L_l will shine when D_k has been sent. The receiver will be called "*ideal with respect to* $\{D_k\}_{k=1,2,3,\ldots,n}$" when $m \geq n$ and, say,

$$p_l(D_k) = \delta_{lk}, \qquad l, k = 1, 2, 3, \ldots, n$$

For simplicity we have assumed here the channel to be ideal, too, i.e., the channel does not change the signals it is transmitting.

The informational value of a single message is the central point. It is defined by the value of the "*information function*," which depends on the probability by which the message is sent. This information function is assumed to be a mapping

$$I: \quad [0, 1] \mapsto [0, \infty]$$

and is determined by one further axiom which shall be described now: Enumerate the letters of the source by two numbers instead of one, i.e., write for them D_{ij} $(i = 1, 2, 3, \ldots, r; j = 1, 2, 3, \ldots, s_i)$. Then assume a coarse receiver with r lamps L_l $(l = 1, 2, 3, \ldots, r)$ such that

$$p_l(D_{i1} \vee D_{i2} \vee D_{i3} \vee \cdots \vee D_{is_i}) = \delta_{li}$$

This receiver is ideal with respect to some decomposition into disjoint subsets of the set of letters, but it is not able to discern between individual elements of them. Let another receiver be equipped with $n = \sum_{i=1}^{r} s_i$ lamps L_{ij} and assume it to be ideal with respect to $\{D_{i1}, D_{i2}, D_{i3}, \ldots, D_{is_i}\}_{i=1,2,3,\ldots,r}$, i.e.,

$$p_{ij}(D_{i'j'}) = \delta_{ii'}\delta_{jj'}$$

The second receiver, obviously, gets more information than the first one. Let w_{ij} denote the probability by which the letter D_{ij} is sent. Hence $\mathbf{w}_i = \sum_{j=1}^{s_i} w_{ij}$ is the probability that a letter from the ith subset is sent. Now let $I(\mathbf{w}_i)$ be the information obtained when some letter from the ith subset has been received. For the surplus information which can only be obtained by the second receiver which detects, say, that D_{ij} is sent, only the letters $\{D_{ij}\}_{j=1,2,3,\ldots,s_i}$ are in question. Therefore, it is natural to base it on the conditional probabilities

$$\mathsf{w}_{ij} := \frac{w_{ij}}{\mathbf{w}_i}$$

This motivates the axiom

$$I(w_{ij}) = I(\mathbf{w}_i) + I(\mathsf{w}_{ij})$$

which is equivalent to

$$I(\mathbf{w}_i \mathsf{w}_{ij}) = I(\mathbf{w}_i) + I(\mathsf{w}_{ij})$$

It can be shown that I is determined up to a constant $c \in \mathbb{R}_+$ by

$$I(\lambda) = -c \log(\lambda)$$

It is called the information function.

3.2. Quantum Measurements

We now apply the concepts of statistical information theory to quantum measurements. The preparative part of an experiment shall be taken for the source and the registrative part for the receiver. Let the preparative part produce an ensemble described by the density operator $W_{\mathcal{O}}$ acting in $\mathcal{H}_0$. One is tempted to consider a decomposition into density operators W_i

$$W_{\mathcal{O}} = \sum_i w_i W_i; \qquad w_i \in (0, 1), \qquad \sum_i w_i = 1$$

and to call W_i a letter sent with the probability w_i. But such a decomposition is in general not unique. The minimal requirement to interpret the components as letters seems to be that there exists a receiver which can discern between them. This motivates the definition: A decomposition into density operators V_i

$$W_{\emptyset} = \sum_i v_i V_i; \qquad v_i \in (0, 1), \qquad \sum_i v_i = 1$$

is called an "*admissible*" one iff there exists a positive operator valued measure F on the power set of $\{1, 2, 3, \ldots, n\}$, $F_i := F(\{i\})$, $\sum_i F_i = 1$, such that

$$\operatorname{tr}(V_i F_k) = \delta_{ik}$$

There may be many admissible decompositions of one and the same density operator.

Now the entropy of a density operator with respect to an admissible decomposition $(V_1, V_2, V_3, \ldots, V_n)$ is defined by the average information sent using V_i as letters:

$$H_W((V_1, V_2, V_3, \ldots, V_n)) := -c \sum_{i=1}^{n} v_i \log v_i, \qquad v_i \neq 0$$

For this expression it can be shown that

$$0 < H_W((V_1, V_2, V_3, \ldots, V_n)) \leq c \log n$$

where the equality holds iff $v_i = 1/n$. Moreover, consider (admissible) decompositions of the V_i, say

$$V_i = \sum_{j=1}^{m_i} u_{ji} U_{ji}$$

such that $(U_{11}, U_{12}, U_{13}, \ldots, U_{nm_n})$ is an admissible decomposition of W; then there holds

$$H_W((V_1, V_2, V_3, \ldots, V_n)) \leq H_W((U_{11}, U_{12}, U_{13}, \ldots, U_{nm_n}))$$

Finally, it can be shown that the usual entropy is just the supremum of the entropies with respect to the admissible decompositions

$$\begin{aligned}\mathbf{H}_W &:= \sup\{H_W(V_1, V_2, V_3, \ldots, V_n) \mid \text{admissible decompositions}\} \\ &= -c \operatorname{tr}(W \log W)\end{aligned}$$

3.3. Disturbed Transmission

We now assume a nonideal channel which disturbs the signal. Such disturbance is most generally described by an affine map K operating on the density operators. If $W = \sum_{i=1}^{n} v_i V_i$ is an admissible decomposition of the density operator W of the source, the statistical ensemble arriving at the receiver will be described by $K(W) = \sum_{i=1}^{n} v_i K(V_i)$ and this decomposition of $K(W)$ is not admissible in general. Let G denote the positive operator valued measure on the power set of $\{1, 2, 3, \ldots, n\}$ by which the receiver analyzes the signal and let $G_i := G(\{i\})$. Then this nonideality is expressed by

$$\mathrm{tr}(G_k K(V_l)) =: p_{kl} \neq \delta_{kl}$$

Since p_{kl} is the probability that the lamp k will shine when the letter l is sent, we have $\sum_{k=1}^{n} p_{kl} = 1$. The conditional probability that V_l has been sent when the lamp k is shining is given by

$$q_{kl} := \frac{v_l \,\mathrm{tr}(G_k K(V_l))}{\mathrm{tr}(G_k K(W))}$$

where $\sum_{l=1}^{n} q_{kl} = 1$. Now we ask for the average loss of information when the nonideal receiver is used instead of the ideal one. Under the hypothesis that the lamp k is shining, the average information $-c \sum_{l=1}^{n} q_{kl} \log q_{kl}$ gets lost. If we average this loss over all lamps remembering that the kth one is shining with the probability $\mathrm{tr}(G_k K(W))$, we get the "equivocation with respect to $(V_1, V_2, V_3, \ldots, V_n)$" as

$$\mathbf{E}_W(\{V_j\}) := -c \sum_{k=1}^{n} \mathrm{tr}(G_k K(W)) \sum_{l=1}^{n} q_{kl} \log q_{kl}$$

The average information correctly transmitted is called the "*transinformation with respect to* $(V_1, V_2, V_3, \ldots, V_n)$" and is the average information produced by the source minus equivocation,

$$\mathbf{T}_W(\{V_j\}) := H_W(\{V_j\}) - \mathbf{E}_W(\{V_j\})$$

In the case that there exists a decomposition

$$K(W) = \sum_{k=1}^{n} u_k U_k$$

such that the operators G_k give rise to an ideal receiver with respect to $(U_1, U_2, U_3, \ldots, U_n)$, then one may also write

$$\mathbf{T}_W(\{V_j\}) = H_{K(W)}(\{U_j\}) - \mathbf{I}_{K(W)}(\{K(V_j)\})$$

where

$$\mathbf{I}_{K(W)}(\{K(V_j)\}) := -c \sum_{l=1}^{n} v_l \sum_{k=1}^{n} p_{kl} \log p_{kl}$$

is called the noise or "*irrelevance*" produced by the nonideal channel.

4. CONCLUSION

It has been demonstrated how concepts of the theory of statistical information can be fitted into the description of quantal measurements. For details and proofs as well as further literature I refer to Singer (1989). Technical applications of such investigations are possible in quantum optical communication systems. I propose they may also give additional insight into the nature of microparticles, but results do not yet seem to exist.

ACKNOWLEDGMENTS

The author thanks Werner Stulpe, Volker Perlick, Utz Grimmer, and Michael Keyl for critical readings of the manuscript.

REFERENCES

Ali, S. T., and Prugovečki, E. (1977*a*). Systems of imprimitivity and representations of quantum mechanics on fuzzy phase space, *Journal of Mathematical Physics*, **18**, 219–228.

Ali, S. T., and Prugovečki, E. (1977*b*). Classical and quantum statistical mechanics in the common Liouville space, *Physica*, **89A**, 501–521.

Bell, J. S., (1964). On the Einstein–Podolsky–Rosen paradox, *Physics*, **1**, 195–200.

Born, M. (1926). Zur quantenmechanik der Stossvorgänge, *Zeitschrift für Physik*, **37**, 863–867.

DeWitt, B., and Graham, N., eds. (1973). *The Many Worlds Interpretation*, Princeton University Press, Princeton, New Jersey.

Einstein, A., Podolsky, B., and Rosen, N. (1935). Can quantum mechanical description of physical reality be considered complete, *Physical Review*, **47**, 777–780.

Jammer, M. (1974). *The Philosophy of Quantum Mechanics*, Wiley, New York.

Jauch, J. M. (1964). The problem of measurement in quantum mechanics, *Helvetica Physica Acta*, **37**, 293–316.

Kraus, K. (1983). *States, Effects, and Operations*, Springer, Berlin.

Ludwig, G. (1985). *Foundations of Quantum Mechanics I*, Springer, New York.

Singer, M. (1989). Zur Informationstheorie statistischer Experimente, Fachbereich Physik der Technischen Universität Berlin, D 83.

Singer, M., and Stulpe, W. (1992). Phase space representations of general statistical physical theories, *Journal of Mathematical Physics*, **33**, 131–142.

Weidlich, W. (1967). Problems of the quantum theory of measurement, *Zeitschrift für Physik*, **205**, 199–220.

Wigner, E. P. (1952). Die Messung quantenmechanischer Operatoren, *Zeitschrift für Physik*, **133**, 101–108.

Wigner, E. P., and Yanase, M. M. (1963). Information contents of distributions, *Proceedings of the National Academy of Science of the USA*, **49**, 910–918.