Physics – Uspekhi **41** (9) 885 – 922 (1998)

REVIEWS OF TOPICAL PROBLEMS PACS number: 03.65.Bz

Basic quantum mechanical concepts from the operational viewpoint

D N Klyshko

Contents

Abstract. The physical meaning of the basic quantum mechanical concepts (such as the wave function, reduction, state preparation and measurement, the projection postulate, and the uncertainty principle) is clarified using realistic experimental procedures and employing classical analogies whenever possible. Photon polarization measurement and particle coordinate and momentum measurement are considered as examples, as also are Einstein – Podolsky – Rosen correlations, Aharonov – Bohm effects, quantum teleportation, etc. Various nonclassicality criteria of quantum models, including photon antibunching and the violation of the Bell inequality, are discussed.

D N Klyshko M V Lomonosov Moscow State University, Physics Department, Vorob'evy Gory, 119899 Moscow, Russia
Tel. (7-095) 939 11 04
E-mail: dnk@klyshko.phys.msu.su

Received 17 February 1998
Uspekhi Fizicheskikh Nauk **168** (9) 975 – 1015 (1998)
Translated by M V Chekhova; edited by L V Semenova

Theoretical cognition is speculative when it relates to an object or certain conceptions of an object which is not given and cannot be discovered by means of experience.
I Kant "Critique of Pure Reason"

1. Introduction

About a hundred years ago, the Planck formula for thermal radiation opened the list of victories of quantum physics. In all known experiments, excellent agreement is observed between the predictions of the quantum theory and the corresponding experimental data. Paraphrasing the famous words of Wigner, one can speak of 'the inconceivable efficiency of the quantum formalism'.

Unfortunately, the efficiency of the formalism is accompanied by difficulties in its interpretation, which have not yet been overcome. In particular, there is still no common viewpoint on the sense of the wave function (WF). Another important notion of quantum mechanics, the WF reduction, is also uncertain. Two basic types of understanding can be distinguished among a variety of viewpoints. A group of physicists following Bohr considers the WF to be a property of each isolated quantum system such as, for instance, a single electron (the orthodox, or Copenhagen, interpretation). The other group, following Einstein, assumes that the WF describes an ensemble of similar systems (the statistical, or

ensemble, interpretation). This question is discussed in more detail in the exhaustive review by Home and Whittaker [1]. In the remarkable textbook by Sudbery [2], there is a chapter named "Quantum metaphysics" where nine different interpretations of the quantum formalism are considered. Among many other studies devoted to methodological problems of quantum physics, it is also worth mentioning Refs [3 – 7].

In the present notes, the sense of some basic notions in nonrelativistic quantum physics is clarified using the *operational* approach, i.e., demonstrating how these notions manifest themselves in experiments. For the quantum models discussed here, the closest classical analogues are considered where possible. The present consideration may be entitled "Classical and quantum probabilities from the viewpoint of an experimenter". Using simple examples, we show common features of quantum and classical probability models and the principal differences between them. As far as possible, a comprehensible style is used and bulky mathematical expressions are avoided. Necessary algebra is given in Appendices.

Four basic topics are considered in the paper: (1) the logical structure of the quantum description; (2) the necessity of distinguishing between a theory and its interpretation; (3) the WF: its sense, preparation, modulation, measurement, and reduction; (4) the 'nonclassicality' of quantum physics, i.e., the impossibility of introducing joint probabilities for non-commuting operators. In this connection, nonclassical optical experiments are discussed.

The paper is organized as follows. In Section 2, the operational approach in physics is described and its significant role in the methodology of quantum physics is emphasized. Further, in Section 3, using classical probability models with dice or coins, we discuss several notions that are important for further consideration and have close analogues in quantum physics. In Section 4, general features of quantum models are considered, basic notions and terms of quantum physics are defined, and the general logical scheme of quantum dynamical experiments on measuring various observables, as well as the WF itself, are presented. Further, in Sections 5 and 6, the general formalism is illustrated using specific examples. These examples are two basic models of quantum mechanics, namely, a two-level system and a nonrelativistic point-like particle moving in one dimension. The simplicity of these models and the existence of their classical optical and mechanical analogues make them ideal objects for introductory courses in quantum physics and for discussing problems of methodology and terminology. Here we only consider some essential aspects of these models that are almost untouched in textbooks. A considerable part of Section 5 is devoted to optical experiments related to photon polarizations and demonstrating essential nonclassicality.

Here we mostly focus on dynamical experiments connected with the evolution of quantum systems in space and time. As a typical example, we consider the Stern – Gerlach experiment where particles with magnetic moment **M** are deflected in an inhomogeneous magnetic field (Fig. 1). Using this example, one can clearly specify the basic elements of a dynamical experiment: the *source* of particles S, the *detectors* D (crystals of silver bromide contained in the photosensitive film), the space between S and D where quantum evolution of the particles takes place, and the filters F_1, F_2. The source S and the collimator F_1 (a screen with a pinhole for spatial selection) form the *preparation* part of the setup. The magnet F_2 provides the inhomogeneous magnetic field that couples

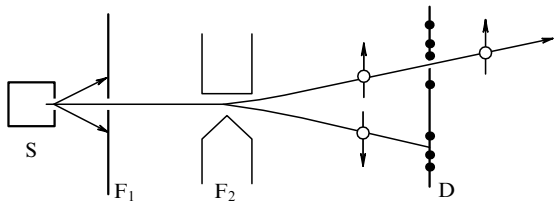

Figure 1. Schematic plot of the Stern – Gerlach experiment. S is the source of particles, F_1 is a screen with a pinhole (collimator), F_2 is a domain with an inhomogeneous magnetic field, D is a photographic plate. The elements F_1 and F_2 perform spatial and magnetic filtering and can be considered as parts of the preparation and measurement sections of the setup, respectively. If D contains a pinhole, then F_2 and D work as a filter, which sometimes transmits particles in the state with definite spin projection.

the spin and kinetic degrees of freedom of a particle. Together with the detectors D, the magnet can be considered as the *measurement* part of the setup. In such a scheme, only the evolution of a particle between the source and the detector is described by the Schrödinger equation accounting for the classical magnetic field. S, F_1, F_2, and D are supposed to be classical devices with known parameters.

In an idealized case, each individual particle is registered. The parameter directly *measured* in this experiment, namely, the classical coordinate x_1 of a black dot on the film, is determined, for instance, with the help of a calibrated ruler. The resulting dimensional value is assumed to be the *a priori* coordinate of the particle, i.e., the coordinate of the particle before it is absorbed by the film. (Of course the accuracy of such a measurement is restricted, for instance, by the size of a silver atom.) Thus, in this case, one can assume the coordinate operator X to be the directly *observable* operator. (This procedure is considered in more detail in Section 6.1.) Hence, for given parameters one can *calculate* the *a priori* projection of the particle moment m_x using the Schrödinger equation and the initial WF of the particle. This is an example of *indirect* measurement of the operator M_x.

If the photosensitive film D is replaced by a screen with a pinhole, we obtain a device that prepares the particle in a state with given moment projection m_x. Note that in this case, the operator M_x is not measured, and the screen with a pinhole plays the role of an additional filter. We see that the procedures of measurement and preparation are not identical as is supposed in the framework of the orthodox approach [2, 3]. However, in principle, it is possible that the particle is detected at a certain point of the plane D without being destroyed. This measurement gives information about the operator M_x of the moving particle. After that, one can measure M_x once more using a second set of devices and observe the correlation between the signals from the two detectors.

As a rule, capital letters $A, B, \ldots$, denote operators (*q-numbers*) and small letters $a, b, \ldots$, denote their eigenvalues and the parameters like mass m, charge e, time t (*c-numbers*, which correspond to classical observables). This rule will be violated in some cases, in order to follow traditional notation; for instance, the photon annihilation operator will be denoted by a. In the description of experiments, capitals correspond to registered values, such as, for instance, the coordinate of a particle, X, and small letters correspond to fluctuating values measured in various trials $(x_1, x_2, \ldots)$.

2. Operational approach

One of the most important, or maybe the most important tool for establishing a clear universal terminology in physics is the approach in which all basic notions are defined by means of appropriate experimental operations (procedures), i.e., the operational approach. Here we mean 'moderate' operationalism where only basic notions are defined via (more or less) realistic experiments. In addition, it is possible to use convenient notions that have only an indirect relation to experiment.

As in any accomplished branch of physics, nonrelativistic quantum physics includes four basic components.

(1) Mathematical models.

(2) Rules of correspondence between mathematical formalism and experiment. The aim of the operational approach, which forms the basis of the present paper, is namely to establish a mapping between two sets: the set of symbols and the set of experimental procedures.

(3) Experiments that either confirm or disprove a mathematical model or the rules of correspondence (see the epigraph to this paper). According to Popper, any scientific statement should admit *falsification* (disproof). Many philosophers reject this viewpoint; however, without such criteria it is difficult to distinguish between science and pseudoscience like parapsychology.

(4) Interpretation of the formalism and the experiment. This includes verbal definitions of symbols and descriptions of idealized models, explicit images and figures. This component is closely related to philosophy, gnoseology, semantics, etc. Here one can specify a group of *metaphysical* notions, which are introduced without any necessity, in spite of the principles laid by Ockham, Newton, and Kant. In our opinion, an example of such a redundant notion, which is useless for quantitative theory, is given by the *partial reduction* of the field WF occurring as a result of detecting one of two correlated photons (see Sections 5.5 and 5.7). This subset of useless notions has no fixed boundaries: some time ago, atoms could also be classified as a metaphysical notion. Metaphysical notions and explicit models play an important role in any theory at the initial stages of its development.

This extremely simplified structurization of physics (and of the professional activity of physicists) is certainly not the only one possible. A lot of efforts have been made in this direction. An interesting approach, which emphasizes the principal role of models, is being developed by Lipkin [7].

Let us consider the uncertainty relation for two arbitrary Hermitian operators A and B,

$$\Delta a \Delta b \geqslant \frac{\left|\langle\psi|[A,B]|\psi\rangle\right|}{2} .$$

This inequality has purely mathematical origin and therefore relates to the first component in the structure introduced above. In the particular case where A and B are the coordinate X and the momentum P of a particle, the inequality takes the familiar form $\Delta x \Delta p \geqslant \hbar/2$. Its operational sense and the corresponding experiments (components 2 and 3) will be considered below in Section 6.3. The fourth component, which is connected with the uncertainty relation, includes speculations on the 'wave–particle' dualism, the complementarity principle, the role of the interaction between the particle and the measurement device and so on. A typical feature of such speculations is the absence of strict unambiguous definitions and testable statements. In this sense, they have much in common with art, which presents an alternative way of reflecting reality.

The operational approach, in our opinion, is only aimed at formulating the experimental sense of certain basic notions and statements. Being defined this way, the operational approach has no relation to philosophy. It consists only in defining a set of basic symbols via appropriate (better realistic) experimental procedures, which is necessary for the comparison between theory and experiment. An operational definition for terms and symbols implies certain instructions given to an experimenter. A theorist who gives a task to an experimenter should say in a language that they both understand: "Do this, and you will obtain the following result...". Such a description should include realistic procedures for preparation and measurement. A typical feature of reliable scientific conclusions is their reproducibility in different laboratories. This requires a possibility to exchange information on the conditions of experiment, which means the existence of the corresponding language.

This approach should be distinguished from the philosophic *operationalism*. Similarly to various versions of *positivism*, philosophic operationalism rejects all notions that have no direct relation to experiment. In quantum physics, most researchers share the so-called *minimal* viewpoint (see Ref. [2]), according to which it is only the efficiency of calculations that is essential. In fact, in this approach, one neglects the necessity of interpretation. Extreme viewpoints of this kind exaggerate the abilities of the axiomatic approach. At the same time, they underestimate the important role played by explicit models in young branches of physics and the convenience of various metaphysical terms for verbal communication and planning new experiments.

A 'naive realist' or a 'metaphysicist' is curious about 'what goes on there in reality?' A 'pragmatist' or an 'instrumentalist' considers this question to have no scientific sense because any answer to it cannot be falsified. In his opinion, this question is similar to the famous problem about the number of angels on a needle point. According to a 'pragmatist', the only aim of a physicist is to construct mathematical models (universal if possible) that reflect some features of the real world (mostly, its symmetry) and test them. In return, a metaphysicist accuses his opponent of *extended solipsism* (see Ref. [2]). The old philosophic problem about the relation between the *essence* and the *appearance* is emphatically revealed in quantum physics. If one defines scientific knowledge as a projection of some part of nature onto another part, onto our consciousness, then, clearly, this projection cannot be complete or precise and the question "What actually goes on there?" makes no sense.

In the framework of the *literary* interpretation of the WF [2], it is assumed that each quantum object can be characterized by its 'true' WF. In the case of a single particle, the WF replaces its classical kinematic parameters, coordinate and velocity. It is often supposed that the WF accompanies a particle as some (complex) field or 'cloud'. In the case of two individual particles, this 'cloud' exists in eight-dimensional space–time and varies there according to the Schrödinger equation. Correspondingly, each measurement giving an observable result a_1 is supposed to 'actually change' this individual WF, that is, to cause its immediate reduction $|\psi\rangle \to |a_1\rangle$, see Sections 4.7, 4.8, 5.5–5.9. (Here a_1 is the measured eigenvalue of the A operator.)

At present, the interpretation of the quantum formalism is chosen according to one's taste. However, in our opinion, one

should still avoid redundant notions like immediate reduction, nonlocality (Section 5.7), teleportation (Section 5.9) at least in order not to promote pseudosciences. On the other hand, operational definitions for the main terms form the basis of any physical theory. They are especially important for teaching quantum physics.

3. Classical probabilities

In this section, we consider classical analogues of some notions and procedures of quantum physics. Using simple classical models, we try to present a clear interpretation of the notion of a quantum *state* (*pure* and *mixed*) and of its *preparation* and *measurement*. We also prove the following two statements that also seem to be valid in the quantum case.

(1) Ascribing a set of probabilities (which will be called 'a state', in analogy with the quantum notation) to an individual system with random properties has clear operational sense in some ideal cases.

(2) There is no principal, qualitative difference between a single trial and an arbitrarily large finite number of uniform trials; in both cases, the experiment does not give reliable result.

3.1 Preparation of a classical state

Throwing an ordinary die, one can get one of six possible outcomes, or *elementary events*: the figure on the upper side may be $n = 1, 2, 3, 4, 5,$ or 6. (Here we mean a 'fair', i.e., sufficiently random throwing of dice with unpredictable results). Let the set of these six possibilities be called the *space of elementary events*. This space consists of discrete numbered points $n = 1, \ldots, N$ $(N = 6)$. To each one of these events, we ascribe, from some physical or other considerations, some probability p_n. Next, we assume Kolmogorov's axioms of non-negativity, $p_n \geqslant 0$, normalization, $\sum p_n = 1$, and additivity (see, for instance, Ref. [8]). The set of probabilities will be called the *state* of this individual die and denoted as $\Psi \equiv (p_1, p_2, p_3, p_4, p_5, p_6) \equiv \{p_n\}$. If the die is made of homogeneous material and has ideal symmetry, it is natural to assume all probabilities to be equal, $p_n = 1/6$.

However, in the general case this is not correct. One can *prepare* a die with shifted center of mass or some more complicated model like a roulette wheel that has, for instance,

$$\Psi \equiv (0.01,\ 0.01,\ 0.01,\ 0.01,\ 0.01,\ 0.95)\,. \qquad (3.1.1)$$

Clearly, each die or each roulette wheel can be characterized by a certain state Ψ, i.e., by six numbers that contain complete probability information about this die and about its asymmetry. The state (the set of probabilities) of this die is determined by its form, construction, the position of its center of mass, and other physical parameters. This state practically does not vary with time. (Hence, according to our definition, the state of the die does not contain information about the throwing procedure; the results of throwing are supposed to be almost completely random and unpredictable.)

The state is often characterized by the set of *moments* $\{\mu_k\}$, i.e., numbers generated by the state according to the rule

$$\mu_k \equiv \langle n^k \rangle \equiv \sum_n n^k p_n\,.$$

Combining the first and the second moments, we obtain the variance $\langle n^2 \rangle - \langle n \rangle^2 \equiv \Delta n^2$. Its square root, Δn, called the *standard* deviation or the *uncertainty*, characterizes deviations from the mean value, i.e., *fluctuations*. For instance, for a regular die, $\langle n \rangle = 3.5$ and $\Delta n = 1.7$, while for state (3.1.1), $\langle n \rangle = 5.85$ and $\Delta n = 0.73$. Having the full set of moments, one can, in principle, reconstruct the state, i.e., the probabilities. (In quantum models, this is not always true, see Sections 4.5, 5.6 – 5.8, 6.4.)

Any possible state of the die can be depicted as a point in the six-dimensional *space of states*. The frame of reference for this space should be given by the axes p_n or $c_n \equiv \sqrt{p_n}$. In the last case, the depicting point belongs, due to the normalization condition, to the multi-dimensional sphere S^5, and the state vector can be written as $\Psi = \{c_n\}$ (for comparison with the *Poincaré* sphere S^2, see Sections 5.3, 5.4).

Now let $N = 2$. One can imagine a coin made of magnetized iron. Due to the magnetic field of the Earth, the probabilities of the heads, p_+, or tails, $p_- = 1 - p_+$, depend on the value and direction of magnetization. Each individual coin can be characterized by a state $\Psi = (p_+, p_-)$.

3.2 Measurement of a classical state

For a state Ψ, which is prepared by means of a certain procedure and therefore known, one can *predict* the outcomes of individual trials. However, these predictions only relate to probabilities, with the exception for the case where one of the components of Ψ equals 1. One can pose the inverse problem of *measuring* the state Ψ.

Clearly, it is impossible to measure Ψ for a given coin in a single *trial*. (Speaking of a trial, we mean a 'fair' throw of the coin with the initial toss being sufficiently chaotic.) For instance, 'tails' can correspond to any initial state except $\Psi_1 = (0, 1)$, where the index of Ψ denotes the number of trials M. One should either throw one and the same coin many times or make a large number of identically prepared coins, a *uniform ensemble*. If the coins remain the same, are not damaged in the course of trials, then all these ways to measure the state are equivalent (the probability model is *ergodic*).

From the viewpoint of measurement, the only way to define the probability is to connect it with the rate of corresponding outcome. Throwing a coin 10 times and discovering 'heads' each time, one can state, with a certain extent of confidence, that $\Psi \approx \Psi_{10} = (1, 0)$. However, it is possible that the next 90 trials the coin will show 'tails'. This time, we will be more or less confident that $\Psi \approx \Psi_{100} = (0.1,\ 0.9)$, — and still we can be mistaken, since the actual state might be, say, $\Psi = (0.5,\ 0.5)$. This example of exclusive bad luck shows that an actual (prepared) state Ψ cannot be measured with full reliability. One can only hope that as M increases, the probability of a large mistake falls and Ψ_M approaches the actual value Ψ. In other words, relative rates of different outcomes almost always manifest *regularity* for increasing M.

Hence, for the case of known ideal *preparation* procedure, the state Ψ (the set of probabilities) can be associated with the chosen individual object. Here the state is understood as the information about the object allowing the prediction of the probabilities of different events. At the same time, for the case of known *measurement* results, the state can be only associated with an ensemble of similarly prepared objects, always with some finite reliability. There is no principal difference between a single trial and a number of trials: the results of experiments are always probabilistic. Similar conclusions can be made in the quantum case.

3.3. Analogue of a mixed state and the marginals

Consider two sets of coins prepared in the states $\Psi' = (p'_+, p'_-)$ and $\Psi'' = (p''_+, p''_-)$. The numbers of coins in the sets are denoted by N' and N'' $(N' + N'' = N)$. If the coins are randomly chosen from both sets and then thrown, the 'heads' and 'tails' will evidently occur with weighted probabilities

$$\rho_+ = \frac{p'_+ N' + p''_+ N''}{N}\,, \qquad \rho_- = \frac{p'_- N' + p''_- N''}{N}\,, \tag{3.3.1}$$

which are determined by both the properties of the coins and the relative numbers of coins in the sets, N'/N and N''/N. In this case, double stochasticity appears: due to the random choice of the coins and due to the random occurring of 'heads' and 'tails'. This is the simplest classical analogue of a *mixed state* in quantum theory (in its first definition, see Section 4.10). Clearly, such a mixed state cannot be associated with an individual system; it is a property of the *ensemble* containing two sorts of coins. In quantum theory, this corresponds to a classical ensemble of similar systems being in various states with some probabilities.

In quantum theory, there also exists another definition of a mixed state. This definition characterizes a part of the degrees of freedom for a quantum object, see Section 4.10; in the classical theory, it corresponds to *marginal* probability distributions, or *marginals*. Marginal distributions are obtained by summing elementary probabilities, in accordance with Kolmogorov's additivity theorem. Hence, they can be also considered as a property of an individual object. For instance, for a die, one can determine the marginal probabilities of odd and even numbers, p_+ and p_-. For the state (3.1.1), we obtain $p_+ = 0.97$ and $p_- = 0.03$.

3.4. Moments and probabilities

Now let two coins from different sets be thrown simultaneously. We introduce two random variables S_1, S_2 taking values $s_1, s_2 = \pm 1$ for 'heads' or 'tails', respectively. The system is described by a set of probabilities $p(s_1, s_2)$ of four different combinations $(\pm 1, \pm 1)$. If the coins do not interact and are thrown independently, then the 'two-dimensional' probabilities $p(s_1, s_2)$ are determined by the products of the corresponding one-dimensional probabilities, $p(s_1, s_2) = p_1(s_1)p_2(s_2)$.

However, let the peculiarities of the throw or the interaction between the magnetic moments of the coins lead to some correlation between the results of the trials. Then the state of the two coins is determined by the set of four elementary probabilities $p(s_1, s_2)$. The marginal probabilities and the moments are obtained by summing,

$$\begin{aligned} &p_k(s_k) = p(s_k, +1) + p(s_k, -1)\,, \\ &\langle S_k \rangle \equiv p_k(+1) - p_k(-1) = 2p_k(+1) - 1 \qquad (k = 1, 2)\,, \\ &\langle S_1 S_2 \rangle \equiv p(+1, +1) + p(-1, -1) - p(+1, -1) - p(-1, +1)\,. \end{aligned} \tag{3.4.1}$$

Hence, $\big|\langle S_k \rangle\big| \leqslant 1$, $\big|\langle S_1 S_2 \rangle\big| \leqslant 1$. In the simple case considered here, one can easily solve the inverse problem, which is called the *problem of moments*. In other words, one can easily express the probabilities in terms of moments,

$$p_k(s_k) = 2^{-1}\big(1 + s_k \langle S_k \rangle\big)\,, \tag{3.4.2}$$

$$p(s_1, s_2) = 2^{-2}\big(1 + s_1\langle S_1 \rangle + s_2 \langle S_2 \rangle + s_1 s_2 \langle S_1 S_2 \rangle\big)\,. \tag{3.4.3}$$

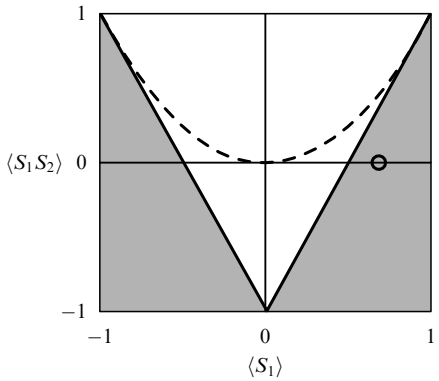

Figure 2. Connection between the correlator $\langle S_1 S_2 \rangle$ and the first moments $\langle S_1 \rangle$, $\langle S_2 \rangle$ (in the case $\langle S_1 \rangle = \langle S_2 \rangle$) for two random variables S_1 and S_2 taking the values $s_k = \pm 1$. In the shaded 'prohibited' area, the probabilities corresponding to the moments take negative values. The dotted line shows the case of independent variables where $\langle S_1 S_2 \rangle = \langle S_1 \rangle^2$. The circle with coordinates $(0.71, 0)$ corresponds to the quantum moments for the Stokes parameters in the case of a photon polarized linearly at an angle 22.5° to the x axis (see Section 5.6).

From Eqn (3.4.3) and the condition $p(s_1, s_2) \geqslant 0$, it follows that the moments are not independent; they must satisfy certain inequalities. Provided that the first moments $\langle S_k \rangle$ are given, the *correlator* $\langle S_1 S_2 \rangle$ cannot be arbitrarily large or small,

$$f_{\min} \leqslant \langle S_1 S_2 \rangle \leqslant f_{\max}\,. \tag{3.4.4}$$

Here

$$\begin{aligned} f_{\min} &\equiv \max\big(-1 - \langle S_1 \rangle - \langle S_2 \rangle,\, -1 + \langle S_1 \rangle + \langle S_2 \rangle\big)\,, \\ f_{\max} &\equiv \min\big(1 + \langle S_1 \rangle - \langle S_2 \rangle,\, 1 - \langle S_1 \rangle + \langle S_2 \rangle\big)\,. \end{aligned}$$

For instance, for $\langle S_1 \rangle = \langle S_2 \rangle$, we have the limitation $2\big|\langle S_1 \rangle\big| - 1 \leqslant \langle S_1 S_2 \rangle \leqslant 1$ (Fig. 2). In particular, the correlator cannot equal zero for $\langle S_1 \rangle > 1/2$ (i.e., for $p_+ > 3/4$).

In the quantum theory, analogous inequalities for quantum moments $\langle F \rangle$, which are obtained by averaging with respect to the WF, $\langle F \rangle \equiv \langle \psi | F | \psi \rangle$, are sometimes violated. Paradoxes of this kind will be discussed in Sections 4.5, 4.6, 5.5–5.8. Note that in such cases, the notion of elementary probabilities has no sense, and the quantum probability model can be called *non-Kolmogorovian*.

4. Quantum probabilities

The classical models described above have little connection with quantum physics. The 'state' of a die can include not only the properties of this die, as we supposed above, but also the parameters of the initial toss. (According to classical dynamics, these parameters unambiguously determine the outcome.) Stochasticity appears here as a result of variations in the value and direction of the initial force. (Under certain additional conditions, such models manifest *dynamical chaos*.) Quantum stochasticity is believed to have a fundamental nature; it is not caused by some unknown *hidden variables*, though Einstein could never admit that "God plays dice".

It is an astonishing feature of quantum probability models that in some cases, there exist marginals but there are no elementary probabilities. This feature can be called the *non-*

Kolmogorovianness of the quantum theory; in the general case, it corresponds to the absence of *a priori values* of the observables, see Sections 4.5, 5.5 – 5.8, 6.4. For instance, one can measure (or calculate using ψ) coordinate and momentum distributions for a particle at some time moment, but their joint distribution cannot be measured. Reconstruction of the joint distribution from the marginals is ambiguous and sometimes leads to negative probabilities. Therefore, it is natural to assume that a particle has no *a priori* coordinates and momenta.

It is also important that classical models have no concept of complex *probability amplitudes* and hence, do not describe *quantum interference* and complex vector spaces of states. There is no classical analogue of *non-commuting* variables, which do not admit joint probability distributions.

Quantum physics presents extremely special procedures for preparation and observation. From the operational viewpoint, a pure state ψ_0 is a detailed coded description of an ideal preparation procedure (history) for a given individual quantum object. However, one can use or check the information contained in the WF only under the condition that there exist several objects prepared similarly. It is only in some special cases that knowing the state of a single particle, one can make (almost) unambiguous predictions concerning the result of a single trial (see the example in Section 6.6). Almost all real experiments result in the preparation of mixed states where additional classical uncertainty is present in the parameters of the pure states. For instance, the coherent state of the field prepared with the help of an ideal laser has a random phase.

An interesting question is: ''in what cases is the quantum theory really necessary?'' It is often supposed that the quantum theory is necessary for describing *microscopic* objects, in contrast to *macroscopic* ones. However, in some cases, macroscopic objects also require a quantum theoretical description. For instance, recent experiments on Bose condensation involve hundreds of thousands of atoms (lithium, sodium, or rubidium) [9 – 12]. The atoms are stored in a magnetooptical trap and cooled, using laser radiation and other methods, to $10^{-6}-10^{-7}$ K. At the same time, the motion of the centers of mass for all atoms is described by a joint WF. This WF describes the collective localization of atoms in a small spatial domain at the center of the trap. Note that here, one can ignore the 'frozen' degrees of freedom relating to atom electrons and the internal structure of the nuclei, nucleons, and quarks. This illustrates the idea of a phenomenological approach in quantum physics and, more generally, the idea of *reductionism*, a hierarchic description of reality.

At present, considerable interest is also attracted to experiments on the interference of composite particles such as atoms and molecules. The interference is determined by the de Broglie wavelength of such particles, $\lambda = h/Mv$ (see Refs [13, 14]). For instance, the interference pattern observed for sodium molecules Na_2 has an oscillation period half that of sodium atoms [15]. Here again the effect is described by a WF relating to the center of mass of the molecule, although the actual sizes of the particles can be much larger than λ. Recently, interference of this type was observed between two groups of Bose-condensed atoms, each group containing 10^6 atoms [12]. This experiment proves that both groups can be described in terms of a two-component WF containing some phase difference. [In this connection, the concept of an *atom laser* has been suggested (see Ref. [12]).]

There are well-known examples of macroscopic quantum phenomena, such as the effects of superfluidity, superconductivity, and the Josephson effect. The wave packet of an electron can occupy macroscopic volume, and an electron manifests itself as a 'point-like' particle only when it is registered, see Section 6.1. In modern optical experiments, the coherence lengths of the fields sometimes exceed several kilometers. In such cases, it is quite sufficient to use a phenomenological description with a small number of parameters and the single-mode approximation for the field, with the atom variables excluded by introducing the linear and nonlinear susceptibilities of matter $\chi^{(n)}$, $n = 1, 2, \ldots$, and so on. For instance, with the help of the quadratic nonlinearity $\chi^{(2)}$, it is convenient to describe the preparation of 'two-photon' or 'squeezed' light by means of coherent nonelastic scattering of ordinary light in transparent piezoelectric crystals (the effect of parametric scattering, or spontaneous parametric down-conversion).

Apparently, all sufficiently cooled and isolated objects can be and should be described by phenomenological quantum equations ignoring the 'frozen' degrees of freedom.

4.1 Classical stages in quantum models

Several crucial problems can be pointed out in the quantum measurement theory. First, this is the fundamental problem of unifying quantum and classical physics, the development of a universal approach to the description of a quantum object and the preparation and measurement devices. This global task is still unsolved. Probably, it cannot be treated in the framework of the standard quantum formalism and requires the creation of some metatheory. Recently, a number of interesting dynamical models have been developed describing reduction and measurement of the WF (for recent results and references, see Refs [13, 16, 17]). However, these models are so far not connected with real experiments, and we will not touch upon this problem. Another important group of problems includes the development, in the framework of the standard quantum theory, of the optimal methods of precise measurements for various applications and methods of suppressing quantum noise [18, 19].

Formally, quantum theoretical description operates only with the WF Ψ of an isolated system that should include both the subsystem under study and the preparation and measurement devices interacting with it. In some considerations, the isolated system also includes the experimenters, their brains, or even the whole Universe. In this sense, a purely quantum model is a thing in itself; it leaves no space for an external observer. Predictions of such models cannot be tested, and therefore, as Bohr has mentioned, one has to use hybrid models including both quantum and classical components.

In order to compare theoretical results with experiment, one should somehow, taking into account additional considerations, restrict the number of degrees of freedom. A correspondence should be postulated between the symbols of the quantum formalism describing the system and the parameters of real classical devices used for preparation and measurement. The terms 'observable' and 'operator' are usually identified; however, for any quantum model, comparison with experiment requires setting certain boundaries between the quantum system and the classical environment. In the chain of interacting subsystems described by the operators $B_1, B_2, \ldots$, some operator B_m (or set of operators $B_m, B'_m, \ldots$) should be chosen as 'the most observable' (the readout observable). It is assumed that the 'measurement'

subsystem interacting with B_m manifests classical properties; it has many degrees of freedom and a continuous spectrum (*an open system*). For a particular model of measurement, calculation with the help of the Schrödinger equation allows the readout observable to be changed, $B_m \to B_{m-1}$. In the case of a formal consideration, the choice of the readout observable is not unique, and the boundary between the two worlds can be set arbitrarily (see Ref. [2]). But formal models of this kind, as we have already mentioned, do not admit quantitative comparison with real experiments, and therefore, for comparison with the experiment one should still choose some readout observable B_m.

At the next stage, the Born postulate should be included in the consideration. This postulate sets a relation between the probabilities of observable events $p(b_m)$ and the WF and therefore, 'legalizes' stochasticity in quantum models (see Section 4.7). This 'measurement' postulate is so far the only 'bridge' connecting the mathematical formalism and the experimental results.

In most modern experiments, the observed 'elementary' quantum events are photocurrent pulses at the output of the detector, a droplet appearing in the Wilson chamber, etc. The 'invisible' world of individual quantum objects seems to reveal itself only by means of such 'clicks'. Observing such an event, one can assign some *a priori* coordinates to the particle that caused the 'click'. The particle is 'localized' in a certain space–time domain, which is determined by the classical dimensions of the detecting device. These dimensions are measured by usual methods, with the help of rulers and clocks.

In the well-known model of photodetection suggested by Glauber [20], the observable event is defined as the transition of one of the atoms of the detector from the ground state $|g\rangle$ into the excited state $|e\rangle$. This event corresponds to the projection operator $|e\rangle\langle e| \equiv B_m$, which plays the role of the readout observable. Due to the amplification in the detector, the event is supposed to manifest itself as a macroscopic current pulse at the output of the detector. In order to describe fast detection, one assumes that the spectrum of the atoms constituting the detector is sufficiently broad. (Probably, it is necessary to use the assumption about the relaxation of the density matrix off-diagonal elements.) Calculating the evolution of the system 'field + atoms' via the Schrödinger equation, one can show that the statistics of the photocurrent pulses $i(t)$ are determined by the correlation functions of the free field $E(\mathbf{r}, t)$. Further, one can assume the field $E(\mathbf{r}_1, t_1)$ at the center of the detector to be the readout observable instead of $|e\rangle\langle e|$. Here $(\mathbf{r}_1, t_1)$ are the classical coordinates in space–time, and they are measured using rulers and clocks. The coordinate $\mathbf{r}_1$ of the center of mass of the detected atom and the time moment of the pulse t_1 are supposed to be c-numbers.

Similarly, in the model of a particle counter (see Section 6.1), the role of the readout observable is played by the potential of the interaction between the detector and the particle, $V(\mathbf{R} - \mathbf{r}_1)$, which depends on the coordinate operator for the particle $\mathbf{R}$. However, note that for justifying some choice of B_m, one should use some particular model of the detector. Certainly, the adequacy of the model should be tested experimentally.

Actually, when describing dynamical experiments (see Fig. 1), one should use the 'semiclassical' approach consisting of two stages. In other words, two boundaries should be set: at the 'input', where one determines the initial state of the quantum system ψ_0 in terms of the classical forces, and at the 'output', where one chooses the operator B_m, which influences the classical measurement device. Between the input and the output, the system develops by itself, and its WF evolves according to the Schrödinger equation: $\psi_0 \to \psi_t$. [Here, classical fields should be taken into account (see Fig. 1).] By choosing ψ_0 and B_m, we exclude the operators of the preparation and measurement devices, respectively. If the preparation of the WF is described in the framework of the classical theory, one can consider the ground (bottom) state ψ_{00}, which is achieved due to relaxation or cooling, and to describe its transformation $\psi_{00} \to \psi_0$ by including the classical field in the Hamiltonian (see the example in Section 5.2). A bright example are the experiments on the Bose condensation of atoms in traps [9–12] where a localized WF is prepared by means of cooling and applying classical fields.

The effect of various *filters*, such as diaphragms, magnetic filters, monochromators, etc., is also described classically. As a rule, it can be included into the preparation or measurement stages. (Still, it is reasonable to distinguish between these procedures.) In quantum optics, spectral filters, beam splitters, polarizers, lenses, etc. are described in terms of classical phenomenological Green's functions, which transform the state of the field (in the Schrödinger approach) or field operators (in the Heisenberg approach) [21]. One can also point out various *modulators*, which change the WF of a prepared system via time-dependent classical fields (see Section 6.6). For instance, in a detector of gravity waves, the WF of a quantum object (a macroscopic oscillator) is modulated by an alternating gravitational field [19]. In the description of parametric scattering, the laser ('pump') field modulating the dielectric function of the crystal can be considered as classical. At the same time, the effect of the pump transforms the scattered field from the vacuum state into a superposition of Fock states with even photon numbers: $|\psi_{00}\rangle = |0\rangle \to |\psi_0\rangle = c_0|0\rangle + c_2|2\rangle + c_4|4\rangle + \ldots$ Note that in the general case, filtration and modulation are described by a nonunitary transformation converting the system into a mixed state [21].

Let us consider once more how the measurement procedure is described in the framework of the standard quantum formalism (for more detail, see Refs [18, 19]). There exist models of direct and indirect quantum measurement [5, 18, 19]. In the first case, the consideration includes a single quantum object A, which is characterized by the WF $\psi(a)$. (For simplicity, we assert that the state is pure and that its WF has a single argument.) In order to describe the interaction with the external world, some operator A is assumed to be the observable. The experimental data are compared with the distribution $p(a) = |\psi(a)|^2$ or with its moments $\langle a^k \rangle$.

In the models of indirect measurement, in addition to the object under study, one introduces at least one more 'sample body' B interacting with A and acting as an interface between A and the macroscopic world. One considers the WF $\psi(a, b)$ of the system A + B, and the correlation between a and b resulting from the interaction of A with B is calculated using the Schrödinger equation. This time, the role of the 'readout observable' is played by the operator B relating to B. As a result of this theory, one gets a joint probability distribution $p(a, b) = |\psi(a, b)|^2$. The correlation between a and b is described by the function $p(a, b)$ and allows one to learn about b from the analysis of a. After classical summation over the probabilities of non-observable events, one obtains the marginal distribution $p(b) = \sum_a p(a, b)$, which can be mea-

sured experimentally and contains information about $p(a)$. A description of quantum correlations can be found in Section 4.8.

The operators A and B can relate to different degrees of freedom for one and the same object. For example, in the Stern–Gerlach experiment (see Fig. 1), $A \equiv M_x$ and $B \equiv X$ are operators of angular momentum projection and transverse coordinate for a single particle; these operators become correlated if the particle moves in inhomogeneous magnetic field [22]. As a result, from the transverse coordinate of the particle, x_1, which is directly observable, for instance, as a spot on the film, one can obtain the spin projection $m_x = m_x(x_1)$ onto the transverse direction for the chosen particle. This projection is obtained indirectly, by means of a theoretical model that describes the influence of the inhomogeneous magnetic field on the WF evolution for a spin particle. The density of spots on the film (Fig. 1) gives a two-peak distribution $p(x)$ containing information about $p(m_x)$.

One can consider a chain of interacting objects $\mathrm{A}, \mathrm{B}_1, \mathrm{B}_2, \ldots, \mathrm{B}_m$, which are described by the operators $A, B_1, B_2, \ldots, B_m$. The quantum formalism allows calculation of the total WF $\psi(a, b_1, b_2, \ldots, b_m)$ and the joint probability distribution:

$$p(a, b_1, b_2, \ldots, b_m) = \left|\psi(a, b_1, b_2, \ldots, b_m)\right|^2 .$$

The last operator in the chain, B_m, is declared to be the observable. After that, one applies the classical probability theory. The marginal distribution $p(a, b_m)$ is obtained by summing the elementary distribution $p(a, b_1, b_2, \ldots, b_m)$ with respect to the 'redundant' variables $b_1, b_2, \ldots, b_{m-1}$. In the Heisenberg representation, the 'output' operators $B(t)$ are expressed via the 'input' ones, $B(t_0)$. Using the relation between them, one can easily calculate the transformation of correlation functions due to the interaction.

Note that the interaction between quantum subsystems, which is described by the Schrödinger equation in the framework of the standard quantum formalism and causes correlations between the subsystems, should be distinguished from the 'real' measurement process. In the description of real measurement, it is necessary to consider the interaction between classical and quantum systems, which is not included in the standard formalism.

4.2 A complete set of operators and the measurement of the wave function

Consider free one-dimensional motion of a nonrelativistic spinless particle. Its observable statistical properties at a fixed time moment are fully described by the state vector $|\psi\rangle$ in some representation. For instance, in the coordinate representation, $\langle x|\psi\rangle \equiv \psi(x)$. In other words, a single coordinate operator X forms a *complete set* of operators that is necessary for specifying the state. The same relates to the momentum $\hbar K$ (sometimes we put $\hbar \equiv 1$), and a state can be given by the state vector in the momentum representation $\langle k|\psi\rangle \equiv \psi(k)$, i.e., by the Fourier transform of $\psi(x)$. At the same time, the energy operator $H = K^2/2m$ does not form a complete set, since it leaves uncertainty in the sign of the momentum. In other words, the energy levels are doubly degenerate, and a complete set can be formed by H and by the operator of the momentum sign.

In order to specify a state, it is sufficient to give eigenvalues for all the operators forming a complete set. For instance, the information $k = k_1$ fully determines the WF: $\psi(x) = \exp(\mathrm{i}k_1 x)$. For the case of operators with discrete spectra, a state is fixed by specifying the *quantum numbers* that enumerate the eigenstates and the eigenvalues. It is known that the states of an electron in a hydrogen atom are conveniently described using spherical coordinates, $\psi = \psi(r, \theta, \phi)$, and the quantum numbers n, l, m, s, which determine the eigenvalues of the energy, angular momentum, its projection, and the spin projection.

Consider now the *measurement* of a state. Repeated measurement of the coordinate by means of an ideal detector gives the WF's absolute value (*the envelope*) $\left|\psi(x)\right|$ (see Section 6.1). At the same time, the phase of the WF $\phi(x) \equiv \arg\left[\psi(x)\right]$ cannot be observed directly; therefore, such an experiment does not provide a complete measurement of the WF, in spite of the fact that X forms a complete set. For complete determination of the WF, additional measurements are required, such as, for instance, measurement of the WF envelope in the momentum representation, $\left|\psi(k)\right|$ (Section 6.1). In real experiments, a state is measured in a set of experiments where different combinations of X and K are measured [23–27, 99].

It is often mentioned that the phase of the WF has no physical sense, is not observable. Here one means the constant *global* phase ϕ_0, which does not depend on the coordinate. At the same time, the *local* phase $\phi(x)$ has a considerable effect on the observed function $\left|\psi(k)\right|$. Observable effects caused by the time dependence of the WF phase $\phi(t)$ are discussed in Section 6.6.

Thus, one should distinguish between specifying the WF in theory, where it is introduced as an eigenfunction for some complete set of operators such as, for instance, X or K, and measuring it in experiment where one has to study, for instance, both coordinate and momentum distributions, i.e., to deal with more than one complete set. Thus, a complete set of operators is incomplete from the viewpoint of measurement.

Another example: for fixing the polarization of a photon, it is sufficient to state, for instance, that it has right circular polarization. In this case, the field has fixed angular momentum m. But in order to check this statement, it is not enough to measure m. Such an experiment should consist of several series of measurements for non-commuting observables (the Stokes parameters) (see Section 5.4).

At present, various methods of preparation and reconstruction of the states of optical fields, atoms, and molecules are attracting considerable attention [23–34].

4.3 Quantum moments

In the classical probability theory, the moments of a random variable A are defined via the probability distribution function $p(a)$: $\mu_n \equiv \langle A^n \rangle \equiv \int \mathrm{d}a\, p(a) a^n$ (the integrals are supposed to converge for all $n = 0, 1, \ldots$). For a discrete random variable, the integral is replaced by the sum (see Sections 3.1, 3.4). In the case of several random variables $A, B, \ldots$, the moments are given by multi-dimensional integrals

$$\mu_{nm\ldots} \equiv \langle A^n B^m \ldots \rangle \equiv \int \ldots \int (\mathrm{d}a\, \mathrm{d}b \ldots)\, p(a, b, \ldots)(a^n b^m \ldots) . \tag{4.3.1}$$

In quantum theory, the moments are defined not via the distribution function $p(a, b, \ldots)$ but via the WF,

$$\mu_{nm\ldots} \equiv \langle A^n B^m \ldots \rangle \equiv \langle \psi | A^n B^m \ldots | \psi \rangle . \tag{4.3.2}$$

It is essential that moments composed from non-commuting operators depend on the ordering of the operators. Consider two non-commuting Hermitian operators, $[A, B] \neq 0$. A question arises: "which moments composed from A and B manifest themselves in an experiment?" Even if we add the requirement that the moments should be real, there still remain many possibilities: $\big(\langle AB\rangle + \langle BA\rangle\big)/2$, $\big(\langle AB\rangle - \langle BA\rangle\big)/2\mathrm{i}$, $\langle ABA\rangle$, $\langle BAB\rangle$, and so on. The answer depends on the particular experimental device and on the parameters measured in the experiment. This problem is especially interesting in the case where non-commuting observables are measured at various time moments (see below). So far, we assume for simplicity that all operators relate to the same moment.

As an example, consider quantum-optical experiments where one measures the energy of the field. Sometimes, it is possible to take into account only a single mode of the field. In this case, the field has the same description as a harmonic oscillator, and the energy operator has the form $H(X, P) = (P^2 + \omega^2 X^2)/2$, where ω is the mode frequency. It is convenient to pass from the operators X, P to the operators a, $a^\dagger$, which are called photon creation and annihilation operators. (Here we use the traditional notation in small letters.) By definition,

$$a \equiv (2\hbar\omega)^{-1/2}(\omega X + \mathrm{i}P)\,, \qquad a^\dagger \equiv (2\hbar\omega)^{-1/2}(\omega X - \mathrm{i}P)\,.$$

From $[X, P] = \mathrm{i}\hbar$, we find $[a, a^\dagger] = 1$ and obtain several equivalent forms for the Hamiltonian:

$$H(a, a^\dagger) = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) = \hbar\omega\left(aa^\dagger - \frac{1}{2}\right)$$
$$= \hbar\omega\left[\alpha\left(a^\dagger a + \frac{1}{2}\right) + \beta\left(aa^\dagger - \frac{1}{2}\right)\right], \qquad (4.3.3)$$

where α is an arbitrary number and $\beta = 1 - \alpha$.

From the models of photodetection, it follows that in the first approximation, the probability of energy transfer from the field to a detecting atom in the ground state is determined not by the whole energy operator but only by its *normally ordered* part, $H - \hbar\omega/2 = \hbar\omega a^\dagger a$. (This probability also depends on the antinormally ordered operator of the detector $DD^\dagger$, where D is the *positive-frequency* part of the atom dipole moment [35].) In other words, the probability of stimulated one-photon 'up' transition for the atom is determined, in the linear approximation, by the photon number operator $N \equiv a^\dagger a$. Choosing N as the observable operator ensures that the term $\hbar\omega/2$ gives no contribution to the excitation probability for the atom. Similarly, the probability of a k-photon 'up' transition for an atom is determined by the operator

$$(a^\dagger)^k a^k = N(N-1)\ldots(N-k+1) \equiv :N^k:\,. \qquad (4.3.4)$$

Here colons denote *normal ordering*, $:(a^\dagger a)^k: \equiv (a^\dagger)^k a^k$.

At the same time, for a correct description of the observable fluctuations of energy near its average value, one should use the non-ordered operator $H^2 = (\hbar\omega N)^2$, which contains the term $X^2P^2 + P^2X^2$ and is proportional to the operator

$$N^2 = a^\dagger aa^\dagger a = a^\dagger a^\dagger aa + a^\dagger a = :N^2: + N\,. \qquad (4.3.5)$$

Hence, the observable variance of the energy is determined by the non-ordered moment

$$\langle \Delta N^2\rangle = \langle N\rangle + \langle :\Delta N^2:\rangle\,. \qquad (4.3.6)$$

Here the term $\langle N\rangle$, which is typical for the variance of a Poissonian random process, describes quantum fluctuations for the energy measurement. They manifest themselves in experiment in the form of *shot* (or *photon*) *noise* [36]. Normally ordered variance $\langle :\Delta N^2:\rangle$, also called the *excess noise*, describes the deviation of the variance from the Poissonian level. For the cases of Fock, coherent, and chaotic states, the variance $\langle \Delta N^2\rangle$ is equal to 0, $\langle N\rangle$, and $\langle N\rangle^2 + \langle N\rangle$, respectively. At $\langle \Delta N^2\rangle < \langle N\rangle$, the statistics are called *sub-Poissonian*, and at $\langle \Delta N^2\rangle > \langle N\rangle$, *super-Poissonian*. (One also uses the terms *antibunching* and *bunching*, respectively.) Note that for sub-Poissonian states of the field, the excess noise $\langle :\Delta N^2:\rangle$ is negative. Distinguishing between the quantum noise and the excess fluctuations has an operational sense: the quantum noise has a 'white' spectrum, while the spectrum of excess noise is determined by the dynamics of the radiation source [36].

Normally ordered moments are also convenient for the description of optical elements with linear absorption. Let η be the transmission coefficient of such an element, then the moments at its input and output are connected by the simple relation:

$$\langle :N^k:\rangle_{\mathrm{out}} = \eta^k \langle :N^k:\rangle_{\mathrm{in}}\,. \qquad (4.3.7)$$

For example, putting $k = 1$ and 2 here, we find

$$\langle \Delta N^2\rangle_{\mathrm{out}} = (1-\eta)\langle N\rangle_{\mathrm{out}} + \eta^2 \langle \Delta N^2\rangle_{\mathrm{in}}\,. \qquad (4.3.8)$$

This formula describes the 'poissonization' of intensity fluctuations as a result of absorption: at $\eta \to 0$, there is only Poissonian noise at the output, regardless of the fluctuations at the input. Assuming η in Eqn (4.3.7) to be the quantum efficiency of a photon counter, we obtain the relation between the statistics of photons and photocounts.

From these examples, it is obvious that the choice of ordering of the operators in quantum moments depends on the particular measurement procedure, which is to be described by these moments. This fact becomes essential for the description of time-of-flight experiments with high time resolution (see Sections 4.9 and 6.2).

4.4 Schrödinger and Heisenberg representations

Let us consider moments as functions of time. The dynamics of a quantum system can be described by means of two mathematically equivalent methods called the Schrödinger and the Heisenberg representations. The solution to the nonstationary Schrödinger equation $\mathrm{i}\hbar\partial\psi/\partial t = H\psi$, with the energy operator H independent of time, can be represented as $\psi(t) = U(t)\psi(0)$. Here we introduced the evolution operator $U(t) \equiv \exp(-\mathrm{i}Ht/\hbar)$. According to the Born postulate, the mean value $\langle A(t)\rangle$ of some observable A at the moment t has the form $\langle A(t)\rangle \equiv \langle \psi(t)|A|\psi(t)\rangle = \langle \psi(0)U(t)^+|A|U(t)\psi(0)\rangle$.

Let us introduce the operator A in the Heisenberg representation, $A(t) \equiv U(t)^+ AU(t)$, then the mean value can also be written as $\langle A(t)\rangle = \langle \psi(0)|A(t)|\psi(0)\rangle$. In the case of two commuting operators measured simultaneously, we also have two equivalent calculation algorithms:

$$\langle A(t)B(t)\rangle = \langle \psi(t)|AB|\psi(t)\rangle = \langle \psi(0)|A(t)B(t)|\psi(0)\rangle\,. \qquad (4.4.1)$$

However, *multitemporal* moments (correlation functions) are defined only in the Heisenberg representation. For example, a correlation function of two observables has the form

$$\langle A(t)B(t')\rangle = \langle\psi(0)|A(t)B(t')|\psi(0)\rangle\,. \qquad (4.4.2)$$

In order to calculate this function in terms of the Schrödinger parameters for $t \neq t'$, one needs additionally the evolution operator,

$$\langle A(t)B(t')\rangle = \langle\psi(t)|AU(t-t')B|\psi(t')\rangle\,. \qquad (4.4.3)$$

In some simple cases, operators in the Heisenberg representation depend on time in the same way as classical variables. For instance, for a free nonrelativistic particle, $H = P^2/2m$ and $[X, P] = i\hbar$; it follows that $X(t) = X + (P/m)t$, $P(t) = P$. In addition, the Heisenberg representation admits an explicitly covariant formulation of the theory [2].

In quantum optics, calculations are usually more simple in the Heisenberg representation, where the field operators $E(t)$, $H(t)$ in linear problems depend on time in the same way as classical fields. This allows one to exploit useful classical analogues and to use classical Green's functions for the description of optical elements, such as diaphragms, lenses, mirrors, etc. As a result, the quantum description of the field evolution in a linear optical tract, including the relation between the observable correlation functions at the input and at the output, coincides with the classical description. The only difference is contained in the procedure of averaging with respect to the initial state, which can be quantum or classical [37].

For our consideration, it is essential that the evolution of a system can be equivalently described both in terms of varying operators $A(t)$ and in terms of varying WF $\psi(t)$. Therefore, the evident representation of a quantum object in terms of some propagating 'field' $\psi(t)$ accompanying it or as a vector in the configuration space is not the only one possible. Here again we have a senseless question: "what actually does take place there, is it the WF or the operators that vary?" Note that possible observable manifestations of the projection postulate and the WF reduction should be described in the Heisenberg representation [see Wigner's formula (4.9.1)].

4.5 Quantum problem of moments

In Section 3.4, we obtained a formula that expressed the probabilities via the moments and imposed certain restrictions on the moments (see Fig. 2). In the case of continuous variables, this inverse problem in mathematics is called the problem of moments. A well-known example of a restriction imposed on the moments due to the non-negativity of probability is the Cauchy – Schwarz inequality $|\langle fg\rangle|^2 \leqslant \langle f^*f\rangle\langle g^*g\rangle$.

For a set of quantum moments μ, it is natural to pose the problem of constructing the corresponding probability distribution p. But in the case of non-commuting operators, this procedure, first, is ambiguous and second, gives functions taking negative or complex values. Such functions are called quasi-probabilities or quasi-distributions. Well-known examples are the Wigner function $W(x, p)$ (Section 6.4) and the Glauber – Sudarshan function $P(\alpha)$ (α is the complex amplitude of oscillations in a single mode, see Section 4.6). Thus, the quantum problem of moments in some cases has no solutions. One can say that quantum probability models are in the general case *non-Kolmogorovian* [38]. The absence of a non-negative joint distribution for non-commuting observables can be naturally interpreted as the impossibility of these observables having *a priori values*. In other words, it is not reasonable to suppose that each particle 'actually' has some fixed coordinate and momentum before the measurement but our rough devices spoil everything and do not allow their simultaneous observation.

In some models, the incompatibility of classical and quantum viewpoints can be demonstrated experimentally. Bell's inequalities [39, 40] and the Kochen – Specker theorem [41] relate to such models. As a rule, such models include several observables with discrete spectra (for example, spin projections or photon numbers in different modes). Non-commuting variables are measured in different trials. Such experiments with polarization-correlated photon pairs and triples will be considered in Sections 5.7, 5.8.

In several experiments, mostly optical, predictions of quantum models for the moments have been confirmed and violation of the Bell classical inequalities has been demonstrated. However, there still remain 'loopholes' in the interpretation of experimental results. These 'loopholes' initiate further theoretical and experimental research in this direction [42].

The statement about the incompatibility of certain classical and quantum probability models is sometimes called *Bell's theorem* or *Bell's paradox*. It is commonly supposed that Bell's paradox demonstrates 'quantum nonlocality', since one usually speaks about the correlation between events separated by spacelike intervals (photocounts in two remote detectors). However, the term *quantum nonlocality*, which implies some mysterious, telepathy-like connection between remote devices, cannot be considered helpful for the solution of Bell's paradox.

It seems more consistent to assume that the quantum mechanics is *non-Kolmogorovian*: it admits the absence of joint distributions and *a priori* values for non-commuting observables [38]. For example, the quantum theory allows calculation of moments of the form $\langle xp\rangle$; however, in the general case, there exists no corresponding joint non-negative distribution $w(x, p)$. Therefore, there is no sense in introducing *a priori* values for non-commuting observables. The absence of elementary probabilities in combination with the existence of marginal probabilities and moments (i.e., the absence of the solution to the problem of moments) can be considered as a characteristic feature of a non-Kolmogorovian probability model. Such a classification gives a general approach to various 'nonclassical' effects and quantum paradoxes [38].

4.6 Nonclassical light

Bell's inequalities and other similar constructions are in fact restrictions (similar to the Cauchy – Schwarz inequality) imposed on the moments by the non-negativity of the joint distribution. In other words, they follow from rather general mathematical considerations. In quantum optics, there exists another model, which is less general but also demonstrates that classical probability concepts cannot be applied to electromagnetic waves. This model is based on the well-known Mandel formula, which gives a relation between measured probabilities of photocounts and the Glauber – Sudarshan quasi-distribution $P(\alpha)$. The function $P(\alpha)$ plays the role of a classical distribution function for the complex

amplitude of a monochromatic field $\alpha = x + ip$; there is a one-to-one correspondence between this function and the WF of the field. However, for all pure states except the coherent one, $P(\alpha)$ takes negative values or is irregular [43]. For instance, for the Fock n-photon states, $P(\alpha)$ is given by a combination of nth-order derivatives of the δ-function. Such states of the field are called *nonclassical*.

For nonclassical fields, observable values like moments and probabilities of photocounts do not satisfy certain restrictions that follow from the non-negativity of $P(\alpha)$ [44, 45]. Similar nonclassical optical effects have been observed in numerous experiments. This confirms the adequacy of simple phenomenological models in quantum optics and shows that the concept of a probability distribution cannot be applied to a wave amplitude. As the most well-known and important example, one can mention the effect of photon antibunching, which consists in the decrease of photocurrent fluctuations below the shot-noise (photon) level [46, 47]. This level is called the *standard quantum limit* [18, 19]. Another 'nonclassical' optical effect, *two-photon interference*, can be classified as *intensity interference* with the visibility exceeding 50% (see Section 5.5 and Ref. [37]). Such a high visibility also contradicts the description of a light field in terms of a non-negative regular distribution.

The concept of nonclassical light is closely connected with the attempts to describe photodetection within the framework of the *semiclassical theory of radiation*, in which the field is described classically and the substance, which interacts with the field, is considered as quantum. Let monochromatic light with fixed intensity I (an ideal laser in the classical approximation) be incident on a detector . It is natural to assume that the excitation probability $\mathrm{d}p_1$ for any atom of the detector photocathode during a small time interval $\mathrm{d}t$ is independent of time and proportional to I: $\mathrm{d}p_1/\mathrm{d}t = kI$. (The factor k characterizes the quantum efficiency of the detector.) This model adds stochasticity to the dynamical theory: any number of pulses m $(m = 0, 1, 2, \ldots)$ can appear during some finite time T, and the probability of this event is given by the Poisson distribution, $p_m(I) = \mu^m \exp(-\mu)/m!$, $\mu \equiv kTI$.

Let us take into account that the intensity of light can be stochastic. Let T be much less than the characteristic time of intensity variation. Additional averaging of $p_m(I)$ with respect to the intensity distribution $p(I)$ results in the Mandel formula:

$$p_m = \frac{1}{m!}\int_0^\infty \mathrm{d}I\, p(I)(kTI)^m \exp(-kTI)\,. \tag{4.6.1}$$

In the quantum theory, one can obtain a similar expression, with the only difference that the function $p(I)$ is expressed in terms of the Glauber–Sudarshan function, $p(I) \propto P(|\alpha|)$, where $|\alpha|^2 \sim I$, and can therefore take negative values.

It follows from Eqn (4.6.1) that $m!p_m$ can be considered as moments of some distribution $p(I)\exp(-kTI)$. The condition $p(I) \geqslant 0$ leads to certain restrictions on the set of probabilities $\{p_m\}$ [44]. For example,

$$mp_m^2 \leqslant (m+1)p_{m-1}\,p_{m+1} \qquad (m = 1, 2, \ldots)\,. \tag{4.6.2}$$

This inequality is violated for some states of the field. In particular, for the case of 'two-photon light' consisting of photon pairs and for 100% efficiency of the detector, $p_1 = p_3 = 0$, $p_2 \neq 0$, so that Eqn (4.6.2) is violated for $m = 2$.

Further, it follows from Eqn (4.6.1) that the *factorial moments* of photocounts

$$G_k \equiv \langle m(m-1)\ldots(m-k+1)\rangle$$

are given by the relation

$$G_k = \int_0^\infty \mathrm{d}I\, p(I)(kTI)^k\,,$$

i.e., G_k are proportional to ordinary moments for the intensity $\langle I^k\rangle$. Hence, we obtain another set of nonclassicality criteria for the light [45],

$$G_k^2 \leqslant G_{k-1}G_{k+1} \qquad (k = 1, 2, \ldots)\,. \tag{4.6.3}$$

In particular, putting $k = 1$, we obtain $G_1^2 \leqslant G_2$, or $\langle \Delta m^2\rangle \geqslant \langle m\rangle$. Thus, the sub-Poissonian statistics of photocounts contradicts the semiclassical theory. Note that the criteria of nonclassicality (4.6.2), (4.6.3) have a clear geometric interpretation: for example, $\ln(G_k)$ plotted versus k, according to inequality (4.6.3), has a concave form [44]. There also exist other observable criteria of light 'nonclassicality' [44, 45].

Hence, the semiclassical Mandel formula (4.6.1) for the statistics of photocounts gives several observable criteria of nonclassicality for the light. Nonclassical light cannot be considered as a variety of waves whose random intensities obey some non-negative distribution $P(I)$. The observable criteria of nonclassicality are directly related to the well-known mathematical problem of moments.

Let us trace once again the initial controversies between quantum and semiclassical descriptions of photodetection. In quantum models, the energy transfer from an excited system to a nonexcited one is determined by normally ordered moments relating to the first system (or by antinormally ordered moments relating to the second system). Normally ordered moments are not 'true' moments of some non-negative distribution; therefore, in contrast to ordinary moments, they do not obey general relations like the Cauchy–Schwarz inequality. It is this difference that allows one to point out a class of states that have no classical analogues.

4.7 Projection postulate and the wave function reduction

One should distinguish between the two meanings associated with the terms *projection postulate* and *reduction*. They are connected, respectively, with the postulates of Born (1926) and Dirac (1930).

(1) The Born postulate. In order to calculate the probability of observing a certain eigenvalue a_1 of an operator A at the moment t_1, one should find the *projection* of the vector $|\psi(t_1)\rangle$ on the vector $\langle a_1|$ and take the square of its absolute value,

$$\begin{aligned} p(a_1, t_1) &= \left|\langle a_1|\psi(t_1)\rangle\right|^2 = \left|\langle a_1, t_1|\psi_0\rangle\right|^2 \\ &= \langle\psi_0|P(a_1, t_1)|\psi_0\rangle\,. \end{aligned} \tag{4.7.1a}$$

The last two equalities were obtained using the Heisenberg representation. Here $P(a,t) \equiv |a,t\rangle\langle a,t|$ is the projection operator (*projector*), $|a,t\rangle \equiv U^\dagger(t)|a\rangle$ is an eigenvector of the operator $A(t)$, $U \equiv \exp(-\mathrm{i}Ht/\hbar)$ is the evolution operator, and H is time-independent Hamiltonian. The Born postulate in the Heisenberg representation is also valid for

the case where several commuting operators are measured simultaneously at arbitrary time moments,

$$p(a,t;b,t_2) = \langle\psi_0|P(a,t_1)P(b,t_2)|\psi_0\rangle = \langle\psi_0|P(b,t_2)P(a,t_1)|\psi_0\rangle\,. \qquad (4.7.1b)$$

Symmetric correlation functions of this kind can be called *Born* correlation functions. In contrast to *Wigner* correlation functions, they do not depend on the sign of $t_1 - t_2$ (see Section 4.9).

Thus, Eqns (4.7.1) give an algorithm for the comparison between theory and experiment but does not tell us what happens to a quantum object as a result of its interaction with the measurement devices. One can imagine that as soon as a particle is registered at a point $\mathbf{r}_1$, its WF 'collapses' from the whole space to this point. However, this picture has no operational sense unless one can repeat the experiment with the same particle, see below. Here, the idea of a collapse is an interpretation of the quantum formalism. It is an attempt to describe the events that 'actually' take place in the system.

(2) The Dirac, or projection, postulate (also ascribed to von Neumann) states that registering a value a_1 results in the reduction: the WF of the system $|\psi(t_1)\rangle$ is *projected* onto the vector $|a_1\rangle$,

$$|\psi(t_1)\rangle \to |\psi(t_1)\rangle' = P(a_1,t_1)|\psi_0\rangle \propto |a_1,t_1\rangle \qquad (4.7.2)$$

(the vector $|\psi(t)\rangle'$ is not normalized). Here, in contrast to Eqns (4.7.1), the relation does not describe how the measurement results can be calculated. Instead, it describes what happens to the WF as a result of the observation. According to Eqn (4.7.2), a *measurement* is at the same time the *preparation* of a new WF $|\psi'(t)\rangle$, which allows, with the help of Eqn (4.7.1b), calculation of the result of a *repeated* measurement carried out at $t > t_1$. (Therefore, the Dirac postulate violates the symmetry of the quantum formalism with respect to time inversion.)

In what follows, the terms 'reduction' and 'projection postulate' will be understood according to the second, 'active' definition (4.7.2). (The first definition, which is given by the Born postulate (4.7.1) can be called the 'passive' one.) According to von Neumann, there are two ways of WF variation with time: a 'legal' one, given by the Schrödinger equation, and some other, special way, which is not described by the equations of the standard quantum theory. It is supposed that the reduction is caused by the interaction between the quantum system and the macroscopic measurement device.

The projection postulate (4.7.2) is sometimes derived from the *repeatability principle* (see Refs [2, 48]): a repeated measurement of A in a rather short time interval should give the same value a_1. Otherwise, the concept of measurement would only relate to the past, to *a priori* properties of the object under measurement. Various dynamic models of reduction have been proposed in order to take into account that the macroscopic device has a large (or infinite) number of degrees of freedom [2, 13, 16, 17]. However, so far these models are not confirmed experimentally.

In many textbooks and monographs, reduction is claimed to be the basic postulate of quantum physics (see, for example, Ref. [2]). Reduction is often treated as a 'real' event [2, 18, 19, 49]. One can imagine the state vector of a particle or other quantum object to turn spasmodically (at the 'instance' of the measurement t_1) in some hypothetic complex multi-dimensional space of states. As a rule, an explicit qualitative description of quantum correlation effects such as the Einstein – Podolsky – Rosen (EPR) paradox or 'quantum teleportation' (see below) is based on this idea. However, postulate (4.7.2) is actually not necessary; it is never used for the *quantitative* description of observable effects (for exceptions, see Sections 4.9, 6.2). In some papers, the concept of reduction and its necessity is considered to be doubtful [50 – 54]. For example, according to Ref. [53], p. 351, "... Von Neumann's projection rule is to be considered as purely mathematical and no physical meaning should be ascribed to it." In Ref. [2], on p. 294, it is noted that the projection postulate is not needed if one sets a careful distinction between the preparation and measurement procedures.

In accordance with Eqn (4.7.2), it is often stated that a *measurement* is at the same time the *preparation* of a new WF (see, for instance, Refs [2, 3, 18, 19]). However, in real quantum experiments, completely different procedures are used for the preparation of a WF and for its measurement (see examples in Sections 5 and 6). It is reasonable to distinguish between measurement and *filtering* (using a screen with a pinhole or a polaroid). Filters allow some measurement only with the help of a *detector* (see Fig. 1). Here detection is understood as an evidence of the particle existence, such as a click in a Geiger counter or a track in Wilson's chamber [55].

4.8 Partial wave function reduction

Consider the general scheme of an experiment on observing quantum correlations. Two dispersing particles A and B are prepared in the state

$$|\psi\rangle = \frac{|a_1,b_1\rangle + |a_2,b_2\rangle}{\sqrt{2}}\,, \qquad (4.8.1)$$

where a_i, b_i are eigenvalues of the operators A and B. Nonfactorable states of this kind are called *entangled* states. They form the basis for the famous EPR paradox. Observable events, such as the measurement of A at time t yielding the result a_i, and the measurement of B at time t' yielding the result b_j, can be separated by a spacelike interval. Therefore, $[A,B] = 0$, the sequence of measurements plays no role, and one can apply the Born postulate. According to Eqns (4.7.1) and (4.8.1), signals from remote detectors show exact correlation,

$$p(a_m,b_n) \equiv \langle P(a_m)P(b_n)\rangle = |\langle a_m,b_n|\psi\rangle|^2 = \frac{1}{2}\,\delta_{mn} \quad (m,n=1,2)\,. \qquad (4.8.2)$$

In order to explain this correlation effect, one often assumes that at the moment of observing the result a_m, *partial reduction* of the WF takes place: $|\psi\rangle \to \sqrt{2}\langle a_m|\psi\rangle = |b_m\rangle$. Similarly, from the viewpoint of the second observer, $|\psi\rangle \to \sqrt{2}\langle b_m|\psi\rangle = |a_m\rangle$.

However, two questions arise here: "in which one of two equivalent detectors does the reduction take place and how does the second detector 'know' about this event?" One has to speak about mysterious 'quantum nonlocality', which implies some superluminal interaction of a new type. Reduction and nonlocal interaction between remote devices are not necessary for the quantitative calculation of EPR experiments. These notions are introduced *ad hoc* in the attempts to find a clear interpretation for quantum correlations (and also, in connection with Bell's paradox, Section 5.7). Similar correla-

tions exist in classical models, Section 5.5; a working setup of this kind is used for teaching students in one of the laboratories of the Department of Physics in Moscow State University [56]. This paradox of a 'superluminal telegraph' is often resolved with the help of the operational approach: if one considers the actual experimental procedure, it becomes clear that observation of the correlation requires a normal classical information channel introduced between the observers [56, 57].

It is natural to generalize Eqn (4.8.1) in the form

$$|\psi\rangle = \sum_{mn} c_{mn}|a_m, b_n\rangle \,. \tag{4.8.3}$$

Hence, we obtain the joint distribution

$$p(a_m, b_n) = |c_{mn}|^2 \,. \tag{4.8.4}$$

One can also define the conditional probability to discover the observable A to be equal to a_m provided that B is equal to b_n,

$$p(a_m|b_n) \equiv \frac{p(a_m, b_n)}{p(b_n)} = \frac{|c_{mn}|^2}{\sum\limits_k |c_{kn}|^2} \,. \tag{4.8.5}$$

For instance, for the state (4.8.1), we obtain $p(b_n) = 1/2$, and it follows from (4.8.2), (4.8.5) that $p(a_m|b_n) = \delta_{mn}$, i.e., the conditional probabilities for entangled states are equal to 1 or 0. This is another feature of full (ideal) correlation.

The verbal description of this correlation, "if I observed $B = b_1$, then I know immediately that $A = a_1$" is often understood as a proof of the 'nonlocality' of quantum phenomena. (Another proof can be found in Section 6.7.) However, such a correlation is also possible in classical models. An even more delicate property of EPR correlations, their controllability, is not a specific feature of quantum models [56]. (The EPR correlations can be controlled, that is, they depend on the parameters of measurement devices in A and B, such as the orientations of the polaroids, Section 5.5.) Principal differences between quantum and classical correlations can be observed only in special cases, see Sections 4.5, 5.5–5.8.

Consider once again the description of measurement and reduction according to the common viewpoint (see Ref. [2] and Section 4.1). An entangled state of the form (4.8.3) appears as a result of the interaction between any two initially independent quantum systems, A and B. Suppose that A is the observed system and B is a macroscopic measurement device, which is also described by the quantum theory. Let A be the operator of the measured quantum value and B correspond to the macroscopic observed value, such as the position of a voltmeter pointer. In addition, let $c_{mn} = \delta_{mn}$, then Eqn (4.8.3) describes a one-to-one EPR correlation between the observed value and the measured one. However, in each separate trial, the pointer shows at a fixed value (let us denote it by the subscript 1). Therefore, we have to postulate the following transformation [for comparison, see Eqn (4.7.2)]:

$$|\psi\rangle = \sum_m c_{mm}|a_m, b_m\rangle \to c_{11}|a_1, b_1\rangle \equiv c_{11}|a_1\rangle \otimes |b_1\rangle \,,$$

i.e., all coefficients c_{mm} except one for some unknown reason turn to zero. The coefficient c_{11} should turn to unity due to the normalization of the new WF. This stage can be called transformation of the *possible* into the *real*, and it is one of the most difficult problems in the quantum measurement theory. As a result, the WF of the whole system factors, and the system again becomes independent of the device, so that they can be considered separately. It is commonly assumed that such reasoning justifies the Dirac postulate (4.7.2), i.e., proves that the WF reduction exists as a result of measurement.

4.9 Wigner correlation functions

Consider the case where the observable Heisenberg operators in Eqn (4.7.1b) do not commute, $\big[P(t_1), P(t_2)\big] \neq 0$. One can easily see that in this case, the standard algorithms of the quantum theory are not valid for calculating $p(t_1, t_2) \equiv p(a, t_1; b, t_2)$. The point is that the operator $P_1(t_1)P_2(t_2)$ is not Hermitian and the Born correlation function $\langle\psi_0|P_1(t_1)P_2(t_2)|\psi_0\rangle$ contains an imaginary term, equal to $\langle\psi_0|[P_1(t_1), P_2(t_2)]|\psi_0\rangle/2\mathrm{i}$, and therefore cannot be used for calculating $p(t_1, t_2)$. The standard formula for the transition probability based on the Born postulate is also useless since it operates with a single time moment t, which is a parameter of the WF ψ_t, and it cannot give the two-time function $p(t_1, t_2)$. We also recall that the 'pure' quantum dynamics, similarly to the classical dynamics, is invariant with respect to the sign of $t_1 - t_2$. It does not reflect causality and irreversibility, which should be introduced additionally, by setting the rules of going around the poles and taking into account dissipation.

In order to improve this situation, let us start from the Dirac projection postulate (4.7.2), i.e., let us assume that the first (in time) measurement of the observable $P(a, t_1)$ causes the reduction

$$|\psi(t_1)\rangle \to |\psi'(t_1)\rangle \equiv P(a_1, t_1)|\psi_0\rangle \,.$$

Hence, the second measurement device 'sees' a changed WF $|\psi'(t_1)\rangle$, and the averaging of $P(b, t_2)$ in the Born postulate should be done with respect to this new WF. Thus, using first Eqn (4.7.2) and then (4.7.1), we obtain the Wigner formula for the joint distribution of two variables [3],

$$\begin{aligned} p(a, t_1; b, t_2) &= \langle\psi'(t_1)|P(b, t_2)|\psi'(t_1)\rangle \\ &= \langle\psi_0|P(a, t_1)P(b, t_2)P(a, t_1)|\psi_0\rangle \,. \end{aligned} \tag{4.9.1a}$$

It is supposed that $t_0 < t_1 < t_2$, i.e., a 'time arrow' is introduced. This correlation function is asymmetric with respect to the sign of $t_1 - t_2$; such functions can be called *Wigner* correlation functions. Sometimes, equations like Eqn (4.9.1a) require additional summation over the non-observable variables. Evidently, Eqn (4.9.1a) can be generalized to the case where an arbitrary number of operators $P_1, \ldots, P_m$ are observed in a sequence [3],

$$\begin{aligned} &p(t_1, \ldots, t_m) = \langle\psi_0|P_1 \ldots P_{m-1}P_mP_{m-1} \ldots P_1|\psi_0\rangle \\ &(m = 1, 2, \ldots; \quad t_0 < t_1 < t_2 \ldots < t_m) \,. \end{aligned} \tag{4.9.1б}$$

From the operational viewpoint, this formula can be compared with experiment only as a whole, the reduction itself cannot be observed. Therefore Eqn (4.9.1b) can be assumed as the basic *measurement postulate*. In fact, it is a generalization of the Born postulate (4.7.1) to the measurement of non-commuting operators. If all operators in Eqn (4.9.1b) commute, then, due to the property $P_m^2 = P_m$, this

formula coincides with Born's definition (4.7.1b) of multi-time correlation functions $\langle\psi_0|P_1 \dots P_m|\psi_0\rangle$.

For a mixed initial state described by the density operator ρ_0, Eqn (4.9.1b) takes the form

$$p(t_1, \dots, t_m) = \mathrm{Sp}(P_m \dots P_1 \rho_0 P_1 \dots P_m)$$
$$(t_0 < t_1 < t_2 < \dots < t_m)\,. \qquad (4.9.2)$$

Although the concept of reduction is quite convenient for the verbal description of certain experiments, there is no sense in the question: "whether reduction actually takes place or not?" (Of course, the appearance of new facts may change this situation.) For clarity, one can admit that a track in the Wilson chamber appears due to a chain of reductions: each 'seeding' atom starting a droplet of water prepares a new WF for the next atom. In such a model, each droplet should correspond to a pair of projectors P_k in Eqn (4.9.1b). It should be stressed, however, that this is nothing but a possible interpretation. Actually, the concept of reduction is not necessary for the quantitative description of a track (see the calculation based on the Born postulate in Ref. [55]). The same relates to the most part of observed quantum effects, including the EPR effects and quantum teleportation (Sections 4.8, 5.5 – 5.9). All of them are actually described by Born's correlations.

Still, the projection postulate in the form (4.9.1) seems to be indeed necessary for the quantitative description of some narrow class of experiments [58]. (Here we mean practical calculations that can be compared with experiment, in contrast to the abstract models of the quantum measurement theory or speculations about the WF of the whole setup or the whole Universe.) Such experiments should satisfy three conditions: each trial contains measurements of two or more operators in sequence; these operators do not commute for some t_1 and t_2 in the Heisenberg representation; and the measurements are carried out with sufficiently fine time resolution. The last condition is not satisfied for the case of the Wilson chamber. The second condition is not satisfied in EPR experiments and experiments on quantum teleportation, therefore, one does not need reduction for their description, and the observed effects can be explained in terms of ordinary correlation functions. Examples of using the Wigner formula will be given in Sections 6.2 and 6.3.

Thus, one should distinguish between two types of correlation effects, depending on whether the corresponding Heisenberg operators $A(t)$ and $B(t')$ commute or not. In the first case, one can use standard ('Born') symmetrical correlation functions

$$\big\langle A(t)B(t')\big\rangle = \big\langle B(t')A(t)\big\rangle\,, \qquad (4.9.3)$$

while in the second case, the time sequence of measurements is essential, and the 'Wigner' correlation functions of the form

$$\big\langle A(t)B(t')A(t)\big\rangle\theta(t'-t) + \big\langle B(t')A(t)B(t')\big\rangle\theta(t-t') \qquad (4.9.4)$$

should be used. These correlation functions can be interpreted in terms of reduction.

4.10 Mixed states

As a rule, a system cannot be prepared in a *pure state*, i.e., in a state described by some WF ψ_0. In each trial, different WFs are prepared, and one can only give probabilities $p_1, p_2, \dots p_j, \dots$ of preparing the system in different pure states from the set $\psi_1, \psi_2, \dots \psi_j, \dots$ (This set is not necessarily complete or orthogonal.) For comparison with experiment, quantum moments and distributions obtained for all ψ_j should be additionally averaged with respect to the classical distribution p_j. As a result, we obtain a combination of classical and quantum probability models with double stochasticity: $\langle A\rangle = \sum p_j\langle\psi_j|A|\psi_j\rangle$. For instance, it seems reasonable to assume that a macroscopic source of particles heated to temperature T emits particles in pure states $|v_j\rangle$ with definite momenta mv_j; the corresponding probabilities p_j are given by the Maxwell distribution with temperature T.

It is convenient to introduce the *density operator* $\rho \equiv \sum p_j|\psi_j\rangle\langle\psi_j|$. Let $\{|n\rangle\}$ be some complete set of vectors, i.e., $I = \sum |n\rangle\langle n|$, then we can define the *density matrix* in the n-basis, $\rho_{mn} = \sum_j p_j\langle m|\psi_j\rangle\langle\psi_j|n\rangle$. The mean value can be written as $\langle A\rangle = \sum_{mn}\rho_{mn}A_{nm}$. In the invariant form (regardless of the basis), $\langle A\rangle \equiv \mathrm{Sp}(\rho A)$, where Sp denotes the sum of diagonal elements. If the initial set $\{\psi_j\}$ forms a complete orthogonal basis, the density matrix in this basis has a diagonal form: $\rho_{jj'} = p_j\delta_{jj'}$. In the particular case of a pure state, $|\psi\rangle = |k\rangle$, the density operator has the form $\rho = |k\rangle\langle k|$, and the density matrix has a single nonzero element equal to 1: $\rho_{mn} = \delta_{mk}\delta_{nk}$. In this case, $\rho^2 = \rho$.

Additional classical averaging of $\sum p_j\langle\psi_j|A|\psi_j\rangle$ can be performed at the very end of the calculation; however, classical stochasticity is usually introduced at the beginning. Then the *state* of the system is understood as an element in the corresponding extended space of states, i.e., a set of Hermitian non-negative normalized matrices ρ_{mn}. Such states are called *mixed* states. The time dependence of the density operator can be obtained by replacing the basis $|n, t_0\rangle$ by $|n, t\rangle$. Taking into account the Schrödinger equation, we come to the von Neumann equation:

$$\mathrm{i}\hbar\,\frac{\mathrm{d}\rho_t}{\mathrm{d}t} = [H, \rho_t]\,.$$

Let us mention another model where the notions of a mixed state and the density matrix ρ are very helpful. Let a system described by two independent operators A and B be in an arbitrary pure state $\psi(a, b) = \langle a, b|\psi\rangle$, $\rho = |\psi\rangle\langle\psi|$. Here $|a, b\rangle \equiv |a\rangle|b\rangle$ is the direct product of two vector spaces where A and B are defined $\big([A, B] = 0\big)$. Suppose that we are interested only in the operator B or in its functions $f(B)$. One can easily check that the 'marginal' moments $\big\langle f(B)\big\rangle \equiv \big\langle\psi\big|f(B)\big|\psi\big\rangle$ can be represented as $\big\langle f(B)\big\rangle = \mathrm{Sp}\big\{\rho_b f(B)\big\}$. This is done by introducing the following definition for the density operator ρ_b, which relates only to the system B:

$$\rho_b \equiv \mathrm{Sp}_a\{\rho\} \equiv \sum_a \langle a|\psi\rangle\langle\psi|a\rangle\,.$$

The operator ρ_b is defined only in the space of the operator B; unnecessary variables a are excluded beforehand. (This definition is analogous to the definition of marginal probabilities in probability theory, Section 3.3.) In the general case, the operator ρ_b is not diagonal.

As a rule, the second definition of the density operator is used in cases where the operators A and B relate to two different objects, for instance, to interacting subsystems in the models of measurement or to correlated particles in EPR experiments. In the general case, the state of the whole system ψ is not factorable [see Eqn (4.8.1)], and each separate system cannot be described by an individual WF; a correlation

between the observable parameters of the two systems can exist even for a large distance between them.

5. Two-level systems

5.1 q-bits

Sometimes, the interaction of an atom with resonant monochromatic radiation can be described taking into account only two nondegenerate energy levels of the atom. In this case, an arbitrary state of the atom can be represented as a combination of two base vectors: $|\psi\rangle = \alpha|g\rangle + \beta|e\rangle$, where the letters g and e relate to the ground and excited states, respectively. Hence, an arbitrary state of a two-level atom is given by a pair of complex numbers (α, β). Taking into account the normalization $|\alpha|^2 + |\beta|^2 = 1$ and ignoring the common phase of α and β, we obtain that a state is given by two real parameters. For instance, these may be the spherical coordinates (θ, ϕ) of a point on the *Bloch sphere* (see, for instance, Ref. [35]). The polarization state of a classical monochromatic wave or of a photon can be analogously depicted on the *Poincaré sphere*. A single-mode polarized field in a cavity interacting with a two-level atom [59–62] can exist in a superposition of vacuum and single-photon states. This means that the space of states for the cavity has the same structure, $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. Analogous geometry is typical for the space of spin states for a particle with spin $1/2$. In terms of group theory, such a space is called SU(2) space.

During the last few years, considerable interest has been attracted to the idea of *quantum computers* (see Refs [13, 17, 59, 60]), where electronic cells with a dichotomous spectrum of states $(0, 1)$ could be replaced by systems described by vectors in SU(2) space (correlated two-level atoms, photons). All cells of the computer should be in a joint pure entangled state Ψ. Such a device is supposed to increase drastically the computing rate for some classes of problems. The information contained in the numbers (α, β) or (θ, ϕ) is called a q-bit.

If an atom interacts with a single-mode field in a cavity, an inverse exchange of q-bits can take place:

$$|\Psi\rangle = \big(\alpha|g\rangle + \beta|e\rangle\big)|0\rangle \to |\Psi'\rangle = |g\rangle\big(\alpha|0\rangle + \beta|1\rangle\big)\,. \quad (5.1.1)$$

This process, as well as further transfer of a q-bit from the field to the second atom, has been recently observed in Ref. [61]. Moreover, an entangled EPR state of two atoms has been prepared using the interaction between the atoms via the field [61]:

$$|e_1, g_2, 0\rangle \to \frac{\big(|e_1, 0\rangle + |g_1, 1\rangle\big)|g_2\rangle}{\sqrt{2}} \to \frac{\big(|e_1, g_2\rangle + |e_2, g_1\rangle\big)|0\rangle}{\sqrt{2}}\,. \quad (5.1.2)$$

In the process of *quantum teleportation* (Section 5.9), a q-bit is irreversibly transferred from one photon to another.

5.2 An example of quantum state preparation

With the help of modern laboratory equipment, a single atom can be confined in a limited spatial domain (magneto-optical trap) and cooled to superlow temperatures of about 10^{-7} K. In this case, one knows for sure that the atom gets into the ground state $|g\rangle$. Let a short laser pulse with definite amplitude and duration τ (a π-pulse, see Ref. [35]) be incident on the atom at time $t_0 \equiv -\tau$. Intense laser radiation to a good approximation can be considered as classical. Let the laser frequency coincide with the Bohr frequency $\omega_e = (E_e - E_g)/\hbar$ of the transition between the ground state $|g\rangle$ and one of the excited states $|e\rangle$. According to semiclassical theory, a laser pulse drives the atom into the state $|\psi\rangle = \alpha|g\rangle + \beta|e\rangle$, where the coefficients α, β are determined by the phase of optical oscillations and the 'square' of the laser pulse, i.e., the product of the amplitude E_0 and the duration τ. This method is used in modern experiments [61].

Thus, at $t_0 \equiv 0$, an atom is prepared in a given quantum state, similarly to a coin or a die (Section 3.1). Further, this state varies in time in accordance with the Schrödinger equation,

$$|\psi_t\rangle = \alpha|g\rangle + \beta|e\rangle \exp(-\mathrm{i}\omega_e t)\,.$$

Because of the inevitable fluctuations of the laser amplitude and phase, in a series of repeated measurements the state of the ensemble of atoms should be considered as a mixed state.

Note that no properties of the quantum object are measured in the course of this procedure. In other words, preparation does not necessarily coincide with measurement, as it is traditionally supposed [2, 3, 18, 19]. It is essential here that the laser field, which plays the role of a given external force, is described classically, and the atom is supposed to be prepared in the ground state $|g\rangle$ due to relaxation. Similarly to the measurement stage, the preparation stage should be described after introducing a physically reasonable boundary between the classical and quantum worlds. A lot of successful previous experiments make one confident that this heuristic model is correct.

Up to now, we have neglected the interaction of the atom with nonexcited (vacuum) modes of the field. This assertion holds true only at sufficiently short time intervals. Taking into account the interaction between the atom and the vacuum field modes, one comes to the spontaneous emission of a photon (more accurately, an exponential wave packet with central frequency ω_e and duration $\tau_e = 1/w_{ge}$, which is determined by the probability of spontaneous transition per unit time, w_{ge}). The point on the Bloch sphere, which depicts the state of the atom, spirals from the north pole to the south pole [35]. In a time much larger than τ_e, the atom, with high probability, gets to the south pole, into the ground state, while the field gets into the one-photon state $|1\rangle$. Thus, the model suggests an example where both the atom and the field are prepared in a definite state.

One can see that modern equipment provides rather reliable methods for preparing given states of atoms and fields. This possibility of 'WF engineering' is widely used for the experimental verification of many interesting effects in the interaction of a field with matter predicted by the quantum theory [13, 14, 62]. As we have already mentioned, this technique also attracts attention in connection with the idea of quantum computing.

5.3 Polarization of light

Let us recall the classical description of polarization (see Ref. [63]). The field of a plane quasi-monochromatic wave in vacuum can be represented as

$$\mathbf{E}(z, t) = 2\,\mathrm{Re}\big[\mathbf{E}_0 \exp(\mathrm{i}kz - \mathrm{i}\omega t)\big]\,,$$

where the complex vector $\mathbf{E}_0 = \mathbf{x}E_x + \mathbf{y}E_y$ gives the intensity and polarization properties of the wave; and E_x and E_y are projections of the field onto the directions $\mathbf{x}, \mathbf{y}$. For an ideal

monochromatic wave, the vector $\mathbf{E}_0$ is constant; however, in real experiments it always varies, $\mathbf{E}_0 = \mathbf{E}_0(t)$. (Of course, this variation is slow in comparison with ωt.)

The 'instant' (nonaveraged) Stokes parameters are introduced as

$$S_0 \equiv |E_x|^2 + |E_y|^2\,, \qquad S_1 \equiv |E_x|^2 - |E_y|^2\,,$$
$$S_2 \equiv 2\,\mathrm{Re}(E_x^* E_y)\,, \qquad S_3 \equiv 2\,\mathrm{Im}(E_x^* E_y)\,. \tag{5.3.1}$$

The parameter $S_0(t)$ gives the total intensity of the wave at fixed time and the direction of the vector $\mathbf{S}(t) \equiv (S_1, S_2, S_3)$ characterizes the instant polarization. The norm of the vector $S(t) = (S_1^2 + S_2^2 + S_3^2)^{1/2}$, by virtue of Eqns (5.3.1), is equal to $S_0(t)$. The parameter $S_3(t)$ is proportional to the angular momentum of the wave.

Let us introduce a unit vector $\boldsymbol{\sigma} \equiv \mathbf{S}/S$. The set of vectors $\boldsymbol{\sigma}$ belongs to the *Poincaré sphere*. Each point of the sphere corresponds to a definite *type of polarization*. It is convenient to introduce the spherical coordinates

$$\sigma_1 = \cos\theta\,, \qquad \sigma_2 = \sin\theta\cos\phi\,, \qquad \sigma_3 = \sin\theta\sin\phi$$
$$(\theta = 0 - \pi\,, \qquad \phi = 0 - 2\pi)\,. \tag{5.3.2}$$

We also define a unit complex *polarization vector* $\mathbf{e} \equiv (\alpha, \beta)$ (the Jones vector) with the components

$$\alpha \equiv \frac{E_x}{\sqrt{S_0}} = \cos\frac{\theta}{2}\,, \qquad \beta \equiv \frac{E_y}{\sqrt{S_0}} = \exp\mathrm{i}\phi\sin\frac{\theta}{2} \tag{5.3.3}$$

(this vector is defined up to an arbitrary phase multiplier). The inverse transformations have the form

$$\sigma_1 = |\alpha|^2 - |\beta|^2\,, \qquad \sigma_2 = 2\,\mathrm{Re}(\alpha^*\beta)\,, \qquad \sigma_3 = 2\,\mathrm{Im}(\alpha^*\beta)\,. \tag{5.3.4}$$

Thus, an instant polarization state can be given either by spherical coordinates (θ, ϕ) on the Poincaré sphere or by a pair of complex numbers (α, β). For instance, the vectors $\boldsymbol{\sigma} = (\pm 1, 0, 0)$, $\mathbf{e} = \mathbf{x} = (1, 0)$, $\mathbf{e} = \mathbf{y} = (0, 1)$ correspond to linear polarization along x or y, while $\boldsymbol{\sigma} = (0, 0, \pm 1)$, $\mathbf{e} = \mathbf{e}_\pm = (1, \pm\mathrm{i})/\sqrt{2}$ describe right and left circular polarization.

In the case of partially polarized light, all these parameters vary slowly and the depicted point $\boldsymbol{\sigma}(t)$ moves on the Poincaré sphere. The statistics of the field are assumed to be stationary and ergodic; therefore, time and ensemble averaging are equivalent. In terms of the ensemble approach, the Poincaré sphere represents the space of random events, and the space of states consists of various distribution functions $p(\theta, \phi)$, which satisfy the conditions of normalization and non-negativity. One can imagine that the Poincaré sphere is covered by points representing members of the ensemble. The 'density' of points determines the function $p(\theta, \phi)$. Averaging the definitions (5.3.1) gives the ordinary Stokes parameters $\langle S_n \rangle$ $(n = 0, 1, 2, 3)$. The ratio $\left(\langle S_1 \rangle^2 + \langle S_2 \rangle^2 + \langle S_3 \rangle^2\right)/\langle S_0 \rangle \equiv P$ is called the degree of polarization.

The effect of ideal polarization transformers (phase plates, or retardation plates) can be represented as a rotation of the Stokes vector $\mathbf{S}$ described by the Müller matrix $\mathbf{M}$, or as a linear transformation of the polarization vector $\mathbf{e} \equiv (\alpha, \beta)$ by means of the Jones matrix $\mathbf{T}$,

$$\alpha' = t^*\alpha + r^*\beta\,, \qquad \beta' = -r\alpha + t\beta\,. \tag{5.3.5}$$

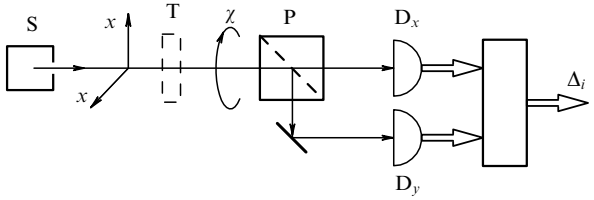

Figure 3. Schematic plot of the measurement of the Stokes parameters. S is the light source, P is a Nicol prism, χ denotes its orientation with respect to the x axis: 0°, for measuring S_1 and 45°, for measuring S_2. T is an additional quarter-wave plate for the measurement of S_3, D_x and D_y are photodetectors, Δi is the difference of their currents.

In vector notation, $\mathbf{S}' = \mathbf{M}\cdot\mathbf{S}$, $\mathbf{e}' = \mathbf{T}\cdot\mathbf{e}$. Here, absorption and reflections are neglected, therefore, S_0 and P do not vary. These parameters are the invariants of the transformation.

Consider measurement of the functions $S_n(t)$ and the Stokes parameters $\langle S_n \rangle$. Let the photodetectors and the registering electronics be sufficiently fast, i.e., their transmission bands be much broader than the spectrum of the field. In this case, one can measure instant values of all parameters. In order to observe all four Stokes parameters simultaneously, one can divide the initial light beam among three detecting devices, see Fig. 3. In the first device, a prism separates the x- and y-polarized components of the beam ($\chi = 0$, there is no retardation plate), so that the difference between the currents from the two detectors, $\Delta i(t)$, is proportional to $S_1(t)$: $\Delta i = kS_1$. In the second device, the prism is rotated by the angle $\chi = 45°$, so that $\Delta i' = kS_2$. In the third device, a quarter-wave plate oriented at 45° is placed before the prism, and $\Delta i'' = kS_3$. (The proportionality coefficients k are assumed to be equal for all three devices.) The sum of the two currents in each device is proportional to the total intensity of the beam: $i_1 + i_2 = kS_0$. For an arbitrary polarization transformer T inserted before the prism, the difference between the currents is proportional to the projection of the vector $\mathbf{S}$ onto a fixed direction in the Poincaré space [63].

Such a device enables one to observe fluctuations of the Stokes parameters $S_n(t)$ near their mean values $\langle S_n \rangle$. The mean values of photocurrents give the ordinary Stokes parameters $\langle S_n \rangle$. They can be measured in turn using a single detector, since the field is supposed to be stationary.

The ultimate accuracy of this measurement is limited by the quantum (photon, shot) noise, $\langle \Delta S_k^2 \rangle_{\mathrm{quant}} = \langle S_0 \rangle$. However, for some states of the field, called *polarization-squeezed* states [64], this noise can be reduced, $\langle \Delta S_k^2 \rangle < \langle S_0 \rangle$. Let us mention here the effect of *hidden polarization* [63], where $P = 0$ but the current fluctuations and the correlation between them depend on the parameters of the polarization transformers. These effects can be described phenomenologically in terms of the higher-order Stokes parameters introduced in Ref. [63].

5.4 Measurement of photon polarization

Let us consider the polarization properties of a single photon and their measurement in the optical range. For the description of real experiments, one should use quasi-stationary states, i.e., superpositions of Fock one-photon states with close energies,

$$|\psi(t)\rangle = \int \mathrm{d}\mathbf{k}\, f(\mathbf{k}) \exp(-\mathrm{i}\omega_{\mathbf{k}} t) a_{\mathbf{k}}^\dagger |0\rangle\,.$$

An obvious classical counterpart of such a state is a *single-photon wave packet*, that is, a quasi-monochromatic field with the spectrum $f(\mathbf{k})$, which is localized in space and time [37]. However, for simplicity, here we speak about photons and use the single-mode description.

As was shown for the example in Section 5.2, modern laser technique permits rather reliable preparation of the field in a one-photon state, i.e, generation of a single photon with a fixed polarization. In the general case, this state is given by the vector

$$|\psi\rangle = \alpha|x\rangle + \beta|y\rangle \equiv \alpha|1,0\rangle + \beta|0,1\rangle\,. \tag{5.4.1}$$

Here, $|\alpha|^2 + |\beta|^2 = 1$, and $|x\rangle \equiv |1,0\rangle \equiv |1\rangle_x \otimes |0\rangle_y$ denotes the state with a single photon polarized along the x direction. The second mode y is in the vacuum state. The state $|\psi\rangle$, up to the phase, can be described by two real parameters θ, ϕ, which are the spherical coordinates of the point on the Poincaré sphere [for comparison, see Eqn (5.3.3)],

$$\alpha \equiv \cos\frac{\theta}{2}\,, \qquad \beta \equiv \exp\,\mathrm{i}\phi \sin\frac{\theta}{2}$$
$$(\theta = 0 - \pi\,, \quad \phi = 0 - 2\pi)\,. \tag{5.4.2}$$

The parameters α, β and θ, ϕ play the same role as in the case of a classical polarized wave. For $\alpha = \beta = 1/\sqrt{2}$, the photon is polarized at an angle 45° to the x axis, and for $\alpha = \mathrm{i}\beta = 1/\sqrt{2}$, it has right circular polarization.

Let the photon be polarized linearly in the plane (x, y) at the angle $\theta/2$ to the x axis, i.e., $\alpha = \cos\theta$, $\beta = \sin\theta$. (This can also be written as $|\psi\rangle = |\theta\rangle$.) Let the photon be detected by an ideal photon counter with 100% quantum efficiency regardless of the photon polarization. Then a current pulse (a photocount) appears at the output of the detector with probability equal to unity. There is no stochasticity here, the probability of photon detection is $p = 1$. (Here we pay no attention to the stochasticity at the moment t_1 of appearance of the photocurrent pulse; it is only essential that the pulse appears within the wave packet duration.)

Stochasticity appears only in the case where some polarizing device is inserted before the detector. This can be, for instance, a polarizing beam splitter (see Fig. 3). Let the prism axis be directed along x ($\chi = 0$), and the two output beams of the prism be fed into a pair of ideal detectors (photon counters). Then the counters 'click' with probabilities $p_1 = \cos^2\theta$ and $p_2 = \sin^2\theta$. (This is analogous to the space of states for a magnetized coin, see Section 3.1.) Each trial results in only a single photocount observed in one of the two detectors. Rotation of the prism by the angle $\chi = \theta/2$ restores the regularity, $p_1 = 1$, $p_2 = 0$. Note that the angle θ can be determined from the measured dependencies $p_n(\chi)$ with a certain ambiguity, and it is necessary to repeat the measurement for a different position of the prism.

To consider a more general case, let us define the Stokes operators in terms of photon creation and annihilation operators $a^\dagger$, a in two polarization modes, by analogy with Eqn (5.3.1):

$$S_0 \equiv a_x^\dagger a_x + a_y^\dagger a_y\,, \qquad S_1 \equiv a_x^\dagger a_x - a_y^\dagger a_y\,,$$
$$S_2 \equiv a_x^\dagger a_y + a_x a_y^\dagger\,, \qquad S_3 \equiv \frac{a_x^\dagger a_y - a_x a_y^\dagger}{\mathrm{i}}\,. \tag{5.4.3}$$

Note that the operators S_1, S_2, S_3 do not commute, for instance, $[S_1, S_2] = 2\mathrm{i}S_3$. Performing the averaging with the help of Eqn (5.4.1), we obtain

$$\langle S_0\rangle = 1\,,$$
$$\langle S_1\rangle = |\alpha|^2 - |\beta|^2 = \cos\theta\,,$$
$$\langle S_2\rangle = 2\,\mathrm{Re}(\alpha^*\beta) = \sin\theta\cos\phi\,,$$
$$\langle S_3\rangle = 2\,\mathrm{Im}(\alpha^*\beta) = \sin\theta\sin\phi \tag{5.4.4}$$

[for comparison, see Eqn (5.3.2)]. Now, $\langle\mathbf{S}\rangle = \big(\langle S_1\rangle, \langle S_2\rangle, \langle S_3\rangle\big)$ is a unit vector: $\langle S_1\rangle^2 + \langle S_2\rangle^2 + \langle S_3\rangle^2 = \langle S\rangle_0 = 1$. Hence, $P = 1$, and a one-photon single-mode field in a pure state is completely polarized, like a classical monochromatic field. Note that $\langle S_k\rangle^2 \neq \langle S_k^2\rangle = 1$ and $\langle S_1^2\rangle + \langle S_2^2\rangle + \langle S_3^2\rangle = 3$. The Stokes vector of a photon $\langle\mathbf{S}\rangle$ can be depicted as a point on the Poincaré sphere. Like the polarization vector $\mathbf{e} = (\alpha, \beta)$, it characterizes the degree of polarization of the photon. In other words, $\langle S_k\rangle$ characterizes correlations between the properties of the field in two modes. Therefore, S_k can be called *correlation* operators.

Using retardation plates, one can change the polarization parameters and turn the initial photon with fixed polarization (5.4.1) into a photon with any given polarization: linear, circular, or elliptic. Such transformations can be conveniently described in terms of the Jones matrices $\mathbf{T}$ acting on a two-dimensional polarization vector $\mathbf{e} \equiv (\alpha, \beta)$ [see Eqn (5.3.5)].

Consider the operators S_k acting on the vector (5.4.1). With the help of Eqns (5.4.3), we obtain

$$S_0|\psi\rangle = \alpha|x\rangle + \beta|y\rangle\,, \qquad S_1|\psi\rangle = \alpha|x\rangle - \beta|y\rangle\,,$$
$$S_2|\psi\rangle = \beta|x\rangle + \alpha|y\rangle\,, \qquad S_0|\psi\rangle = -\mathrm{i}\beta|x\rangle + \mathrm{i}\alpha|y\rangle\,. \tag{5.4.5}$$

It follows that the action of the operators S_k on the state vector of a photon, $|\psi\rangle$, is equivalent to the action of the 2×2 Pauli matrices on the photon polarization vector, $\mathbf{e}$, so that $S_1 \sim \sigma_z$, $S_2 \sim \sigma_x$, $S_3 \sim \sigma_y$. In experiments, such transformations can be performed using retardation plates. This demonstrates that observables in physics have a dual character: they describe both the values being measured and the transformations of the states.

The state of a one-photon field can be measured in three stages described above: for $\chi = 0$, for $\chi = 45°$, and inserting an additional quarter-wave plate. In these stages, the operators S_1, S_2, and S_3 are measured in turn. Therefore, the operators S_k can be considered as observables. Their eigenvalues are $s_k = \pm 1$. In each trial, it is assumed that a 'click' in the upper or lower detector indicates that $s_{ki} = +1$ or -1. (For comparison, one can recall the classical case where s_k are determined by the photocurrent differences Δi and have continuous spectrum.) Naturally, the Nicol prism and the retardation plates are described classically. It is also assumed, according to the model of photodetection, that the probability of a photocount is proportional to $\langle\psi|a'^\dagger a'|\psi\rangle$, where the operators $a'^\dagger$, a' relate to the field on the detector. In a large number of trials N, one can measure the mean value of one of the Stokes parameters for the state ψ: $\langle S_k\rangle \approx N^{-1}\sum s_{ki}$. Further, with the help of Eqns (5.4.2) one can determine the parameters of the state α, β [or, equivalently, the components of the photon polarization vector $\mathbf{e} \equiv (\alpha, \beta)$]. With a proper choice of retardation plates, one can measure the projection of the Stokes vector $\langle\mathbf{S}\cdot\mathbf{n}\rangle$ on any given direction $\mathbf{n}$ with respect to the Poincaré sphere [63].

Measurements according to the scheme shown in Fig. 3 enable one to find out whether the initial state of the field is a

one-photon state or not. In the first case, each trial gives exactly one photocount registered by one of the two detectors. (For simplicity, the detectors are assumed to be ideal.)

Note that it is better to consider the polaroid (the Nicol prism) inserted into the path of the photon as a *filter* and not as a detector. To measure the photon polarization, the polaroid (the Nicol prism) should be completed by a detector in combination with a retardation plate, and the procedure should be repeated many times for different positions of the polaroid and the plate. With no detector inserted, the polaroid only *prepares* photons with certain polarization and uncertain time of birth t_0. Formally, the action of a polaroid (unlike the Nicol prism) is described as a nonunitary transformation of the WF or the field operators [21]. This transformation leads to additional stochasticity, and the photons at the output of a polaroid should be described in terms of mixed states.

Thus, once again a certain state ψ [or polarization vector $\mathbf{e} = (\alpha, \beta)$] determined by the preparation devices is ascribed to an individual quantum object (the field in a given trial). Again, in order to check this information, one needs a large number of repeated measurements, see Section 3.2. In other words, one cannot measure the *a priori* polarization of a given photon, since a photocount in the detector with, say, x-polarization can be caused by a photon with any kind of polarization [with the exception for the set $\mathbf{e} = (0, 1)$ with zero measure]. Even if a measurement does not destroy the photon (this could be done using *non-destructive measurements* suggested by V B Braginsky with collaborators [19]), it would vary the initial polarization, according to the Wigner formula (4.9.1). Thus, repeated measurements of a single photon polarization are useless.

Recently, this feature of quantum measurements found a surprising application in *cryptography* where it can be used to discover 'eavesdroppers' [16, 17, 65 – 69]. In quantum cryptography, one sends coded messages using polarization modulation of very weak (better, single-photon) light flashes. One can also use frequency modulation and take advantage of the impossibility of measuring the *a priori* 'color of a photon', i.e., the spectrum of a one-photon wave packet [69].

The properties of photon polarization demonstrate that quantum probability models have a specific feature (are 'non-Kolmogorovian'). Namely, there are no elementary joint probabilities for non-commuting variables, while there exist marginal probabilities (see Sections 5.6 – 5.8).

Analyzing photon polarization, we observe another principal feature of quantum stochasticity: it depends on the parameters of the measurement devices and vanishes for some particular cases. In the case of linear polarization, this occurs if the axis of the polarizing prism is parallel to the initial polarization of the photon, $\chi = \theta/2$. This means that the system (the optical field) is prepared in the eigenstate $|\xi\rangle$ of the operator under measurement. In terms of projecting operators, in this case, one measures the operator $|\psi\rangle\langle\psi|$.

Let us briefly consider the polarization of two-photon states $|\psi\rangle_2$, which are generated in the process of parametric scattering [70]. In the general case, $|\psi\rangle_2 = \alpha|2, 0\rangle + \beta|1, 1\rangle + \gamma|0, 2\rangle$, where $|\alpha|^2 + |\beta|^2 + |\gamma|^2 = 1$. Now, the polarization vector $\mathbf{e} = (\alpha, \beta, \gamma)$ has three components and the projective space is a sphere S^3 in four-dimensional space. One can also define the fourth-order Stokes parameters and the degree of polarization P_4 characterizing this state [63]. For $\beta = 1$, the field is nonpolarized in the usual sense ($P_2 = 0$) but there is *hidden polarization* ($P_4 = 1$). In addition, the state $|1, 1\rangle$ is *polarization-squeezed* [63, 64].

5.5 Correlated photons

Consider two light beams A and B, each one containing a single photon. The beams can differ in frequency and/or directions. In the general case, the state of such a four-mode field can be represented as $|\psi\rangle = \sum_{i,j} c_{ij}|Ai, Bj\rangle$, see for comparison Eqn (4.7.5). Here the subscripts $i, j = x, y$ denote polarization, $|Ai, Bj\rangle \equiv |a\rangle_{Ai} \otimes |1\rangle_{Bj}$ is the state with one photon in the mode Ai and one photon in the mode Bj. If two or more coefficients c_{ij} are nonzero, the WF does not factor ($\psi \neq \psi_A\psi_B$) and no WF exists for a single photon. In other words, the separate photons do not have definite polarization but there is a correlation in their polarizations. Nonfactorable states of this kind, also called entangled states, demonstrate the EPR – Bohm paradox [22, 38 – 40, 56].

Let, for instance,

$$|\psi\rangle = \frac{|Ax, By\rangle - |Ay, Bx\rangle}{\sqrt{2}} \,. \tag{5.5.1}$$

In a more detailed notation, this state can be represented in the form $\big(|10, 01\rangle - |01, 10\rangle\big)/\sqrt{2}$.

Consider the experimental scheme in Fig. 4. Unlike the scheme in Fig. 3, it includes an additional detector for the measurement of the Stokes parameters and a two-photon source. This scheme allows one to measure the Stokes parameters S_{Zk} for two photons ($Z = A, B$; $k = 0 - 3$). From Eqn (5.5.1), it follows that $\langle S_{Z0}\rangle = 1$, $\langle S_{Zk}\rangle = 0$, $P_Z = 0$, i.e., the photons are completely depolarized, and in each beam, repeated trials give photocounts randomly in one of the two detectors, no matter what polarization transformers T_A and T_B are installed before the polarization prisms.

At the same time, from Eqn (5.5.1) it follows that $S_{Ak}S_{Bk}|\psi\rangle = -|\psi\rangle$, i.e., $|\psi\rangle$ is an eigenvector for all three products of operators $S_{Ak}S_{Bk}$. Hence,

$$\langle S_{Ak}S_{Bk}\rangle = -1 \qquad (k = 1, 2, 3) \,. \tag{5.5.2}$$

The form of Eqn (5.5.1) and the property (5.5.2) are invariant in any basis; for instance, in a circular basis, $|\psi\rangle = \big(|A^+, B^-\rangle - |A^-, B^+\rangle\big)/\sqrt{2}$. According to Eqn (5.5.2),

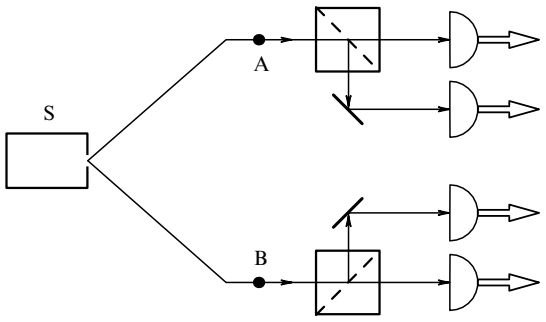

Figure 4. Schematic plot of an experiment demonstrating the absence of *a priori* polarization for single photons. The source S sends pairs of polarization-correlated photons to two detecting devices, A and B. In each trial, one of the two detectors in each device gives a photocount with equal probabilities. Photocounts in the two devices manifest a certain correlation. This correlation cannot be quantitatively described in terms of classical probability models.

there is complete (anti)correlation: if some component of the Stokes vector is measured in both beams, i.e., the polarization transformers are identical, $T_A = T_B$, then $s_{Ak} = -s_{Bk}$. For instance, if observer A measures S_{A1} and obtains $s_{A1} = +1$, i.e., a photocount occurs in the upper detector in Fig. 3, then observer B surely obtains $s_{B1} = -1$. Opposite detectors, the upper one in A and the lower one in B, and vice versa, always 'click' simultaneously, regardless of the basis where the Stokes vector is defined. In other words, the corresponding conditional probabilities are equal to unity (see Section 4.8).

It is often supposed that this correlation demonstrates 'action at a distance' or 'nonlocality' in quantum phenomena. According to the hypothesis of partial reduction (Section 4.8) and the concept 'measurement is preparation', if a photon is registered by detector A, which measures the x-polarization, then the state of the field (5.5.1) is projected onto the vector $\langle Ax|$, so that photon B is instantaneously put into a state with a certain polarization $|By\rangle$. At the same time, if detector A measures right circular polarization, then photon B instantaneously gets left-polarized, i.e., the vector $|\psi\rangle = \big(|A^+, B^-\rangle - |A^-, B^+\rangle\big)/\sqrt{2}$ is projected onto $\langle A^+|$. We stress that this is nothing but an explicit interpretation of the property (5.5.2), which can be neither verified nor disproved experimentally. Note that in order to observe the correlation, one needs an additional information channel between A and B, therefore, the correlation cannot be used for information exchange.

Now, let observer A measure S_1 and B measure S_2. According to Eqn (I.1), $S_{A1}S_{B2}|\psi\rangle = \big(|Ax, Bx\rangle + |Ay, By\rangle\big)/\sqrt{2}$. This vector is orthogonal to $|\psi\rangle$, therefore, $\langle S_{A1}S_{B2}\rangle = 0$. There is no correlation in the signals from the detectors; the upper and lower detectors in Fig. 4 'click' independently. Similarly, $\langle S_{A2}S_{B1}\rangle = 0$. Thus, both observers can control the correlation if they independently change the parameters of T_A and T_B. This result seems at first sight unusual: how can rotation of a plate at point A influence the signal at point B?

However, classical models can give qualitatively similar controlled correlations. The difference between classical and quantum analogues is only quantitative [56]. Recall that the classical description of a nonpolarized wave (Section 5.3) implies that the wave can be considered as completely polarized over short time intervals. At short intervals, the Stokes vector $\mathbf{S}(t)$ has definite directions, and it is time averaging that makes $\langle \mathbf{S}(t) \rangle = 0$. Let us consider for clarity that both photons in Fig. 4 are emitted with certain polarizations, described by the Stokes vectors $\mathbf{S}_Z$ or by the polarization vectors $\mathbf{e}_Z$, and these polarizations change chaotically from trial to trial. Since, in accordance with Eqn (5.5.2), all three components of the Stokes vector have opposite signs for the beams A and B, the two Stokes vectors are oppositely directed, $\mathbf{S}_A = -\mathbf{S}_B$, and the polarization vectors are orthogonal, $(\mathbf{e}_A^* \cdot \mathbf{e}_B) = 0$. The two depicting points on the Poincaré sphere are placed oppositely, i.e., if the A photon is x-polarized, then the B photon is y-polarized, and so on.

Radiation with similar properties can be obtained by means of two ideal lasers A and B generating polarized beams with equal intensities. Both beams pass through polarization transformers, which are controlled by a common random number generator, in a way that provides orthogonality: $\mathbf{e}_B(t) \perp \mathbf{e}_C(t)$. As a result, the points depicting polarization cover the whole Poincaré sphere; any polarization of the beams A or B has equal probability, but the Stokes vectors for A and B are oppositely directed, $\mathbf{S}_A = -\mathbf{S}_B$. There is complete 'anticorrelation' of random polarizations. Thus, we obtained a classical analogue for the property (5.5.2) of the state (5.5.1). (Here, instead of averaging over $|\psi\rangle$, we used classical time or ensemble averaging.)

However, this classical statistical model with the *a priori* photon polarization is inconsistent with quantitative predictions of the quantum theory and with interference experiments. Consider a simplified version of the scheme in Fig. 4, where polarization prisms are replaced by polaroids (analyzers) completely absorbing light with a certain polarization. The photons are registered by two detectors with polaroids at their inputs. Both polaroids are oriented at the same angle $\chi_A = \chi_B \equiv \chi$ to the x axis. From $(\mathbf{e}_A^* \cdot \mathbf{e}_B) = 0$, it follows that each time, only one of the detectors 'clicks', A or B, i.e., the probability of a photocount coincidence $p_{AB}(\chi_A, \chi_B) \equiv p_{AB}(\chi) = \langle m_A m_B \rangle$ is equal to zero for any χ. (The parameter m is set to unity if a photocount occurs and zero if there is no photocount.)

At the same time, a complete absence of coincidences contradicts the classical principles. Indeed, let, for instance, $\chi_A = \chi_B = 0$, then, from time to time, photons with polarizations $\theta_A/2 = +45^\circ$ and $\theta_B/2 = -45^\circ$ should *both* pass through the polaroids. This paradox can be considered as a consequence of our assertion that each photon has *a priori* polarization; as a result, the probability of a photon passing through a polaroid obeys the classical Malus law $p_+ = \cos^2(\theta_A/2 - \chi)$.

Now let $\chi_B - \chi_A = \phi$. From $\langle m_A m_B \rangle = \big|\langle \chi_A, \chi_B | \psi \rangle\big|^2$ and $|\chi\rangle = \cos\chi\,|x\rangle + \sin\chi\,|y\rangle$, it follows that in the state (5.5.1), $\langle m_A m_B \rangle = (1/2)\sin^2\phi = (1/4)(1 - \cos 2\phi)$. This is an example of polarization *two-photon interference* with the visibility V equal to 100%. In particular, for $\phi = 0$ we obtain $\langle m_A m_B \rangle = 0$, i.e., coincidences are completely absent.

On the other hand, analogous classical models for intensity interference lead to a photocurrent correlation of the form $\langle i_A i_B \rangle \sim 1 - V_{clas}\cos 2\phi$, where $|V_{clas}| \leqslant 1/2$. The point is that the interference visibility V is determined by the relation between the moments G_{xx}, G_{yy}, and G_{xy} at the input of the interferometer [37]. In the classical theory, $G_{xx} \equiv \langle a^{*2}a^2 \rangle$, $G_{yy} \equiv \langle b^{*2}b^2 \rangle$, $G_{xy} \equiv \langle a^* a b^* b \rangle$, and the existence of the joint probability distribution for the field amplitudes a, b leads to the Cauchy–Schwarz inequality $(G_{xy})^2 \leqslant G_{xx}G_{xy}$ limiting the maximal visibility V_{clas}. In the quantum case, one should operate with the normally ordered moments $G_{xx} \equiv \langle \psi | a^\dagger a^\dagger a a | \psi \rangle$, $G_{xy} \equiv \langle \psi | a^\dagger b^\dagger a b | \psi \rangle$, which cannot be put into correspondence with some joint probability distribution. In particular, for the state (5.5.1), the corresponding inequality for quantum moments is violated, $G_{xy} = 1$ and $G_{xx} = G_{xy} = 0$. In this case, one speaks of nonclassical light (Section 4.6). Thus, the paradox of coincidence suppression, similarly to many other quantum paradoxes, is caused by the absence of joint probabilities for non-commuting operators.

In Section 5.4, we came to the conclusion that a single photon can be considered as having a certain polarization, i.e., for the case of pure one-photon state preparation, the concept of *a priori* polarization has an operational sense (although it is impossible to measure the polarization of a single photon). For the experiment with two polarization-correlated photons, this is not so. Here, only the whole two-wave field is in the pure state (5.5.1); the polarization of a single wave is described in terms of a mixed state (of the second type, see Section 4.10). In each beam, all three observable Stokes parameters are equal to zero,

$\langle S_{Z1}\rangle = \langle S_{Z2}\rangle = \langle S_{Z3}\rangle = 0$, i.e., the radiation is completely depolarized. In the case of a ordinary mixed one-photon state, this could be explained by random variations of the polarization parameters from trial to trial; however, a 100% visibility of the interference pattern contradicts this explanation. Here, the concept of *a priori* polarizations of two photons has no sense.

One can also prepare photon pairs correlated in frequency and wavefront structure [57, 71, 72]. In this case, it makes no sense to speak about the *a priori* spatiotemporal form of the wave packets corresponding to separate photons. Thus, an evident concept of a photon as a 'real' wave packet with a certain form and polarization contradicts the quantum theory in the case of 'nonclassical' light.

5.6 Negative and complex 'probabilities'

Let us return to Fig. 3. Let the input field be periodically prepared in some one-photon state. The prism is oriented at the angle $\chi = 0$ to the x axis, i.e., one measures the first Stokes parameter S_1. In each trial, there is a 'click' in the upper or lower detector. Let the random variable S_1 take the value $s_1 = +1$ for a 'click' in the upper detector and the value $s_1 = -1$ for a 'click' in the lower one. One can imagine a colored lamp that is governed by the output pulses of the detectors and flashes green at $s_1 = +1$ and red at $s_1 = -1$. Let us forget about the photons and polarizations and try to make a phenomenological description of the events. The detector is considered as a 'black box' with an input aperture and a lamp. A sender generates a sequence of signals with stationary statistics characterized by some probability $p_1(s_1)$. For a set of trials, the result can be written as a sequence of bits of the form $(+\,+\,-\,+\,-\,-\,-\,+\ldots)$. From this sequence, one can determine the probabilities $p_1(+1)$, $p_1(-1) = 1 - p_1(+1)$ and the Stokes parameter $\langle S_1\rangle = p_1(+1) - p_1(-1) = 2p_1(+1) - 1$.

According to the theories of hidden variables, there is no actual stochasticity in reality. Therefore, each signal contains information about the forthcoming value of s_1 and about the detector that should 'click' in this trial. Similarly, the trajectory of a Brownian particle is considered as pre-defined by the initial conditions and the dynamic equations of motion for separate molecules. Since the variables are 'hidden', we only can introduce some probability $p_1(s_1)$, which 'actually' results from averaging over a variety of hidden variables. [One can also assume the hidden variables to be random, then $p_1(s_1)$ plays the role of a marginal distribution given by summation over the multi-dimensional distributions for the hidden variables [38].] Thus, we suppose that the signals have some *a priori* property S_1, which randomly varies from trial to trial. Let us call this assumption the postulate of *a priori* existing observables.

Let us now rotate the prism in Fig. 3 by 45°, i.e., let us measure S_2. A set of trials gives a new distribution $p_2(s_2)$, which, evidently, also characterizes the signals sent to the detector. Indeed, the parameters of the sending device did not vary while we rotated the prism. (In principle, the prism can be rotated after the signal leaves the sender. This is the well-known method of *delayed choice* suggested by Wheeler.) From the viewpoint of hidden variables theory, the signals have at least two *a priori* properties S_1 and S_2. These properties determine which detector 'clicks' in each of the two options in Fig. 3. (Each polarization transformer should be characterized by its own random value but for our purposes, two variables are enough.) The signals should carry information about the outcome of any possible trial. For instance, the source can send the command $(+-)$. This command makes a type-1 detector (measuring S_1) to give a green flash, and a type-2 detector (measuring S_2), a red flash. A series of signals is a sequence of commands of the form $(++)$, $(+-)$, $(-+)$, or $(--)$. For a stationary source, these commands should occur with certain probabilities $p(s_1, s_2)$ (for comparison, see the model with two coins in Section 3.4). Both variables, s_1 and s_2, have certain values $+1$ or -1 no matter whether they are observed or not.

However, this joint distribution cannot be measured: in each trial, only a single polarization transformer is inserted before the detector, and one of the two commands is ignored. If the signals were not single-photon ones, they could be 'cloned', i.e., divided between two channels and measured by two independent setups. It is also impossible to perform repeated measurements on a single photon: according to the projection postulate (4.7.2) or the Wigner formula (4.9.1), the first measurement changes the state of the field, so that one cannot observe the *a priori* values s_1 and s_2 in a single trial. Similarly, the spin projections σ_k of a single particle cannot be measured simultaneously.

Thus, with the help of an experiment with two types of detectors, one can formally introduce the concept of a nonmeasurable joint distribution $p(s_1, s_2)$ for two random variables S_1, S_2 observed in turn. This corresponds to the common classical viewpoint, which implies that all observable properties of objects exist *a priori*, i.e., before the measurement.

It is also impossible to calculate four probabilities $p(s_1, s_2)$ in the framework of the quantum theory: there is no appropriate algorithm for such a calculation. The operators S_1 and S_2 do not commute; hence, they have no common eigenstates and the Born postulate (4.7.1) is not valid. However, one can first calculate the averaged products of non-commuting operators (quantum moments) and then express the probabilities in terms of moments using the algorithms of classical probability theory [38, 56]. It should be noted that the product of two non-commuting Hermitian operators is non-Hermitian; therefore, the moments and the 'probabilities' calculated via the moments can take complex values.

This procedure can be illustrated by a simple example. Let the field be prepared in a one-photon state with arbitrary polarization, $|\psi\rangle = \alpha|x\rangle + \beta|y\rangle$. Suppose that in each trial, a photon has *a priori* components of the Stokes vector s_1, s_2, equal to $+1$ or -1. Let us introduce two random variables S_1, S_2 taking these values with some probabilities $p(s_1, s_2)$. In the classical theory, elementary probabilities are related to the moments by Eqn (3.4.3), which leads to the constraint (3.4.4). In particular, the following inequality should be satisfied:

$$\langle S_1 S_2\rangle \geqslant \langle S_1\rangle + \langle S_2\rangle - 1\,. \tag{5.6.1}$$

Using (I.1), we find the nonzero moments in the state $|\psi\rangle$:

$$\begin{aligned}
&\langle S_2 S_3\rangle = \langle S_3 S_2\rangle^* = \langle S_1 S_2 S_3\rangle = \langle S_1 S_3 S_2\rangle^* = \mathrm{i}\langle S_1\rangle\,,\\
&\langle S_3 S_1\rangle = \langle S_1 S_3\rangle^* = \langle S_2 S_3 S_1\rangle = \langle S_2 S_1 S_3\rangle^* = \mathrm{i}\langle S_2\rangle\,,\\
&\langle S_1 S_2\rangle = \langle S_2 S_1\rangle^* = \langle S_3 S_1 S_2\rangle = \langle S_3 S_2 S_1\rangle^* = \mathrm{i}\langle S_3\rangle\,.
\end{aligned} \tag{5.6.2}$$

Here, the Stokes parameters $\langle S_k\rangle$ are defined in Eqn (5.3.4) in terms of α, β. For instance, let $\phi = 0$, $\theta = 45°$ (the light is polarized linearly at an angle 22.5° to the x axis). Then $\langle S_1\rangle = \langle S_2\rangle = 1/\sqrt{2}$, $\langle S_1 S_2\rangle = \langle S_3\rangle = 0$ (see Fig. 2). As a

result, inequality (5.6.1) is not satisfied and the probability $p(-1,-1)$, according to Eqn (3.4.3), is negative, $p(-1,-1) = (1/4)(1-\sqrt{2}) \approx -0.1$.

In the general case, a question arises: "which sequence of operators in quantum moments gives a proper correspondence with the classical theory?" According to Eqn (5.6.2), the symmetrized expression turns to zero,

$$\frac{1}{2}\big(\langle S_1 S_2\rangle + \langle S_2 S_1\rangle\big) = \mathrm{Re}\big(\langle S_1 S_2\rangle\big) = 0\,.$$

The normally ordered moment $\langle :S_1 S_2: \rangle = \langle S_1 S_2\rangle - \mathrm{i}\langle S_3\rangle$ is also equal to zero. As a result, Eqn (3.4.3) takes the form

$$p(s_1,s_2) = \frac{1}{4}\big(1 + s_1\langle S_1\rangle + s_2\langle S_2\rangle\big)\,. \qquad (5.6.3)$$

Then the probability of the event $(s_1 = s_2 = -1)$ in the state $|\psi\rangle$ should be

$$\begin{aligned} p(-1,-1) &= \frac{1}{4}\big(1 - \langle S_1\rangle - \langle S_2\rangle\big) \\ &= \frac{1}{4}(1 - \cos\theta - \sin\theta\cos\phi)\,. \end{aligned} \qquad (5.6.4)$$

This expression takes negative values for certain types of photon polarization; therefore, it cannot be considered as a probability.

If we choose the antisymmetrized expression

$$\frac{\langle S_1 S_2\rangle - \langle S_2 S_1\rangle}{2\mathrm{i}} = \mathrm{Im}\big(\langle S_1 S_2\rangle\big) = \langle S_3\rangle\,,$$

Eqn (3.4.3) takes the form

$$p(s_1,s_2) = \frac{1}{4}\big(1 + s_1\langle S_1\rangle + s_2\langle S_2\rangle + s_1 s_2\langle S_3\rangle\big)\,. \qquad (5.6.5)$$

This expression also takes negative values for some polarizations. Generalization of this result to the case of two or more photons belonging to several beams gives a general approach and a 'minimal solution' to certain paradoxes known in quantum optics [38].

According to this example, it makes no sense to suppose that in each trial, a one-photon field contains information about the outcomes of any possible experiments of measuring its polarization, and about the detector in Fig. 3 that would 'click' in this trial. Such a viewpoint contradicts the quantum theory since it leads to negative probabilities. In the general case, quantum moments do not correspond to any elementary probability distribution; in this sense, they are not 'proper' moments and the model is non-Kolmogorovian. Recall that the operator S_1 has the sense of the photon number difference and the operator S_2, for the case of one-photon states, corresponds to the cosine of the phase difference for the amplitudes of fields in polarization modes. Therefore, these properties of the fields cannot be assumed to have *a priori* values. If the state of a photon is given by the polarization vector (α,β) or by the Stokes vector $\big(\langle S_1\rangle, \langle S_2\rangle, \langle S_3\rangle\big)$, i.e., the preparation procedure is known, this does not mean that one knows the properties of the photon. It is only possible to predict the statistics of future experiments with a large number of identically prepared photons. These statistics cannot be described in terms of joint probabilities $p(s_1,s_2)$ or $p(s_1,s_2,s_3)$.

However, one cannot confirm this conclusion experimentally since non-Hermitian operators are non-observable. In the forthcoming sections, it will be shown that for 4-mode and 6-mode models, a similar controversy between the classical concepts and the quantum theory can be formulated in terms of Hermitian operators and observable moments.

5.7 Bell's paradox for the Stokes parameters

With the help of the experimental setup shown in Fig. 4, violation of the famous Bell inequality [38–40, 42] can be demonstrated. The source simultaneously sends photons to two remote detectors A and B. At each trial, the field is prepared in the two-photon state (5.5.1), so that there exists a correlation between the properties of the photons from each pair [see Eqn (5.5.2)].

Let the detector A measure either $A \equiv S_{\mathrm{A1}}$ $(\chi = 0)$ or $A' \equiv S_{\mathrm{A2}}$ $(\chi = 45^\circ)$, and the detector B either $B \equiv 2^{-1/2}(S_{\mathrm{B1}} + S_{\mathrm{B2}})$ $(\chi = 22.5^\circ)$ or $B' \equiv 2^{-1/2}(S_{\mathrm{B1}} - S_{\mathrm{B2}})$ $(\chi = -22.5^\circ)$. Four series of experiments are carried out, each one containing N trials. In each trial, one measures one of the four pairs (A,B), (A',B), (A,B'), and (A',B'). Each trial yields a pair of numbers (a_i,b_i), where $a_i, b_i = \pm 1$. From the $4N$ numbers obtained in this experiment, one calculates the following N numbers:

$$f_i \equiv \frac{1}{2}(a_i b_i + a'_{N+i} b_{N+i} + a_{2N+i} b'_{2N+i} - a'_{3N+i} b'_{3N+i})\,. \qquad (5.7.1)$$

Further, one finds the arithmetic mean,

$$\langle F\rangle_N \equiv N^{-1}\sum_{i=1}^{N} f_i\,.$$

The measurement procedure and a computer simulation of it are described in detail in Ref. [56]; real optical experiments are described and analysed in Ref. [42].

For $N \to \infty$, one can assume $\langle F\rangle_N \to \langle F\rangle$, where

$$\begin{aligned} F &\equiv \frac{1}{2}(AB + A'B + AB' - A'B') \\ &= \frac{1}{2}\big[A(B+B') + A'(B-B')\big] \\ &= 2^{-1/2}(S_{\mathrm{A1}}S_{\mathrm{B1}} + S_{\mathrm{A2}}S_{\mathrm{B2}})\,, \end{aligned} \qquad (5.7.2)$$

and the angular brackets denote averaging either with respect to the WF (5.5.1) or with respect to some classical set of probabilities $p(a,b,a',b')$.

In the quantum case, from Eqns (5.5.2) and (5.7.2) it follows directly that

$$\langle F\rangle_{\mathrm{quant}} = \frac{1}{2}\langle\psi|AB + A'B + AB' - A'B'|\psi\rangle = -\sqrt{2}\,. \qquad (5.7.3)$$

Note that in the quantum theory, $\langle\psi|F|\psi\rangle$ does not correspond to the standard definition of an observable mean value given by the Born postulate (4.7.1), since the operator F contains non-commuting variables that can be only measured in different trials.

According to the classical hypothesis of hidden variables, it is assumed that there exist $2^4 = 16$ elementary probabilities $p(a,b,a',b')$ determined by the properties of the light source. The mean value of a random variable F can be expressed in terms of the elementary probabilities as

$$\langle F\rangle_{\mathrm{clas}} = \sum_{a,b,a',b'=\pm 1} p(a,b,a',b')\,f(a,b,a',b')\,. \qquad (5.7.4)$$

Here

$$f(a,b,a',b') \equiv \frac{1}{2}(ab + a'b + ab' - a'b') = \frac{1}{2}\left[a(b+b') + a'(b-b')\right]. \qquad (5.7.5)$$

This function of four arguments takes only the values ± 1: for $b = b'$, $f = ab = \pm 1$, and for $b = -b'$, also $f = a'b = \pm 1$. Therefore, the random variable $f(a,b,a',b')$ can only take values $f_{\min} \equiv -1$ or $f_{\max} \equiv +1$, in contrast to the measured variable $f_i = 0, \pm 1, \pm 2$. For any classical random variable F, the mean value lies within the interval $[f_{\min}, f_{\max}]$, i.e., $f_{\min} \leqslant \langle F\rangle_{\text{clas}} \leqslant f_{\max}$. This leads to the Bell inequality in the form suggested by Clauser and Horne [40]: $\left|\langle F\rangle_{\text{clas}}\right| \leqslant 1$. (Indeed, the modulus of a sum cannot exceed the sum of the moduli, and hence, it follows from $f = \pm 1, p \geqslant 0$ and $\sum p = 1$ that $\left|\langle F\rangle_{\text{clas}}\right| \leqslant \sum p|f| = \sum p = 1$.)

The quantum value $\left|\langle F\rangle_{\text{quant}}\right| \equiv \left|\langle\psi|F|\psi\rangle\right| = \sqrt{2}$ in the state (5.5.1) exceeds the classical limit (unity) by 41%. This result can be formulated in a more general form: for some quantum models, certain combinations of moments $\langle F\rangle_{\text{quant}} = \langle\psi|F|\psi\rangle$ violate inequalities of the form [see Eqn (3.3.4)]

$$f_{\min} \leqslant \langle F\rangle_{\text{quant}} \leqslant f_{\max}, \qquad (5.7.6)$$

where $[f_{\min}, f_{\max}]$ is the interval of values taken by the corresponding classical variable. The Bell inequality

$$\left|\langle F\rangle_{\text{clas}}\right| = \frac{1}{2}\left|\langle AB\rangle + \langle A'B\rangle + \langle AB'\rangle - \langle A'B'\rangle\right| \leqslant 1$$

is one of the classical restrictions imposed on the moments by the requirement that the elementary probabilities corresponding to the chosen set of moments should be non-negative [for comparison, see Eqns (3.4.4), (5.6.4)]. For the conditions (5.7.5) to be violated, the operator function F can be chosen in several different forms, apart from Eqn (5.7.2) [38–40]. In all versions, violation of Bell inequalities demonstrates that the quantum description is incompatible with the classical one. This paradox is sometimes called Bell's theorem. In the next section, another example of restrictions imposed on the 'true' moments in the quantum model is considered.

How can one solve Bell's paradox? Let us consider the following three possibilities.

(1) One rejects the assertion that non-commuting operators have *a priori* values. In this case, the elementary probabilities $p(a,b,a',b')$ and the mean value $\sum pf$ have no physical sense. At the same time, there exist marginals of the form $p(a,b)$, which indicates that the quantum model is non-Kolmogorovian.

(2) One can admit the existence of negative probabilities. This removes the restriction $f_{\min} \leqslant \langle F\rangle_{\text{clas}} \leqslant f_{\max}$. The probabilities $p(a,b,a',b')$ can be formally expressed via the set of quantum moments in the state (5.5.1) by means of classical algorithms like (3.4.3) [38, 56]. Some of the probabilities obtained this way indeed turn out to be negative, for instance, $p(+1,+1,-1,-1) = -2^{-7/2}$. However, negative probabilities have no operational sense.

(3) One can assume that the orientation of the prism at point A influences the photocounts at detector B, and vice versa. It is often supposed, in accordance with a popular approach to EPR correlations (see Sections 4.8 and 5.5), that detecting a photon at point A causes the reduction of the WF, so that the resulting state depends on the position of the prism. The reduction is supposed to be equivalent to changing the properties of the photon B. This *nonlocality* assumption doubles the number of arguments of F. Now, F should be written in the form $F \equiv (1/2)(AB + A'B' + A''B'' - A'''B''')$. In this case, $f = 0, \pm 1, \pm 2$, so that $-2 \leqslant \langle F\rangle_{\text{clas}} \leqslant +2$, and the quantum mean value $\langle F\rangle_{\text{quant}} = -\sqrt{2}$ lies within the classical interval $[-2, +2]$.

The 'minimal', i.e., the least speculative interpretation may be the first possibility (see above). From this viewpoint, violation of Bell inequalities indicates that both the notion of elementary probabilities $p(s_{\text{A1}}, s_{\text{B1}}, s_{\text{A2}}, s_{\text{B2}})$ and the notion of *a priori* values for the Stokes operators in four modes A_x, A_y, B_x, B_y have no physical sense. In contrast to the two-mode case (Section 5.6), here the violation of the Bell inequality can be demonstrated experimentally.

5.8 Greenberger – Horne – Zeilinger paradox for the Stokes parameters

Let us add a third channel to the scheme in Fig. 4. The setup demonstrates the well-known Greenberger – Horne – Zeilinger (GHZ) paradox [73] (see also Refs [38, 74]). In each trial, the source simultaneously sends a photon to each one of the three remote detectors A, B, and C. The detectors measure either S_1 ($\chi = 0$) or S_2 ($\chi = 45°$). Let $A \equiv S_{\text{A1}}$, $A' \equiv S_{\text{A2}}$, and similarly for the channels B, C. Each trial results in the registration of three values. For instance, for measuring $(A, B', C') \equiv (S_{\text{A1}}, S_{\text{B2}}, S_{\text{C2}})$, i.e., for $\chi_{\text{A}} = 0$, $\chi_{\text{B}} = \chi_{\text{C}} = 45°$, the result of the trial may be $(a, b', c') = (++-)$, so that the product $ab'c'$ is equal to -1. Note that the operators S_{Z1} and S_{Z2} do not commute and are measured in different trials.

Let the field be prepared in a three-photon state [for comparison, see Eqn (5.5.1)],

$$|\psi\rangle = \frac{|+\rangle_{\text{A}}|+\rangle_{\text{B}}|+\rangle_{\text{C}} + |-\rangle_{\text{A}}|-\rangle_{\text{B}}|-\rangle_{\text{C}}}{\sqrt{2}}, \qquad (5.8.1)$$

then the quantum theory (see Appendix I) predicts the following correlation:

$$ab'c' = -1, \quad a'bc' = -1, \quad a'b'c = -1, \quad abc = +1. \qquad (5.8.2)$$

At the same time, all the first moments $\langle S_k\rangle$ are equal to zero, i.e., separate photons are not polarized. Equations (5.8.2) describe full correlation between the triples of measured values. This means, for instance, that the observable $AB'C' \equiv S_{1\text{A}}S_{2\text{B}}S_{2\text{C}}$ does not fluctuate, i.e., in each trial, the product of three numbers $ab'c'$ is always equal to -1. Repeated measurements with the same positions of the prisms always give even numbers of 'pluses' (even numbers of green flashes), i.e., the following triples occur with equal probabilities:

$$(a, b', c') = (++-), (+-+), (-++), (---).$$

The same is observed for trials where one measures A', B, C' and A', B', C. On the other hand, when measuring ABC (for all three detectors measuring S_1), one always obtains an odd number of 'pluses',

$$(a', b', c') = (--+), (-+-), (+--), (+++).$$

Let us try to describe this experiment, which has not been carried out as yet, from the viewpoint of 'common sense', i.e.,

in the framework of the classical probability model, with the properties of the signals existing *a priori*. According to this model, there exist six random dichotomous variables A, A', B, B', C, C', which take the values $a, a', b, b', c, c' = \pm 1$. The symbol $\langle \ldots \rangle$ denotes classical averaging with respect to some six-dimensional probability distribution $p(a, b, c, a', b', c')$. In each trial, the source sends full information, i.e., a set of six numbers ± 1. All six variables a, b, c, a', b', c' have some definite values $+1$ or -1 no matter whether they are observed or not. At *each* trial, these six numbers should satisfy the quantum prediction of full correlation between the observed triples. (Indeed, one can choose the position of the prism while the signal is on its way to the detector.)

Thus, according to the theory of hidden variables, the information sent by the light source should satisfy four requirements (5.8.2) imposed by the quantum model, though each one of the numbers $a', b, c, \ldots$ takes values with equal probabilities. One can easily see that Eqns (5.8.2) are inconsistent. Let, for instance, the signal be $(a, b, c, a', b', c') = (- - - + + +)$, then the first three equalities in Eqns (5.8.2) are satisfied while the last is not. Moreover, no combination of the six signs can satisfy all four observed correlations (5.8.2). This fact becomes clear if one multiplies the right-hand parts and the left-hand parts of all the equalities in Eqns (5.8.2). The product of the left-hand parts contains all multipliers twice:

$$a'bcab'cabc'a'b'c' = (abca'b'c')^2 = +1\,. \qquad (5.8.3)$$

At the same time, the product of the right-hand parts gives $(-1)^3(+1) = -1$. To solve this paradox $+1 = -1$, it is sufficient to assume that the Stokes parameters S_{Z1} and S_{Z2} have no *a priori* values, no matter whether they are observed or not, and to take into account that all four equalities (5.8.2) are tested in different trials (with different positions of the polarizing prisms).

As a result, the distribution $p(a, b, c, a', b', c')$ also has no sense. In the classical model, $2^6 = 64$ numbers $p(a, b, c, a', b', c')$ form the set of elementary probabilities; all lower-dimensional marginal probabilities $p(a)$, $p(a, b'), \ldots$, are obtained from the elementary probabilities by means of summation. Hence, if we assume that $p(a, b, c, a', b', c')$ do not exist, then all marginal probabilities are undefined. However, all one-dimensional probabilities like $p(a)$ do have physical sense since they can be measured directly. Hence, the above-considered experiment cannot be described in terms of the Kolmogorov model: one can measure marginal probabilities and the corresponding moments while the initial six-dimensional elementary probability distribution has no sense. In other words, the quantum moments $\langle S_{1A} S_{1B} S_{1C} \rangle$, $\langle S_{1A} S_{2B} S_{2C} \rangle, \ldots$ are not 'true' moments of some non-negative distribution; the problem of moments has no solution, and the model is 'non-Kolmogorovian'. A formal calculation of $p(a, b, c, a', b', c')$ via the quantum moments using the standard classical algorithm is ambiguous and leads to negative probabilities, which have no operational sense (see Section 5.6 and Refs [38, 56]).

Note that quite a different reasoning is commonly used for solving paradoxes of this kind. It is believed that such paradoxes prove the existence of 'quantum nonlocality': for instance, one assumes that the position of the prism at point A influences the values measured by the detectors B and C (see Section 5.7).

Two interesting features distinguish the GHZ paradox from Bell's paradox. First, here one observes a full correlation between the observables, and that is why no angular brackets are used in Eqns (5.8.2). Second, instead of the violation of a classical inequality, here we obtain violation of a classical equality.

Formally, the GHZ paradox can also be defined as a violation of the classical restriction for the moments (5.7.6), $f_{\min} \leqslant \langle F \rangle_{\text{clas}} \leqslant f_{\max}$

$$F \equiv F_1 F_2 F_3 F_4 \equiv AB'C' \cdot A'BC' \cdot A'B'C \cdot ABC\,. \quad (5.8.4)$$

The classical model is based on commutative algebra, so that $F = (ABCA'B'C')^2 = 1$, i.e., the classical variable F takes the only value $f_{\min} = f_{\max} \equiv f_0 = 1$ and $\langle F \rangle_{\text{clas}} = F = 1$. On the other hand, according to Eqn (I.3), the mean value for the corresponding quantum operator F in the state (5.8.1) is $\langle F \rangle_{\text{quant}} \equiv \langle \psi | F | \psi \rangle = -1 \neq f_0$. (For more detail, see Ref. [38].)

5.9 'Teleportation' of photon polarization

A surprising possibility of copying the quantum state of an individual system and passing it to another system, isomorphic to the first, has been discovered by Bennett et al. [75]. In contrast to the reversible exchange of q-bits (5.1.1), here the initial system influences the final one by means of a classical control channel. In fact, the 'quantum teleportation' suggested in Ref. [75] is a method of *preparing* an individual quantum system in a given state. It is essential that the information about the state to be prepared exists in the *quantum* form, i.e., it is encoded in the state of another system. This means that the copying does not reveal all the information about the system; otherwise, some part of the information would be lost. (Recall that one cannot measure the polarization of a photon.) Therefore, only some part of the information is transformed into the classical data consisting of observable macroscopic events. The idea suggested in Ref. [75] was further developed in Refs [76–81]. The first experiments in this direction are described in Refs [80, 81].

To explain the effect of polarization copying, let us consider a simplified and idealized scheme of the experiment performed in Ref. [80] (Fig. 5). In three quasi-monochromatic light beams A, B, and C that are fed to the input of the optical system, photons (denoted by circles) appear simultaneously, in 'triples'. The photons to be copied (A) are fully polarized. The photons B and C are depolarized but there is an ideal correlation between their polarizations (Section 5.5). The beams A and B are mixed by the nonpolarizing beamsplitter BS with transmission 50%, therefore all three beams become polarization correlated. This correlation is analysed by means of two polarization transformers T_A, T_C, the polarizing prism P_C in the C beam, three photon counters D_Z, and the coincidence circuit CC.

In fact, this scheme is a modification of the polarization intensity interferometer operating in the photon counting regime. (For comparison, see the description of two-photon interferometry, Section 5.5.) In the experiment [80], the rate of triple coincidences N was studied as a function of T_A and T_C, and the time delay τ in one of the three channels. (The interference visibility decreases with the increase of τ, as usual.)

The three-photon interference observed for the scheme in Fig. 5 manifests a remarkable feature: the rate of triple coincidences N depends similarly on the orientations of T_A and T_C, as if both transformers were placed one after another

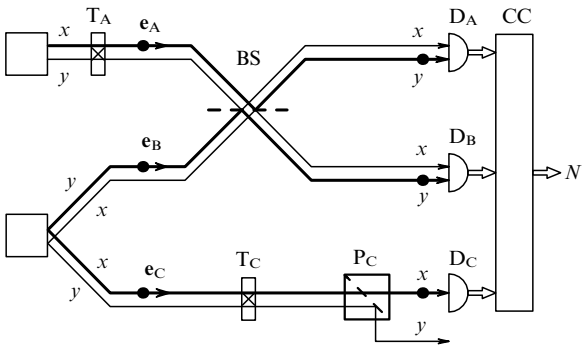

Figure 5. Simplified and idealized scheme of experiment [80] and the explicit model of the polarization copying $\mathbf{e}_A \to \mathbf{e}_C$. At the input, each of the beams A, B, C contains a single photon (circles). Bold lines show the trajectories of the photons. Photon A has an arbitrary polarization $\mathbf{e}_A$. The base vectors $\mathbf{e}_x$ and $\mathbf{e}_y$ are chosen so that $\mathbf{e}_x = \mathbf{e}_A$. Simultaneous detection of photons by the detectors D_A and D_B means that photons A and B did not interfere at the beamsplitter since they had orthogonal polarizations, $\mathbf{e}_A \perp \mathbf{e}_B$. Photons B and C are prepared in states with orthogonal polarizations, $\mathbf{e}_B \perp \mathbf{e}_C$, therefore, $\mathbf{e}_C \perp \mathbf{e}_A$. Here, T_A, T_C are polarization transformers, P_C is a polarizing prism, D_Z are detectors, CC is a triple coincidence circuit, N is the number of triple coincidences during some time interval, and x and y denote polarization.

in one of the beams or as if the polarization of photon A at the output of T_A were transferred to photon C at the input of T_C: $\mathbf{e}_A \to \mathbf{e}_C$. In other words, the phase and visibility of the interference pattern observed in triple coincidences is determined by the product of the Jones matrices $\mathbf{T}_C\mathbf{T}_A$. If T_C corresponds to the transformation inverse to T_A, $(\mathbf{T}_C\mathbf{T}_A = 1)$, then N depends neither on T_A nor on T_C, and the interference visibility is equal to zero.

One can consider the detectors D_{Cx}, D_{Cy} and the transformer T_C as a device for measuring the polarization of C-photons (see Fig. 3), but the counting rate of the 'singles' in D_{Cj} is independent of the polarization since the C-photons are not polarized. However, 'conditional' photocounts in D_{Cj} (photocounts simultaneous with those in D_A, D_B) manifest full polarization. For instance, by means of the dependence of N_3 on T_C, one can measure the polarization vector for A-photons $\mathbf{e}_A = (\alpha_x, \alpha_y)$, or the Stokes vector $\langle \mathbf{S}_A \rangle$, which is almost the same. From time to time, both photons are registered by one and the same detector D_A or D_B. In such cases, the polarization is not copied (see Appendix II).

In order to turn the interferometer into a device preparing photons with the polarization copied from the initial photons, an optical gate (modulator) M removing 'spare' C-photons should be added to the setup. The gate can be controlled by pulses from the detectors D_{Aj} and D_{Bj}; in fact, it can replace the triple coincidence circuit since it 'blocks' C-photons in 'bad' cases where two photons are fed to a single detector (and also in cases where the nonideal detectors D_{Aj} and D_{Bj} 'miss' photons). One can also use an additional polarization transformer T'_C that would 'improve' the C-photon polarization in certain cases and thus increase the proportion of 'good' events from 25% to 50%. As a result, all photons passing the gate and the transformer T'_C (one half of all photons) have polarizations coinciding with the initial polarization of the A-photons: $\mathbf{e}_C = \mathbf{e}_A$. One can say that the device prepares single photons with unknown polarization repeating the polarization of single A-photons.

There exists a more primitive analogous device preparing photons with a given (but not copied) polarization using a two-photon source and a gate. In this device, a detector registering a photon opens the gate for the second photon [82–84]. Note that the action of the amplitude modulator M is described by a nonunitary transformation, in contrast to T_k and BS, so that the photons passing through M should be considered to be in a mixed state. The 'teleportation' $\mathbf{e}_A \to \mathbf{e}_C$ occurs at best in 50% of all trials. The imperfection of the detectors and other elements makes this proportion still lower.

For the effect to be observed, the input field should be prepared in a partially factored three-photon state $|\psi\rangle = |\psi\rangle_A|\psi\rangle_{BC}$. Photon A should be in a pure state,

$$|\psi\rangle_A = \alpha|Ax\rangle + \beta|Ay\rangle\,, \tag{5.9.1}$$

so that a certain polarization vector $\mathbf{e}_A \equiv (\alpha, \beta)_A$ and the Stokes vector $\langle \mathbf{S} \rangle_A$ could be ascribed to it. These vectors can be varied by means of the polarization transformer T_A. Here $|Ax\rangle \equiv |1\rangle_{Ax}|0\rangle_{Ay}$ is a state with one photon per mode Ax. Photons B and C should be prepared in an entangled (nonfactored) state with full polarization correlation [see Eqn (5.5.2)],

$$|\psi\rangle_{BC} = \frac{|Bx, Cy\rangle - |By, Cx\rangle}{\sqrt{2}}\,. \tag{5.9.2}$$

(This was done in an experiment using spontaneous parametric scattering [85–87].) Each one of the photons B and C has no *a priori* polarization, and they should be described in terms of mixed states. Such two-photon states are called *Bell* or *EPR states* [76]. Modes A and B should have equal central frequencies, while mode C can have any frequency. All three photons should be correlated in time at the point where they reach the beamsplitter. Due to the initial correlation between B- and C-photons and the action of the beamsplitter, all three photons become polarization-correlated. The information about the initial state of the A-photon, i.e., its polarization vector $\mathbf{e}_A = (\alpha, \beta)$, is encoded in the triple correlation at the output of the scheme.

In each trial, ideal detectors register the numbers of photons n_k equal to 0, 1, or 2. The total number of detected photons is equal to 3, one of these photons being detected by detector C. Any repeated trial results in one of 16 elementary events with three photons randomly distributed among six detectors. The probabilities of these events are obtained by a standard calculation (see Appendix II). It is convenient to describe the observed correlations in terms of the Schrödinger representation and the effective WF for the C beam proportional to the projection of the output WF $|\psi'\rangle$ on the vector $\langle Ax, Ay|$,

$$|\psi\rangle_{C\,\mathrm{eff}} \equiv \sqrt{8}\,\langle Ay, Bx|\psi'\rangle = \alpha|Cx\rangle + \beta|Cy\rangle\,. \tag{5.9.3}$$

Here $|\psi'\rangle$ is the WF of the whole six-mode field accounting for the beamsplitter [see Eqn (II.10)]. (For simplicity, we do not take into account T_B.) Note that in a consistent theory, the C beam should be described by a mixed state and no WF can be associated with it (Section 4.10). Projecting $|\psi\rangle_{C\,\mathrm{eff}}$ onto the vector $\langle Cx|$, we obtain

$$p(Ay, Bx, Cx) = \frac{\left|\langle Cx|\psi\rangle_{C\,\mathrm{eff}}\right|^2}{8} = \frac{|\alpha|^2}{8}\,.$$

Comparison between Eqns (5.9.3) and (5.9.1) shows that in the chosen subset of events, the field in the C beam has the same polarization properties as the initial A-photon. The state of the A-photon (its polarization vector) seems to be transferred from the A beam to the C beam, $\mathbf{e}_A \to \mathbf{e}_C$. The action of the transformer T_C on the chosen subset of events can be described in terms of the effective function (5.9.3) assuming that T_C transforms the vectors $|Cj\rangle$ [see Eqn (II.9)].

The same 'teleportation' takes place for another 1/8 of the trials, where photons A and B are registered by the detectors Ax and By [see Eqn (II.8a)]. The total share of 'good' events can be increased to 1/2 if a controlled polarization transformer T'_C is inserted into the beam C [75]. This element should provide the transformation $\beta \to -\beta$ [described by the Pauli matrix σ_x, see Eqn (5.4.5)] if there are 'clicks' in the detectors D_{Ax} and D_{Ay} or D_{Bx} and D_{By}. Now only for one half of all events, where two photons get into a single output mode Ax, Ay, Bx, or By, there is no 'polarization copying'.

This experiment can be also described in the Heisenberg representation: $p'_{klm} = \mu'_{klm} = \langle\psi|N'_k N'_l N'_m|\psi\rangle$ (see Appendix II). The optical scheme including the beamsplitter and the polarization transformers is described by a *transformation matrix* (the spectral Green's function), $a_k \to a'_k = \sum G_{km} a_m$. The optical scheme transforms the input moments, which can be written as $\mu \to \mu' = G^6 \mu$. The matrix G_{km} coincides with the corresponding matrix in the classical theory; therefore, transformation of the field statistics $\mu \to \mu'$ is described in the quantum theory the same way as in the classical theory. The only difference between the quantum and classical descriptions is contained in the relative values of the input moments μ. This leads to the limited visibility of two-beam intensity interference in the classical case, $V < 50\%$ (see Section 5.5 and Ref. [37]).

The original paper by Bennett et al. [75] was entitled "Teleportation of an unknown quantum state via dual classic and EPR channels". In the formalism used in Ref. [75], Bell states similar to Eqn (5.9.2) were chosen as base vectors. Such an approach ignores the cases where two photons get into a single detector; it is not taken into account that a nonunitary operation is necessary for excluding these cases. It is stressed in the paper that the initial state is not measured but the information about the state is split into two parts, the quantum one and the classical one. In the optical case (Fig. 5), the classical part includes the signal that controls the transformer T'_C in case of certain events; the quantum part includes the photon C. The information sender is traditionally called Alice. In addition to the input A photons, Alice has a source of EPR-correlated photons B and C, the beamsplitter BS and the detectors at the output of the beamsplitter. To Bob, who operates the transformer T'_C, Alice sends the photon C (using the quantum channel) and classical signals with commands to set $T'_C = 1$ or $T'_C = \sigma_x$. Photon C, which gets the state of photon A as a result of reduction and the action of T'_C, is sent further by Bob.

If it were not for the 'bad' events where two photons are fed to the same detector, the limiting efficiency of the scheme would be 100%. (This situation is specific for the optical case; for fermions, such 'bad events' are forbidden by the Pauli principle.) The idea of 'teleporting' the state of a two-level system can be generalized to the case of more complicated systems [60, 75–79]. One can 'move' information from the field to atoms and back, see Eqns (5.1.1), (5.1.2). One can expect that these possibilities will find applications in quantum computing and quantum cryptography.

There are two possible interpretations for the effect considered here.

(1) The effect of three-photon interference and teleportation is usually considered in terms of WF partial reduction: simultaneous 'clicks' in two detectors D_A and D_B cause the reduction of the three-photon state into a one-photon one, $|\psi'\rangle_{ABC} \to |\psi\rangle_C$. The field or the detectors in the C beam instantaneously know this due to some superluminal interaction [80]; therefore, the effect is believed to be evidence for 'quantum nonlocality'. It is supposed that the mathematical procedure of projecting $|\psi'\rangle$ onto $\langle Ay, Ax|$ in Eqn (5.9.3) following from the Born postulate corresponds to some real event caused by detecting photons at points D_{Ay} and D_{Bx}. The necessity for a gate or a coincidence circuit is ignored.

However, if a coincidence circuit is used, the detectors in the beams A, B, and C are equivalent: one can assume that the reduction occurs first in the detectors D_{Cj} (in the absence of the modulators M and T'_C). At present, the hypothesis of instantaneous reduction is confirmed neither theoretically nor experimentally. It is hardly consistent with the special relativity theory, since the detectors can be placed at any distance. In Appendix II, it is shown that the copying effect is fully described in terms of a standard formalism with Born's correlation functions (4.7.1b). In other words, the idea that the effect is caused by instantaneous reduction is redundant; it is nothing more than a possible interpretation.

(2) In the framework of the 'minimal' interpretation, the effect can be considered as a manifestation of the specific correlation between the three light beams. Certainly, all observable events do not violate the special relativity theory. One can assume the information $(\alpha, \beta)_A$ to be carried from the input of the optical scheme to its output either by the WF of the field ($\psi \to \psi'$, the Schrödinger representation) or by the field operators ($a \to a'$, the Heisenberg representation), similarly to the case of a single polarized photon. If the modulators T'_C and M are not used, then the time sequence for the three detectors is not essential, since the detectors are separated by spacelike intervals (see Section 4.8). There is no reason for selecting two stages in a triple photocount occurring in the three detectors (first, two photocounts cause the WF reduction, and then the new WF influences the third detector). Neither the consistent theory nor the experiment confirms this interpretation. Instead of introducing *ad hoc* 'nonlocality', it is more consistent to assume that the quantum theory is non-Kolmogorovian and to neglect the existence of *a priori* values (see Sections 4.5, 5.6–5.8). If modulators are used, one can speak of the preparation of a photon state in the C beam.

In its simplest version, the effect of polarization copying, $\mathbf{e}_A \to \mathbf{e}_C$, has a clear (but not strict) explanation (see Fig. 5). It follows from two well-known effects: the intensity anticorrelation in the output beams of the beamsplitter, A′ and B′, and the polarization anticorrelation for the photons in the initial beams, B and C (Section 5.5). Assume for simplicity that all three photons A, B, and C have some *a priori* polarizations $\mathbf{e}_Z$ and the corresponding Stokes vectors $\mathbf{S}_Z$. These polarizations vary randomly from trial to trial. Property (5.5.2) can be understood as an anticorrelation between the directions of the Stokes vectors, $\mathbf{S}_A = -\mathbf{S}_B$, or orthogonality of the polarization vectors, $\mathbf{e}_B \perp \mathbf{e}_C$ (see Section 5.5). Let us choose a basis with $\mathbf{e}_x = \mathbf{e}_A$ and assume that according to Eqn (5.9.2), there exist only two events with equal probabilities: either $\mathbf{e}_B = \mathbf{e}_x$ and $\mathbf{e}_C = \mathbf{e}_y$ or $\mathbf{e}_B = \mathbf{e}_y$ and $\mathbf{e}_C = \mathbf{e}_x$. In the first case, where $\mathbf{e}_B = \mathbf{e}_A$, one should observe

an anticorrelation between the photocounts in D_A and D_B, since both photons should get into the same detector, either into D_A or into D_B. (This is the so-called effect of two-photon interference or intensity anticorrelation at the output of a beamsplitter, see Ref. [37].) Hence, if 'clicks' are observed in both detectors D_A and D_B, then the second case takes place, i.e., $\mathbf{e}_C = \mathbf{e}_x = \mathbf{e}_A$ (see the bold lines in Fig. 5).

Using this explanation, one can suggest a similar classical method of copying polarization from one light beam to another without measuring this polarization. Consider three ideal lasers A, B, C generating polarized light beams, so that beams A and B have equal intensities I_0 and frequencies ω_0. The beams B and C are transmitted through polarization transformers controlled by a common random number generator in such a way that the orthogonality condition $\mathbf{e}_B(t) \perp \mathbf{e}_C(t)$ is always satisfied (see Section 5.5). As a result, all polarizations of the beams B and C have equal probabilities but their Stokes vectors have opposite directions, $\mathbf{S}_B = -\mathbf{S}_C$.

Further, let the beams A and B be mixed at a beamsplitter and the intensities $I'_A(t)$ and $I'_B(t)$ in the output beams A′ and B′ be measured by two analogue detectors. Because of the fluctuations of the vector $\mathbf{e}_B(t)$, there are also fluctuations in the intensities $I'_A(t)$ and $I'_B(t)$. These fluctuations are anticorrelated, since the total intensity is preserved, $I'_A(t) + I'_B(t) = 2I_0$. At the moments when $I_A(t)$ is equal to $I_B(t)$, to a given accuracy $\Delta I/I_0$, there is no interference of the beams at the input of the beamsplitter, and this means that their polarizations are orthogonal, $\mathbf{e}_A \perp \mathbf{e}_B$. In this case, $\mathbf{e}_B \perp \mathbf{e}_C$ also, so that $\mathbf{e}_A = \mathbf{e}_C$. At such time moments, the gate blocking the C beam is automatically opened. As a result, we obtain light pulses with random intervals and random durations but with the frequency ω_C and polarization $\mathbf{e}_A$.

A principal drawback of this model distinguishing it from the quantum one is the limited accuracy of copying, which is inversely proportional to ΔI and to the relative time of gate opening (the efficiency). In the quantum case, an ideal setup provides exact copying.

6. A particle in one dimension

In Section 5.2, we considered a trapped atom and showed how its *internal* degrees of freedom can be prepared in a given state by means of cooling and a resonant laser pulse. For an atom cooled in a trap, its *external* (kinetic) degrees of freedom are also prepared in a definite (ground) stationary state $\psi_0(\mathbf{r} - \mathbf{r}_0)$, with the shape and the length of the packet a_0 being determined by the trap potential $V(\mathbf{r} - \mathbf{r}_0)$. Here $\mathbf{r}_0$ is the classical coordinate of the trap center. Switching off the trapping potential at the moment t_0, one prepares the free particle in the state $\psi(\mathbf{r}, t_0) = \psi_0(\mathbf{r} - \mathbf{r}_0)$ with some localized form of the packet, with known moment of preparation t_0, and localization domain $\mathbf{r}_0 \pm a_0$. The state is no more stationary and the packet starts to 'diffuse'. The mean energy E and the momentum of the particle are equal to zero but they can be increased using classical fields (gravitational, electric, or optical fields). We stress once more that in this process, no *measurement* is performed on the quantum system: the particle (or its WF) is influenced but its back action on the measurement devices is not observed. An experimenter measures (via comparison with references) only numerical values of the classical parameters $\mathbf{r}_0$, t_0, $V(\mathbf{r} - \mathbf{r}_0)$ for the preparation device.

As another example, one can consider a short field pulse with $E = 1$ keV applied to a metal. This pulse causes cold emission of electrons with relatively well-defined energies. Additional filtering in space and velocity allows one to prepare free particles in sufficiently well-defined (but mixed) states.

6.1 Coordinate or momentum measurement

How does one actually observe signals from the microworld in a laboratory? For detecting single particles, one uses scintillators, photosensitive films, Wilson chambers, Geiger counters, ionization detectors of atoms, photomultipliers (PMTs) and similar devices. Probably, a common feature of all these devices is the transfer of an energy quantum from the particle to the atoms of the detector and further amplification, an 'explosive' process leading to a macroscopic event [4], which can be the appearance of a droplet in a super-cooled vapor due to the thermodynamic instability or the appearance of an electron avalanche in a PMT due to the accelerating field. It seems reasonable to place the border between the quantum and classical parts of the setup (Section 4.1) after some 'seeding' atom in the detector and to consider the energy of this atom as the 'readout observable'. In this approach, the macroscopic event registered in the experiment is supposed to be caused by the excitation or ionization of one of the atoms of the detector. The well-known Glauber model for the optical photons detection [20], which is successfully used in quantum optics, is also based on this scheme.

Some devices detect only the space coordinate $\mathbf{r}_1$ or a sequence of coordinates (a track) for macroscopic objects, such as droplets, silver particles, and so on. Devices with fine time resolution generate short electric pulses and this way fix the moment t_1 of detecting a particle. Thus, one can state that only some events $(\mathbf{r}_1, t_1)$ in space–time are actually registered. These events are measured using macroscopic rulers and clocks and assumed to be the coordinates of the particle under study. (Of course, any measurement of continuous parameters has restricted 'laboratory' accuracy, which should be distinguished from the principal quantum uncertainty.) Further, using these coordinates, one determines (indirectly, from theoretical considerations) other parameters of the particle, such as the energy, momentum, spin, etc. (see the scheme in Fig. 1).

Following Glauber, let us consider the model of detecting a charged particle, for instance, an electron. Let a massive detecting atom be placed at a fixed point x_1. Due to the large mass of the detecting atom, the coordinate of its center of mass x_1 can be considered as a c-number. Suppose that the particle is prepared in some pure state $\langle x|\psi_0\rangle = \psi(x, t_0)$, where t_0 is the time moment of preparation. The detecting atom is in the ground state $|g\rangle$, and the state of the whole system (particle + atom) has a factored form, $|\Psi_0\rangle = |\psi_0\rangle|g\rangle$. In each trial, time is measured with respect to a new moment $t_0 \equiv 0$. As the 'readout observable' (see Section 4.1), we choose the operator of projection on the excited state of the detecting atom, $P_1(e) \equiv |e\rangle\langle e|$. According to the Born postulate (4.7.1), this Heisenberg operator averaged with respect to the initial WF gives the probability of the event 'the atom at point x_1 at time t_1 is in the excited state':

$$p_1(e, t_1) \equiv \langle \Psi_0 | P_1(e, t_1) | \Psi_0 \rangle . \qquad (6.1.1)$$

This function can easily be calculated in the first order of the perturbation theory with respect to the interaction energy

for the particle and the atom, V_1 (for more detail, see Ref. [58]). Further, suppose there are many levels (or many atoms) with different transition frequencies ω_e. We integrate the probability $p_1(e, t_1)$ over the transition frequency ω_e under the assumption that the frequency band $\Delta\omega_e$ is infinitely broad. This classical procedure of probability summation describes a broad-band detector with infinitely fast response and indirectly takes into account relaxation processes.

It is convenient to introduce the differential probability of the detector excitation per unit time (the transition rate), $w \equiv \mathrm{d}p/\mathrm{d}t$. The excitation rate for a detector at point x_1 passed by a particle takes the form

$$w(x_1, t_1) = \eta_1 \langle \psi_0 | V_1^2(X, t_1) | \psi_0 \rangle = \eta_1 \int \mathrm{d}x \, |\psi(x, t_1) V_1(x)|^2 . \quad (6.1.2)$$

Here η_1 is the detector efficiency, $V_1(x)$ is the potential of the interaction between the particle and the atom. This potential has either a maximum or a minimum at point $x = x_1$ and plays the role of the 'instrumental function' determining the inaccuracy of the measurement $\Delta x = a_1$, where a_1 is the width of $V_1(x)$.

Note that Eqn (6.1.2) could be obtained at once from the formula $w = \langle \psi_0 | P_1 | \psi_0 \rangle$ [see Eqn (6.1.1)] by choosing the integral

$$P_1 \equiv \sqrt{\eta_1} \int \mathrm{d}x \, V_1(x) |x\rangle\langle x| ,$$

i.e., a weighted sum of elementary projectors $|x\rangle\langle x|$, as the 'readout observable'. However, this choice has to be verified, which is done by means of the present calculation.

For $a_1 \ll a_0$, where a_0 is the width of the initial packet $\psi(x, t_0)$, one can assume $V_1^2(x) = \delta(x - x_1)$, with all unessential constants included into the efficiency η_1, so that Eqn (6.1.2) takes the form

$$w(x_1, t_1) = \eta_1 |\psi(x_1, t_1)|^2 . \quad (6.1.3)$$

We see that the WF absolute value $|\psi|$ (the envelope of the wave packet) can be measured, i.e., it is an operationaly-defined parameter.

Expression (6.1.3) resembles the Born postulate stating the probability meaning of the WF. However, in our case, it follows from Eqn (6.1.1), and the arguments x_1, t_1 play the role of directly measurable classical parameters of the quantum theory. Let us stress that $t_1 \equiv t_1' - t_0$ is an argument of the distribution function, which is obtained by processing experimental data. It is not an arbitrary measurement time chosen by an experimenter, as is usually supposed in the quantum measurement theory. In the ith trial, the moment t_i of a pulse appearing at the output of the detector is unpredictable up to the duration of the particle wave packet, $\Delta t_0 = a_0/v_0$.

In a real experiment, the operator of the particle coordinate X cannot be measured directly, and the position of the particle is always identified with the classical coordinate x_1 of a massive fixed detector (a microcrystal in a photosensitive film, a water droplet in a super-cooled vapor, etc.), up to some uncertainty $\pm a_1$. When an excited atom is registered, it is natural to conclude that the passing particle has the coordinate $x_1 \pm a_1$ at the moment of the pulse $t_1 \pm \tau$. (Here τ is the time constant of the detector, which is assumed to be zero in our model.) This procedure relates the mathematical symbols $X(t)$ or $\psi(x, t)$ to our 'actual' space–time (x_1, t_1), which is measured by means of rulers and clocks.

Consider now the simplest model of measuring the distribution of the longitudinal momentum for a charged particle, $p = mv \equiv \hbar k$. Let a domain with constant magnetic field H_0 be placed before the detector. In this domain, the trajectory of the particle is bent: the particle moves along a circle with radius $r = cp/eH_0$. Measuring the transverse coordinate of the particle, one finds r and calculates p and $k = p/\hbar$. Repeated many times, this experiment allows one to measure the distribution $w(k)$, the mean value k_0, and the uncertainty Δk. Under the assertion that the measurement is accurate, one can assume the projector $P_k = |k\rangle\langle k|$ to be the readout observable. Then, from the Born postulate, one obtains $w(k) = \langle \psi_0 | P_k | \psi_0 \rangle = |\langle k | \psi_0 \rangle|^2 = |\psi(k, t_0)|^2$. Since free motion conserves momentum, the vectors $|k\rangle$ have only phase variations, and the moment of measurement is not essential.

6.2 Time-of-flight experiment

There exists another method of velocity measurement, with a high-energy particle passing by two fast detectors in sequence, for instance, two Geiger counters (Fig. 6a). Only trials where both detectors 'click' are taken into account (the coincidence method). As a result, one can measure the joint distribution for two events, $p_{12} \equiv p(x_1, t_1; x_2, t_2)$ (Fig. 6b). This is the scheme of the *time-of-flight* experiment, which is widely used for measuring velocities of particles. The distance between the detectors, $x_2 - x_1 \equiv L$, divided by the time delay between the two pulses, $t_2 - t_1 \equiv T$, gives the *a priori* group velocity of the particle wave packet $v_0 = L/T$. (The energy loss in the first detector is not taken into account.) Let the particle be prepared each time in a pure state with the momentum sufficiently well-defined, so that it is described by a long wave packet. As a result, one observes fluctuations in the detection moments t_1, t_2, with respect to some preparation moment $t_0 \equiv 0$, and in the time delay T. Repeating the procedure many times, one can measure the distribution p_{12}.

As it is shown in Section 4.9, standard algorithms of the quantum theory cannot be used for calculating p_{12}. The point is that the Heisenberg operators $P_1(t_1)$ and $P_2(t_2)$, which describe the responses from the detectors at time moments t_1 and t_2, do not commute, since the detectors interact with the particle. The only way to calculate the probability p_{12} in the case of a time-of-flight experiment seems to be to use the Wigner formula (4.9.1). In addition, one should assume the time moments t_n in Eqn (4.9.1) to be random. (In the quantum measurement theory, it is supposed that the moments of measurement are arbitrarily chosen by the experimenter and that the reduction takes place at these moments.)

The differential probability (transition rate) $w_{12} \equiv \partial^2 p_{12}/\partial t_1 \partial t_2$ can be calculated using the above-described model of a broad-band detector in the second order perturbation theory [58] [for comparison, see Eqn (6.1.2)],

$$w_{12} = \eta_1 \eta_2 \langle \psi_0 | V_1(X, t_1) V_2^2(X, t_2) V_1(X, t_1) | \psi_0 \rangle . \quad (6.2.1)$$

Here V_n are the potentials of the interaction between the particle and the detectors, and η_n are the detector efficiencies. Since only coincident counts are registered, the parameter characterizing the interaction of the particle with the detectors, $\eta_n V_n$, can be considered as small.

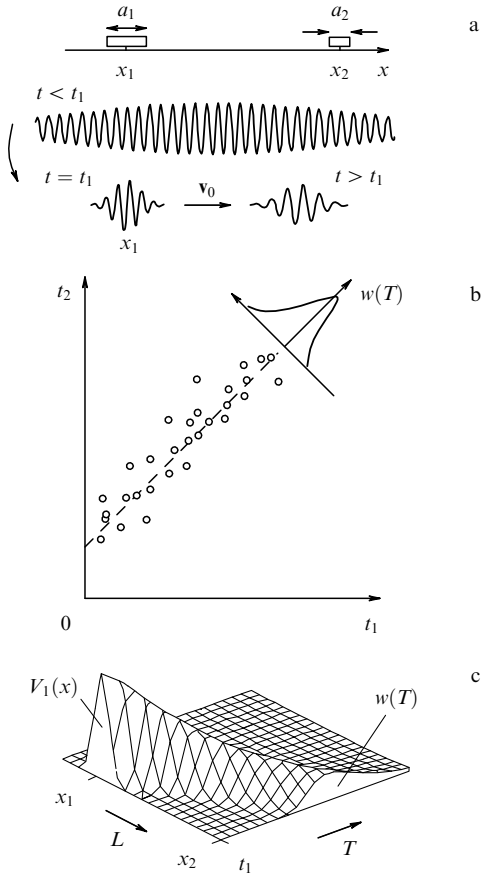

Figure 6. Time-of-flight experiment. (a) Schematic plot and the interpretation. A particle with definite momentum passes two detectors (rectangulars) with sizes a_1 and $a_2 \ll a_1$ placed at points x_1 and $x_2 = x_1 + L$. According to Eqn (6.2.2), the initial wavefunction of the particle, which has the shape of a long sinusoid, seems to collapse, at the detection moment t_1, into a short wave packet of length a_1. Further, this packet moves towards the second detector with group velocity $\mathbf{v}_0$ and gradually diffuses. (b) Experimental results. Each point with coordinates (t_{1i}, t_{2i}) denotes the detection moments of both detectors in the ith trial. The particle is prepared in a state with definite momentum; therefore, any time moment t_{1i} is possible. The second pointlike detector registers the particle at an arbitrary moment $t_{2i} > t_{1i}$. The dotted line shows a linear regression corresponding to the group velocity v_0. At the top right, the measured delay distribution $w(T)$ is shown, with $T \equiv t_2 - t_1$. (c) The observed delay distribution $w(T)$ is determined by the shape of the potential $V_1(x)$ of the first detector, in accordance with the propagation law for a free particle, as if the reduction of the WF happened at the moment t_1: $\exp(\mathrm{i}k_0 x) \to V_1(x)\exp(\mathrm{i}k_0 x)$.

Let the particle be prepared in a state with definite momentum mv_0 and the potentials $V_n(x)$ have Gaussian shapes with widths a_n such that $a_2 \ll a_1$. Then Eqn (6.2.1) takes the form

$$w_{12} \equiv w(T) = \frac{\eta_1\eta_2}{\sqrt{1 + (T/T_a)^2}} \exp\left\{-\frac{(L - v_0 T)^2}{2a_1^2\left[1 + (T/T_a)^2\right]}\right\}. \tag{6.2.2}$$

Here $T \equiv t_2 - t_1 > 0$, $L \equiv x_2 - x_1$, and $T_a \equiv ma_1^2/\hbar$ is the typical time of packet diffusion. The observed detection moments t_{1i}, t_{2i} have a uniform distribution over the time axis, but the delay between them has a distribution (6.2.2) with the maximum at $T = L/v_0$. Equation (6.2.2) describes the mapping of the Gaussian function $V_1(x)$ onto the observed delay distribution $w(T)$ (Fig. 6c).

This distribution can be understood in the following way. The initial broad wave packet is 'cut' at the moment t_1, so that its size becomes equal to the size a_1 of the first detector, $\exp(\mathrm{i}k_0 x) \to \psi_{\text{eff}}(x) = V_1(x)\exp(\mathrm{i}k_0 x)$ (Fig. 6a). This is the so-called *back action* of the detector on the particle as a result of their interaction [18, 19]. For this term, we obtained an operational definition. At $t > t_1$, the center of the effective wave packet $\psi_{\text{eff}}(x, t)$ moves with the group velocity v_0. In the near-field zone ($L \ll k_0 a_1^2 = v_0 T_a$), its envelope is constant, but in the far-field zone (the inverse inequality) it broadens proportionally to L/a_1. With the help of the second detector, one can measure the envelope $\left|\psi_{\text{eff}}(x)\right|$ [see Eqn (6.1.3)]. Of course, this is nothing but a convenient interpretation for the calculated result and not the 'actual picture' of the events.

Let us stress that here the 'reduction moment' t_1 is the argument of the distribution function but not one of the actual detection moments t_{1i}, which fluctuate from trial to trial (Fig. 6b). This fact seems to have principal importance. According to the traditional viewpoint, the reduction $\psi \to \psi'$ takes place in each trial at the moment t_{1i}, while Eqn (4.9.1) used when deriving Eqn (6.2.2) corresponds to the reduction at some moment t_1, which is not related to any physical event.

By means of the experiment described above, two effects can be observed directly: manifestation of the projection postulate and wave packet diffusion caused by the vacuum dispersion $\omega \sim k^2$.

6.3 The uncertainty relation and experiment

Let us consider the operational meaning of the uncertainty relation. It can manifest itself in two types of experiments. (Other possibilities are discussed in Ref. [88].)

Experiments of the first type contain two series of measurements. For instance, in the first series, one measures the coordinate of the particle, X. From the set obtained x_i, one finds the quantum uncertainty Δx, which is determined by the WF. In the second series, the momentum P is measured and its uncertainty Δp is calculated. (All measurements are assumed to be ideal.) Models for such measurements have been considered in Sections 6.1 and 6.2. As a result, the obtained uncertainties should obey the inequality $\Delta x \Delta p \geqslant \hbar/2$. This example illustrates how the uncertainty relation can be directly observed in experiment. The term *directly observed* admits a quantitative criterion: one can state that a direct observation allows the upper bound for the Planck constant to be chosen from the measured set of numbers. (This criterion is sometimes not satisfied by examples given in textbooks.)

In experiments of the second type, in *each* trial, one first measures X and then P. Apparently, a quantitative description of such experiments is only possible using the Wigner formula (4.9.1), as in the example considered above. The first detector discovers the particle at point $x_1 \pm a_1$, and then the second detector (or a set of detectors with various transverse coordinates) placed after the domain with the magnetic field measures its longitudinal momentum $P = \hbar k$. Let us register only coincident counts of both detectors. Observing a large number of such coincidences for identically prepared particles and different x_1, one can measure the distribution $w(x_1, k)$.

As previously, we assume the momentum measurement to be exact and describe it by the projector $P_2 = |k\rangle\langle k|$. The time moment of the second measurement is not essential, but it should be stressed that the coordinate and momentum are not measured simultaneously. (This is practically the only possible way to measure these two parameters.) For simpli-

city, let $t_1 = t_0$ and $\eta = 1$. Using Eqn (4.9.1), we find the coincidence probability as a function of the parameters x_1 and k of the devices,

$$w(x_1, k) = \langle\psi_0|P_1P_2P_1|\psi_0\rangle = (2\pi)^{-1}\left|\int \mathrm{d}x\, V_1(x)\psi_0(x)\exp(-\mathrm{i}kx)\right|^2 . \quad (6.3.1)$$

Again, this result can be understood as a manifestation of the back action. Indeed, the interaction of the particle with the first detector is taken into account by multiplying its initial WF by the interaction potential,

$$\psi_0(x) \to \psi_{\mathrm{eff}}(x, t_1) = V_1(x)\psi_0(x) .$$

Note that the inverse sequence of measurements would give quite a different result, namely, the initial momentum distribution,

$$\langle\psi_0|P_2P_1P_2|\psi_0\rangle \propto \left|\langle k|\psi_0\rangle\right|^2 = w(k, t_0) .$$

The second (coordinate) measurement plays no role here.

Let the particle be prepared in a state with definite momentum $\hbar k_0$, i.e., $|\psi_0\rangle = |k_0\rangle$ or

$$\psi_0(x) \equiv \langle x|\psi_0\rangle = (2\pi)^{-1/2}\exp(\mathrm{i}k_0x) .$$

Then Eqn (6.3.1) takes the form

$$w(x_1, k) \propto \left|\int \mathrm{d}x\, V_1(x)\exp\left[\mathrm{i}(k_0 - k)x\right]\right|^2 . \quad (6.3.2)$$

Now, the momentum distribution obtained in the correlation experiment is given by the Fourier transform of the coordinate detector 'instrumental function' $V_1(x)$; the width of this distribution is of order of $1/a_1$ and its uncertainty satisfies the Fourier relation $a_1\Delta k \geqslant 1/2$. This demonstrates an important operational feature of the uncertainty relation, the so-called *intervention of the device*, which limits the product of the accuracies for successively measured non-commuting variables. The first measurement, with accuracy $\Delta x = a_1$, limits the accuracy of the second measurement to the value $\Delta k_{\min} = 1/2\Delta x$. Note that here Δx and $\Delta k_{\min}$ are not related to the variances of the observables in the initial state ψ_0, as in the experiments of the first type.

6.4 Wigner's distribution

At first sight, the above-considered experiments with one or two detectors can be described trivially and explicitly in terms of classical subensembles of particles. Suppose that 'actually', the WF only describes the statistics of a classical ensemble of particles with some distribution of initial coordinates and velocities. The effective WF ψ_{eff} in the theory of time-of-flight experiments simply results from the *selection* of some particles by the first detector, the velocities of these particles being determined by its position x_1 and detection time t_1. This simple interpretation does not require the projection postulate and the mysterious reduction process. However, such reasoning leads to certain principal difficulties, even if the effects of particle interference are not taken into account.

From the classical viewpoint, a pointlike particle has only one state, which is determined by its coordinate x_t and momentum $p_t = mv_t$ at a given time moment. The space of events (*the phase space*) is a plane R^2 with coordinates x and p. The distribution function has the form

$$w(x, p, t) = \delta(x - x_t)\delta(p - p_t) .$$

The evolution of the state in time is given by the Hamilton equations

$$\dot{x} = \frac{\mathrm{d}H}{\mathrm{d}p} , \qquad \dot{p} = -\frac{\mathrm{d}H}{\mathrm{d}x} , \qquad H = \frac{p^2}{2m} + V(x) .$$

In the absence of external forces, the potential $V(x) = 0$, so that

$$x(t) = x_0 + v_0t , \qquad p(t) = p_0 ,$$
$$w(x, p, t) = \delta(x - x_0 - v_0t)\delta(p - p_0)$$

(here $v_0 \equiv p_0/m$).

In order to introduce stochasticity, consider a set of identical independent particles differing by random initial parameters x_{0i}, p_{0i}, where i enumerates the particles. This corresponds to a set of points on the phase plane. Their distribution density $w(x, p, t)$ is proportional to the number of points in a small domain near x, p. The space of states is given by a set of various distribution functions $w(x, p, t)$ satisfying the conditions

$$\int \mathrm{d}x\,\mathrm{d}p\, w(x, p, t) = 1 , \qquad w(x, p, t) \geqslant 0 .$$

The function $w(x, p, t)$ allows one to calculate the moments $\langle x^m p^n\rangle$ and, more generally, the mean value $\langle A(x, p, t)\rangle$ for any function of x and t. The marginal distributions for the coordinate and momentum have the forms

$$w(x, t) = \int \mathrm{d}p\, w(x, p, t) , \qquad w(p, t) = \int \mathrm{d}x\, w(x, p, t) . \quad (6.4.1)$$

In the case of free motion, the momenta of the particles are conserved, $w(p) = \mathrm{const}$, therefore, the dependence of the state on time is taken into account by a trivial argument shift,

$$w(x, p, t) = w\left(x - \frac{pt}{m} , p, 0\right) . \quad (6.4.2)$$

In the differential form, we obtain the Liouville equation:

$$\frac{\partial w}{\partial t} + v\frac{\partial w}{\partial x} = 0 \qquad \left(v = \frac{p}{m}\right) .$$

This transformation describes classical diffusion of wave packets (see Fig. 7). (It should be distinguished from true 'diffraction in time', i.e., the envelope variation caused by vacuum dispersion for nonrelativistic particles.) Naturally, under certain initial conditions, both classical diffusion and quantum delocalization can be preceded by localization or 'focusing' of the packet.

Let us now pass to the quantum theory where the pure state of a particle is given by some complex function $\psi(x)$. This function determines the probability distribution for the coordinate, $w(x) \equiv |\psi(x)|^2$, the mean values $\langle(x)\rangle$, and also the moments

$$\mu_{m0} \equiv \langle x^m\rangle \equiv \langle\psi|x^m|\psi\rangle = \int \mathrm{d}x\, w(x)x^m .$$

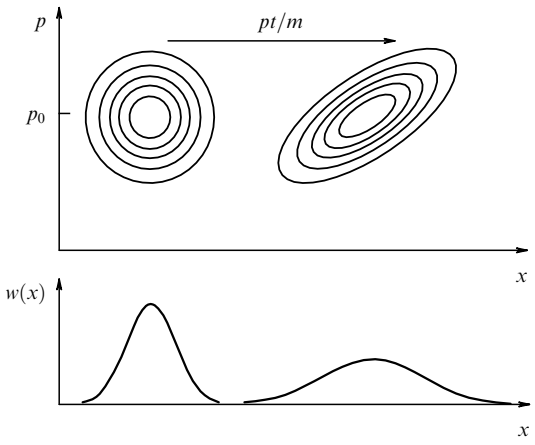

Figure 7. Distribution function for the coordinate and momentum $w(x,p)$ at two time moments (top) and the resulting diffusion of the coordinate distribution $w(x)$ with time (bottom).

Its Fourier conjugate determines the probabilities for the momentum, $|\psi(p)|^2 \equiv w(p)$, and the moments $\mu_{0n} \equiv \langle p^n \rangle$. The simultaneous joint distribution $w(x,p,t) \geqslant 0$ in the general case can be neither measured nor calculated. At the same time, the moments of the form $\mu_{mn} \equiv \langle x^m p^n \rangle \equiv \langle \psi | x^m p^n | \psi \rangle$ can be calculated. Naturally, one could try to define the corresponding function $w(x,p)$ such that

$$\mu_{mn} = \int \mathrm{d}x\, \mathrm{d}p\, w(x,p) x^m p^n \,.$$

(This problem is called the problem of moments, see Section 4.5.) In this approach, one could suppose that the coordinate and momentum of a given particle have definite *a priori* values. However, this approach has two principal obstacles: first, in the quantum theory, all moments $\langle x^m p^n \rangle, \langle x^{m-1} p^n x \rangle, \ldots, \langle p^n x^m \rangle$ are different; second, the function $w(x,p)$ can take negative values, so that it cannot have the operational sense of a probability.

To overcome the first obstacle, one can choose some fixed order of the operators. For instance, the Wigner function $W(x,p,t)$, which determines the moments symmetrized according to some rule [89, 90], can be expressed in terms of the WF as follows:

$$W(x,p,t) \equiv (2\pi\hbar)^{-1} \int \mathrm{d}y\, \exp\left(\frac{\mathrm{i}py}{\hbar}\right) \psi^*\left(x + \frac{y}{2}, t\right) \times \psi\left(x - \frac{y}{2}, t\right) . \tag{6.4.3}$$

For fixed t, this transformation defines the function of two variables $W(x,p)$ via the function of one variable $\psi(x)$. Calculating the marginals, one can easily verify that the *consistency conditions* are satisfied,

$$w(x,t) \equiv \int \mathrm{d}p\, W(x,p,t) = \left|\psi(x,t)\right|^2 ,$$
$$w(p,t) \equiv \int \mathrm{d}x\, W(x,p,t) = \left|\psi(p,t)\right|^2 . \tag{6.4.4}$$

In combination with the Schrödinger equation, Eqn (6.4.3) leads to the equation of motion for W, i.e., the quantum Liouville equation [89, 90]. In the case of free motion, it has a 'classical' form

$$\frac{\partial W}{\partial t} + v \frac{\partial W}{\partial x} = 0 \,, \qquad v = \frac{p}{m} \,,$$

i.e., the argument is transformed as $x \to x_t \equiv x - pt/m$. Thus, the Wigner function W for a free particle depends on time the same way as the classical distribution function w [see Eqn (6.4.2) and Fig. 7].

For instance, consider a Gaussian packet

$$\psi(x,t) = \frac{1}{\pi^{1/4}\sqrt{a_0 + \mathrm{i}\hbar t/ma_0}} \times \exp\left[\mathrm{i}k_0 x - \mathrm{i}\omega_0 t - \frac{(x - v_0 t)^2}{2(a_0^2 + \mathrm{i}\hbar t/m)}\right] ,$$

$$\left|\psi(x,t)\right|^2 = \frac{1}{\sqrt{\pi}\, a_t} \exp\left[-\frac{(x - v_0 t)^2}{a_t^2}\right] . \tag{6.4.5}$$

Here

$$k_0 \equiv \frac{p_0}{\hbar} \,, \qquad \omega_0 \equiv \frac{\hbar k_0^2}{2m} \,, \qquad v_0 \equiv \frac{p_0}{m} \,,$$
$$a_t \equiv a_0 \left(1 + \frac{t^2}{T_a^2}\right)^{1/2} , \qquad T_a = \frac{ma_0^2}{\hbar} \,,$$

and a_0 is the minimal length of the packet at $t = 0$. According to Eqns (6.4.5), the envelope $\left|\psi(x,t)\right|$ conserves its functional form moving with the group velocity $v_0 \equiv p_0/m$ and broadening with the growth of $|t|$. Substituting Eqn (6.4.5) into (6.4.3), we obtain

$$W(x,p,t) = \frac{1}{\pi\hbar} \exp\left[-\frac{(x - pt/m)^2}{a_0^2} - \frac{(p - p_0)^2 a_0^2}{\hbar^2}\right] . \tag{6.4.6}$$

This function is positive and can be considered as a probability. For $t \neq 0$, it describes the correlation between the coordinate and the momentum (see Fig. 7).

However, the Wigner functions of all other pure states take negative values [89, 90] and do not conserve their shapes in the course of propagation. As an example, consider a packet with a rectangular envelope,

$$\psi(x,0) = \Pi(x,a) \exp(\mathrm{i}k_0 x) \,, \tag{6.4.7}$$

where $\Pi(x,a) \equiv \theta(x + a)\theta(a - x)$ and $\theta(x)$ is the step function. From Eqns (6.4.3) and (6.4.7), we find the Wigner function:

$$W(x,p,0) = \Pi(x,a)\, \frac{\sin\left[(k - k_0)(a - 2|x|)\right]}{\pi\hbar a(k - k_0)} \,, \tag{6.4.8}$$

which definitely takes negative values (Fig. 8). Thus, the properties of the state (6.4.7) cannot be described in terms of some joint probability distribution for the coordinate and momentum.

Note that the *coherence length* a_{coh}, which is actually measured in experiments on particle interference, is usually determined not by the true length of the packet a_t but by its 'nonuniformity length', since the beam contains different particles with a classical velocity distribution (see Fig. 7). (Here one can find an analogy with the nonuniform broadening of spectral lines.) Using higher time resolution and applying other techniques, one can increase the observed

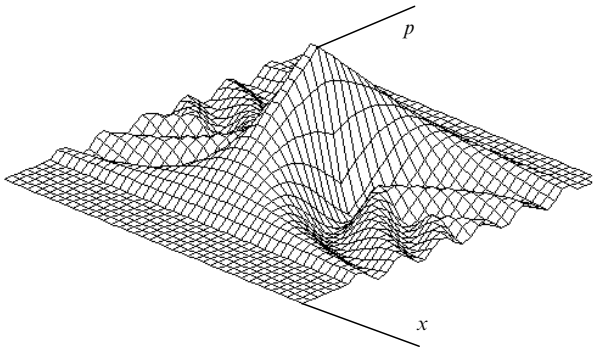

Figure 8. Wigner's function $W(x, p, 0)$ for a wave packet with a rectangular envelope.

coherence length. Real experiments with particle beams should be described in terms of mixed states (of the first type, Section 4.10). In the simplest case, each particle of the beam can be considered in a pure state depending on classical random parameters. For instance, let $p_0 = \hbar k_0$ from Eqns (6.4.5) or (6.4.7) be such a parameter, then the observed 'smoothed' distribution can be found by an additional classical averaging over $w(p_0)$,

$$w(x,t) = \int \mathrm{d}p_0\, w(p_0) \left|\psi(x,t;\, p_0)\right|^2 . \qquad (6.4.9)$$

Here $w(p_0)$ can be determined by the Maxwell velocity distribution with some given temperature. This equation describes two processes: the classical 'diffusion' of individual wave packets caused by the group velocity p_0/m dispersion (see Fig. 7) and the diffraction in time, i.e., variation of the shape of each packet caused by the vacuum dispersion.

6.5 Model of alpha-decay

A simple one-dimensional model provides a remarkably accurate description of alpha-decay [91]. Consider the motion of a pointlike particle in the presence of the two-hump potential $V(x)$ shown in Fig. 9a. Suppose that we know the moment of birth t_0 for the chosen nucleus, say, ${}_{88}\mathrm{Ra}^{226}$. In the framework of a primitive model, one can assume that the α-particle is prepared in a quasi-stationary state $\psi(x, t_0)$ localized inside the nucleus and determined by the potential well $V(x)$ (Fig. 9a). Further evolution of this state is shown in Figs 9b, c. The solution to the equation

$$\mathrm{i}\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial x^2} + V(x)\psi$$

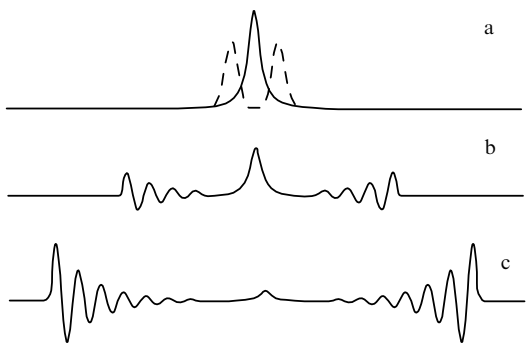

Figure 9. Evolution of an α-particle wave packet during radioactive decay.

describes the gradual tunneling of the particle through both barriers. The estimate of transmission coefficient gives the relation between the α-particle velocity and the half-life $T_{1/2}$. This relation is in qualitative agreement with the experimental data [91].

Using the Schrödinger equation, one can calculate the shape of the wave packet describing an individual α-particle for any $t > t_0$. According to this simple model, the length of the packet at $T_0 \gg T_{1/2}$ is of order of $a_0 = v_0 T_{1/2}$. Certainly, to check this information with considerable reliability, one should have a sufficient number of identical nuclei. Still, even for a single nucleus, it is possible to predict the 'probable' distance between the α-particle and the nucleus at any time moment, $x \approx v_0(t - t_0 \pm T_{1/2})$.

Let us consider a well-known optical experiment on *quantum jump* observation combined with the time-of-flight method (Section 6.2). A radioactive atom is trapped and illuminated by resonant laser radiation exciting one of its electron transitions. The energy of the laser light is partly re-emitted by the atom in the form of resonance fluorescence. Note that modern equipment allows the detection of resonance fluorescence from single atoms. Hence, as soon as the alpha-decay takes place, the electron levels are reorganized, the resonance fluorescence stops, and this moment can be detected and identified with the moment t_{1i} of the α-particle escape from the nucleus. (The inverse is also possible: the appearance of the resonance fluorescence indicates that the nucleus is created.) Suppose that a Geiger counter placed at a distance x_2 detects an α-particle at time t_{2i}. Hence, the group velocity is $v_{0i} = x_2/(t_{2i} - t_{1i})$. Repeated trials allow one to observe the distribution of the moments and this way to study the shape of the packet (see Section 6.2).

6.6 Modulation of the wave function

A lot of interesting effects are connected with the phase of the WF, such as, for instance, the Josephson effect and magnetic flux quantization in superconductors. Various fine effects, like the Aharonov–Bohm effect for electrons (see Section 6.7) and its neutron analogues, the Sagnac effect for neutrons, the influence of the gravitational field on the phase of the WF for slow neutrons and atoms, the geometric Berry phase, are studied using electron, neutron, and atomic interferometers (see Refs [13–15]). Such experiments are described by taking into account the dependence of the WF on classical quasi-stationary fields: electric, magnetic, gravitational, and inertial fields. In the case of atoms or molecules, one can additionally modulate the WF by means of an optical quasi-resonant field. The action of this field on the motion of an atom can be described in terms of some effective potential $V(x)$.

Let us try to use the effect of the WF *phase modulation* for proving the statement that a WF can be associated with a given individual particle. Consider a two-beam Mach–Zehnder interferometer for particles. The flux of particles at the input of the interferometer is made sufficiently weak, so that the particles enter the interferometer one by one, without influencing each other. For each particle, the phase difference ϕ between the two components of its WF in the two arms of the interferometer can be controlled, for instance, using an electric or magnetic field. The interference visibility in real interferometers for electrons, neutrons, atoms, or molecules can be close to 100%. This means that for some phase $\phi = \phi_0$, a particle is almost surely directed to one of the output ports of the interferometer, while for the phase $\phi_0 + \pi$, it almost surely gets to the other port. In other words, the WF

amplitude in one of the output beams can be turned to zero by the experimenter. Using classical language, the particle is directed along one or another path. Thus, by varying the phase delay, one can control the WF of any particle at the output of the interferometer. Note that here we mean the WF of a given individual particle. Since the system is prepared in one of the pure states of the observable (the path of the particle), the outcome is fixed, $p_1 = 0$ or 1, and in the ideal case, there is no need for repeated trials.

Let ϕ now take intermediate values. Then the particle is discovered either in the first output beam or in the second one, with the probabilities $p_1 = \cos^2(\phi - \phi_0)$ and $p_2 = \sin^2(\phi - \phi_0)$, i.e., the interferometer works like a beamsplitter with transmittance $T_1 = p_1$. The phase can depend on time, $\phi = \phi(t)$, and in this case, the probabilities $p_1(t)$ and $p_2(t)$ also vary (with opposite phases). Using radiotechnical language, one can say that the interferometer operates as a phase detector: it transforms the phase modulation of the WF $\phi(t)$ into the amplitude modulation of the classical probabilities $p_k(t)$ at the output.

The interferometer performs a unitary transformation of the particle state, with the normalization of the total two-component WF being invariant. A pure state at the input is transformed into another pure state at the output. However, it should be stressed that if only a single output beam is considered, with the second one ignored, it should be described in terms of mixed states (of the second type, see Section 4.10): the output state is a mixture of the one-photon pure state, $|1\rangle$, and the vacuum pure state, $|0\rangle$, weighted with the classical probabilities p_1 and $1 - p_1$, respectively. Thus, if the second output beam is not considered, the interferometer works as an absorber with the transmission coefficient $\eta = p_1$.

Is there a possibility for WF amplitude modulation, with the normalization (number of particles) conserved, in free space without an interferometer? From our consideration, it is clear that the effect of an absorber with transmittance η on an electron or a neutron should be described in terms of mixed states, using classical probabilities. Indeed, at the output of an absorber, one finds an incoherent mixture of the states $|1\rangle$ and $|0\rangle$ with the probabilities η and $1 - \eta$, respectively. In other words, an absorber or a nontransparent screen can be phenomenologically described as performing a nonunitary transformation of the particle state. (This problem was studied in detail in quantum optics [21].) An obturator, which periodically blocks a beam of particles, can be described by a time-dependent absorption $\eta(t)$; in addition, this case is characterized by the classical probability for a wave packet to get into the obturator 'window'. Any absorber modulates not the WF amplitude but the classical probability of one-particle state preparation. On the other hand, the effect of a semi-transparent reflecting screen can be described by a unitary transformation retaining the WF normalization. This case is analogous to the case of the two-beam interferometer, and similarly, here the reflected beam should be taken into account.

Consider the phase (frequency) modulation of the particle WF in free space in the one-dimensional approximation. A time-dependent phase $\phi(t)$, in contrast to the global (constant) phase of the WF, can lead to observable effects. Let a plane monochromatic WF with frequency ω_0 pass through a thin phase modulator placed at $x = 0$. In the case of harmonic modulation at frequency $\Omega \ll \omega_0$ with the frequency deviation $\beta\Omega$, the WF at the output has the form

$$\psi(0, t) = \exp\left[-i\omega_0 t - i\beta \sin(\Omega t)\right].$$

Hence, for $x > 0$,

$$\psi(x, t) = \sum_{n=-\infty}^{\infty} J_n(\beta) \exp(ik_n x - i\omega_n t),$$

$$\omega_n \equiv \omega_0 + n\Omega, \qquad k_n \equiv \left(\frac{2m\omega_n}{\hbar}\right)^{1/2} \tag{6.6.1}$$

with J_n denoting the Bessel functions. Thus, harmonic phase modulation leads to the appearance of new frequency components $\omega_0 \pm n\Omega$. Due to the dispersion $\omega \sim k^2$, these components propagate with different velocities. Therefore, a propagating WF acquires an amplitude modulation in addition to the phase modulation, i.e., there appear slow beats of the wave packet amplitude in space – time (Fig. 10). The time period of these beats is $\Delta t = 2\pi/\Omega$, while their space period is approximately $\Delta x \approx 2\pi/(k_1 - k_0) \approx v_0 \Delta t$, where $v_0 \equiv \hbar k_0/m$ is the group velocity.

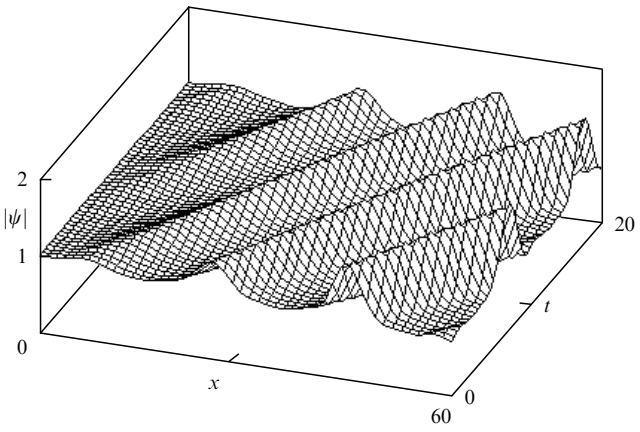

Figure 10. Transformation of the frequency modulation of the wavefunction $\psi(x, t)$ (for $x = 0$) into amplitude modulation ($x > 0$) according to Eqn (6.6.1). The modulation frequency is $\Omega = \omega_0/10$; the frequency deviation is $\beta = \Delta\omega/\Omega = 1$, the coordinate x is normalised by $\lambda_0 = 2\pi/k_0$, and the time t by $T_0 = 2\pi/\omega_0$.

These beats can be observed using synchronous detection. Periodic transformation of the phase modulation into the amplitude one has been recently studied for rubidium atoms [92]; the effects observed in this work were analogous to the optical echo effect [35]. Phase modulation can be used to control wave packets of finite length, to shorten or extend them.

For light waves, transformation of the phase modulation into the amplitude modulation (*chirping*) due to dispersion is widely used for obtaining supershort high-power laser pulses. Similar effects are predicted for beams of slow neutrons under various types of modulation [93, 94].

6.7 Quantum magnetometers and the Aharonov – Bohm paradox

Suppose that the above-considered interferometer for charged particles, say, electrons, contains a source of static magnetic field $\mathbf{B}(\mathbf{r})$. (The influence of the spin is neglected.) This field can be equivalently described in terms of the vector

potential $\mathbf{A}(\mathbf{r})$ given by the relation $\operatorname{rot}\mathbf{A}(\mathbf{r}) \equiv \mathbf{B}(\mathbf{r})$. This definition is ambiguous: all potentials of the form $\mathbf{A}$ and $\mathbf{A}' \equiv \mathbf{A} + \operatorname{grad}\chi$, where $\chi(\mathbf{r})$ is an arbitrary scalar field, give the same magnetic field $\mathbf{B}(\mathbf{r})$ and therefore, are indistinguishable.

Assume that the field $\mathbf{A}(\mathbf{r})$ does not vary much within the cross section of the electron WF in each arm of the interferometer. Then it follows from the Schrödinger equation with the Hamiltonian $(\mathbf{P} - e\mathbf{A}/c)^2/2m$ that each component of the WF has a phase shift given by the line integral:

$$\phi_n = \frac{e}{\hbar c}\int_{C_n} \mathbf{A}(\mathbf{r})\,\mathrm{d}\mathbf{r}\,, \tag{6.7.1}$$

where C_n is the path along the electron trajectory in the nth arm of the interferometer between the input and the output beamsplitters. (As in geometric optics, the trajectory is understood as the path along a bounded beam.) Hence, the phase difference between the WF components at the output beamsplitter is

$$\phi \equiv \phi_1 - \phi_2 = \frac{e}{\hbar c}\oint_C \mathbf{A}(\mathbf{r}_C)\,\mathrm{d}\mathbf{r}_C = \frac{e}{\hbar c}\iint_S \mathbf{B}\,\mathrm{d}S \equiv \frac{e}{\hbar c}\,\Phi_S\,. \tag{6.7.2}$$

(Note that the result is independent of χ.) This value is observable since it determines the probabilities $p_{1,2}$ for the electron to be discovered in the two output beams of the interferometer, see the previous section. Here $C = C_1 - C_2$ is a closed contour coinciding with the electron trajectory in both arms of the interferometer, S is the surface bounded by this contour, and Φ_S is the magnetic flux through this surface.

Thus, an interferometer can be used to measure magnetic fluxes. A similar principle is used in Josephson-transition magnetometers [95]. This demonstrates how a classical variable can be measured by means of a quantum effect. (Such devices are used in *quantum metrology*.)

It is typical for this effect that ϕ depends only on the integral (*global*) parameter, Φ_S, which is determined by the contour C and the field. The electron velocity does not play any role but different paths can lead to equal phase differences. For this reason, such effects are called *topological*, or *geometrical*.

According to Eqn (6.7.1), the phase shift is not defined for an open contour, since the potential $\mathbf{A}$ is ambiguous. Still, one can define the phase $\phi(\mathbf{r}, \mathbf{r}_0)$ at each point $\mathbf{r}$ of some trajectory with respect to the phase at a fixed point $\mathbf{r}_0$. This can be done by closing the trajectory from $\mathbf{r}$ to $\mathbf{r}_0$ along a curve orthogonal to $\mathbf{A}(\mathbf{r})$ at each point, i.e., a curve belonging to an equipotential surface of $\mathbf{A}(\mathbf{r})$. In accordance with Eqn (6.7.1), this closure does not influence the phase. The obtained 'moving' phase is additive, $\phi(\mathbf{r}_2, \mathbf{r}_0) = \phi(\mathbf{r}_2, \mathbf{r}_1) + \phi(\mathbf{r}_1, \mathbf{r}_0)$. For the case of several sources of field, this notion has interesting topological properties.

The Aharonov–Bohm effect (see Refs [13, 95–98]) is observed in the case where the magnetic field is equal to zero along the whole electron trajectory. (More precisely, it is equal to zero in the whole domain where the WF of the particles have noticeable values.) According to Eqn (6.7.2), the phase difference can be nonzero in this case if $\mathbf{B}(\mathbf{r}_C) = 0$, but at the same time, $\Phi_S \neq 0$. These conditions can be satisfied using magnetic screening, a thoroidal magnet or a long solenoid placed between the arms of the interferometer. Interpretation of the Aharonov–Bohm effect is connected with an interesting paradox. The Lorenz force acting on the electron at a point $\mathbf{r}_C$ is determined by the field $\mathbf{B}(\mathbf{r}_C)$; therefore, it is commonly supposed that only $\mathbf{B}$ is a 'real' field, while the potential $\mathbf{A}$ is an auxiliary mathematical notion. But in the case considered here, $\mathbf{B}(\mathbf{r}_C) = 0$, so one has to accept 'nonlocality' or 'action at a distance', since the static field $\mathbf{B}_0$ inside the solenoid somehow influences the electron 'at a distance'. Expressing $\mathbf{B}_0$ via the current I_0, one can speak about the 'action at a distance' of this current. This conclusion can be avoided if the potential $\mathbf{A}(\mathbf{r}_C)$ is claimed to be a 'real' field, but this potential is defined at each point $\mathbf{r}_C$ with a certain ambiguity. Both alternatives contradict the traditional viewpoint. Note that the condition $\mathbf{B}(\mathbf{r}_C) = 0$ is actually not necessary for formulating the paradox, since in the general case, the integral formula (6.7.2) also describes the global, i.e., nonlocal, action of the field $\mathbf{B}(\mathbf{r})$.

On the other hand, the term 'action' implies a dynamic effect, i.e., variation of the observed phase difference $\Delta\phi$ as a result of the current ΔI variation. Clearly, any change in the WF phase (*the phase modulation*, see the previous section) would be delayed in time, in accordance with the solution to the Maxwell equations for the classical field of a given source. No instant 'action at a distance' would be observed. Probably, it is more consistent to consider the field source, such as, for instance, a heavy particle with a dipole magnetic moment, as a quantum system. In this case, we are dealing with the interaction between two quantum systems and no question arises whether it is the field or the potential which is 'real'.

This reasoning demonstrates that the question of which is 'more real', $\mathbf{B}(\mathbf{r}_C)$ or $\mathbf{A}(\mathbf{r}_C)$, makes no sense from the operational viewpoint. It relates to the group of 'what-actually-goes-on-there' questions. One can only state that the formalism based on the potential A_μ instead of the fields $F_{\mu\nu}$ is usually more compact and symmetric for solving relativistic problems. In the *electrodynamics calibration theory*, it is namely the potential that plays the most important role, and its existence is supposed to follow from the charge conservation law.

7. Conclusions

In this paper, we tried to find out the operational meaning of the basic terms used in nonrelativistic quantum physics. Our consideration was based on the techniques applied in laboratories and on the observable experimental data. Of course, the situation may drastically change with time; for instance, some new metatheory may appear, bringing together quantum and classical physics, or new experimental facts may be obtained. The essence of quantum notions was also clarified by means of comparison with the closest explicit models based on classical stochasticity.

Our attention was mostly drawn to the central object in quantum physics, to the WF. We found reasons for ascribing a definite WF ψ to a particle or to any other individual quantum object prepared in the course of an ideal preparation procedure. This position agrees with the orthodox viewpoint. It was shown that the notion of a pure state of a given individual system has a strict operational meaning, since it is determined by the preparation procedure. Knowing the state of a system, one can calculate the probabilities of possible observations.

However, to convincingly check the information contained in the symbol ψ, one needs a set of identically prepared

systems. This conclusion agrees with the ensemble viewpoint. Thus, two basic approaches to the WF, the orthodox and the ensemble approach, correspond to two different experimental procedures, to preparation and to measurement. This fact removes the seeming contradiction between the two approaches. Both approaches have operational verifications, each its own.

At the same time, there is no principal difference between a single trial and an arbitrarily large finite number of trials. As a result of some measurement, one can ascribe the WF to a single system or to, say, a hundred identically prepared systems (Section 3.2). In both cases, the theory gives only probability predictions, with the only exception for the case where the system is prepared in an eigenstate of an observed operator. Even for the case of a single measurement, knowing the preparation procedure for the WF of some individual system, one has certain *a priori* information about the outcome of the future experiment. For instance, after a two-slit screen, a given particle 'almost for sure' will not get into 'sufficiently close' vicinity of the WF knot. On the other hand, even a hundred measurements of the particle coordinate giving the statistical mean $\langle x\rangle_{100} = (1 \pm 0.1)$ cm, do not guarantee that the true mean value $\langle\psi|x|\psi\rangle$ is not 2 cm.

For a classical mixture of pure states (Sections 3.3 and 4.10), the ensemble meaning is implied at the very stage of preparation. If the preparation conditions are not known, then the question of whether an individual particle (for instance, from cosmic rays) possesses a WF is a rhetorical one. This question has no operational sense, and the answer is chosen according to one's taste.

Modern techniques suggest surprising possibilities for preparing atoms in given states ψ, and for controlling and monitoring these states (bright recent results can be found in [99]). The 'reality' of an individual WF seems to be clearly demonstrated in Section 6.6 using a realistic experiment with phase modulation of a particle in a two-beam interferometer. On the other hand, dynamic experiments of this kind can also be described in the framework of the Heisenberg representation (Section 4.4), with the time dependence included into the coordinate operator $X(t)$ instead of the WF. In this case, it makes no sense to imagine a moving particle as a (complex) wave packet $\psi(x, t)$ changing its shape in the course of propagation.

Another group of problems discussed in the paper relates to nonclassical effects (Sections 4.5, 4.6, 5.6–5.8, 6.4). Comparing the predictions of certain quantum and classical probability models, one comes to paradoxes demonstrating that these models are incompatible. These paradoxes are commonly solved with the help of the term 'quantum nonlocality'. However, a more conservative viewpoint is possible admitting that the quantum theory is 'non-Kolmogorovian' and it makes no sense to ascribe *a priori* values to non-commuting variables. In this approach, several features of quantum probability models can easily be obtained. Recall that 'nonlocally controlled' EPR correlations have rather close classical analogues [56] and the contradiction between classical and quantum predictions is only quantitative.

Considerable attention was also paid to one of the most contradictory notions in quantum physics, WF reduction as a result of measurement. In real experiments, the measurement procedure is never used for the preparation of a quantum state. The preparation and measurement procedures are essentially different, in spite of the common viewpoint dating from the thirties. One of the few 'dissidents', W Lamb, in his paper [100] entitled "Operational interpretation of nonrelativistic quantum mechanics" writes that "although some authors confuse preparation and measurement of a state, these notions are essentially different, both physically and logically."

At present, probably, all known experiments can be described using the standard algorithms of the quantum theory and the Born postulate (4.7.1). According to this postulate, the value to be compared with experiment is the projection of the state of the system onto some vector, which is determined by the experimental procedure. Then Dirac's statement (4.7.2) that the measurement creates a new WF is not necessary. As far as we know, at present there exist no experimental facts that could confirm or disprove the reduction hypothesis and various models of the measurement process. Despite all efforts, they remain completely isolated from experiment. Again and again, new results confirm only the adequacy of the quantum formalism (provided that the model is chosen correctly) and the Born postulate. It is remarkable that the projection postulate (4.7.2), in contrast to Born's postulate (4.7.1), seems never be used in *quantitative* descriptions of real experiments. Like the notion of partial reduction (Section 4.8), it is only used in qualitative speculations.

Thus, the notion of WF reduction at the moment of measurement is so far redundant, it is only convenient for an obvious interpretation of the observed effects. It is an explanation of 'what actually goes on', i.e., it relates to the fourth component of the quantum physics, to its *interpretation* (Section 2). The choice of interpretation is a matter of taste. (This is the difference between an *interpretation* and a *theory*.) Note, however, that describing quantum correlation effects in terms of reduction and using the terminology related to it (*nonlocality, teleportation*), one comes to (pseudo)paradoxes like a superluminal telegraph. This fills physics with an unnecessary atmosphere of mystery and provides grounds for various pseudosciences. It seems useful to return, from time to time, back to the beginning and to try to build the axiomatic structure within the given branch of physics, distinguishing between the necessary and redundant notions with the help of an operational approach.

On the other hand, it is not reasonable to reject convenient but not strictly defined notions; it is better to clarify their status. The obviousness of reduction and other model notions of physics promotes the planning of new experiments, the development of intuition, and the discovery of new effects. It is worth mentioning the positive role of the alchemists' ideas, Faraday's lines, various models of the ether, Dirac's concept of positrons as 'holes' in a sea of particles with negative energy or their definition as electrons moving backward in time, given by Wheeler and Feynman. Several times in history, 'metaphysics' has turned into 'physics' (atoms, antiparticles, quarks). It is possible that reduction will manifest itself in future experiments (with timelike-separated events).

In Section 6.1, we presented a simple model for the measurement of a particle's longitudinal coordinate. This model allows one to set a relation between the parameters of the measurement devices x_1, t_1, directly measurable by means of rulers and clocks, and the basic construction of the quantum formalism, the function $\psi(x, t)$.

In Sections 4.9 and 6.2, we tried to prove the statement that the reduction postulate in the form of the Wigner formula (4.9.1) is actually necessary only for describing a narrow group of correlation experiments like time-of-flight

experiments, where one measures two or more operators that do not commute due to the interaction. If such experiments were carried out, they would probably provide direct evidence for the fact that the Wigner formula gives a correct description of repeated measurements. From the operational viewpoint, the Wigner formula (4.9.1) obtained by combining the Born postulate (4.7.1) with the Dirac postulate (4.7.2) can be considered as the basic measurement postulate of the quantum theory, a generalization of the Born postulate. In our opinion, it is essential that according to the Wigner formula in one of its modifications (Section 6.2), 'reduction' occurs not at one of the numerous moments t_i when a particle is registered by the first detector but at some abstract moment t_1. This parameter is the argument of the distribution function obtained by statistical averaging of a large series of measurements, see Fig. 6.

Acknowledgements. The author is grateful to P V Elyutin for constant interest, numerous discussions, and valuable comments throughout the work over this subject, to Yu I Vorontsov for constructive criticism, to A I Lipkin for fruitful discussions concerning the role of models in physics, and also to the anonymous reviewer for several critical remarks which were taken into account.

The present work was supported in part by the Russian Foundation for Basic Research (grant numbers 96-02-16334-a, 96-15-96673) and the state program "Fundamental Metrology".

8. Appendices

Appendix I. Eigenvectors of the Stokes operators and the Greenberger – Horne – Zeilinger paradox

For the description of photon polarization, two bases are convenient: the one formed by the eigenvectors $|x\rangle$, $|y\rangle$ of the S_1 operator and the one formed by the eigenvectors $|\pm\rangle = \big(|x\rangle \pm \mathrm{i}|y\rangle\big)/\sqrt{2}$ of the S_3 operator. According to Eqn (5.4.3),

$$\begin{aligned} &S_1|x\rangle = |x\rangle\,, \quad && S_1|y\rangle = -|y\rangle\,, \quad && S_1|\pm\rangle = |\mp\rangle\,,\\ &S_2|x\rangle = |y\rangle\,, \quad && S_2|y\rangle = |x\rangle\,, \quad && S_2|\pm\rangle = \pm\mathrm{i}|\mp\rangle\,,\\ &S_3|x\rangle = \mathrm{i}|y\rangle\,, \quad && S_3|y\rangle = -\mathrm{i}|x\rangle\,, \quad && S_3|\pm\rangle = \pm|\pm\rangle\,. \end{aligned} \tag{I.1}$$

Let the incident field be prepared in the three-photon state

$$|\psi\rangle = \frac{|+\rangle_\mathrm{A}|+\rangle_\mathrm{B}|+\rangle_\mathrm{C} + |-\rangle_\mathrm{A}|-\rangle_\mathrm{B}|-\rangle_\mathrm{C}}{\sqrt{2}}\,, \tag{I.2}$$

the indices A, B, C relating to the three beams. According to Eqns (I.1), $\langle S_{1Z}\rangle = \langle S_{2Z}\rangle = \langle S_{3Z}\rangle = 0$, i.e., the radiation is completely depolarized, and measurement of the Stokes parameters for each beam gives the values $s_{kZ} = \pm 1$ with equal probabilities. However, one can easily see from Eqns (I.1) that $|\psi\rangle$ is an eigenvector for some products of three Stokes operators,

$$\begin{aligned} &S_{1\mathrm{A}}S_{1\mathrm{B}}S_{1\mathrm{C}}|\psi\rangle = |\psi\rangle\,, \quad && S_{1\mathrm{A}}S_{2\mathrm{B}}S_{2\mathrm{C}}|\psi\rangle = -|\psi\rangle\,,\\ &S_{2\mathrm{A}}S_{1\mathrm{B}}S_{2\mathrm{C}}|\psi\rangle = -|\psi\rangle\,, \quad && S_{2\mathrm{A}}S_{2\mathrm{B}}S_{1\mathrm{C}}|\psi\rangle = -|\psi\rangle\,. \end{aligned} \tag{I.3}$$

Hence,

$$\begin{aligned} &\langle\psi|S_{1\mathrm{A}}S_{1\mathrm{B}}S_{1\mathrm{C}}|\psi\rangle = 1\,, \quad && \langle\psi|S_{1\mathrm{A}}S_{2\mathrm{B}}S_{2\mathrm{C}}|\psi\rangle = -1\,,\\ &\langle\psi|S_{2\mathrm{A}}S_{1\mathrm{B}}S_{2\mathrm{C}}|\psi\rangle = -1\,, \quad && \langle\psi|S_{2\mathrm{A}}S_{2\mathrm{B}}S_{1\mathrm{C}}|\psi\rangle = -1\,. \end{aligned} \tag{I.4}$$

Thus, experiments where the Nicol prisms have $\chi_\mathrm{A} = \chi_\mathrm{B} = \chi_\mathrm{C} = 0$ must give an ideal correlation between the photocounts observed in the three detectors, $s_{1\mathrm{A}}s_{1\mathrm{B}}s_{1\mathrm{C}} = \langle S_{1\mathrm{A}}S_{1\mathrm{B}}S_{1\mathrm{C}}\rangle = 1$. Similarly, for $\chi_\mathrm{A} = 0$ and $\chi_\mathrm{B} = \chi_\mathrm{C} = 45^\circ$, ideal anticorrelation must be observed, $\langle S_{1\mathrm{A}}S_{2\mathrm{B}}S_{2\mathrm{C}}\rangle = -1$. This result leads to the GHZ paradox (see Section 5.8).

Appendix II. To the theory of 'quantum teleportation'

The effect of the beamsplitter BS and the transformer $\mathrm{T_C}$ in Fig. 5 can be described in the Heisenberg representation by the following unitary transformations [see Eqn (5.4.5)]:

$$\begin{aligned} &a_j' = \frac{a_j + b_j}{\sqrt{2}}\,, \qquad && c_x' = t_\mathrm{C}^* c_x + r_\mathrm{C}^* c_y\,,\\ &b_j' = \frac{-a_j + b_j}{\sqrt{2}}\,, \qquad && c_y' = -r_\mathrm{C} c_x + t_\mathrm{C} c_y\,. \end{aligned} \tag{II.1}$$

Here a, b, c are photon creation operators for the beams A, B, C, $j = x, y$, and $|t_\mathrm{C}|^2 + |r_\mathrm{C}|^2 = 1$. These relations define the 6×6 transformation matrix G_{mn} for the whole optical system. The matrix G_{mn} relates the output operators to the input ones and hence, allows one to express the statistics of the output field in terms of the input statistics, which are given by the WF $|\psi\rangle$ of the initial field.

According to the Born postulate (4.7.1), the probabilities $p' = |q'|^2$ (below, the primes of p and q are omitted) of discovering the given numbers of photons in six output modes are determined by the projections of $|\psi\rangle$ on the corresponding Fock states,

$$q(n_1, \ldots, n_6) = \langle n_1, \ldots, n_6|\psi\rangle = \frac{\langle 0|a_1'^{n_1} \ldots a_6'^{n_6}|\psi\rangle}{\sqrt{n_1! \ldots n_6!}}\,. \tag{II.2}$$

Note that in the case of one-photon states, the probabilities coincide with the corresponding moments,

$$p(110010) = \mu(Ax, Ay, Cx) \equiv \langle\psi|N'_{Ax}N'_{Ay}N'_{Cx}|\psi\rangle\,. \tag{II.3}$$

(The modes are numbered in the following order: Ax, Ay, Bx, By, Cx, Cy.)

Substitution of Eqns (II.1) into (II.2) gives the probabilities of all observable events. For instance, the probability amplitude of detecting three photons in the output modes Ax, By, Cx is

$$\begin{aligned} q(Ax, By, Cx) &= \langle 0|a_x' b_y' c_x'|\psi\rangle\\ &= \frac{1}{2}\big\langle 0\big|(a_x + b_x)(-a_y + b_y)(t_\mathrm{C}^* c_x + r_\mathrm{C}^* c_y)\big|\psi\big\rangle\,. \end{aligned} \tag{II.4}$$

'Teleportation' takes place under the condition that all but two matrix elements entering Eqns (II.4) are zero. Let there be a single photon in each input beam, then $\langle 0|a_x a_y|\psi\rangle = \langle 0|b_x b_y|\psi\rangle = 0$. Let, in addition, $\langle 0|a_x b_y c_y|\psi\rangle = \langle 0|a_y b_x c_x|\psi\rangle = 0$, then Eqn (II.4) takes the form

$$q(Ax, By, Cx) = \frac{1}{2}\big\langle 0\big|(t_\mathrm{C}^* a_x b_y c_x - r_\mathrm{C}^* a_y b_x c_y)\big|\psi\big\rangle\,. \tag{II.5a}$$

According to this expression, the transformer $\mathrm{T_C}$ has the same effect on the polarization of photons in both beams A and C.

On the other hand, the probability amplitude of detecting two photons in the output mode Ax and one photon in the

output mode Cx has the form

$$q(2Ax,Cx)=\frac{1}{2\sqrt{2}}\langle 0|(a_x+b_x)^2(t_C^*c_x+r_C^*c_y)|\psi\rangle$$
$$=\frac{1}{2\sqrt{2}}\langle 0|2a_xb_x(t_C^*c_x+r_C^*c_y)|\psi\rangle$$
$$=\frac{1}{\sqrt{2}}\langle 0|(r_C^*a_xb_xc_y)|\psi\rangle\,. \qquad \text{(II.5b)}$$

In this case, there is no teleportation effect (no dependence on t_C). Hence, such outcomes should be excluded using a coincidence circuit or a gate.

Let us specify the input state. Let

$$|\psi\rangle=|\psi\rangle_A|\psi\rangle_{BC}\,, \qquad \text{(II.6a)}$$

$$|\psi\rangle_A=\alpha|Ax\rangle+\beta|Ay\rangle\,, \qquad \text{(II.6b)}$$

$$|\psi\rangle_{BC}=\frac{1}{\sqrt{2}}\big(|Bx,Cy\rangle-|By,Cx\rangle\big)\,. \qquad \text{(II.6c)}$$

Here $|Ax\rangle\equiv|10***\rangle$ is the state with a single x-polarized photon in mode A, $|Bx,Cy\rangle\equiv|**1001\rangle$ is the state with a single photon in each of the modes Bx, Cy, and so on. The state $|\psi\rangle_{BC}$ has the necessary properties $b_xc_y|\psi\rangle_{BC}=-b_yc_x|\psi\rangle_{BC}=|0\rangle/\sqrt{2}$ and $b_xc_x|\psi\rangle_{BC}=b_yb_y|\psi\rangle_{BC}=0$, which provide the transition from Eqn (II.4) to (II.5).

In the case (II.6), using Eqn (II.4) and analogous relations, we find

$$q(Ay,Bx,Cx)=-q(Ax,By,Cx)=\frac{1}{\sqrt{8}}(t_C^*\alpha+r_C^*\beta), \qquad \text{(II.7a)}$$

$$q(Ax,Ay,Cx)=-q(Bx,By,Cx)=\frac{1}{\sqrt{8}}(-t_C^*\alpha+r_C^*\beta), \qquad \text{(II.7b)}$$

$$q(2Ax,Cx)=-q(2Bx,Cx)=\frac{r_C^*\alpha}{2}\,, \qquad \text{(II.7c)}$$

$$q(2By,Cx)=-q(2Ay,Cx)=\frac{t_C^*\beta}{2}\,. \qquad \text{(II.7d)}$$

Here $|2Ax\rangle$ is the state with two x-polarized photons in beam A. The amplitudes of the form $q(*,*,Cy)$ can be found by replacing $t^*\to -r$, $r^*\to t$, then, $p(*,*,Cx)+p(*,*,Cy)=1$; the amplitudes $q(Ax,Bx,*)$ and $q(Ay,By,*)$ equal zero. As a result, the probabilities of all 16 observable events are

$$p(Ax,By,Cx)=p(Ay,Bx,Cx)$$
$$=\frac{1}{8}\left[|t_C\alpha|^2+|r_C\beta|^2+2\,\mathrm{Re}(t_Cr_C^*\alpha^*\beta)\right],$$

$$p(Ax,By,Cy)=p(Ay,Bx,Cy)$$
$$=\frac{1}{8}\left[|r_C\alpha|^2+|t_C\beta|^2-2\,\mathrm{Re}(t_Cr_C^*\alpha^*\beta)\right], \qquad \text{(II.8a)}$$

$$p(Ax,Ay,Cx)=p(Bx,By,Cx)$$
$$=\frac{1}{8}\left[|t_C\alpha|^2+|r_C\beta|^2-2\,\mathrm{Re}(t_Cr_C^*\alpha^*\beta)\right],$$

$$p(Ax,Ay,Cy)=p(Bx,By,Cy)$$
$$=\frac{1}{8}\left[|r_C\alpha|^2+|t_C\beta|^2+2\,\mathrm{Re}(t_Cr_C^*\alpha^*\beta)\right], \qquad \text{(II.8b)}$$

$$p(2Ax,Cx)=p(2Bx,Cx)=\frac{1}{4}\,|r_C\alpha|^2\,,$$

$$p(2Ax,Cy)=p(2Bx,Cy)=\frac{1}{4}\,|t_C\alpha|^2\,,$$

$$p(2Ay,Cx)=p(2By,Cx)=\frac{1}{4}\,|t_C\beta|^2\,,$$

$$p(2Ay,Cy)=p(2By,Cy)=\frac{1}{4}\,|r_C\beta|^2\,. \qquad \text{(II.8c)}$$

According to Eqns (II.7a) or (II.8a), the transformer T_C acts on the four events $p(Ax,By,*)$ and $p(Ay,Bx,*)$ (which take place in 25% of all trials) in the same way as if it were placed in beam A at the input of the system. The joint action of T_A and T_C on these events is described by the product of the Jones matrices $\mathbf{T}_C\mathbf{T}_A$; varying T_C, one can measure α, β. From the operational viewpoint, this is the essence of the observed effect.

For another four events, $p(Ax,Ay,*)$ and $p(Bx,By,*)$, the dependence on T_A and T_C can be made the same. For this purpose, after such an event occurs, one should perform, in accordance with Eqns (II.8b), the additional unitary transformation $T_C'=\sigma_z$, which changes the sign of β before T_C [75].

At the same time, 8 events (II.8c) where two photons are fed to the same detector manifest no 'teleportation' effect. (Such events occur in 50% of all trials.) They can be excluded by means of an optical gate (Fig. 5). In this case, the C beam is completely polarized, $P_C=1$.

Summing all probabilities of the form $p(*,*,Cx)$, we find the marginal probability of detecting a Cx-photon, $p(Cx)=\sum p(*,*,Cx)=\langle N_{Cx}'\rangle=1/2$. Similarly, $p(Cy)=\sum p(**Cy)=\langle N_{Cy}\rangle=1/2$ i.e., C-photons stay completely depolarized, as one should expect. Thus, the transformers T_A and T_C have no influence on the unconditional counts of both detectors D_{Cj} in beam C.

Let us find the degree of polarization for beam C in the presence of the controlled transformer T_C' but without the gate excluding the events (II.8c). According to (II.8), taking the inverse sign of β in (II.8b), we find the Stokes parameters for beam C: $S_{0C}=1$, $S_{1C}=0$, $S_{2C}=\mathrm{Re}(\alpha^*\beta)=(1/2)S_{2A}$, $S_{3C}=\mathrm{Im}(\alpha^*\beta)=(1/2)S_{3A}$. As a result, the degree of polarization P_C for beam C is $P_C=(1/2)\sin(\theta_A)$, where $\theta_A=\arctan|\beta/\alpha|$ is the polar angle for the point mapping the state of photon A on the Poincaré sphere [see Eqns (5.4.2)]. For instance, for linear polarization of A-photons, $P_C=1/2$, and for circular polarization, $P_C=0$. Thus, a controlled unitary transformation T_C' equal to 1 or σ_x is not sufficient for exact copying of the A-photon polarization. A gate is needed even in the case where all detectors and other elements are ideal.

Let us briefly consider the same calculation in the Schrödinger representation. Vector transformations equivalent to (II.1) are given by the matrices T^+,

$$|A_j\rangle=\frac{|A_j'\rangle-|B_j'\rangle}{\sqrt{2}}\,, \qquad |C_x\rangle=t_C|C_x'\rangle-r_C^*|C_y'\rangle\,,$$

$$|B_j\rangle=\frac{|A_j'\rangle+|B_j'\rangle}{\sqrt{2}}\,, \qquad |C_y\rangle=r_C|C_x'\rangle+t_C^*|C_y'\rangle\,. \qquad \text{(II.9)}$$

Here primed letters denote output modes. Substituting these expressions into Eqns (II.6), we obtain the WF of the field at the output of the scheme, $|\psi'\rangle$. Let, for simplicity, $T_C=1$,

then

$$|\psi'\rangle = \frac{1}{\sqrt{8}}\big[\big(|Ay', Bx'\rangle - |Ax', By'\rangle\big)\big(\alpha|Cx'\rangle + \beta|Cy'\rangle\big) + \big(|Bx', By'\rangle - |Ay', Ax'\rangle\big)\big(\alpha|Cx'\rangle - \beta|Cy'\rangle\big) + \sqrt{2}\,\beta\big(|2By'\rangle - |2Ay'\rangle\big)|Cx'\rangle + + \sqrt{2}\,\alpha\big(|2Ax'\rangle - |2Bx'\rangle\big)|Cy'\rangle\big]\,. \quad \text{(II.10)}$$

Here $|2Ax'\rangle = (a_x'^{+})^2|0\rangle/\sqrt{2}$ is the state with two photons in the mode Ax'.

According to the Born postulate in the Schrödinger representation,

$$q(Ay, Bx, Cx) = \langle 0|a_y b_x c_x|\psi'\rangle = \langle Ay', Bx', Cx'|\psi'\rangle\,,$$

etc. Here, all operators commute, and therefore their order plays no role. However, for calculating the dependence of the output moments on T_A and T_C, it is convenient first to find the projections of $|\psi'\rangle$ on the vectors describing detection of photons only in beams A′ and B′. For example,

$$\langle Ay', Bx'|\psi'\rangle = \frac{1}{\sqrt{8}}\big(\alpha|Cx'\rangle + \beta|Cy'\rangle\big) \equiv |\psi\rangle_{C\,\mathrm{eff}}\,. \quad \text{(II.11)}$$

This is a (non-normalized) vector in the space C. We also define here the normalized effective WF $|\psi\rangle_{C\,\mathrm{eff}}$ for the field C. This WF describes the influence of T_A and T_C on the probability of the subset of events $(0110**)$. It is essential that the vector $|\psi\rangle_{C\,\mathrm{eff}}$ has the same form as the initial WF $|\psi\rangle_A$ for the A beam. Hence, $q(Ax, By, Cx) = \alpha/\sqrt{8}$, $q(Ax, By, Cy) = \beta/\sqrt{8}$. Replacing $|Cj\rangle$ according to Eqns (II.7) or, equivalently, replacing $(\alpha, \beta) \to T_C^*(\alpha, \beta)$, we again obtain Eqns (II.7).

Consider a version of the experiment shown in Fig. 5, with no polarization prisms used and only a single detector inserted into each of the beams A, B. In this version, there is no polarization analysis for the photons A and B. (Apparently, it is this version that was used in Ref. [80].) According to Eqns (II.8b), the probability of discovering a single photon in the A beam and a single photon in the B beam, regardless of their polarizations, and likewise a third photon in the mode Cx, is

$$p(A, B, Cx) = 2p(Ax, By, Cx) = \frac{1}{4}\left[|t_C\alpha|^2 + |r_C\,\beta|^2 + 2\,\mathrm{Re}(t_C r_C^*\,\alpha^*\beta)\right]. \quad \text{(II.12)}$$

These events manifest the effect of copying, in contrast to the events where two photons get into a single output beam A or B, their probabilities being

$$2p(2A, Cx) = 2p(2B, Cx) = 2\big[p(2Ax, Cx) + p(2Ay, Cx) + p(Ax, Ay, Cx)\big] = \frac{1}{4}\left[1 + |r_C\alpha|^2 + |t_C\,\beta|^2 - 2\,\mathrm{Re}(t_C r_C^*\alpha^*\beta)\right]. \quad \text{(II.13)}$$

Again, a gate is necessary to exclude these events. Now, the share of 'good' events is $p(A, B, Cx) + p(A, B, Cy) = 1/4$, half that for the version with four controlling detectors and the transformer T_C'.

References

1. Home D, Whittaker M A B *Phys. Rep.* **210** 223 (1992)
2. Sudbery A Quantum Mechanics and the Particles of Nature (Cambridge: Cambridge Univ. Press, 1986) [Translated into Russian (Moscow: Mir, 1989)]
3. Wigner E P *Am. J. Phys.* **31** 6 (1963)
4. Blokhintsev D I *Printsipial'nye Voprosy Kvantovoĭ Mekhaniki* (Principle Problems of Quantum Mechanics) (Moscow: Nauka, 1966); *Kvantovaya Mekhanika. Lektsii po Izbrannym Voprosam* (Quantum Mechanics. Lectures in Selected Problems) (Moscow: Izd. MGU, 1988); *Principles of Quantum Mechanics* (Reidel: Dodrecht, 1964)
5. Mandel'shtam L I *Lektsii po Optike, Teorii Otnositel'nosti i Kvantovoĭ Mekhanike* (Lectures in Optics, Relativity Theory, and Quantum Mechanics) (Moscow: Nauka, 1972)
6. Markov M A *O Trekh Interpretatsiyakh Kvantovoĭ Mekhaniki* (On the Three Interpretations of Quantum Mechanics) (Moscow: Nauka, 1991)
7. Lipkin A I "Paradoksy kvantovoĭ mekhaniki glazami 'realista – empirika', 'konstruktivista – empirika' i 'konstruktivista – ratsionalista'' (Paradoxes of the Quantum Mechanics from the Viewpoints of a 'Realist-empiric', a 'Constructivist-empiric', and a 'Constructivist-rationalist'), in *Filosofiya Nauki* Vyp. 2 (Moscow: IPhRAN, 1996)
8. Sevost'yanov B A *Veroyatnostnye Modeli* (Probability Models) (Moscow: Nauka, 1992)
9. Anderson M H et al. *Science* **269** 198 (1995)
10. Bradley C C et al. *Phys. Rev. Lett.* **75** 1687 (1995)
11. Davis K B et al. *Phys. Rev. Lett.* **75** 3969 (1995)
12. Andrews M R et al. *Science* **275** 637 (1997)
13. *Fundamental Problems in Quantum Theory* (Ann. N.Y. Acad. Sci., Vol. 755, Eds D M Greenberger, A Zeilinger) (New York: New York Academy of Sciences, 1995)
14. *Atom Interferometry* (Ed. P R Berman) (San Diego: Academic Press, 1997); Special Issue on Optics and Interferometry with Atoms: *Appl. Phys. B* **54** (1992)
15. Pritchard D et. al., in Ref. [13] p. 192
16. *Quantum Communications and Measurement* (Eds V P Belavkin, O Hirota, R L Hudson) (New York: Plenum Press, 1995)
17. *Quantum Communication, Computing, and Measurement* (Eds O Hirota, A S Holevo, C M Caves) (New York: Plenum Press, 1997)
18. Vorontsov Yu I *Teoriya i Metody Makroskopicheskikh Izmereniĭ* (Theory and Methods of Macroscopic Measurements) (Moscow: Nauka, 1989)
19. Braginsky V B, Khalili F Ya *Quantum Measurement* (Cambridge, New York: Cambridge Univ. Press, 1992)
20. Glauber R J *Phys. Rev.* **130** 2529 (1963); Glauber R, in *Quantum Optics and Electronics* (Eds C DeWitt, A Blandin, C Cohen-Tannoudji) (New York: Gordon and Breach, 1965) [Translated into Russian (Moscow: Mir, 1966)]
21. Klyshko D N *Phys. Lett. A* **137** 334 (1989)
22. Bohm D *Quantum Theory* (New York: Prentice-Hall, 1952) [Translated into Russian (Moscow: GIFML, 1961)]
23. Special Issue: *J. Mod. Opt.* **44** (11/12) (1997)
24. Smithey D T et al. *Phys. Rev. Lett.* **70** 1244 (1993)
25. Dunn T J, Walmsley I A, Mukamel S *Phys. Rev. Lett.* **74** 884 (1995)
26. Schiller S et al. *Phys. Rev. Lett.* **77** 2933 (1996)
27. Kurtsiefer Ch, Pfau T, Mlynek J *Nature* (London) **386** 150 (1997)
28. Bertrand J, Bertrand P *Found. Phys.* **17** 397 (1987)
29. Vogel K, Risken H *Phys. Rev. A* **40** 2846 (1989)
30. D'Ariano G M, Macchiavello C, Paris M G A *Phys. Rev. A* **50** 4298 (1994)
31. Munroe M et al. *Phys. Rev. A* **52** R924 (1995)
32. Leonhardt U et al. *Opt. Commun.* **127** 144 (1996)
33. Richter Th *Phys. Lett. A* **221** 327 (1996)
34. Janicke U, Wilkens M *J. Mod. Opt.* **42** 2183 (1995); Leichtle C et al. *Phys. Rev. Lett.* **80** 1418 (1998)
35. Klyshko D N *Fizicheskie Osnovy Kvantovoĭ Électroniki* (Physical Foundations of Quantum Electronics) (Moscow: Nauka, 1986)
36. Klyshko D N, Masalov A V *Usp. Fiz. Nauk* **165** 1249 (1995) [*Phys. Usp.* **38** 1203 (1995)]

37. Klyshko D N *Usp. Fiz. Nauk* **164** 1187 (1994) [*Phys. Usp.* **37** 1097 (1994)]
38. Klyshko D N *Phys. Lett. A* **218** 119 (1996); *Laser Phys.* **6** 1056 (1996)
39. Bell J S *Physics* **1** 195 (1964); Clauser J F et al. *Phys. Rev. Lett.* **23** 880 (1969)
40. Clauser J F, Horne M A *Phys. Rev. D* **10** 526 (1974); Clauser J F, Shimony A *Rep. Prog. Phys.* **41** 1881 (1978)
41. Kochen S, Specker E P *J. Math. Mech.* **17** 59 (1967)
42. Santos E *Phys. Rev. A* **45** 3646 (1992)
43. Hillery M *Phys. Rev. A* **39** 2994 (1989)
44. Klyshko D N *Usp. Fiz. Nauk* **166** 613 (1996) [*Phys. Usp.* **39** 573 (1996)]; *Phys. Lett. A* **213** 7 (1996)
45. Lee C T *Phys. Rev. A* **41** 1569, 1721 (1990)
46. Kimble H J, Dagenais M, Mandel L *Phys. Rev. Lett.* **39** 691 (1977)
47. Paul H *Rev. Mod. Phys.* **54** 1061 (1982)
48. Balian R *Am. J. Phys.* **57** 1019 (1989)
49. Kadomtsev B B, Kadomtsev M B *Usp. Fiz. Nauk* **166** 651 (1996) [*Phys. Usp.* **39** 609 (1996)]
50. Margenau H *Ann. Phys.* (N.Y.) **23** 469 (1963)
51. Home D, Whittaker M A B *Phys. Lett. A* **128** 1 (1988)
52. Ballentine L E *Int. J. Theor. Phys.* **27** 211 (1988)
53. Namiki M, Pascazio S, in Ref. [13] p. 335; *Phys. Rev. A* **44** 39 (1993)
54. *Quantum Mechanics without Reduction* (Eds M Sini, J Levy-Leblond) (Bristol: Hilger, 1990)
55. Schiff L I *Quantum Mechanics* (New York: McGraw-Hill, 1955) [Translated into Russian (Moscow: IL, 1959)]
56. Evdokimov N V et.al. *Usp. Fiz. Nauk* **166** 91 (1996) [*Phys. Usp.* **39** 83 (1996)]
57. Klyshko D N *Usp. Fiz. Nauk* **154** 133 (1988) [*Sov. Phys. Usp.* **31** 74 (1988)]; *Usp. Fiz. Nauk* **158** 327 (1989) [*Sov. Phys. Usp.* **32** 555 (1989)]
58. Klyshko D N *Phys. Lett. A* **123** 179 (1998); *Laser Phys.* **8** 363 (1998)
59. Bennet˚ H *Phys. Today* **48** (10) 24 (1995)
60. Sleator T, Weinfurter H, in Ref. [13] p. 715
61. Maitre X et al. *Phys. Rev. Lett.* **79** 769 (1997)
62. Walter H *Usp. Fiz. Nauk* **166** 777 (1996) [*Phys. Usp.* **39** 727 (1996)]
63. Klyshko D N *Zh. Eksp. Teor. Fiz.* **111** 1955 (1997) [*JETP* **84** 1065 (1977)]
64. Chirkin A S, Orlov A A, Parashchuk D Yu *Kvantovaya Elektron.* **20** 999 (1993) [*Quantum Electron.* **23** 870 (1993)]
65. Bennett C H et al. *J. Crypto* **5** 3 (1992)
66. Bennett C H *Phys. Rev. Lett.* **68** 3121 (1992)
67. Special Issue on Quantum Communication: *J. Mod. Opt.* **41** (12) (1994)
68. Bennett C H, Brassard G, Ekert A K *Sci. Am.* **267** (4) 50 (1992)
69. Klyshko D N *Phys. Lett. A* **227** 1 (1997)
70. Klyshko D N *Zh. Eksp. Teor. Fiz.* **105** 1574 (1994) [*JETP* **78** 848 (1994)]
71. Strekalov D V et al. *Phys. Rev. Lett.* **74** 3600 (1995)
72. Pittman T B et al. *Phys. Rev. A* **53** 2804 (1996)
73. Greenberger D M et al. *Am. J. Phys.* **58** 1131 (1990)
74. Klyshko D N *Phys. Lett. A* **172** 399 (1993)
75. Bennett C H et al. *Phys. Rev. Lett.* **70** 1895 (1993)
76. Braunstein S L, Mann A *Phys. Rev. A* **51** R1727 (1995); **53** 630(E) (1996)
77. Davidovich L et al. *Phys. Rev. A* **50** R895 (1994)
78. Cirac J I, Parkins A S *Phys. Rev. A* **50** R4441 (1994)
79. Braunstein S L, Kimble H J *Phys. Rev. Lett.* **80** 869 (1998)
80. Bouwmeester D et al. *Nature* (London) **390** 575 (1997)
81. Boschi D et al. *Phys. Rev. Lett.* **80** 1121 (1998)
82. Zel'dovich B Ya, Klyshko D N *Pis'ma Zh. Eksp. Teor. Fiz.* **9** 69 (1969) [*JETP Lett.* **9** 40 (1969)]
83. Klyshko D N *Kvantovaya Elektron.* **4** 1056 (1977) [*Sov. J. Quantum Electron.* **7** 591 (1977)]
84. Hong C K, Mandel L *Phys. Rev. Lett.* **56** 58 (1986)
85. Shih Y H, Alley C O *Phys. Rev. Lett.* **61** 2921 (1987)
86. Rubin M H, Klyshko D N, Shih Y H, Sergienko A V *Phys. Rev. A* **50** 5122 (1994)
87. Kwiat P G et al. *Phys. Rev. Lett.* **75** 4337 (1995)
88. Raymer M G *Am. J. Phys.* **62** 986 (1994)
89. Tatarskiĭ V I *Usp. Fiz. Nauk* **139** 587 (1983) [*Sov. Phys. Usp.* **26** 311 (1983)]
90. Hillery M et al. *Phys. Rep.* **106** 121 (1984)
91. Wichmann E H *Quantum Physics* (McGraw-Hill, 1974) [Translated into Russian (Moscow: Nauka, 1977)]
92. Cahn S B et al. *Phys. Rev. Lett.* **79** 784 (1997)
93. Frank A I , Nosov V G *Yad. Fiz.* **57** 1029 (1994)
94. Frank A I, Nosov V G, in Ref. [13] p. 293
95. Feynman R P, Leighton R B, Sands M *Feynman Lectures on Physics* (Reading: Addison-Wesley, 1963) [Translated into Russian (Moscow: Mir, 1967)]
96. Aharonov Y, Bohm D *Phys. Rev.* **115** 485 (1959)
97. Olariu S, Popescu I I *Rev. Mod. Phys.* **57** 339 (1985)
98. Hamilton J *Aharonov – Bohm and Other Cyclic Phenomena* (Springer Tracts in Modern Physics, 139) (Berlin, New York: Springer-Verlag, 1997)
99. Leibfried D, Pfau T, Monroe C *Phys. Today* **51** (4) 22 (1998)
100. Lamb W *Phys. Today* **22** 23 (1969) [Translated into Russian *Usp. Fiz. Nauk* **99** 719 (1969)]