Classical statistical physics and quantum mechanics

D. I. Blokhintsev

Joint Institute for Nuclear Research, Dubna (Moscow region)
Usp. Fiz. Nauk **122**, 745–757 (August 1977)

1. INTRODUCTION

Quantum mechanics is usually regarded as a generalization of classical mechanics. The physical content of this generalization is basically expressed by the complementarity principle of N. Bohr.

As will be explained below a different approach to the fundamentals of quantum mechanics is possible and in certain respects appears to be necessary. In this other approach quantum mechanics is regarded as a generalization of classical statistical mechanics. In such an interpretation of quantum mechanics the basic concept is not the wave function ψ, but the *statistical operator* introduced already in 1932 by von Neumann.[1] It is now usually referred to as the density operator $\hat{\rho}$. It is well known that this operator is the analogue of the density $\rho(q,p)$ in phase space $\mathcal{R}(q,p)$, which is utilized in classical statistical mechanics for the description of motion of systems. Here q are the coordinates of the system being studied, p are the canonically conjugate momenta.[1]

The relationship between classical mechanics and classical statistical mechanics differs in principle from the relationship between quantum mechanics operating with the wave function ψ and quantum mechanics based on the concept of the statistical operator $\hat{\rho}$.

The point is that classical mechanics represents a science which does not at all require statistical mechanics. Within the limits of its applicability it gives maximally complete information concerning the motion of a mechanical system and does not make use of the concept of probability or the concept of some sort of a statistical ensemble. In the domain of quantum phenomena the analogue of classical mechanics is quantum mechanics which makes use of the wave function ψ which also gives maximally complete information concerning the motion of quantum systems compatible with the basis of this theory—the complementarity principle of N. Bohr. However it was established long ago that in the theory of quantum measurement it is not possible to restrict onself to the concept of a wave function—it is necessary to introduce the concept of the density operator $\hat{\rho}$ which has its analogue in statistical mechanics. Thus, in contrast to classical mechanics, quantum mechanics requires quantum statistical mechanics. Speaking more accurately, the two are in fact the same.

After these remarks we turn to classical statistical mechanics. In classical statistical mechanics information is expressed in the language of probabilities.

Probability is a numerical measure of the potential possibility of some particular outcome of events.

The outcome is determined within a certain statistical ensemble of events which must be defined by clearly formulated material conditions. Thus, in thermodynamical statistics the Gibbs ensemble is defined by the temperature of a large heat bath, with which the molecular system under investigation interacts weakly.[2]

Knowledge of the probability of a particular event enables one to predict the mathematical expectation of some particular possible outcome, the average value of observable quantities, the fluctuations in these quantities, etc.

Probability is not a characteristic of an individual mechanical system taken by itself. It belongs to such a system only to the extent that such a system is a member of a definite statistical ensemble. In classical statistical mechanics the ensemble is determined by the probability of finding a mechanical system $d\rho(q,p,t)$ in the neighborhood of a point (q,p) of phase space $\mathcal{R}(q,p)$ at the instant of time t.[3]

Instead of the probability $d\rho$ one usually considers the probability density in phase space $\rho(q,p,t)$:

$$\rho(q, p, t) = \frac{d\rho(q, p, t)}{dq\, dp}. \tag{1}$$

As long as one is dealing with microscopic atomic systems, then in order for such an ensemble to be definite one must specify the external macroscopic conditions within which the development of the ensemble under investigation is realized, for example: the size of the container, the temperature of the walls, external fields, etc.

The probability density $\rho(q,p)$ obeys the equation of motion which states:

$$\frac{\partial \rho}{\partial t} + [\mathcal{H}, \rho]_{q,p} = 0, \tag{2}$$

where $\mathcal{H} = \mathcal{H}(q,p)$ is the Hamiltonian function, while $[\mathcal{H}, \rho]_{q,p}$ is the classical Poisson bracket which for any

---

[1] One should intepret q as the coordinates $q_1, q_2 \ldots, q_f$ of the system under consideration, and $p-p_1, p_2, \ldots p_f$ as the momenta conjugate to them, f is the number of degrees of freedom. In future to avoid awkwardness all the formulas are written out explicitly as if $f=1$. For example, $dqdp = dq_1 \ldots dq_f \cdot dp_1 \ldots dp_f$.

[2] We use the term "ensemble" coined by the founder of thermodynamical statistics Gibbs.[2] Other terms for an ensemble are: statistical collective (Mises),[3] statistical set (von Neumann).[1]

[3] In future we shall not explicitly write out the argument t.

arbitrary dynamical quantities A and B has the form:

$$[A,B]_{q,p} = \frac{\partial A}{\partial p}\frac{\partial B}{\partial q} - \frac{\partial B}{\partial p}\frac{\partial A}{\partial q}. \tag{3}$$

In particular, for a pair of canonically conjugate variables we have

$$[p, q]_{q,p} = 1. \tag{4}$$

Explicitly, for one degree of freedom the Hamiltonian function is equal to

$$\mathcal{H}(q, p) = \frac{p^2}{2m} + V(q); \tag{5}$$

here m is the mass of the particle, $V(q)$ is its potential energy. From (2) and (3) we obtain the equation of motion for the probability density in explicit form

$$\frac{\partial \rho}{\partial t} + \frac{p}{m}\frac{\partial \rho}{\partial q} - \frac{\partial V}{\partial q}\frac{\partial \rho}{\partial p} = 0. \tag{6}$$

This equation expresses the law of conservation of the number of particles in each element of phase space.

The probability of finding the system in the neighborhood of the point q in configuration space $\mathcal{R}(q)$ is equal to

$$\rho(q)\,dq = dq \int \rho(q, p)\,dp, \tag{7}$$

while the probability of the system to have a momentum in the neighborhood of the point p is expressed by the formula

$$\rho(p)\,dp = dp \int \rho(q, p)\,dq. \tag{8}$$

Finally, the average value $\bar{L}$ of any dynamical variable $L(q,p)$ is determined by the formula

$$\bar{L} = \int \rho(q, p)\, L(q, p)\, dq\, dp \tag{9}$$

with the normalization condition

$$\int \rho(q, p)\, dq\, dp = 1. \tag{10}$$

In virtue of the equation of motion (6) this condition is satisfied at any arbitrary instant of time t.

2. CLASSICAL STATISTICAL MECHANICS IN CONFIGURATION SPACE

Classical statistical mechanics can be formulated not only in phase space $\mathcal{R}(q,p)$, as is customary, but also in configuration space $\mathcal{R}(q)$.[4] In order to go over to such a representation of statistical mechanics we must instead of the momenta p take a second arbitrary point q' in configuration space $\mathcal{R}(q)$. In other words instead of the phase space $\mathcal{R}(q,p)$ we go over to the doubled configuration space

$$\mathcal{R}(q, q') \equiv \mathcal{R}(q) \times \mathcal{R}(q'). \tag{11}$$

We shall define the transition from the description in the space $\mathcal{R}(q,p)$ to the description in the space $\mathcal{R}(q,q')$ with the aid of the Fourier transformation, applicable to any dynamical variable $L(q,p)$ defined in the phase space $\mathcal{R}(q,p)$. This transformation states:

$$L(q,q') \equiv L(q, \zeta) = \int L(q, p)\frac{e^{ip\zeta/\hbar^*}}{2\pi\hbar^*}dp, \tag{12}$$

$$L(q, p) = \int L(q, \zeta)\, e^{-ip\zeta/\hbar^*}\, d\zeta, \tag{12'}$$

where $\zeta = q' - q$, while $\hbar^*$ is some constant having the dimensions of action. This quantity is arbitrary within the framework of classical mechanics. Choosing a certain characteristic scale for the coordinate a and a characteristic scale for the momenta b it is natural to set $\hbar^* = ab$. In view of the fact that for a Fourier transformation it is not the absolute values of q and p that are particularly important but the range of their variation, it is useful to set

$$\hbar^* = \sqrt{\overline{\Delta p^2}\,\overline{\Delta q^2}}, \tag{13}$$

where $\overline{\Delta p^2}$ and $\overline{\Delta q^2}$ are the mean square deviations of p and q, say, at $t = 0$.

The transformation for the density $\rho(q,p)$ can be more conveniently defined by:

$$\rho(q, q') \equiv \rho(q, \zeta) = \int \rho(q, p)\, e^{ip\zeta/\hbar^*}\, dp, \tag{14}$$

$$\rho(q, p) = \int \rho(q, \zeta)\frac{e^{-ip\zeta/\hbar^*}}{2\pi\hbar^*}\, d\zeta; \tag{14'}$$

here $L(q,q')$ and $L(q, \zeta)$, as can be seen from the formulas, denote the same quantities. Generally speaking, these quantities turn out to be generalized functions. The rules for dealing with them are now well known. The equation of motion (2) for the probability density can also be rewritten in the space $\mathcal{R}(q,q')$. With the aid of formula (12) we find the representation of the Hamiltonian $\mathcal{H}(q,p)$ (5) in the space $\mathcal{R}(q,q')$. It takes the form:

$$\mathcal{H}(q, q') \equiv \mathcal{H}(q, \zeta) = -\frac{\hbar^{*2}}{2m}\frac{d^2\delta(\zeta)}{d\zeta^2} + V(q)\,\delta(\zeta), \tag{15}$$

where $\delta(\zeta)$ is the usual delta-function. Further, we have the following expression for the Fourier components of the derivatives of the Hamiltonian function $\mathcal{H}$ and the density ρ:

$$\left(\frac{\partial \mathcal{H}}{\partial q}\right)_{q,q'} \equiv \left(\frac{\partial \mathcal{H}}{\partial q}\right)_{q,\zeta} = \frac{\partial V}{\partial q}\delta(\zeta), \tag{16}$$

$$\left(\frac{\partial \mathcal{H}}{\partial p}\right)_{q,q'} \equiv \left(\frac{\partial \mathcal{H}}{\partial p}\right)_{q,\zeta} = -\frac{i\hbar^*}{m}\frac{d\delta(\zeta)}{d\zeta}, \tag{17}$$

$$\left(\frac{\partial \rho}{\partial q}\right)_{q,q'} \equiv \left(\frac{\partial \rho}{\partial q}\right)_{q,\zeta} = \frac{\partial \rho(q, \zeta)}{\partial q}, \tag{18}$$

$$\left(\frac{\partial \rho}{\partial p}\right)_{q,q'} \equiv \left(\frac{\partial \rho}{\partial p}\right)_{q,\zeta} = -\frac{i}{\hbar^*}\zeta\rho(q, \zeta). \tag{19}$$

Substitution of these Fourier components into (2) leads with the aid of (12') to the equation for the density $\rho(q, \zeta)$ in the space $\mathcal{R}(q,q')$:

$$\frac{\partial \rho(q, \zeta)}{\partial t} - \frac{i\hbar^*}{m}\frac{\partial^2 \rho(q, \zeta)}{\partial q\, \partial \zeta} - \frac{1}{i\hbar^*}\frac{\partial V}{\partial q}\zeta\cdot\rho(q, \zeta) = 0. \tag{20}$$

---

[4] This section is based on Ref. 4; cf., also Ref. 5.

This equation replaces in the space $\mathscr{R}(q, q')$ equation (6) in the space $\mathscr{R}(q, p)$. This equation can be concisely written in the form

$$\frac{\partial \rho}{\partial t} + [\mathscr{H}, \rho]_{q, \zeta} = 0, \tag{21}$$

where $[AB]_{q, \zeta}$ is to be interpreted as the Poisson bracket in the space $\mathscr{R}(q, q')$. Naturally, the density ρ in (18) is taken to be in the same space.

Equations (7) and (8) assume the form

$$\rho(q)\, dq = \rho(q,\, q)\, dq, \tag{7'}$$

$$\rho(p)\, dp = dp \int \rho(q,\, \zeta) \frac{e^{-iq\zeta/\hbar^*}}{2\pi\hbar^*}\, dq\, d\zeta. \tag{8'}$$

The normalization condition (10) states:

$$\int \rho(q,\, q)\, dq = 1. \tag{10'}$$

Equation (9) for the determination of the average value of the dynamical quantity $L(q, p)$ is brought by the same methods to the form

$$\bar{L} = \int \rho(q,\, \zeta)\, L^*(q,\, \zeta)\, dq\, d\zeta$$

or, what is the same thing, to the form

$$\bar{L} = \int \rho(q,\, q')\, L^*(q,\, q')\, dq\, dq'. \tag{9'}$$

In addition to these formulas we write out the formulas for the Fourier components of the coordinate q and the momentum p:

$$q_{q'} = q\delta(\zeta), \tag{16'}$$

$$p_{q'} = -i\hbar^* \frac{d\delta(\zeta)}{d\zeta}. \tag{17'}$$

These expressions coincide completely with the well-known quantum mechanical expressions for the operators $\hat{q}$ and $\hat{p}$ in the coordinate representation.

However, one should remember that the quantities (16') and (17') undergo multiplication as Fourier components (and not as matrices).

In concluding this section I introduce two simple examples of the solution of Eq. (20).

a) Free motion

In this case $V(q) = 0$. We represent $\rho(q, \zeta, t)$ in the form of the Fourier integral:

$$\rho(q,\, \zeta,\, t) = \int \widetilde{\rho}(\alpha,\, \beta)\, e^{i\omega(\alpha,\, \beta)\, t - i(\alpha q + \beta \zeta)}\, d\alpha\, d\beta. \tag{22}$$

Substitution of $\rho(q, \zeta, t)$ in this form into (20) leads to the relation

$$\left[\omega(\alpha,\, \beta) + \frac{\hbar^*}{m} \alpha\beta \right] \widetilde{\rho}(\alpha,\, \beta) = 0, \tag{23}$$

from which it follows that

$$\omega(\alpha, \beta) = -\frac{\hbar^*}{m}\alpha\beta. \tag{24}$$

Setting the right hand side of (23) equal to unity, we obtain the Fourier transform of the Green's function

$$G(\alpha, \beta) = \frac{1}{\omega(\alpha,\, \beta) + (\hbar^*/m)\, \alpha\beta \pm i\varepsilon} \tag{25}$$

for the equation (17) for the free motion of particles.

b) The harmonic oscillator

In this case $V(q) = (m\omega_0^2/2)q^2$, where ω_0 is the frequency of the oscillator, while m is its mass. In this case the function $\rho(q, \zeta, t)$ can be sought in the form $\rho(q, \zeta, t) = e^{i\omega t}\rho(z)$, where $z = q\zeta/\Lambda^2$, while $\Lambda^2 = \hbar^*/m\omega_0$. In terms of this variable the equation for $\rho(q, \zeta, t)$ can be easily brought to the form

$$\frac{d^2\rho}{dz^2} + \frac{1}{z}\frac{d\rho}{dz} - \left(1 + \frac{\omega}{\omega_0}\frac{1}{z}\right)\rho = 0. \tag{26}$$

The general solution of this equation is expressed in terms of the hypergeometric series $_1F_1$ [5]:

$$\rho(z) = e^{\pm z}{}_1F_1\left(\frac{1}{2} \mp \frac{\omega}{2\omega_0},\; 1,\; \mp 2z\right). \tag{27}$$

3. THE QUANTUM ENSEMBLE

As far back as in 1932 von Neumann introduced the important distinction between pure quantum ensembles (in von Neumann's terminology "einheitliche Gesamtheiten") and mixed ensembles ("gemischte Gesamtheiten"). [1] The former ensembles—ensembles described by a wave function ψ, correspond to the case of maximal information allowable by the laws of quantum mechanics. The latter type of ensembles contains states with different wave functions $\psi_1, \psi_2, \ldots, \psi_s, \ldots$, concerning which only the probabilities of these states $W_1, W_2, \ldots, W_s, \ldots$ are known. Such an ensemble is analogous to the ensembles of classical statistical mechanics, but not to classical mechanics itself, in which there is no place for such an ensemble. A mixed ensemble is described by the density operator $\hat{\rho}$.

The necessity noted in Sec. 1 of the introduction into quantum mechanics of the density operator as of a concept more general than the wave function, is based on the fact that in the quantum domain measurements carried out on systems described by the wave function ψ ("pure" ensemble), bring these systems into states described by a set of wave functions, i.e., into a "mixed" ensemble.

Therefore, if we wish to regard the theory of quantum measurements as a chapter of quantum mechanics then it is not possible to exclude from consideration mixed ensembles which have no analogues in classical mechanics. They are analogues of statistical mechanics. In this point is contained the whole essence of the difference of my conception of quantum mechanics from

---

[5] Cf., for example,: J. Watson, Theory of Bessel Functions, (Russ. Transl. M., 1949, p. 118).

the conception of the Copenhagen school.

N. Bohr clearly preferred to consider the situation when an atomic system is described by a wave function (i. e., a pure ensemble).

In such an approach the process of measurement itself is completely excluded from a quantum mechanical investigation and *a fortiori* can not be the object of a theoretical calculation. The interpretation of measurement in such an approach is restricted to the concept of measurement as a phenomenon of the alteration in the information available. It should be emphasized that within the framework of an analysis concentrated on a pure ensemble such an interpretation of measurement is logically consistent and is the only one possible. But it excludes the possibility which in fact exists on the basis of the same quantum mechanics to investigate and to calculate the phenomenon of measurement. In this connection von Neumann's concept based on the concept of statistical sets appears to be a broader base for understanding quantum mechanics than the concept based on a more restricted concept of wave function. The ideas of von Neumann presented by him in a brilliant but difficult to assimilate book: "Mathematical Principles of Quantum Mechanics" (1932[1]) have exerted in their time considerable influence on L. I. Mandel'shtam, [6] particularly on K. V. Nikol'skii (cf., his monograph[7]) and on myself. In contrast to us these ideas, apparently, did not generate in their time much interest on the part of N. Bohr.

On the basis of the foregoing material it is natural to regard quantum mechanics as *a generalization of classical statistical* mechanics. The representation of classical statistical mechanics in the space $\mathscr{R}(q, q')$ turns out to be a convenient starting point. This representation, as was shown in Sec. 2, deals with the Fourier components of the dynamical variables $L(q, q')$ and of the density $\rho(q, q')$.

We now follow the recipe for the transition from classical mechanics to quantum mechanics due to W. Heisenberg. The essence of its prescription reduces to two points: a) replacement of the Fourier components of dynamical variables by elements of Hermitian matrices and b) replacement of the classical Poisson bracket by the quantum Poisson bracket.

Turning to classical statistical mechanics represented in the space $\mathscr{R}(q, q')$ we realize this program with the aid of the following formulas which express the correspondence between classical and quantum quantities according to Heisenberg[6]:

$$L(q, \zeta) \equiv L(q, q') \rightarrow L(q, q') = L^+(q, q'), \quad (28)$$

$$\rho(q, \zeta) \equiv \rho(q, q') \rightarrow \rho(q, q') = \rho^+(q, q'). \quad (29)$$

This replacement means that the dynamical variables

---

[6]W. Heisenberg had in mind Fourier components reflecting the dependence of dynamical variables on the time. In my case this correspondence principle is extended also to the dependence of the dynamical variables on the coordinates.

and functions of them become *Hermitian operators* which we shall denote by $\hat{L}, \hat{\rho}$ etc. $\hat{L}^+$ and $\hat{\rho}^+$ denote Hermitian-conjugate operators. Further, according to point b),

$$[A, B]_{\text{class}} \rightarrow [\hat{A}, \hat{B}]_{\text{qu}} = -\frac{1}{i\hbar}(\hat{A}\hat{B} - \hat{B}\hat{A}). \quad (30)$$

In particular,

$$[\hat{p}, \hat{q}] = 1. \quad (31)$$

Following (28), (29) and (30) we replace in (9') $\rho(q, q')$ and $L(q, q')$ by the corresponding elements of the Hermitian matrices $\hat{\rho}$ and $\hat{L}$. We then obtain

$$\bar{L} = \text{Sp}(\hat{\rho}\hat{L}), \quad (32)$$

where Sp denotes the trace of a matrix. Formulas (7'), (8') and (10') take on the form

$$\rho(q)\, dq = \rho(q, q)\, dq, \quad (33)$$

$$\rho(p)\, dp = \rho(p, p)\, dp \quad (34)$$

and

$$\text{Sp}\,\hat{\rho} = 1. \quad (35)$$

Concurrently in all the formulas the constant $\hbar^*$ is now fixed and is equal to the Planck constant $\hbar$. This removes the arbitrariness in the scale in the choice of the constant $\hbar^*$ which is characteristic for classical statistical mechanics.

Equation (20) for the density $\rho(q, q')$ in accordance with (30) is replaced by an operator equation with the quantum Poisson brackets:

$$\frac{\partial\hat{\rho}}{\partial t} + [\hat{\mathscr{H}}, \hat{\rho}]_{\text{qu}} = 0. \quad (36)$$

The operator for the Hamiltonian function $\hat{\mathscr{H}}$ has the form (in the simplest case)

$$\hat{\mathscr{H}} = \frac{1}{2m}\hat{p}^2 + V(\hat{q}). \quad (37)$$

We now write out Eq. (36) in the space $\mathscr{R}(q, q')$. In this space the matrices for the operators $\hat{p}$ and $\hat{q}$ satisfying Eq. (31) have the form

$$p(q, q') = -i\hbar\frac{\partial}{\partial q}\delta(q - q'), \quad (38)$$

$$q(q, q') = q\delta(q - q'). \quad (39)$$

These expressions are identical with the expressions for the Fourier components of the quantities p and q in classical statistical mechanics represented in the space $\mathscr{R}(q, q')$. In order to verify this one should recall that in formulas (16') and (17') the quantity $\zeta = q' - q$.

Utilizing the rules for multiplication of matrices, Eqs. (36), (37), (38) and (39) yield Eq. (36) in the space $\mathscr{R}(q, q')$:

$$\frac{\partial\rho(Q, \zeta)}{\partial t} - \frac{i\hbar}{m}\frac{\partial^2\rho(Q, \zeta)}{\partial Q\,\partial\zeta} - \frac{1}{i\hbar}\left[V\left(Q + \frac{\zeta}{2}\right) - V\left(Q - \frac{\zeta}{2}\right)\right]\rho(Q, \zeta) = 0, \quad (40)$$

where $Q = \frac{1}{2}(q + q')$, $\zeta = q' - q$. Comparison with (20)

shows that the equation of classical statistical mechanics with $\hbar^* = \hbar$ can be regarded as an approximation to the exact quantum equation (36) under the condition of sufficiently smooth potentials $V(q)$ and smooth distributions $\rho(q, \zeta)$. In this case in (36) one can set $V(Q + \zeta/2) - V(Q - \zeta/2) \approx (\partial V/\partial q)\zeta + \dots$. Then (40) coincides with (20).

4. EIGENVECTORS

We turn first of all to the special case when the density operator $\hat{\rho}$ obeys the special condition

$$\hat{\rho}^2 = \hat{\rho}. \tag{41}$$

We inquire what other conditions must be satisfied by this operator in order that the statistical ensemble described by it should have the characteristic that a certain specified dynamical variable L in it should have *one* and only one definite value.

Let the average value of this quantity be $\overline{L}$, then the mean squared deviation is equal to

$$\overline{\Delta L^2} = \overline{(L - \overline{L})^2}. \tag{42}$$

Let $\hat{\mathcal{L}}$ be the Hermitian operator representing the quantity L. Then, in accordance with the basic formula (32), $\overline{\Delta L^2}$ is defined by the formula

$$\overline{\Delta L^2} = \mathrm{Sp}\,\{\hat{\rho}(\hat{\mathcal{L}} - \overline{L})^2\} \tag{43}$$

and the requirement that the quantity L should have only the single value $L = \lambda$, reduces to the condition

$$\mathrm{Sp}\,\{\hat{\rho}\,(\hat{\mathcal{L}} - \lambda)^2\} = 0. \tag{44}$$

We denote $\hat{\rho}(\hat{\mathcal{L}} - \lambda) = \hat{C}$, then $(\hat{\mathcal{L}} - \lambda)\hat{\rho} = \hat{C}^+$. Using (43) and the possibility of permuting the operators preceded by the symbol Sp, we obtain from (44)

$$\mathrm{Sp}\,\{\hat{C}\hat{C}^+\} = 0. \tag{45}$$

This condition can be satisfied only in the case when the operator $\hat{C} = 0$, and consequently also $\hat{C}^+ = 0$. Thus, we arrive at the equations for the operator $\hat{\rho}$:

$$(\hat{\mathcal{L}} - \lambda)\,\hat{\rho} = 0, \quad \hat{\rho}\,(\hat{\mathcal{L}} - \lambda) = 0. \tag{46}$$

We investigate these equations in the coordinate representation, i.e., for the matrix elements $\rho(q, q')$ of the operator $\hat{\rho}$. In the first of these equations the operator $\hat{\mathcal{L}}$ acts on the argument q of the element $\rho(q, q')$, and in the second the operator $\hat{\mathcal{L}}^*$ acts on the argument q' of the same matrix element. Therefore from the first equation of (46) it follows that the element $\rho(q, q')$ is proportional to the eigenfunction $\psi_\lambda(q)$ of the operator $\hat{\mathcal{L}}$, which belongs to the eigenvalue λ; this function obeys the equation

$$\mathcal{L}\,\hat{\psi}_\lambda(q) = \lambda\psi_\lambda(q). \tag{47}$$

In a similar manner from the second equation it follows

that $\rho(q, q')$ is proportional to $\psi_\lambda^*(q')$:

$$\hat{\mathcal{L}}^*\psi_\lambda^*(q') = \lambda\psi_\lambda^*(q'). \tag{47'}$$

It is well known that the functions $\psi_\lambda(q)$ form an orthogonal system. We can take it to be orthonormalized. Then

$$\rho(q, q') = \psi_\lambda(q)\,\psi_\lambda^*(q'). \tag{48}$$

It is not difficult to verify that condition (47) holds and equations (46) are satisfied.

The eigenfunctions of Hermitian operators form a system of basis vectors in Hilbert space. Any other arbitrary vector in this space $\varphi(q)$ can be represented in the form

$$\varphi(q) = \sum_\lambda c_\lambda \psi_\lambda(q), \tag{49}$$

where c_λ are its components.

Applying the operator $\hat{\rho}_\lambda$ to φ, we obtain from (48) and (49)

$$\hat{\rho}_\lambda \varphi = c_\lambda \psi_\lambda(q). \tag{50}$$

From this it follows that the operator $\hat{\rho}_\lambda$ is the projection operator on the λ axis: $\hat{\rho}_\lambda \equiv \hat{P}_\lambda$. In the more general case

$$\rho(q, q') = \varphi(q)\,\varphi^*(q') \tag{48'}$$

is the projection operator on the vector φ (regarding it to be normalized to unity).

It is not difficult to show that the vectors $\varphi(q)$ in Hilbert space satisfy the Schrödinger equation

$$i\hbar\frac{\partial\varphi}{\partial t} = \hat{\mathcal{H}}\varphi. \tag{51}$$

To prove this it is sufficient to substitute $\rho(q, q')$ in the form (48') into equation (36) and to divide the result by $\varphi(q)\varphi^*(q')$. This substitution leads to the relation

$$\left[\frac{\partial\varphi(q)}{\partial t} + \frac{1}{i\hbar}\hat{\mathcal{H}}\varphi(q)\right]\varphi^{-1}(q) + \left[\frac{\partial\varphi^*(q')}{\partial t} - \frac{1}{i\hbar}\hat{\mathcal{H}}\varphi^*(q')\right][\varphi^*(q')]^{-1} = 0.$$

From this it follows that $[\partial\varphi(q)/\partial t + \cdots] = ic\varphi(q)$, where c is a real constant which can be incorporated into $\hat{\mathcal{H}}$ and compensated for by a shift of the reference point for energy.

Thus, starting with (28) and (36) we obtain all the equations of linear wave mechanics.

The statistical ensemble described by the density matrix which is a projection operator is referred to as a *pure* ensemble. A pure ensemble corresponds to the description of quantum phenomena with the aid of a single wave function.

The initial equations (32) and (36) enable us to consider an ensemble with a density matrix $\hat{\rho}$ of a more general form, in particular

$$\hat{\rho} = \sum_{\lambda} W_{\lambda} \hat{\rho}_{\lambda}, \qquad (52)$$

where

$$W_{\lambda} \geqslant 0 \quad \text{and} \quad \sum_{\lambda} W_{\lambda} = 1. \qquad (53)$$

Such an ensemble in the terminology of von Neumann is said to be *mixed*. The quantities W_{λ} indicate the probability of finding the system being studied in the state λ, belonging to a pure ensemble described by the density operator $\hat{\rho}_{\lambda}$. From (52) and (53) we obtain

$$\hat{\rho}^2 = \sum_{\lambda} W_{\lambda}^2 \hat{\rho}_{\lambda} \leqslant \hat{\rho}. \qquad (54)$$

In concluding this section we indicate the formula for the matrix elements of $\hat{\rho}_{\lambda}$ in its own λ-representation:

$$\rho_{\lambda \lambda'} = \delta_{\lambda \lambda'}. \qquad (55)$$

In the same representation for the mixed ensemble we have

$$\rho_{\lambda \lambda'} = W_{\lambda} \delta_{\lambda \lambda'}. \qquad (56)$$

5. MEASUREMENTS AND IRREVERSIBILITY

We preface our discussion of measurements by some comments concerning the relationship of the density operator $\hat{\rho}$ to thermodynamical statistics. This connection was pointed out in the same monograph by von Neumann. [11] J. von Neumann proposed a formula for the entropy S of a system generalizing the well-known formula due to L. Boltzmann, i.e.,

$$S = -k \, \mathrm{Sp} \, \{\hat{\rho} \ln \hat{\rho}\}; \qquad (57)$$

here k is the Boltzmann constant. From this formula it follows at once that for a pure ensemble $S = 0$. In order to verify this it is sufficient to bring the operator $\hat{\rho}$ to diagonal form. According to (55) the eigenvalue of such an operator is equal to 1, and $\ln 1 = 0$.

This feature of a pure ensemble is an expression of the fact that a pure ensemble is a statistical set of systems *which are all in the same state*.

For a mixed ensemble in the same representation on the basis of (56) we obtain

$$S = -k \sum_{\lambda} W_{\lambda} \ln W_{\lambda} > 0. \qquad (58)$$

Therefore the entropy of a mixed ensemble is always greater than the entropy of a pure ensemble. This result was also proved by von Neumann. [11]

If the system under consideration is in thermal equilibrium with a large heat bath of temperature Θ, then, according to the theory of a Gibbs thermodynamic ensemble

$$W_{\lambda}(\Theta) = e^{(F - E_{\lambda})/\Theta}, \qquad (59)$$

where $\Theta = kT$, k is the Boltzmann constant, E_{λ} are the eigenvalues of the energy operator $\hat{H}$, F is the free energy. Substitution of (59) into (58) leads to the relation well known from thermodynamics

$$F = E - TS, \qquad (60)$$

where E is the average value of the energy of the system

$$E = \sum_{\lambda} E_{\lambda} W_{\lambda}(\Theta) \qquad (61)$$

with the usual normalization condition: $\sum_{\lambda} W_{\lambda}(\Theta) = 1$.

After these remarks we turn to the process of measurement. The process of measurement is based on the physical process of the interaction of a microsystem with a macroscopic system MA-measuring apparatus. This apparatus must necessarily be a macroscopically unstable system. If this were not so a microsystem could not activate it. It does not have sufficient energy and momentum to accomplish this. Theoreticians have not paid sufficient attention to this important circumstance which is trivial for experimenters. At an early stage of development of quantum mechanics the attention of theoreticians was more concentrated on a new circumstance—on the effect of the measurement on the state of the quantum system.

The significance of this circumstance for the understanding of the process of measurement was shown in papers by the author of the present article. [8]

In these papers it was shown that the process of measurement begins at a microscopic quantum mechanical level and in virtue of the macroscopic instability of the measuring apparatus (MA) is converted into a macroscopic process. Therefore the process of measurement has the nature of an explosion initiated by the microsystem being measured.

For the mathematical description of this process the use of the apparatus of the density operator $\hat{\rho}$ is completely necessary. This necessity arises from the circumstance that the macroscopic apparatus being from the microscopic point of view, a complex system cannot be described by a wave function. In order to include the apparatus into the quantum mechanical description it is necessary to employ the concept of a mixed ensemble and, consequently, the apparatus of the density operator $\hat{\rho}$. We denote the dynamical variables describing the state of the system undergoing measurement by the letter x, and the dynamical variables describing the macroscopic apparatus by the letter q. These variables can be very numerous. Generally speaking, the apparatus can also be ascribed a certain temperature Θ. The density operator $\hat{\rho}$ of the compound system will depend on the variables x, q, the time, and, possibly, on the temperature Θ in such a manner that the matrix element of the operator $\hat{\rho}$ in the representation of the variables x and q will be $\rho(x, q; x', q', t)$. We expand this operator in terms of the eigenfunctions $\psi_n(x)$ of the operator $\hat{\mathcal{L}}$, which represents the quantity being measured, which for simplicity we shall assume to be discrete (n is the number of the eigenvalue $\lambda = L_n$):

$$\rho(x, q; x', q', t) = \sum_{n,m} W_{nm}(q; q', t)\, \psi_n(x)\, \psi_m^*(x'). \tag{62}$$

Among the variables q only a few will be observable. For the sake of definiteness we assume that there is only one such variable: $q = Q$.[7]

Taking a diagonal element with respect to all the q, with the exception of $q = Q$, we integrate (62) over these variables; as a result we obtain a matrix with respect to the observable variables Q:

$$\rho(x, x', Q; Q', t) = \sum_n W_{nm}(Q, Q', t)\, \psi_n(x)\, \psi_m^*(x'). \tag{63}$$

In virtue of the macroscopic nature of the apparatus the nondiagonal elements with respect to Q are vanishingly small. Thus, in (63) $Q' = Q$. The device will be a measuring one if as $t \to \infty$ [8] the quantities $W_{nm} = 0$ for $n \neq m$, while W_{nn} differs from zero only in the case if the observable variable Q lies in the region $Q \in \Omega_n$ (Ω_n does not intersect with Ω_m, $n \neq m$). Taking into account all possible results of measurement: $Q \in \Omega_1, Q \in \Omega_2, \ldots,$ $Q \in \Omega_n, \ldots$, we integrate (63) over all the possible Q. We then obtain

$$\rho(x, x', t) = \sum_n W_{nn}(t)\, \psi_n(x)\, \psi_n(x'), \tag{64}$$

where

$$W_{nn}(t) = \int_{\Omega_n} W_{nn}(Q, t)\, dQ \tag{65}$$

and

$$W_{nm}(t) = 0. $$

From these formulas it can be seen that the coherence of the different particular states of the microsystem $\psi_n(x), \psi_m(x), \ldots$ is destroyed as a result of the interaction with the macroscopic measuring apparatus. This destruction of coherence is due to the indefiniteness of the microscopic variables of the apparatus q and to the macroscopic nature of the observed variables Q.

We consider now the nature of the ensemble arising after measurement. We assume that the initial ensemble was pure:

$$\rho(x, x') = \varphi(x)\, \varphi^*(x') \tag{66}$$

and

$$\varphi(x) = \sum_n c_n \varphi_n(x), \tag{67}$$

where $\varphi_n(x)$ are the eigenfunctions of the operator $\hat{\mathcal{L}}$ which represents the dynamical variable L. If this quantity is not altered in the process of measurement

(for this it is necessary that the operator for the interaction energy of the system with the apparatus $\hat{W}$ should commute with the operator $\hat{\mathcal{L}}$), then it can be shown that as $t \to \infty$ the probability is proportional to $|c_n|^2$:

$$W_{nn}(t)_{t \to \infty} \approx |c_n|^2. \tag{68}$$

Therefore with suitable normalization in the L-representation the elements of the matrix $\rho(x, x', t)$ are equal to

$$\rho_{nm} = |c_n|^2\, \delta_{nm}. \tag{69}$$

If after the measurements one collects the representatives of the system with $L = L_n$ into the n-th box, those with $L = L_m$ into the m-th box, etc., then the set of particles in each box is a pure ensemble described by the density matrix $\rho(x, x') = \varphi_n(x)\varphi_n^*(x')$. We shall be dealing with several independent pure ensembles. The entropy of each of them is equal to zero. From information theory it is known that the volume of information can be measured by entropy.[9,10] In the method of carrying out measurements described above the volume of information is not changed. Only the form of information is altered. It was maximal in the initial ensemble and remains maximal in the set of ensembles after measurement.[9]

But if all the systems after measurement are collected into a single box, then we shall have a mixed ensemble described by the matrix

$$\rho(x, x') = \sum_n |c_n|^2\, \varphi_n(x)\, \varphi_n^*(x'), \tag{70}$$

$$\sum_n |c_n|^2 = 1, \tag{70'}$$

the entropy of which is equal to

$$S = -k \sum_n |c_n|^2 \ln |c_n|^2. \tag{71}$$

This increase in entropy corresponds to the loss of information which was available previously with respect to the initial ensemble and was maximal—it was equal to zero.

Thus, the question of the nature of the ensemble arising after measurement is resolved differently depending on the manner in which the systems subjected to measurement are assembled. It can be a mixed ensemble, but it can also be a pure ensemble or several pure ensembles. The latter case arises when the observer studying the new ensemble has information concerning *which* of the "boxes" is occupied by the system with a given value of $L = L_n$.

[1]J. von Neumann, Mathematischen Grundlagen der Quantenmechanik, Berlin, J. Springer, 1932 (Russ. Transl. Fizmatgiz, M., 1963).
[2]J. Gibbs, Fundamental Principles of Statistical Mechanics

[7]Q is the position of the conventional "indicator" of the measuring apparatus.
[8]We interpret $t \to \infty$ as the duration of the process of measurement.

[9]We do not consider the possible and usually occurring increase in the entropy of the measuring apparatus.

689   Sov. Phys. Usp. 20(8), Aug. 1977

D. I. Blokhintsev   689

(Russ. Transl. Gostekhizdat, M., 1946).

[3]R. V. Mises, Wahrscheinlichkeit, Statistik und Wahrheit, Berlin, J. Springer, 1928.

[4]D. Blokhintsev and Ch. Briskina, Vestn. Mosk. Univ. No. 10 (1942).

[5]D. I. Blokhintsev, Printsipial'nye voprosy kvantovoǐ mekhaniki (Fundamental Problems in Quantum Mechanics), M., Nauka, 1966.

[6]L. I. Mandel'shtam, Collected works, v.5, M.-L. pub. by Academy of Sciences of the USSR, 1950.

[7]K. V. Nikol'skiǐ, Kvantovye protsessy (Quantum Processes),

M., Gostekhizdat, 1940.

[8]D. I. Blokhintsev, Usp. Fiz. Nauk 95, 75 (1968) [Sov. Phys. Usp. 11, 320 (1969)] cf., also Ref. 5 and Osnovy kvantovoǐ mekhaniki (Fundamentals of Quantum Mechanics), 5th ed., M., Nauka, 1976.

[9]R. P. Poplavskiǐ, Usp. Fiz. Nauk 115, 465 (1975) [Sov. Phys. Usp. 18, 222 (1975)].

[10]A. M. Yaglom and I. M. Yaglom, Veroyatnost' i informatsiya (Probability and Information), M., Fizmatgiz, 1960.

Translated by G. Volkoff