The tree of paradox

V. V. Mityugov

Institute of Applied Physics, Russian Academy of Sciences, Nizhnii Novgorod (Submitted 9 March 1993) Usp. Fiz. Nauk 163, 103-114 (August 1993)

The specific nature of the randomness arising when quantum subsystems interact (collide) is analyzed. It is shown that the Birkhoff-Khinchin ergodic theorem—the key theorem for classical statistics—or its analog is absent, in principle, in the quantum theory. Thus quantum probabilities cannot be defined within the ergodic concept. A metric definition of probability, based on von Neumann's theory of measurement, is proposed as a measure of comparison of *a posteriori* physical situation with the *a priori* situation. The workability of the adopted approach is demonstrated for random walk problems and the theory of thermal equilibrium.

...The scientist is possessed by the sense of universal causation. The future, to him, is every whit as necessary and determined as the past. A. Einstein^{1a)}

...The man who regards his own life and that of his fellow creatures as meaningless is not merely unhappy but hardly fit for life. A. Einstein^{1b)}

The two phrases presented in the epigraph probably express the quintessence of the great physicist's emotional anguish. The contradiction between dynamical causality and the significance of human will, the problem of the origin of randomness, and the basis of thermodynamics have disturbed people for hundreds of years. Meanwhile, the famous Einstein–Podolsky–Rosen paradox¹ is already a kind of seed, capable of producing a fruit-bearing tree of answers to many agonizing physical-philosophical questions. The fact that A. Einstein himself recoiled, in confusion, from this mysterious abyss does not diminish his greatness, and it should not stop us. Besides, the quantum theory has matured.

The logical basis of the paradox is the analysis of the detailed quantum description of a collision of two particles which are initially in pure states (with definite momenta). The first fundamental conclusion is that it is impossible to associate a wave function to the state of each particle after the collision. The very peculiar statistical nature of quantum laws, which has no classical analog, is revealed here. The second important point arises when the observer makes an exact measurement of the momentum of one of the collision partners. The possibility, ensuing from a conservation law, of instantaneous indirect reduction of the state of the second particle comprises the crux of the paradox (apparent paradox, as is now clear).

The potential richness of the above-described physical construction is contained in the diverse naturalness of its generalizations. Analysis of two arbitrary physical subsystems, instead of a pair of particles, in arbitrary initial states is the starting point of the theory of indirect measurements, the foundation for which was later laid by J. von Neumann.² Extending the model by increasing the number of interacting subsystems sheds light on many formerly controversial questions concerning the foundation of statistical thermodynamics. Finally, analysis of these problems provides hope for extrication from the impasse of the fundamental mathematical problem of defining the concept of physical probability.

The cornerstone of probability in classical physics is the ergodic hypothesis and the closely associated microcanonical ensemble. It can be shown that the fundamental ergodic theorem of Birkhoff and Khinchin or its analog does not exist, in principle, in a systematically quantum theory. As a consequence, within the ergodic concept quantum probabilities are a purely intuitive notion. On the other hand, however, von Neumann's axiomatic theory of measurement makes it possible to introduce probability independently as a metric characteristic in the space of quantum states of a physical system, as a measure of comparison of the *a posteriori* physical situation to the *a priori* situation. The consequences of adopting such a definition of probability are worked out in detail in the second half of this paper.

Physicists have long been unsatisfied with the state of affairs concerning the concept of probability. Among the extensive literature concerning this question we call attention to only one of the recent publications, Ref. 3, where many sore points are analyzed and where further references can be found. The paper by Demutskii and Polovin on the methodological problems of quantum mechanics are in the same vein. Historical retrospection shows that in science different logical constructions coexist under the general pseudonym "probability." It is difficult to think of greater trouble for a scientific concept. It now seems obvious to me that the solution to the question of randomness—such a necessary element of the human awareness of the universe—is contained in the quantum theory.

1. MIXED STATES

Proceeding now to analysis of the logical scheme of the paradox in its modern interpretation, we introduce the necessary concepts and make some clarifications. Let $L = L_1 \otimes L_2$ be the direct (Kronecker; see Ref. 5) product of the Hilbert spaces of the states of two noninteracting physical subsystems. Let X and Y be Hermitian operators which operate in L_1 and L_2 , respectively. We introduce, in the standard fashion, the eigenvectors and eigenvalues

$$\mathbf{X}|x\rangle = x|x\rangle, \quad \mathbf{Y}|y\rangle = y|y\rangle.$$
 (1)

If the subsystems are not one dimensional, we take X and Y to be complete (for them) sets of operators of the physical quantities and we interpret x and y as the corresponding sets of quantum numbers. On the basis of the arguments presented in Ref. 6, we assume that the sets $\{x\}$ and $\{y\}$ are discrete.

Consider a state $|\Psi\rangle \in L$. If the subsystems interacted in the past, then $|\Psi\rangle$ does not, generally speaking, factorize into a product of states $|\psi_1\rangle \in L_1$ and $|\psi_2\rangle \in L_2$. The brackets $\langle x, y | \Psi \rangle$ in the old "pre-Dirac" notation is simply the function $\Psi(x, y)$.

Since they do not have definite wave functions, the states of the subsystems are characterized by (mixed) density matrices:

$$\langle x | \rho | x' \rangle = \sum_{y} \langle x, y | \Psi \rangle \langle \Psi | x', y \rangle, \qquad (2)$$

$$\langle y | \sigma | y' \rangle = \sum_{x} \langle x, y | \Psi \rangle \langle \Psi | x, y' \rangle.$$
(3)

Here we observe a fundamentally nonclassical randomization of the states of the subsystems in that the system as a whole has a wave function. One would think that a timely and full acknowledgment of this fact would dissuade from searches for primary fundamental randomness in continuous classical nonlinear dynamical models. The other point is that these models are certainly applicable to macrosystems.

The mixed states (1) and (3) of the post-collision subsystems are associated with interesting reciprocity properties, which will be discussed below. For the time being we focus our attention on methods for representing the density matrix $\langle x | \rho | x' \rangle$. Its diagonal elements $\langle x | \rho | x \rangle$ are the probabilities p(x) of observing the values (or complete sets) of x in a physical measurement of the variable X. It is convenient to write these probabilities in an invariant manner with the help of projection operators $\mathbf{P}_x = |x\rangle \langle x|$:

$$P(x) = \operatorname{Sp} \hat{\rho} \mathbf{P}_{x} = \langle P_{x} \rangle, \tag{4}$$

where $\hat{\rho}$ is the operator corresponding to the matrix (2) in the x-representation. Being Hermetian, this matrix can be diagonalized by a unitary transformation

$$\sum_{x,x'} \langle n | V^+ | x \rangle \langle x | \rho | x' \rangle \langle x' | V | m \rangle = \rho_n \delta_{nm}, \qquad (5)$$

where $\mathbf{V}^+\mathbf{V}=\mathbf{1}$ in L_1 . The values of ρ_n are, once again, the probabilities p(n) of obtaining the results of a physical measurement of N, whose eigenvectors are $|n\rangle$ from Eq. (5) with the obvious properties of orthogonality $\langle n|m\rangle = \delta_{nm}$ and completeness $\Sigma |n\rangle \langle n| = 1$ in L_1 . In other words, $[\hat{\rho}, N] = 0$.

The diagonalization (5) makes it possible to represent the state $\hat{\rho}$ in the form of a probabalistic ensemble (mixture)

$$\hat{\rho} = \sum_{n} p(n) |n\rangle \langle n| \tag{6}$$

and to interpret it now as a random collection of orthogonal states $|n\rangle$, which can be constructed, for example, by throwing dice. It is easy to show, however, that there exists an uncountable set of different representations of the same operator as the mixture

$$\hat{\rho} = \sum p(\alpha) |\alpha\rangle \langle \alpha|$$

of some states $|a\rangle$ which are not mutually orthogonal. The completeness of the statistical description contained in the density operator precludes formulation of the inverse problem of reconstructing the original physical-probabalistic mixture from the results of measurements.

The nonuniqueness indicated above has an exceedingly important fundamental significance, making it impossible to transfer classical concepts and theorems about ergodic ensembles into the quantum theory. In order to prove this we recall some required information about the theory of stationary random processes in mathematical and physical interpretations.

Let a probability measure W(F(t)) be given on a set of time-dependent functions $\{F(t)\}$.

Definition 1. An ensemble is stationary if the measure W is invariant under an arbitrary time shift:

$$W(F(t)) = W(F(t+\tau)).$$
⁽⁷⁾

Definition 2. A stationary ensemble is ergodic if it cannot be represented as a sum of two stationary subensembles with measures different from 0 or 1.

In other words, for an ergodic ensemble there does not exist a representation of the form

$$W = p_1 W_1 + p_2 W_2, \quad p_1 + p_2 = 1$$
 (8)

with $p_{1,2} \neq 0,1$ under conditions such that W_1 and W_2 are stationary distributions.

According to the Birkhoff-Khinchin theorem (see Refs. 7 and 8) for any physical quantity (interpreted as a single- or many-time combination of instantaneous values of the function F) the time average of any realization F(t) does not differ from the average over an ergodic ensemble. Thus, each member of the set $\{F(t)\}$ is typical and contains absolutely all information about the statistical properties of the ensemble. In the case of a multidimensional

random process F(t) must be interpreted as a vector in a suitable space, and everything said above remains in force.

In the quantum theory, to the classical functions F(t)there are associated vectors of some nonstationary states $|\alpha\rangle$ of a physical model of a random process or, in the operator representation, the projectors $|\alpha\rangle\langle\alpha|$. The density operator constructed for the mixture from these states with the probabilities $p(\alpha)$ makes it possible to calculate physical averages in a manner similar to the way classical averaging over an ensemble W is performed. A fundamental difference is that the vectors $|\alpha\rangle$ are themselves not observable, while in the classical theory all (without exception) significant characteristics of a state are observable simultaneously.

The main object of the application of the ergodic theory in classical physics is a system of \mathcal{N} identical particles. The instantaneous state of such a system is represented by a point in 6 \mathcal{N} -dimensional phase space, and correspondingly F(t) is a vector of the same dimension. The requirement that the system of particles be ergodic is extremely important in order to justify the validity of the microcanonical ensemble, which is the logical foundation of Gibbs' classical statistics (see Ref. 9): In other words, it is important that the phase trajectory of the system return repeatedly arbitrarily close to a neighborhood of each point with the same energy. It is very difficult to give a rigorous proof of ergodicity of the behavior of a system of particles interacting according to a prescribed law. For this reason, in most textbooks on statistical physics the required property of an ensemble is formulated as the ergodic hypothesis.

The above-noted nonuniqueness of the representation of the density operator as a probabilistic mixture of pure states changes the situation radically. It turns out that the fundamental relationship of classical statistics between the structure of a stationary ensemble and the properties of typical phase trajectories does not exist in principle in the quantum theory. We shall now prove this.

Consider a finite-dimensional physical system with the Hamiltonian H. Let the density operator $\hat{\rho}$ describe a mixed stationary state of this system. Stationarity means that $[\hat{\rho}, \mathbf{H}] = 0$. The average value of any physical quantity (besides artificial constructions that are explicitly time-dependent) in such a state is time independent (see Ref. 10). We define in the standard manner the energies E_n and vectors $|n\rangle$ of the stationary pure states

$$\mathbf{H}|n\rangle = E_n|n\rangle. \tag{9}$$

Some E_n can have the same values. By assigning to them, nonetheless, different numbers n we eliminate the need for studying energy degeneracy separately and we thereby simplify the notation.

We write $\langle n | \rho | m \rangle$ and $\langle n | H | m \rangle$ in a representation in which are both diagonal. Then $\langle n | \rho | m \rangle = \rho_n \delta_{nm}$, and ρ_n is the probability of observing the energy E_n (reducing the state to $|n\rangle$) with a suitable measurement. Such an ensemble can be represented as a mixture of the form (6) consisting of pure stationary states $|n\rangle$ with the probabilities $p(n) = \rho_n$. In the mixed state $\rho_n \neq 1$ for any *n* (the entropy of the state is nonzero). For this reason, such a probabilistic mixture can be divided in at least one way into two subensembles, each of which is stationary. Therefore the classical definition of an ergodic stationary ensemble and the Birkhoff-Khinchin theorem cannot be transferred into the quantum theory. As I have already mentioned, together with this, there does exist an infinite set of representations of the same operator as a mixture of nonstationary pure states (comparable to classical trajectories).

We shall see below that for the reason mentioned above the quantum probabilities p(x) themselves are an intuitive notion, remaining outside the ergodic concept. At least, this is how the situation appears from the standpoint of the traditional classical scheme.

2. QUANTUM PROBABILITIES

The question of the logical dilemma of the theory of probability and of causality and randomness is much older than the idea of a wave function. In a brilliant philosophical-mathematical essay J. Littlewood¹¹ showed that physical probability cannot, in principle, be defined within the theory of limits. Repeated attempts to break the vicious circle by reexamining the very concept of a mathematical limit or on the basis of the theory of relative frequencies of appearance were unsuccessful. An alternative axiomatic approach to constructing a theory of probability-measure theory-appeared to be more successful, but this has no direct bearing on natural science. Professing the dogmas of the faith of pure mathematics, the most consistent members of this school dismiss, in principle, physical applicability of their logical constructions. They have to do so.

In the applied theory of probability the situation is saved by the agreement to study only ergodic random processes. Correspondingly, ergodic sources of messages appear in the theory of signal transmission. Adopting ergodicity as an additional postulate, one can write the probability density of an instantaneous value f of a random function F(t) as the average of a projector $P_f = \delta(F - f)$. Then the Birkhoff-Khinchin theorem guarantees that over an infinitely long time the relative frequencies of events will converge to probabilities.

The problem of convergence is solved similarly for time series of discrete symbols, if an event of a discrete alphabet $\{f\}$ in discrete time is considered instead of continuous random functions F(t). All required definitions and theorems remain in force. We note that misuse of the actual infinity due to operation with continuous quantities and continuous sets is, in general, characteristic for modern mathematics. Of course, there are very serious logicalhistorical reasons for this, but it is this circumstance that too often forces (it also helps out) mathematicians and, especially physicists "to sweep the dust under the rug" (in the colorful words of R. Feynman).

Returning to the discussion of quantum probabilities, we see that, on the strength of everything said above, they remain a purely intuitive notion, unsupported by mathematical definitions. On the other hand, the axiomatics of J. von Neumann already incorporates the possibility of an independent definition of physical probability, not based on the ergodic theorem.

According to Ref. 2, any complete direct measurement is an orthogonal decomposition of unity $1=\Sigma |x\rangle \langle x|$ in the Hilbert space of the states of the physical system. The probability of obtaining a result x in a measurement on the state $|\psi\rangle$ is $|\langle x|\psi\rangle|^2$.

The states $|x\rangle$ and $|\psi\rangle$ are logically equivalent. Indeed, the vector $|\psi\rangle$ is necessarily normalized by the condition $\langle \psi | \psi \rangle = 1$, and it is always possible to construct for it a set of orthogonal complements, having together with the initial state $|\psi\rangle$ the property of mutual completeness with respect to the basis $\{|x\rangle\}$. Based on this we formulate the following definition:

Definition 3. The squared modulus of the scalar product of the state vector in a unitary unimodular Hilbert space is the probability of a reductive transition from one state into another accompanying a suitable measurement.

The term "reductive" means that the transformation $|\psi\rangle \rightarrow |x\rangle$ occurs in the process of measurement (reduction of the wave vector; see Ref. 2).

We now discuss the formal and gnosiological consequences of this definition of probability. First, we note that the concept "probability" itself now appears as an unnecessary tautology of metric relations in unitary space. On the other hand, the physical description of the situation is, in the axiomatic sense, exhaustive. From the procedural standpoint, the adopted definition means that all operations of physical prediction must be constructed using the same formulas as in the case of probability in the previous heuristic sense.

We are essentially talking about the postulation of spectral measure,¹² organically consistent with the formalized procedures of physical measurements. On any prescribed orthogonal basis this quantity exhibits all required properties of probabilistic measure: additivity, positivedefiniteness, and normalization. We shall show below that in a number of fundamental problems of physical kinetics and statistical thermodynamics the assumption of such a measure leads to standard results and gives them a natural formal justification.

There arise here a number of specific problems dictated by the application of the above-formulated scheme to models of successive measurements on interacting physical subsystems. In particular, in the derivation of the reciprocity relations we shall also consider the problem of indirect measurements, which arises in the problem, mentioned at the very beginning of the paper, of the interaction of two subsystems. There arises here also the possibility of formulating the problem of parametrizing the proper time of the observer as a metric property for a chain of successive quantum measurements.

The previously sacred question of convergence of the relative frequencies to probabilities is no longer fundamental. The question is not completely eliminated, but it is relegated to a very specific class of homogeneous infinite models of successive measurements. It is entirely logical to expect that for such models the properties of convergence can and must be proved on the basis of group-theoretical classification of quantum states or combinatorial asymptotic relations in the spirit of limit theorems. However, this question is of secondary importance for the foundations of the theory itself.

In other cases it is impossible to require anything similar to such convergence and it is not necessary to do so, if the problem concerns inhomogeneous models or, especially, single events, which usually comprise the historical process in the macro- and microworld. But it should not be assumed that some formal structure will give an automatic solution at a pivotal moment. The problem is that it is the probabilistic approach that is usually employed for constructing computer algorithms of searching for an optimal action. But, as a purely logical construct the concept of an optimal solution is hardly informative for unique situations. Moreover, human free will can still precede nonoptimally. But then again, to proceed automatically is also a willful decision.

Acknowledgment of the fact that metric relations in Hilbert space are the only reality of predictive knowledge does not lead to any logically absurd consequences and agrees well with existing laws of quantum and statistical physics. The disagreement between the proposed scheme and the presently adopted paradigm concerns only speculative questions that cannot be checked experimentally.

Representations of a quantum ensemble as probabilistic mixtures are also found to be meaningless. There remains only the completeness of the physical description, consisting of the density operator and realized in terms of its projection $\langle x | \rho | x \rangle$ on definite orthogonal bases. There are no observable consequences to denying ontological meaning to the indicated representations.

3. RECIPROCITY RELATIONS

We now return to the discussion of our initial model of two subsystems. We shall consider the correspondence between quantum randomness arising in the presence of interaction and the definition 3. Consider some pure initial states $|k\rangle \in L_1$ and $|\chi\rangle \in L_2$. The interaction (collision) of the subsystems subjects the state $|k,\chi\rangle$ of the complete system consisting of these substates to a unitary transformation: $\mathbf{U}|k,\chi\rangle = |\Psi\rangle$. We shall clarify, in the spirit of the adopted definition, the probabilistic meaning of the description of the resulting states with the help of the density matrices (2) and (3).

We write the joint probability p(x,y) of the result of a simultaneous measurement of X and Y in the complete system (obviously, [X,Y]=0):

$$p(x,y) = |\langle x,y | \Psi \rangle|^2 = |\langle x,y | U | k,\chi \rangle|^2.$$
(10)

Since here X and Y are arbitrary physical variables of the subsystems (or complete sets of variables), the equation (10) now incorporates everything necessary for further discussion.

We introduce in the standard manner¹³ the absolute and conditional probabilities:

$$p(x) = \sum_{y} p(x,y) = \langle x | \rho | x \rangle, \qquad (11)$$

$$p(y) = \sum_{x} p(x,y) = \langle y | \sigma | y \rangle, \qquad (12)$$

$$p(y|x) = \frac{p(x,y)}{p(x)}, \quad p(x|y) = \frac{p(x,y)}{p(y)}, \quad (13)$$

as well as the entropy

$$S(x) = -\sum_{x} p(x) \ln p(x), \qquad (14)$$

$$S(y) = -\sum_{y} p(y) \ln p(y),$$
 (15)

$$S(x,y) = -\sum_{x,y} p(x,y) \ln p(x,y).$$
 (16)

Shannon's inequality¹⁴ is satisfied:

$$S(x) + S(y) \ge S(x,y). \tag{17}$$

The absolute probabilities p(x) and p(y) are given by the diagonal elements of the matrices (2) and (3) for any choice of variables in L_1 and L_2 , respectively. We choose the bases $\{|n\rangle\} \in L_1$ and $\{|v\rangle\} \in L_2$, in which these matrices are diagonal $\langle n | \rho | m \rangle = \rho_n \delta_{nm}$, $\langle \mu | \sigma | v \rangle = \sigma_v \delta_{\mu v}$. The functionals

$$S_1 = -\sum_n \rho_n \ln \rho_n = -\operatorname{Sp} \hat{\rho} \ln \hat{\rho}, \qquad (18)$$

$$S_2 = -\sum_{\nu} \sigma_{\nu} \ln \sigma_{\nu} = -\operatorname{Sp} \hat{\sigma} \ln \hat{\sigma}$$
(19)

are called the quantum entropies of the subsystems in the states (2) and (3). According to the lemma of O. Klein (see Refs. 15 and 16)

$$S(x) \ge S_1, \quad S(y) \ge S_2.$$
 (20)

The complete system initially resided in a pure quantum state, characterized by the density operator $R = |k,\chi\rangle\langle k,\chi|$ and zero quantum entropy S=0. The latter fact did not change, naturally, even after the interaction (the quantum entropy of the complete system is an invariant of any unitary transformation in L), while S_1 and S_2 increased.

We note especially that the situation arising here is unique not only for modern physics with its short thirtyyear history. The fact that the subsystems must be described probabilistically while the complete system is in a definite physical state contradicts the more than thousandyear tradition of natural philosophy. The apparent paradoxical nature of the situation is dictated by the possibility of representing the state vector as a linear superposition of two or more physically completely distinguishable states. In this sense all prequantum descriptions of systems would correspond to a scheme which contain only orthogonal quantum states. The principle of mutual projection of physical states is a substantially new concept not only for physics but for logic also. This circumstance was first pointed out by Yu. F. Orlov (Yu. F. Orlov, "Wave mathematical analysis based on wave logic" (unpublished paper)), though the idea that classical logic must be supplemented with new categories is itself already contained in the works of J. von Neumann.¹⁷

The principle of linear superposition of wave vectors presupposes the use of a metric measure in the construction of any predictive strategy. Any physical theory is only a formalized variant of one such strategy.

In this sense the appearance of the concept of probability even within (more precisely, parallel to) the classical picture of the world is instructive. The point is that the deterministic geometric-dynamical models of all of the old physics (including the theory of relativity) turned out to be in striking contradiction to everyday human experience—in contradiction to the existence of free will at least for the observer-subject himself. The concept of randomness and probabilistic measure with an unclear physical meaning served temporarily as a palliative measure, mitigating somewhat the acuteness of this contradiction.

The formal situation in today's physics is completely similar (and genetically related) to the situation in the mathematical understanding of combinatorics. The binomial coefficients and other factorial combinations first arose in mathematics purely on the basis of games. They were soon found to be useful also for physics, modeling some real-world numerical relations. However, their application to physical prediction once again required additional assumptions, such as equal probability, complete mixing, or other constructions equivalent to the ergodic hypothesis.

In parallel with this, it was found unexpectedly that all combinatorial constructions which are well-known to mathematicians arise naturally in the matrix elements of the theory of linear representations of finite and unitary groups.¹⁸ On the other hand, it gradually became apparent that the transformation symmetries of different physical states are of determining significance for their observable properties. It is sufficient to recall at least the difference between the quantum statistics of gases consisting of identical particles with different permutation symmetry. It is now almost obvious that this is where combinatorial relations originate in models in physical science.

From this standpoint, the construction of the quantum theory in its modern form is itself logically interpreted as a natural step in F. Klein's famous erlang program. This program is based on the ideal of constructions in mathematics (and essentially, the natural sciences) in which different branches of the formal sciences would appear as methods for studying groups of transformations and their representations (symbol isomorphisms). It is difficult to miss the fact that it is this type of construction that is going on in today's fundamental quantum physics.

We now continue our discussion of two statistically coupled subsystems. The subsystem, having lost as a result of the collision part of its "memory" about its previous state, seemingly replaces this state with information about the state of its partner. The quantum entropies of the resulting states satisfy the reciprocity relation:¹⁶

$$S_1 = S_2. \tag{21}$$

The colliding subsystems need not be identical in order for this equality to be valid (for example, the particles need not be identical). Conversely, the collision partners can have different dynamical properties and numbers of degrees of freedom. In order for Eq. (21) to hold in any model it is sufficient that the state of the complete system be pure (S=0).

Equation (21) can be proved by direct calculation, using the property of unitarity of the operator U. However, we shall proceed somewhat differently, and in the process we shall study the question of indirect measurements. We construct an *a posteriori* distribution over $\{y\}$ according to the result of the measurement of X. We represent the conditional probabilities p(y|x) from Eq. (13) as the diagonal elements of the conditional density matrix $\langle y | \sigma^{(x)} | y' \rangle$:

$$p(y|x) = \langle y | \sigma^{(x)} | y \rangle = \frac{|\langle x, y | U | k, \chi \rangle|^2}{\langle x | \rho | x \rangle}.$$
 (22)

According to the rules adopted above this matrix contains the complete physical description of the state of the second subsystem, arising after the measurement of the given variable X in the first subsystem. Indeed, since Y is still arbitrary, an exhaustive prediction of subsequent measurements over the second subsystem is given by the matrix $\langle y | \sigma^{(x)} | y' \rangle$ and all its representations in L_2 .

It is easy to show that $\langle y | \sigma^{(x)} | y' \rangle$ describes a pure state:

$$\langle y | \sigma^{(x)} | y' \rangle = \frac{\langle x, y | U | k, \chi \rangle \langle k, \chi | U^+ | x, y' \rangle}{\langle x | \rho | x \rangle}.$$
 (23)

This state should be compared to the conventional wave vector $|\varphi(x)\rangle$ (result of indirect reduction)

$$|\varphi(x)\rangle = \frac{\langle x | U | k, \chi \rangle}{(\langle x | \rho | x \rangle)^{1/2}},$$
 (24)

which already has an internal meaning independently of the choice of basis $\{|y\rangle\}$.

We now calculate the scalar product of conventional vectors, reduced with different results x (but, as before, with fixed X):

$$\langle \varphi(x') | \varphi(x) \rangle = \sum_{y} \langle \varphi(x') | y \rangle \langle y | \varphi(x) \rangle$$
$$= \frac{\langle x | \rho | x' \rangle}{(\langle x | \rho | x \rangle \langle x' | \rho | x' \rangle)^{1/2}}, \qquad (25)$$

where $\{|y\rangle\}$ is any orthonormalized basis in L_2 . It is interesting that according to Eq. (25) the set $\{|\varphi(x)\rangle\}$ is, in general, nonorthogonal.¹⁶ We shall show below that this significant fact opens the way to understanding nonstationary indirect measurements, comparable to the classical observations of phase trajectories.

Orthogonal indirect reduction is realized only if $X \neq N$, when $\langle n | \rho | m \rangle = \rho_n \delta_{nm}$. This is achieved by measuring the variable N, which commutes with $\hat{\rho}$. We now use the properties of the states generated by such a measurement. The set { $|\varphi(n)\rangle$ } is not necessarily complete in L_2 though it is orthogonal. Supplementing it in an orthogonal manner up to $\{|\nu\rangle\} \ni \{|\varphi(n)\rangle\}$, it is easy to see that the density matrix $\langle \mu | \sigma | \nu \rangle$ of an absolutely mixed state of the second subsystem is diagonalized precisely in this basis:

$$\langle \mu | \sigma | \nu \rangle = \sum_{n} \rho_{n} \langle \mu | \varphi(n) \rangle \langle \varphi(n) | \nu \rangle = \sigma_{\nu} \delta_{\mu\nu}.$$
 (26)

This is possible only if the operators $\hat{\rho}$ and $\hat{\sigma}$ have identical sets of nonzero eigenvalues. In particular, there follows hence the equality

$$\operatorname{Sp} \hat{\rho}^{K} = \operatorname{Sp} \hat{\sigma}^{K}, \tag{27}$$

where K is any positive integer, and the power is taken according to the standard rules of operator multiplication. Finally, for the same reason, all nonzero terms in the sums (18) and (19) are equal pairwise; this proves Eq. (21).

The reciprocity relations (21) and (27) turn out to be simply truisms, if on interaction any additive physical quantity is conserved, and the initial states $|k\rangle$ and $|\chi\rangle$ are its eigenvectors in the subsystems. We now demonstrate this for the example of energy conservation for subsystems with nondegenerate spectra.

Let \mathbf{H}_1 and \mathbf{H}_2 be the Hamiltonians of the subsystems and \mathbf{H}_{int} the interaction between the subsystems. The total energy is conserved in the interaction if $[\mathbf{H}_{int}, \mathbf{H}_1 + \mathbf{H}_2] = 0$. We note that for this possibility the spectrum of the operator $\mathbf{H}_1 + \mathbf{H}_2$ must be degenerate even under our assumption that the spectra of \mathbf{H}_1 and \mathbf{H}_2 are nondegenerate.

We assume that the initial states $|k\rangle$ and $|\chi\rangle$ are pure and stationary:

$$\mathbf{H}_1|k\rangle = E_k|k\rangle, \quad \mathbf{H}_2|\chi\rangle = E_{\chi}|\chi\rangle. \tag{28}$$

The unitary operator of the evolution of the complete system in the interaction representation is $\mathbf{U} = \exp(-i\mathbf{H}_{int}t)$, where t is the interaction time (we employ the scale in which $\hbar = 1$). Under the assumptions made, the resulting density matrices $\langle n | \rho | m \rangle$ and $\langle \mu | \sigma | \nu \rangle$ describe the stationary mixed states of subsystems and are diagonal in the energy representation.

We now study in L the subspace of states $|n,v\rangle$ with identical $E_n + E_v$, equal to the initial energy $E_k + E_{\chi}$. Then the probability $p(n) = \rho_n$ of the value E_n is identical to p(v) with $E_v = E_k + E_{\chi} - E_n$ simply by virtue of conservation of energy. Thus the nature of the relations (21) and (27) is now absolutely clear. In this sense, the general variant of the indicated relations actually means that for any unitary rotation in L it is always possible to construct an additive invariant (to within unitary unphysical rotations in L_1 and L_2).

We now discuss in the same model the case of a nonstationary measurement. This means that the quantity X measured in the first subsystem does not commute with H_1 . Correspondingly, its eigenvectors $|x\rangle$ describe nonstationary states. The results of indirect reduction $|\varphi(x)\rangle$ the *a posteriori* states of the second subsystem—will also be nonstationary. According to Eq. (25), difference states $|\varphi(x)\rangle$ are still not orthogonal to one another. We now explain why this is significant.

In order to establish the correspondence between the quantum equations and the classical laws of motion it is necessary to study wave packets which are not mutually orthogonal. This is why such states were widely discussed at the dawn of the quantum theory. Nonspreading Gaussian packets for a quantum oscillator (coherent states-Ref. 19) give an adequate description of a monochromatic wave field with known amplitude and phase. They were found to be necessary not only for semiclassical interpretation of quantum optics, but they also made it possible to simplify significantly some exact calculations. The coherent vectors $|\alpha\rangle$ are not mutually orthogonal, though they do form a complete (more precisely, supercomplete) set in the Hilbert space of the states of the oscillator. To classical observation of an oscillator it would be reasonable to associate reduction of its states to such vectors. It is clear from what was said above that the logical link necessary for this possibility incorporates nonstationary indirect measurements.

We emphasize that with respect to the initial mixed state the *a posteriori* states in themselves are not potentially nonstationary or stationary—everything depends on the measurement method. They simply do not exist at all *a priori* as some real entities hidden from the observer. This once again elucidates the inapplicability of the ergodic concept to observation of quantum systems.

Completing our analysis of the coupling of two subsystems, we present without proof an extension of some entropy inequalities to the case of mixed initial states. The genesis of such states must once again be understood as resulting from a previous interaction of subsystems with some external environment, but we shall not delve into this aspect of the question here.

Thus, assume that instead of the initial state $|k,\chi\rangle$ the initial states are described by the operators $\hat{\rho}_0$ and $\hat{\sigma}_0$ and are independent of one another. Following the same scheme as in Eqs. (18) and (19), we compare their initial quantum entropies S_{01} and S_{02} , setting $S=S_{01}+S_{02}$. Designating, as before, the resulting quantum entropies of the subsystems after the interaction as S_1 and S_2 , we present the inequality (see Ref. 16)

$$|S_1 - S_2| \leqslant S. \tag{29}$$

Since $(S,S_1,S_2) \ge 0$, this inequality generalizes the relation (21).

In addition, it follows from Shannon's inequality (17) and Klein's lemma in the form $S(x,y) \ge S$ that^{15,20}

$$S_1 + S_2 \ge S. \tag{30}$$

We shall see below that Eq. (30) is of key significance for the foundation of statistical thermodynamics.

4. RANDOM WALK

Physical kinetics has turned out to be an extensive sphere of application of probability theory in natural sciences. The problem of successive evolutions of a physical system under the random action of external forces is the crux. It is now known that the kinetics is only one (though important) branch of the general theory of open quantum systems.²¹ L. Boltzmann essentially viewed a gas molecule as an open system. The evolutionary equation for the single-particle distribution function (Boltzmann's kinetic equation) forms the basis of this theory. For completeness of our analysis, it is appropriate to trace how a description of random walk in a quantum system which agrees with the standard description is achieved within the framework of a formally metric concept of probability.

We introduce the basic concepts required for describing an open system.^{21,22} Let the surrounding medium be a collection of identical physical subsystems (uniform flux), with which the open system (object) of interest to us interacts successively. Let $\hat{\rho}_0$ and $\hat{\sigma}_0$ be the density operators of the initial states of the object and one of the external subsystems (representative of the flux), assumed to be statistically independent. We shall study successive interactions between the representatives of the flux and the object, and we shall show that under certain assumptions the evolution of its states can be described in the formalism of Markov chains.

Let H_s and H_r be the Hamiltonians of the object and the representative, respectively, and let *n* and *v* be the numbers of their eigenvalues (we assume that the spectra are nondegenerate). We introduce again the interaction H_{int} imposing the condition $[H_{int}, H_s + H_r) = 0$. On the basis of the formula from the preceding section, we write the change in the state of the object for one act of interaction as

$$\langle n | \rho_1 | m \rangle = \sum_{\nu} \langle n, \nu | U \rho_0 \sigma_0 U^+ | m, \nu \rangle.$$
(31)

In the particular case when a light beam interacts with the object continuously, the time t of one interaction act must be taken as the coherence time of the beam.

We now enumerate with the index j the successive interactions of the object with the representatives of the flux and designate by ρ_j the density operator of its state after the *j*th action. Assuming the initial states of all representatives to be identical and independent of one another, we obtain with the help of Eq. (31)

$$\langle n | \rho_{j+1} | m \rangle = \sum_{p,q} (n,m | p,q) \langle p | \rho_j | q \rangle$$
(32)

where

$$(n,m|p,q) = \sum_{\mu,\nu,\lambda} \langle n,\mu | U | p,\nu \rangle \langle \nu | \sigma_0 | \lambda \rangle \langle q,\lambda | U^+ | m,\mu \rangle.$$
(33)

We call the matrix (33) the transitional characteristic of a generalized Markov chain. We note that the elements (n,m|p,q) are not, generally speaking, transitional probabilities. They have the obvious property

$$\sum_{n} (n,n|p,q) = \delta_{pq}, \qquad (34)$$

which follows from the unitarity of the operator U and the normalization Sp $\hat{\sigma}_0 = 1$.

We now require that the initial states be stationary: $[\hat{\rho}_0, \mathbf{H}_s] = 0$, $[\hat{\sigma}_0, \mathbf{H}_r] = 0$. Then in the energy representation $\hat{\rho}_0$ and $\hat{\sigma}_0$ are diagonal matrices. Moreover, with the assumed conservation of total energy $[\mathbf{H}_{int}, \mathbf{H}_s + \mathbf{H}_r] = 0$, all $\langle n | \rho_j | m \rangle = \rho_n^{(j)} \delta_{nm}$ will remain diagonal (see Ref. 23), and instead of Eq. (32) we obtain

$$\rho_n^{(j+1)} = \sum_m p(n|m) \rho_m^{(j)}$$
(35)

where p(n|m) = (n,n|m,m) are the standard transitional probabilities of a classical Markov chain,²⁴ whose alphabet is the energy spectrum of the object. The required normalization for p(n|m) is obviously guaranteed by the formula (34).

The coefficients in Eqs. (32) and (35) do not depend on whether a measurement is performed on the object after a successive interaction. Of course, the initial state for calculating successive evolutions of the object after each measurement (reduction) must be chosen according to the result obtained. The chain (35) is essentially the kinetic equation with discrete time (0,1,...,j,...), where p(n|m) is expressed in terms of Eq. (33) and is structurally similar to the kernel of the collision integral. Similarly, Eq. (32) is the quantum kinetic equation, in which the transitional characteristic (n,m|p,q) does not necessarily have a probabilistic meaning.²⁵

The random walk of an object under the action of a purely stationary flux $\vec{\sigma}_0 = |\chi\rangle \langle \chi|$, where $\mathbf{H}_r |\chi\rangle = E_{\chi} |\chi\rangle$, is fundamentally nonclassical. Then the observed randomness is entirely engendered directly under the action of the flux on the object and not in previous interactions with some different physical environment. In this case the walk is once again described by the simple chain (35), where

$$p(n|m) = \sum_{\nu} |\langle n, \nu | U | m, \chi \rangle|^2.$$
(36)

Thus within a quite natural physical model of an open system there is no need for any additional postulates for substantiating the Markov character of the walk. There is also no need for explicit or implicit use of hidden sources of randomness, besides the principle of linear superposition of quantum states and the metric rule adopted with the *definition 3*.

5. THERMAL EQUILIBRIUM

The traditional approach to the statistical foundations of phenomenological thermodynamics consists of studying a closed system of particles. The classical works of Zermelo, Poincaré, and Gibbs (see, for example, Ref. 26) contain a series of very profound and demonstrative assertions, but they do not solve the problem completely. Analytical papers on the ergodic hypothesis still attract the attention of applied mathematicians and theoretical physicists.

The model of an isolated system of particles occupies a somewhat untypical position in physics. It cannot be classified among exactly solvable models, on the basis of which the development of the theory usually precedes more confidently. Historically such an approach dates back to specialists on celestial mechanics, and it is unproductive for the logical development of science.

Meanwhile, within the theory of open systems the canonical ensemble as some limiting type of mixed state is obtained on the basis of strict inequalities. The logical equivalent of the second law of thermodynamics is already contained in the inequality (30). Some metaphysical questions from previous discussions in such an approach indeed cannot be answered, but there is hardly any need for this.

In order to fulfill the indicated program we introduce the concept of a stable state of an open system exposed to a uniform flux.²² (We used the term "stable" state instead of the cliché "stationary state," which in this context has two meanings and is often used in the Russian scientific literature in application to open systems.) Once again we confine our attention to stationary uniform fluxes and energy-conserving interactions. Turning to Eqs. (32) and (35), we require that the object's state $\hat{\rho}$, whose trace is to be taken, after interaction with the flux did not change:

$$\rho_n = \sum_m \rho_m \sum_{\mu,\nu} \sigma_{\mu}^{(0)} |\langle n,\nu | U | m,\mu \rangle|^2.$$
(37)

In terms of the chain (35) this means that stable ρ_n can be found as the solution of a system of linear equations, obtained from the requirement $\rho_n^{(j+1)} = \rho_n^{(j)}$. The problem of possible singularity of the matrix of coefficients p(n|m)merits attention. However, this problem does not arise for maximum-entropy states with which we shall be concerned here.

In its most general formulation an equation of the type (37) potentially incorporates the entire theory of a stable state of an open quantum system. In particular, this refers to the theory of homeostasis—maintenance of stability of a substantially nonequilibrium state of an object due to entropy increase in a nonequilibrium flux. We, however, shall confine our attention to maximum-entropy distributions.

Let the aggregate state of the object and representative of the flux have maximum quantum entropy S with prescribed average total energy $\langle \mathbf{H}_s + \mathbf{H}_r \rangle$. The density operator **R** of a system in such a state is obtained from Elsasser's variational equation²⁷

$$\ln \mathbf{R} = \ln \hat{\rho} + \ln \hat{\sigma}_0 = -(1+\eta)\mathbf{1} - \alpha(\mathbf{H}_{\rm s} + \mathbf{H}_{\rm r}). \quad (38)$$

It is easy to see that this gives simply a Gibbs energy distribution. The undetermined Lagrange multipliers α and η have a simple physical meaning: $\alpha^{-1} = T$ is the energy temperature and $\exp(1+\eta) = Z$ is the usual partition function. For most physical systems studied in different applications the maximum entropy S_{max} is a monotonic function of $\langle \mathbf{H}_{s} + \mathbf{H}_{r} \rangle$.

We shall show that under such initial conditions the state $\hat{\rho}$ has the property of absolute stability, i.e., it is a solution of Eq. (37), independently of the specific form of the interaction operator \mathbf{H}_{int} or the evolution operator U. Indeed, by virtue of Eq. (30) and the conservation of energy the sum of the quantum entropies of the subsystems cannot change, and by virtue of the uniqueness of the solution (38) the states themselves also do not change. Thus

a stable state of the object, which is described by a Gibbs distribution at the same temperature, is established asymptotically under the action of a uniform equilibrium flux (thermostat). If, however, the flux is not in equilibrium, then according to Eq. (30) the total entropy can only increase. It is this separateness of the mixed states with maximum entropy that has made possible equilibrium thermodynamics (more precisely, thermostatics). In cases when, together with energy, other additive invariants of interaction exist, the notion of statistical equilibrium can be generalized and the class of absolutely stable states enlarged.²²

The description, using the same language, of energy exchange in quasiequilibrium systems makes it possible to associate it to a change in entropy, thereby substantiating the traditional phenomenological description of heat transfer. For this, it is once again necessary to examine the Gibbsian initial states of the subsystems, but with different temperatures T_s and T_r . Assuming the interaction to be quite weak (the first order of perturbation theory for $\langle n, \nu | U | m, \mu \rangle$), under quite general other assumptions the following equality can be proved (see Ref. 22):

$$\Delta S = S_{\rm s} + S_{\rm r} - S = \Delta \langle H_{\rm s} \rangle \left(\frac{1}{T_{\rm s}} - \frac{1}{T_{\rm r}} \right), \tag{39}$$

where $\Delta \langle \mathbf{H}_s \rangle$ is the change in the average energy of the object in the course of the interaction, and S_s and S_r are the resulting quantum entropies of the subsystems. The formula (39) justifies the formal interpretation of quasiequilibrium energy transfer as heat exchange $\delta Q = T dS$, which forms the foundation of axiomatic thermodynamics.

The ideology of Eq. (37) also suggests a natural approach to problems of nonequilibrium thermodynamics. In particular, it is easy to construct examples that make it possible to check, using the exact methods of quantum theory, the validity of Prigogine's principle²⁹ of minimum energy production in a stable state.

6. CONCLUSIONS

Thus the proposed metric scheme gives a substantiated formalism for solving many important problems in physical stochastics. In each case it is important only to construct an adequate model of the initial state and the physical interaction, after which the measurement procedure can be formalized unequivocally (in the theory of noise and fluctuations the latter is often not done, and this can lead to misunderstandings). In all other respects, the pseudographic interpretations of randomness, often appealing to outdated fallacious forms of "classical common sense," must be treated with care.

The obvious genetic relationship between the models discussed and the Einstein–Podolsky–Rosen paradox is instructive. Einstein's subtle intuition apparently sensed the maze of age-old contradictions. But the principle of superposition of quantum states at that time did not yet form the foundation of physical thought, and even now it is still not completely recognized. The prejudice of the geometricdynamical ideal of the universe, dating back to the axiomatics of Euclid, in the ancient sources of Hellenic natural philosophy, is not so easily overcome. All our natural sciences, following mathematics, are possessed with the metaphysical conviction of the continuity of the universe, though even purely logically it is obvious that it is not reasonable to make a fetish of continuous sets and they are not necessary for knowledge. It can be conjectured that in future theories Hilbert space itself will be reduced to its discrete analog on the basis of the theory of finite groups.

The conviction of coming changes in the very foundations of the natural sciences is also supported by the state of affairs in modern mathematics. The attempt, continuing over many decades, to construct an autonomous (outside of human experience) edifice of formal sciences is ending with a complete ideological collapse.²⁹ One of the most impressive results of the reassessment of values currently going on is the Lewenheim-Skolem theorem about the possibility of discrete-mathematical modeling of any axiomatic system. At the same time, it is becoming clear that it is in principle necessary to employ natural-science factors external to mathematics in order to choose a mathematical model. It is now virtually obvious that mathematics can only be one of the languages of sense perception, and nothing more, and science itself and its language are faced with an even greater task of revising the reliability of its points of support.

The continual mentality almost inevitably leads to philosophical dynamism, to the elimination of creative volitional beginning from the nature of life. Here there are many questions to which there are no complete answers. The mystery is great, but there is a gleam of hope.

On a more general level, there is also discernable feedback of science with the spiritual yearnings of Euroamerican civilization over the last three decades. This theme is yet to be investigated, and the investigators will have to face a problem of unusual depth—to rethink competently and harmoniously the fundamental trinity of the oldest branches of development of the human spirit: culture, science, and religion.

^(a)A. Einstein, *Ideas and Opinions*, Dell Publishing Co., N.Y., 1973, p. 50; ^(b)*ibid.*, p. 23.

¹A. Einstein, *Collected Scientific Works* [Russian translation], Nauka, Moscow, 1966, Vol. 3, pp. 604–611.

²J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton, 1955.

³Yu. I. Alimov and Yu. A. Kravtsov, Usp. Fiz. Nauk 162, 149 (1992) [Sov. Phys. Usp. 35, 606 (1992)].

⁴V. P. Demutskii and R. V. Polovin, Usp. Fiz. Nauk 162, 93 (1992) [Sov. Phys. Usp. 35, 857 (1992)].

⁵M. Hamermesh, Group Theory and Its Applications to Physical Problems, Addison-Wesley Publishing Company, Redding, MA, 1962.

⁶A. S. Krest'yaninov and V. V. Mityugov, Radiotekh. Elektron. 31, No. 5, 891 (1986).

⁷A. N. Kolmogorov, Usp. Matem. Nauk, No. 5, 52 (1938).

⁸N. Wiener, *Prediction Theory—Modern Mathematics for Engineers* [Russian translation], Inostr. lit., Moscow, 1958, pp. 185–215.

⁹L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Pergamon Press, New York, 1980.

¹⁰L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, Pergamon Press, New York, 1977.

¹¹J. Littlewood, Mathematical Miscellany [in Russian; title not verified], Nauk, Moscow, 1978, pp. 56-59.

- ¹²M. B. Menskiĭ, Method of Induced Representations. Space-Time and the Concept of Particles [in Russian], Nauka, Moscow, 1976.
- ¹³W. Feller, An Introduction to Probability Theory and Its Applications, Wiley, New York, 1957.
- ¹⁴C. Shannon, Works on Information Theory and Cybernetics [in Russian], Inostr. lit., Moscow, 1963, pp. 433-460.
- ¹⁵ V. V. Mityugov, Problemy upravleniya i teorii informatsii 2, 243 (1973).
- ¹⁶ V. V. Mityugov, Physical Foundations of Information Theory [in Russian], Sov. radio, Moscow, 1976.
- ¹⁷ J. Von Neumann, General and Logical Theory of Automata; A. Turing, Can a Machine Think? [in Russian], GIFML, Moscow, 1960.
- ¹⁸N. Ya. Vilenkin, Special Functions and the Theory of Representations of Groups [in Russian], Nauka, Moscow, 1965.
- ¹⁹R. Glauber in *Quantum Optics and Quantum Radiophysics* [Russian translation], Mir, Moscow, 1966.
- ²⁰R. L. Stratonovich, Izv. Vyssh. Uchebn. Zaved. Radiofiz. 8, 116 (1965).

- ²¹E. B. Davies, Quantum Theory of Open Systems, AP, London, 1976.
- ²² V. V. Mityugov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. 32, 436 (1989).
- ²³G. Mackey, Mathematical Foundations of Quantum Mechanics, [Benjamin, N.Y., 1963].
- ²⁴A. N. Kolmogorov, Basic Concepts of Probability Theory [in Russian], Nauka, Moscow, 1974.
- ²⁵ E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics*, Pergamon Press, New York, 1981.
- ²⁶Yu. B. Runer and M. Sh. Ryvkin, *Thermodynamics, Statistical Physics, and Kinetics* [in Russian], Nauka, Moscow, 1977.
- ²⁷ W. M. Elsasser, Phys. Rev. 52, 987 (1937).
- ²⁸I. Prigogine, Introduction to the Thermodynamics of Irreversible Processes, Wiley, New York, 1961.
- ²⁹ M. Kline, *Mathematics: The Loss of Certainty*, Oxford University Press, N.Y., 1980.

Translated by M. E. Alferieff