

Basics of Quantum Theory for Quantum-Like Modeling Information Retrieval



Andrei Khrennikov

Abstract This chapter contains a brief introduction to the mathematical formalism and axiomatics of quantum mechanics (QM). Recently quantum mathematics and methodology started to be widely used for modeling decision making for humans and AI-systems, including quantum-like modeling information retrieval. Experts in such areas do not go deeply into the details of quantum theory. Moreover, typically such consumers of quantum theory do not use all its components. Quantum measurement theory is the most useful for application, including information retrieval. The main issue is the quantum treatment of incompatible observables represented mathematically by noncommuting Hermitian operators. At the level of statistical data incompatibility is represented as interference of probabilities, in the form of modification of the formula of total probability by adding the interference term.

1 Introduction

Recently the mathematical formalism and methodology of QM, especially theory of quantum measurement, started to be widely applied outside of physics¹: to cognition, psychology, economics, finances, decision making, AI, game theory, and information retrieval (for the latter, see, e.g., [24, 51–53, 56, 58, 59] and the chapters in this book). This chapter contains a brief introduction to the mathematical formalism and axiomatics of QM. It is oriented to non-physicists. Since *QM is a statistical theory* it is natural to start with the classical probability

¹See, for example, [1–10, 15, 27, 28, 32, 38–41, 43–48, 55, 57].

A. Khrennikov (✉)

Linnaeus University, International Center for Mathematical Modeling in Physics and Cognitive Sciences, Växjö, Sweden

e-mail: andrei.khrennikov@lnu.se

© Springer Nature Switzerland AG 2019

D. Aerts et al. (eds.), *Quantum-Like Models for Information Retrieval and Decision-Making*, STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health, https://doi.org/10.1007/978-3-030-25913-6_4

model (Kolmogorov [49]). Then we present basics of theory of Hilbert spaces and Hermitian operators, representation of pure and mixed states by normalized vectors and density operators. This introduction is sufficient to formulate the axiomatics of QM in the form of five postulates. The projection postulate (the most questionable postulate of QM) is presented in a separate section. We distinguish sharply the cases of quantum observables represented by Hermitian operators with nondegenerate and degenerate spectra, the von Neumann's and Lüders' forms of the projection postulate. The axiomatics is completed by a short section on the main interpretations of QM. The projection postulate (in Lüders' form) plays the crucial role in the definition of quantum conditional (transition) probability. By operating with the latter we consider interference of probabilities for two incompatible observables, as a modification of the formula of total probability by adding the interference term. This viewpoint to interference of probabilities was elaborated in a series of works of Khrennikov (see, e.g., [29–38]).

Since classical probability theory is based on the Boolean algebra of events a violation of the law of total probability can be treated as the probabilistic sign of a violation of the laws of the Boolean logics. From this viewpoint, quantum theory can be considered as representing a new kind of logic, the so-called quantum logic. The latter is also briefly presented in a separate section.

We continue this review with a new portion of “quantum mathematics,” namely the notion of the tensor product of Hilbert spaces and the tensor product of operators. After the section on Dirac's notation with ket- and bra-vector, we discuss briefly the notion of qubit and entanglement of a few qubits. This chapter is finished with the presentation of the detailed analysis of the probabilistic structure of the two-slit experiment, as a bunch of different experimental contexts. This contextual structure leads to a violation of the law of total probability and the non-Kolmogorovean probabilistic structure of this experiment.

We hope that this chapter would be interesting for newcomers to quantum-like modeling. May be even experts can find something useful, say the treatment of interference of probabilities as a violation of the law of total probability. In any event, this chapter can serve as the mathematical and foundational introduction to other chapters of this book devoted to the concrete applications.

2 Kolmogorov's Axiomatics of Classical Probability

The main aim of QM is to provide probabilistic predictions on the results of measurements. Moreover, statistics of the measurements of a single quantum observable can be described by classical probability theory. In this section we shall present an elementary introduction to this theory.

We remark that classical probability theory is coupled to experiment in the following way:

- Experimental contexts (system's state preparations) are represented by probabilities.
- Observables are represented by random variables.

In principle, we can call probability a state and this is the direct analog of the quantum state (Sect. 6.1, the ensemble interpretation). However, we have to remember that the word “state” has the meaning “statistical state,” the state of an ensemble of systems prepared for measurement.

The *Kolmogorov probability space* [49, 50] is any triple

$$(\Omega, \mathcal{F}, \mathbf{P}),$$

where Ω is a set of any origin and \mathcal{F} is a σ -algebra of its subsets, \mathbf{P} is a probability measure on \mathcal{F} . The set Ω represents random parameters of the model. Kolmogorov called elements of Ω *elementary events*. This terminology is standard in mathematical literature. Sets of elementary events are regarded as *events*. The key point of Kolmogorov's axiomatization of probability theory is that not any subset of Ω can be treated as an event. For any stochastic model, the system of events \mathcal{F} is selected from the very beginning. The key mathematical point is that \mathcal{F} has to be a σ -algebra. (Otherwise it would be very difficult to develop a proper notion of integral. And the latter is needed to define average of a random variable.)

We remind that a σ -algebra is a system of sets which contains Ω and empty set, it is closed with respect to the operations of countable union and intersection and to the operation of taking the complement of a set. For example, the collection of all subsets of Ω is a σ -algebra. This σ -algebra is used in the case of finite or countable sets:

$$\Omega = \{\omega_1, \dots, \omega_n, \dots\}. \quad (1)$$

However, for “continuous sets,” e.g., $\Omega = [a, b] \subset \mathbf{R}$, the collection of all possible subsets is too large to have applications. Typically it is impossible to describe a σ -algebra in the direct terms. To define a σ -algebra, one starts with a simple system of subsets of Ω and then consider the σ -algebra which is generated from this simple system with the aid of aforementioned operations. In particular, one of the most important for applications σ -algebras, the so-called *Borel σ -algebra*, is constructed in this way by starting with the system consisting of all open and closed intervals of the real line. In a metric space (in particular, in a Hilbert space), the Borel σ -algebra is constructed by starting with the system of all open and closed balls.

Finally, we remark that in American literature the term σ -field is typically used, instead of σ -algebra.

The probability is defined as a *measure*, i.e., a map from \mathcal{F} to nonnegative real numbers which is σ -additive:

$$\mathbf{P}(\cup_j A_j) = \sum_j \mathbf{P}(A_j), \quad (2)$$

where $A_j \in \mathcal{F}$ and $A_i \cap A_j = \emptyset, i \neq j$. The probability measure is always normalized by one:

$$\mathbf{P}(\Omega) = 1. \quad (3)$$

In the case of a discrete probability space, see (1), the probability measures have the form

$$\mathbf{P}(A) = \sum_{\omega_j \in A} p_j, \quad p_j = \mathbf{P}(\{\omega_j\}).$$

In fact, any finite measure μ , i.e., $\mu(\Omega) < \infty$, can be transformed into the probability measure by normalization:

$$\mathbf{P}(A) = \frac{\mu(A)}{\mu(\Omega)}, \quad A \in \mathcal{F}. \quad (4)$$

A (real) random variable is a map $\xi : \Omega \rightarrow \mathbf{R}$ which is measurable with respect to the Borel σ -algebra \mathcal{B} of \mathbf{R} and the σ -algebra \mathcal{F} of Ω . The latter means that, for any set $B \in \mathcal{B}$, its preimage $\xi^{-1}(B) = \{\omega \in \Omega : \xi(\omega) \in B\}$ belongs to \mathcal{F} . This condition provides the possibility to assign the probability to the events of the type “values of ξ belong to a (Borel) subset of the real line.” The probability distribution of ξ is defined as

$$\mathbf{P}_\xi(B) = \mathbf{P}(\xi^{-1}(B)). \quad (5)$$

In the same way we define the real (and complex) vector-valued random variables, $\xi : \Omega \rightarrow \mathbf{R}^n$ and $\xi : \Omega \rightarrow \mathbf{C}^n$.

Let ξ_1, \dots, ξ_k be real-valued random variables. Their joint probability distribution $\mathbf{P}_{\xi_1, \dots, \xi_k}$ is defined as the probability distribution of the vector-valued random variable $\xi = (\xi_1, \dots, \xi_k)$. To determine this probability measure, it is sufficient to define probabilities

$$\mathbf{P}_{\xi_1, \dots, \xi_k}(\Gamma_1 \times \dots \times \Gamma_k) = \mathbf{P}(\omega \in \Omega : \xi_1(\omega) \in \Gamma_1, \dots, \xi_k(\omega) \in \Gamma_k),$$

where $\Gamma_j, j = 1, \dots, k$, are intervals (open, closed, half-open) of the real line.

Suppose now that random variables ξ_1, \dots, ξ_k represent observables a_1, \dots, a_k . For any point $\omega \in \Omega$, the values of the vector ξ composed of these random variables are well defined $\xi(\omega) = (\xi_1(\omega), \dots, \xi_k(\omega))$. This vector represents a *joint measurement* of the observables and $\mathbf{P}_{\xi_1, \dots, \xi_k}$ represents the probability distribution for the outcomes of these jointly measured observables. Thus classical probability theory is applicable for jointly measurable observables, compatible observables in the terminology of QM (Sect. 5).

A random variable is called discrete if its image consists of finite or countable number of points, $\xi = \alpha_1, \dots, \alpha_n, \dots$. In this case its probability distribution has the form

$$\mathbf{P}(B) = \sum_{\alpha_j \in B} P_{\alpha_j}, \quad P_{\alpha_j} = \mathbf{P}(\omega \in \Omega : \xi(\omega) = \alpha_j). \quad (6)$$

The mean value (average) of a real-valued random variable is defined as its integral (the Lebesgue integral)

$$E\xi = \int_{\Omega} \xi(\omega) dP(\omega). \quad (7)$$

For a discrete random variable, its mean value has the simple representation:

$$E\xi = \sum_{\alpha_j \in B} \alpha_j P_{\alpha_j}. \quad (8)$$

In the Kolmogorov model the conditional probability is *defined* by the *Bayes formula*

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(B \cap A)}{\mathbf{P}(A)}, \quad \mathbf{P}(A) > 0. \quad (9)$$

We stress that other axioms are independent of this definition.

We also present the *formula of total probability* (FTP) which is a simple consequence of the Bayes formula. Consider the pair, a and b , of discrete random variables. Then

$$\mathbf{P}(b = \beta) = \sum_{\alpha} \mathbf{P}(a = \alpha) \mathbf{P}(b = \beta | a = \alpha). \quad (10)$$

Thus the b -probability distribution can be calculated from the a -probability distribution and the conditional probabilities $\mathbf{P}(b = \beta | a = \alpha)$. These conditional probabilities are known as *transition probabilities*.

This formula plays the crucial role in Bayesian inference. It is applicable to the plenty of phenomena, in insurance, finances, economics, engineering, biology, AI, game theory, decision making, and programming. However, as was shown by the author of this review, in quantum domain FTP is violated and it is perturbed by the so-called interference term. Recently it was shown that even data collected in cognitive science, psychology, game theory, and decision making can violate classical FTP [1–10, 18, 38–41, 44–48].

3 Quantum Mathematics

We present the basic mathematical structures of QM and couple them to quantum physics.

3.1 Hermitian Operators in Hilbert Space

We recall the definition of a complex Hilbert space. Denote it by H . This is a complex linear space endowed with a scalar product (a positive-definite nondegenerate Hermitian form) which is complete with respect to the norm corresponding to the scalar product, $\langle \cdot | \cdot \rangle$. The norm is defined as

$$\|\phi\| = \sqrt{\langle \phi | \phi \rangle}.$$

In the finite-dimensional case the norm and, hence, completeness are of no use. Thus those who have no idea about functional analysis (but know essentials of linear algebra) can treat H simply as a finite-dimensional complex linear space with the scalar product.

For a complex number $z = x + iy$, $x, y \in \mathbf{R}$, its conjugate is denoted by \bar{z} , here $\bar{z} = x - iy$. The absolute value of z is given by $|z|^2 = z\bar{z} = x^2 + y^2$.

For reader's convenience, we recall that the scalar product is a function from the Cartesian product $H \times H$ to the field of complex numbers \mathbb{C} , $\psi_1, \psi_2 \rightarrow \langle \psi_1 | \psi_2 \rangle$, having the following properties:

1. Positive-definiteness: $\langle \psi | \psi \rangle \geq 0$ with $\langle \psi | \psi \rangle = 0$ if and only if $\psi = 0$.
2. Conjugate symmetry: $\langle \psi_1 | \psi_2 \rangle = \overline{\langle \psi_2 | \psi_1 \rangle}$
3. Linearity with respect to the second argument²: $\langle \phi | k_1 \psi_1 + k_2 \psi_2 \rangle = k_1 \langle \phi | \psi_1 \rangle + k_2 \langle \phi | \psi_2 \rangle$, where k_1, k_2 are the complex numbers.

A reader who does not feel comfortable in the abstract framework of functional analysis can simply proceed with the Hilbert space $H = \mathbf{C}^n$, where \mathbf{C} is the set of complex numbers, and the scalar product

$$\langle u | v \rangle = \sum_i u_i \bar{v}_i, u = (u_1, \dots, u_n), v = (v_1, \dots, v_n). \quad (11)$$

In this case the above properties of a scalar product can be easily derived from (11). Instead of linear operators, one can consider matrices.

²In mathematical texts one typically considers linearity with respect to the first argument. Thus a mathematician has to pay attention to this difference.

We also recall a few basic notions of theory of linear operators in complex Hilbert space. A map $\hat{a} : H \rightarrow H$ is called a linear operator, if it maps linear combination of vectors into linear combination of their images:

$$\hat{a}(\lambda_1\psi_1 + \lambda_2\psi_2) = \lambda_1\hat{a}\psi_1 + \lambda_2\hat{a}\psi_2,$$

where $\lambda_j \in \mathbf{C}$, $\psi_j \in H$, $j = 1, 2$.

For a linear operator \hat{a} the symbol \hat{a}^* denotes its *adjoint operator* which is defined by the equality

$$\langle \hat{a}\psi_1 | \psi_2 \rangle = \langle \psi_1 | \hat{a}^* \psi_2 \rangle. \quad (12)$$

Let us select in H some orthonormal basis (e_i) , i.e., $\langle e_i | e_j \rangle = \delta_{ij}$. By denoting the matrix elements of the operators \hat{a} and \hat{a}^* as a_{ij} and a_{ij}^* , respectively, we rewrite the definition (12) in terms of the matrix elements:

$$a_{ij}^* = \bar{a}_{ji}.$$

A linear operator \hat{a} is called *Hermitian* if it coincides with its adjoint operator:

$$\hat{a} = \hat{a}^*.$$

If an orthonormal basis in H is fixed, (e_i) , and \hat{a} is represented by its matrix, $A = (a_{ij})$, where $a_{ij} = \langle \hat{a}e_i | e_j \rangle$, then it is Hermitian if and only if

$$\bar{a}_{ij} = a_{ji}.$$

We remark that, for a Hermitian operator, all its eigenvalues are real. In fact, this was one of the main reasons to represent quantum observables by Hermitian operators. In the quantum formalism, the spectrum of a linear operator (the set of eigenvalues while we are in the finite-dimensional case) coincides with the set of possibly observable values (Sect. 4, **Postulate 3**). We also recall that eigenvectors of Hermitian operators corresponding to different eigenvalues are orthogonal. This property of Hermitian operators plays some role in justification of the projection postulate of QM, see Sect. 5.1.

A linear operator \hat{a} is *positive-semidefinite* if, for any $\phi \in H$,

$$\langle \hat{a}\phi | \phi \rangle \geq 0.$$

This is equivalent to positive-semidefiniteness of its matrix.

For a linear operator \hat{a} , its trace is defined as the sum of diagonal elements of its matrix in any orthonormal basis:

$$\text{Tr } \hat{a} = \sum_i a_{ii} = \sum_i \langle \hat{a}e_i | e_i \rangle,$$

i.e., this quantity does not depend on a basis.

Let L be a subspace of H . The orthogonal projector $P : H \rightarrow L$ onto this subspace is a Hermitian, idempotent (i.e., coinciding with its square), and positive-semidefinite operator³:

- (a) $P^* = P$;
- (b) $P^2 = P$;
- (c) $P \geq 0$.

Here (c) is a consequence of (a) and (b). Moreover, an arbitrary linear operator satisfying (a) and (b) is an orthogonal projector—onto the subspace PH .

3.2 *Pure and Mixed States: Normalized Vectors and Density Operators*

Pure quantum states are represented by normalized vectors, $\psi \in H : \|\psi\| = 1$. Two colinear vectors,

$$\psi' = \lambda\psi, \lambda \in \mathbf{C}, |\lambda| = 1, \quad (13)$$

represent the same pure state. Thus, rigorously speaking, a pure state is an equivalence class of vectors having the unit norm: $\psi' \sim \psi$ for vectors coupled by (13). The unit sphere of H is split into disjoint classes—pure states. However, in concrete calculations one typically uses just concrete representatives of equivalent classes, i.e., one works with normalized vectors.

Each pure state can also be represented as the projection operator P_ψ which projects H onto one dimensional subspace based on ψ . For a vector $\phi \in H$,

$$P_\psi\phi = \langle\phi|\psi\rangle\psi. \quad (14)$$

The trace of the one dimensional projector P_ψ equals 1:

$$\text{Tr } P_\psi = \langle\psi|\psi\rangle = 1.$$

We summarize the properties of the operator P_ψ representing the pure state ψ . It is

- (a) Hermitian,
- (b) positive-semidefinite,
- (c) trace one,
- (d) idempotent.

³To simplify formulas, we shall not put the operator-label “hat” in the symbols denoting projectors, i.e., $P \equiv \hat{P}$.

Moreover, any operator satisfying (a)–(d) represents a pure state. Properties (a) and (d) characterize orthogonal projectors, property (b) is their consequence. Property (c) implies that the projector is one dimensional.

The next step in the development of QM was the extension of the class of quantum states, from pure states represented by one dimensional projectors to states represented by linear operators having the properties (a)–(c). Such operators are called *density operators*. (This nontrivial step of extension of the class of quantum states was based on the efforts of Landau and von Neumann.) The symbol $D(H)$ denotes the space of density operators in the complex Hilbert space H .

One typically distinguishes pure states, as represented by one dimensional projectors, and mixed states, the density operators which cannot be represented by one dimensional projectors. The terminology “mixed” has the following origin: any density operator can be represented as a “mixture” of pure states (ψ_i) :

$$\rho = \sum_i p_i P_{\psi_i}, \quad p_i \in [0, 1], \quad \sum_i p_i = 1. \quad (15)$$

(To simplify formulas, we shall not put the operator-label “hat” in the symbols denoting density operators, i.e., $\rho \equiv \hat{\rho}$.) The state is pure if and only if such a mixture is trivial: all p_i , besides one, equal zero. However, by operating with the terminology “mixed state” one has to take into account that the representation in the form (15) is not unique. The same mixed state can be presented as mixtures of different collections of pure states.

Any operator ρ satisfying (a)–(c) is diagonalizable (even in infinite-dimensional Hilbert space), i.e., in some orthonormal basis it is represented as a diagonal matrix, $\rho = \text{diag}(p_j)$, where $p_j \in [0, 1]$, $\sum_j p_j = 1$. Thus it can be represented in the form (15) with mutually orthogonal one dimensional projectors. The property (d) can be used to check whether a state is pure or not.

We point out that *pure states are merely mathematical abstractions; in real experimental situations, it is possible to prepare only mixed states*. The degree of *purity* is defined as

$$\text{purity}(\rho) = \text{Tr}\rho^2.$$

Experimenters are satisfied by getting this quantity near one.

4 Quantum Mechanics: Postulates

We state again that H denotes complex Hilbert space with the scalar product $\langle \cdot, \cdot \rangle$ and the norm $\| \cdot \|$ corresponding to the scalar product.

Postulate 1 (The Mathematical Description of Quantum States) *Quantum (pure) states (wave functions) are represented by normalized vectors ψ (i.e., $\|\psi\|^2 = \langle \psi, \psi \rangle = 1$) of a complex Hilbert space \mathcal{H} . Every normalized vector $\psi \in H$ may*

represent a quantum state. If a vector ψ corresponding to a state is multiplied by any complex number c , $|c| = 1$, the resulting vector will correspond to the same state.⁴

The physical meaning of “a quantum state” is not defined by this postulate, see Sect. 6.1.

Postulate 2 (The Mathematical Description of Physical Observables) *A physical observable a is represented by a Hermitian operator \hat{a} in complex Hilbert space H . Different observables are represented by different operators.*

Postulate 3 (Spectral) *For a physical observable a which is represented by the Hermitian operator \hat{a} we can predict (together with some probabilities) values $\lambda \in \text{Spec}(\hat{a})$.*

We restrict our considerations by simplest Hermitian operators which are analogous to discrete random variables in classical probability theory. We recall that a Hermitian operator \hat{a} has *purely discrete spectrum* if it can be represented as

$$\hat{a} = \alpha_1 P_{\alpha_1}^a + \cdots + \alpha_m P_{\alpha_m}^a + \cdots, \quad \alpha_m \in \mathbf{R}, \quad (16)$$

where $P_{\alpha_m}^a$ are orthogonal projection operators related to the orthonormal eigenvectors $\{e_{km}^a\}_k$ of \hat{a} corresponding to the eigenvalues α_m by

$$P_{\alpha_m}^a \psi = \sum_k \langle \psi, e_{km}^a \rangle e_{km}^a, \quad \psi \in H. \quad (17)$$

Here k labels the eigenvectors e_{km}^a which belong to the same eigenvalue α_m of \hat{a} .

Postulate 4 (Born’s Rule) *Let a physical observable a be represented by a Hermitian operator \hat{a} with purely discrete spectrum. The probability $\mathbf{P}_\psi(a = \alpha_m)$ to obtain the eigenvalue α_m of \hat{a} for measurement of a in a state ψ is given by*

$$\mathbf{P}_\psi(a = \alpha_m) = \|P_{\alpha_m}^a \psi\|^2. \quad (18)$$

If the operator \hat{a} has nondegenerate (purely discrete) spectrum, then each α_m is associated with one dimensional subspace. The latter can be fixed by selecting any normalized vector, say e_m^a . In this case orthogonal projectors act simply as

$$P_{\alpha_m}^a \psi = \langle \psi, e_m^a \rangle e_m^a. \quad (19)$$

Formula (18) takes a very simple form

$$\mathbf{P}_\psi(a = \alpha_m) = |\langle \psi, e_m^a \rangle|^2. \quad (20)$$

⁴Thus states are given by elements of the unit sphere of the Hilbert space H .

It is Born's rule in the Hilbert space formalism.

It is important to point out that *if state ψ is an eigenstate of operator \hat{a} representing observable a , i.e., $\hat{a}\psi = \alpha\psi$, then the outcome of observable a equals α with probability one.*

We point out that, for any fixed quantum state ψ , each quantum observable \hat{a} can be represented as a classical random variable (Sect. 2). In the discrete case the corresponding probability distribution is defined as

$$\mathbf{P}(A) = \sum_{\alpha_m \in A} \mathbf{P}_\psi(a = \alpha_m),$$

where $\mathbf{P}_\psi(a = \alpha_m)$ is given by Born's rule.

Thus each concrete quantum measurement can be described by the classical probability model.

Problems (including deep interpretational issues) arise only when one tries to describe classically data collected for a few incompatible observables (Sect. 5).

By using the Born's rule (18) and the classical probabilistic definition of average (Sect. 2), it is easy to see that the average value of the observable a in the state ψ (belonging to the domain of definition of the corresponding operator \hat{a}) is given by

$$\langle a \rangle_\psi = \langle \hat{a} \psi, \psi \rangle. \quad (21)$$

For example, for an observable represented by an operator with the purely discrete spectrum, we have

$$\langle a \rangle_\psi = \sum_m \alpha_m \mathbf{P}_\psi(a = \alpha_m) = \sum_m \alpha_m \|P_m^a \psi\|^2 = \langle \hat{a} \psi, \psi \rangle.$$

Postulate 5 (Time Evolution of Wave Function) *Let $\hat{\mathcal{H}}$ be the Hamiltonian of a quantum system, i.e., the Hermitian operator corresponding to the energy observable. The time evolution of the wave function $\psi \in H$ is described by the Schrödinger equation*

$$i \frac{d}{dt} \psi(t) = \hat{\mathcal{H}} \psi(t) \quad (22)$$

with the initial condition $\psi(0) = \psi$.

5 Compatible and Incompatible Observables

Two observables a and b are called *compatible* if a measurement procedure for their joint measurement can be designed, i.e., a measurement of the vector observable $d = (a, b)$. In such a case their joint probability distribution is well defined.

In the opposite case, i.e., when such a joint-measurement procedure does not exist, observables are called *incompatible*. The joint probability distribution of incompatible observables has no meaning.

In QM, compatible observables a and b are represented by commuting Hermitian operators \widehat{a} and \widehat{b} , i.e., $[\widehat{a}, \widehat{b}] = 0$; consequently, incompatible observables a and b are represented by noncommuting operators, i.e., $[\widehat{a}, \widehat{b}] \neq 0$. Thus in the QM-formalism compatibility–incompatibility is represented as commutativity–noncommutativity.

Postulate 4a (Born’s Rule for Joint Measurements) *Let observables a and b be represented by Hermitian operators \widehat{a} and \widehat{b} with purely discrete spectrum. The probability to obtain the eigenvalues α_m and β_k in a joint measurement of a and b in a state ψ —the joint probability distribution—is given by*

$$\mathbf{P}_\psi(a = \alpha_m, b = \beta_k) = \|P_k^b P_m^a \psi\|^2 = \|P_m^a P_k^b \psi\|^2. \quad (23)$$

It is crucial that the spectral projectors of commuting operators commute, so the probability distribution does not depend on the order of the values of observables. This is a classical probability distribution (Sect. 2). Any pair of compatible observables a and b can be represented by random variables: for example, by using the joint probability distribution as the probability measure.

A family of compatible observables a_1, \dots, a_n is represented by commuting Hermitian operators $\widehat{a}_1, \dots, \widehat{a}_n$, i.e., $[\widehat{a}_i, \widehat{a}_j] = 0$ for all pairs i, j . The joint probability distribution is given by the natural generalization of rule (23):

$$\begin{aligned} \mathbf{P}_{a_1, \dots, a_n; \psi}(\alpha_{1m}, \dots, \alpha_{nk}) &\equiv \mathbf{P}_\psi(a_1 = \alpha_{1m}, \dots, a_n = \alpha_{nk}) \\ &= \|P_k^{a_n} \dots P_m^{a_1} \psi\|^2 = \dots = \|P_m^{a_1} \dots P_k^{a_n} \psi\|^2, \end{aligned} \quad (24)$$

where all possible permutations of projectors can be considered.

Now we point to one distinguishing feature of compatibility of *quantum observables* which is commonly not emphasized. The relation of commutativity of operators is the pairwise relation, it does not involve say triples of operators. Thus, for joint measurability of a group of quantum observables a_1, \dots, a_n , their pairwise joint measurability is sufficient. Thus if we are able to design measurement procedures for all possible pairs, then we are always able to design a joint-measurement procedure for the whole group of quantum observables a_1, \dots, a_n . This is the specialty of quantum observables. In particular, if there exist all pairwise joint probability distributions $\mathbf{P}_{a_i, a_j; \psi}$, then the joint probability $\mathbf{P}_{a_1, \dots, a_n; \psi}$ is defined as well.

The Born’s rule can be generalized to the quantum states represented by density operators (Sect. 9.1, formula (40)).

5.1 Post-Measurement State From the Projection Postulate

The projection postulate is one of the most questionable and debatable postulates of QM. We present it in the separate section to distinguish it from other postulates of QM, Postulates 1–5, which are commonly accepted.

Consider pure state ψ and quantum observable (Hermitian operator) \hat{a} representing some physical observable a . Suppose that \hat{a} has nondegenerate spectrum; denote its eigenvalues by $\alpha_1, \dots, \alpha_m, \dots$ and the corresponding eigenvectors by $e_1^a, \dots, e_m^a, \dots$ (here $\alpha_i \neq \alpha_j, i \neq j$.) This is an orthonormal basis in H . We expand the vector ψ with respect to this basis:

$$\psi = k_1 e_1^a + \dots + k_m e_m^a + \dots, \quad (25)$$

where (k_j) are complex numbers such that

$$\|\psi\|^2 = |k_1|^2 + \dots + |k_m|^2 + \dots = 1. \quad (26)$$

By using the terminology of linear algebra we say that the pure state ψ is a *superposition* of the pure states e_j . The von Neumann projection postulate describes the post-measurement state and it can be formulated as follows:

Postulate 6VN (Projection Postulate, von Neumann) *Measurement of observable a resulting in output α_i induces reduction of superposition (25) to the basis vector e_i^a .*

The procedure described by the projection postulate can be interpreted in the following way:

Superposition (25) reflects uncertainty in the results of measurements for an observable a . Before measurement a quantum system “does not know how it will answer to the question a .” The Born’s rule presents potentialities for different answers. Thus a quantum system in the superposition state ψ does not have propensity to any value of a as its objective property. After the measurement the superposition is reduced to the single term in the expansion (25) corresponding to the value of a obtained in the process of measurement.

Consider now an arbitrary quantum observable a with purely discrete spectrum, i.e., $\hat{a} = \alpha_1 P_{\alpha_1}^a + \dots + \alpha_m P_{\alpha_m}^a + \dots$. The Lüders projection postulate describes the post-measurement state and it can be formulated as follows:

Postulate 6L (Projection Postulate, Lüders) *Measurement of observable a resulting in output α_m induces projection of state ψ on state*

$$\psi_{\alpha_m} = \frac{P_{\alpha_m}^a \psi}{\|P_{\alpha_m}^a \psi\|}.$$

In contrast to the majority of books on quantum theory, we sharply distinguish the cases of quantum observables with nondegenerate and degenerate spectra. von

Neumann formulated Postulate 6VN only for observables with nondegenerate spectra. Lüders “generalized” von Neumann’s postulate to the case of observables with degenerate spectra. However, for such observables, von Neumann formulated [60] a postulate which is different from Postulate 6L. The post-measurement state need not be again a pure state.

We remark that Postulate 6L can be applied even to quantum states which are represented by density operators (Sect. 9.1, formula (41)).

6 Interpretations of Quantum Mechanics

Now we are going to discuss one of the most important and complicated issues of quantum foundations, the problem of an interpretation of a quantum state. There were elaborated numerous interpretations which can differ crucially from each other. This huge diversity of interpretations is a sign of the deep crises in quantum foundations.

In this section, we briefly discuss a few basic interpretations. Then in Sect. 9.1 we shall consider the Växjö (realist ensemble contextual) interpretation. Its presentation needs additional mathematical formulas. Therefore we placed it into a separate section.

6.1 Ensemble and Individual Interpretations

The Ensemble Interpretation *A (pure) quantum state provides a description of certain statistical properties of an ensemble of similarly prepared quantum systems.*

This interpretation is upheld, for example, by Einstein, Popper, Blokhintsev, Margenau, Ballentine, Klyshko, and recent years by, e.g., De Muynck, De Baere, Holevo, Santos, Khrennikov, Nieuwenhuizen, Adenier, Groessing, and many others.

The Copenhagen Interpretation *A (pure) quantum state provides the complete description of an individual quantum system.*

This interpretation was supported by a great variety of members, from Schrödinger’s original attempt to identify the electron with a wave function solution of his equation to the several versions of the Copenhagen interpretation [12–14, 53, 54] (for example, Heisenberg, Bohr, Pauli, Dirac, von Neumann, Landau, Fock, and recent years by, e.g., Greenberger, Mermin, Lahti, Peres, Summhammer). Nowadays the individual interpretation is extremely popular, especially in quantum information and computing.

Instead of Einstein’s terminology “*ensemble interpretation*,” Ballentine [7–9] used the terminology “*statistical interpretation*.” However, Ballentine’s terminology is rather misleading, because the term “*statistical interpretation*” was also used

by von Neumann [60] for individual randomness! For him “statistical interpretation” had the meaning which is totally different from the Ballentine’s “ensemble-statistical interpretation.” J. von Neumann wanted to emphasize the difference between deterministic (Newtonian) classical mechanics in that the state of a system is determined by values of two observables (position and momentum) and quantum mechanics in that the state is determined not by values of observables, but by probabilities. We shall follow Albert Einstein and use the terminology *ensemble interpretation*.

We remark that following von Neumann [60] the supporters of the individual interpretation believe in *irreducible quantum randomness*, i.e., that the behavior of an individual quantum system is irreducibly random. Why does it behave in such a way? Because it is quantum, so it can behave so unusually. Nowadays this von Neumann’s claim is used to justify superiority of the quantum technology over the classical technology. For example, superiority of quantum random generators.

6.2 Information Interpretations

The quantum information revolution generated a variety of information interpretations of QM (see, for example, [16, 17, 19, 20]). By these interpretations the quantum formalism describes special way of information processing, more general than the classical information processing. Roughly speaking one can forget about physics and work solely with probability, entropy, and information. Quantum Bayesianism (QBism) [25, 26] can be considered as one of such information, in its extreme form: the quantum formalism describes very general scheme of assignments of subjective probabilities to possible outcomes of experiments, assignment by human agents.

7 Quantum Conditional (Transition) Probability

In the classical Kolmogorov probabilistic model (Sect. 2), besides probabilities one operates with the conditional probabilities defined by the *Bayes formula* (see Sect. 2, formula (9)). The Born’s postulate defining quantum probability should also be completed by a definition of the conditional probability. We have remarked that, for one concrete observable, the probability given by Born’s rule can be treated classically. However, the definition of the conditional probability involves two observables. Such situations cannot be treated classically (at least straightforwardly, cf. Sect. 2). Thus conditional probability is really a quantum probability.

Let physical observables a and b be represented by Hermitian operators with purely discrete (may be degenerate) spectra:

$$\hat{a} = \sum_m \alpha_m P_{\alpha_m}^a, \quad \hat{b} = \sum_m \beta_m P_{\beta_m}^b. \quad (27)$$

Let ψ be a pure state and let $P_{\alpha_k}^a \psi \neq 0$. Then the probability to get the value $b = \beta_m$ under the condition that the value $a = \alpha_k$ was observed in the preceding measurement of the observable a on the state ψ is given by the formula

$$\mathbf{P}_\psi(b = \beta_m | a = \alpha_k) \equiv \frac{\|P_{\beta_m}^b P_{\alpha_k}^a \psi\|^2}{\|P_{\alpha_k}^a \psi\|^2}. \quad (28)$$

One can motivate this definition by appealing to the projection postulate (Lüders' version). After the a -measurement with output $a = \alpha_k$ initially prepared state ψ is projected onto the state

$$\psi_{\alpha_k} = \frac{P_{\alpha_k}^a \psi}{\|P_{\alpha_k}^a \psi\|}.$$

Then one applies Born's rule to the b -measurement for this state.

Let the operator \hat{a} has nondegenerate spectrum, i.e., for any eigenvalue α the corresponding eigenspace (i.e., generated by eigenvectors with $\hat{a}\psi = \alpha\psi$) is one dimensional. We can write

$$\mathbf{P}_\psi(b = \beta_m | a = \alpha_k) = \|P_{\beta_m}^b e_{\alpha_k}^a\|^2 \quad (29)$$

(here $\hat{a}e_{\alpha_k}^a = \alpha_k e_{\alpha_k}^a$). Thus the conditional probability in this case does not depend on the original state ψ . We can say that the memory about the original state was destroyed. If also the operator \hat{b} has nondegenerate spectrum, then we have: $\mathbf{P}_\psi(b = \beta_m | a = \alpha_k) = |\langle e_{\beta_m}^b, e_{\alpha_k}^a \rangle|^2$ and $\mathbf{P}_\psi(a = \alpha_k | b = \beta_m) = |\langle e_{\alpha_k}^a, e_{\beta_m}^b \rangle|^2$. By using symmetry of the scalar product we obtain the following result:

Let both operators \hat{a} and \hat{b} have purely discrete nondegenerate spectra and let $P_k^a \psi \neq 0$ and $P_m^b \psi \neq 0$. Then conditional probability is symmetric and it does not depend on the original state ψ :

$$\mathbf{P}_\psi(b = \beta_m | a = \alpha_k) = \mathbf{P}_\psi(a = \alpha_k | b = \beta_m) = |\langle e_{\beta_m}^b, e_{\alpha_k}^a \rangle|^2. \quad (30)$$

8 Observables with Nondegenerate Spectra: Double-Stochasticity of the Matrix of Transition Probabilities

We remark that classical (Kolmogorov–Bayes) conditional probability is not symmetric, besides very special situations. Thus *QM is described by a very specific probabilistic model.*

Consider two nondegenerate observables. Set $p_{\beta|\alpha} = \mathbf{P}(b = \beta | a = \alpha)$. The matrix of transition probabilities $\mathbf{P}^{b|a} = (p_{\beta|\alpha})$ is not only *stochastic*, i.e.,

$$\sum_{\beta} p_{\beta|\alpha} = 1$$

but it is even *doubly stochastic*:

$$\sum_{\alpha} p_{\beta|\alpha} = \sum_{\alpha} |\langle e_{\beta}^b, e_{\alpha}^a \rangle|^2 = \langle e_{\beta}^b, e_{\beta}^b \rangle = 1.$$

In Kolmogorov's model, stochasticity is the general property of matrices of transition probabilities. However, in general classical matrices of transition probabilities are not doubly stochastic. Hence, double stochasticity is a very specific property of quantum probability.

We remark that statistical data collected outside quantum physics, e.g., in decision making by humans and psychology, violates the quantum law of double stochasticity [38]. Such data cannot be mathematically represented with the aid of Hermitian operators with nondegenerate spectra. One has to consider either Hermitian operators with degenerate spectra or positive operator valued measures (POVMs).

9 Formula of Total Probability with the Interference Term

We shall show that the quantum probabilistic calculus violates the conventional FTP (10), one of the basic laws of classical probability theory. In this section, we proceed in the abstract setting by operating with two abstract incompatible observables. The concrete realization of this setting for the two-slit experiment demonstrating interference of probabilities in QM will be presented in Sect. 16 which is closely related to Feynman's claim [22, 23] on the nonclassical probabilistic structure of this experiment.

Let $H_2 = \mathbf{C} \times \mathbf{C}$ be the two dimensional complex Hilbert space and let $\psi \in H_2$ be a quantum state. Let us consider two dichotomous observables $b = \beta_1, \beta_2$ and $a = \alpha_1, \alpha_2$ represented by Hermitian operators \widehat{b} and \widehat{a} , respectively (one may consider simply Hermitian matrices). Let $e^b = \{e_{\beta}^b\}$ and $e^a = \{e_{\alpha}^a\}$ be two orthonormal bases consisting of eigenvectors of the operators. The state ψ can be represented in the two ways

$$\psi = c_1 e_1^a + c_2 e_2^a, \quad c_{\alpha} = \langle \psi, e_{\alpha}^a \rangle; \quad (31)$$

$$\psi = d_1 e_1^b + d_2 e_2^b, \quad d_{\beta} = \langle \psi, e_{\beta}^b \rangle. \quad (32)$$

By Postulate 4 we have

$$\mathbf{P}(a = \alpha) \equiv \mathbf{P}_{\psi}(a = \alpha) = |c_{\alpha}|^2. \quad (33)$$

$$\mathbf{P}(b = \beta) \equiv \mathbf{P}_{\psi}(b = \beta) = |d_{\beta}|^2. \quad (34)$$

The possibility to expand one basis with respect to another basis induces connection between the probabilities $\mathbf{P}(a = \alpha)$ and $\mathbf{P}(b = \beta)$. Let us expand the vectors e_α^a with respect to the basis e^b

$$e_1^a = u_{11}e_1^b + u_{12}e_2^b \quad (35)$$

$$e_2^a = u_{21}e_1^b + u_{22}e_2^b, \quad (36)$$

where $u_{\alpha\beta} = \langle e_\alpha^a, e_\beta^b \rangle$. Thus $d_1 = c_1u_{11} + c_2u_{21}$, $d_2 = c_1u_{12} + c_2u_{22}$. We obtain the *quantum rule* for transformation of probabilities:

$$\mathbf{P}(b = \beta) = |c_1u_{1\beta} + c_2u_{2\beta}|^2. \quad (37)$$

On the other hand, by the definition of quantum conditional probability, see (28), we obtain

$$\mathbf{P}(b = \beta|a = \alpha) \equiv \mathbf{P}_\psi(b = \beta|a = \alpha) = |\langle e_\alpha^a, e_\beta^b \rangle|^2. \quad (38)$$

By combining (33), (34) and (37), (38) we obtain the *quantum formula of total probability—the formula of the interference of probabilities*:

$$\mathbf{P}(b = \beta) = \sum_{\alpha} \mathbf{P}(a = \alpha)\mathbf{P}(b = \beta|a = \alpha) \quad (39)$$

$$+ 2 \cos \theta \sqrt{\mathbf{P}(a = \alpha_1)\mathbf{P}(b = \beta|a = \alpha_1)\mathbf{P}(a = \alpha_2)\mathbf{P}(b = \beta|a = \alpha_2)}.$$

In general $\cos \theta \neq 0$. Thus the quantum FTP does not coincide with the classical FTP (10) which is based on the Bayes' formula.

We presented the derivation of the quantum FTP only for observables given by Hermitian operators acting in the two dimensional Hilbert space and for pure states. In Sect. 9.1, we give (without proving) the formula for spaces of an arbitrary dimension and states represented by density operators (see [42] for quantum FTP for observables represented by POVMs).

9.1 Växjö (Realist Ensemble Contextual) Interpretation of Quantum Mechanics

The Växjö interpretation [33] is the *realist ensemble contextual interpretation* of QM. Thus, in contrast to Copenhagenists or QBists, by the Växjö interpretation QM is not complete and it can be emergent from a subquantum model. This interpretation is the ensemble interpretation. This interpretation is contextual, i.e., experimental contexts have to be taken into account really seriously.

By the Växjö interpretation *the probabilistic part of QM is a special mathematical formalism to work with contextual probabilities for families of contexts*, which are, in general, incompatible. Of course, the quantum probabilistic formalism is not the only possible formalism to operate with contextual probabilities.

The main distinguishing feature of the formalism of quantum probability is its complex Hilbert space representation and the Born's rule. *All quantum contexts can be unified with the aid of a quantum state ψ* (wave function, complex probability amplitude). A state represents only a part of context, the second part is given by an observable a . Thus the quantum probability model is not just a collection of Kolmogorov probability spaces. These spaces are coupled by quantum states.

Each theory of probability has two main purposes: descriptive and predictive. In classical probability theory its predictive machinery is based on *Bayesian inference* and, in particular, FTP (Sect. 2, formula (10)).

Can the probabilistic formalism of QM be treated as a generalization of Bayesian inference?

My viewpoint is that the quantum FTP with the interference term (Sect. 9, formula (39)) is, in fact, a modified rule for the probability update. QM provides the following inference machinery. There are given a mixed state represented by density operator ρ and two quantum observables a and b represented mathematically by Hermitian operators \hat{a} and \hat{b} with purely discrete spectra. The first measurement of a can be treated as collection of information about the state ρ . The result $a = \alpha_i$ appears with the probability

$$p^a(\alpha_i) = \text{Tr} P_i^a \rho. \quad (40)$$

This is generalization of the Born's rule to mixed states.

Postulate 6L (the projection postulate in the Lüders' form) can be extended to mixed states. Initial state ρ is transferred to the state

$$\rho_{a_i} = \frac{P_i^a \rho P_i^a}{\text{Tr} P_i^a \rho P_i^a}. \quad (41)$$

Then, for each state ρ_{a_i} , we perform measurement of b and obtain probabilities

$$p(\beta_j | \alpha_i) = \text{Tr} P_j^b \rho_{a_i}. \quad (42)$$

These are quantum conditional (transition) probabilities for the initial state given by a density operator (generalization of the formalism of Sect. 7).

We now recall the general form of the quantum FTP [42]:

$$p(b = \beta) = \sum_k p(b = \beta | a = \alpha_k) p(a = \alpha_k) \quad (43)$$

$$+ 2 \sum_{k < m} \cos \phi_{j;k,m} \sqrt{p(b = \beta | a = \alpha_k) p(a = \alpha_k) p(b = \beta | \alpha_m) p(a = \alpha_m)}.$$

Thus we can predict the probability of the result β_j for the b -observable on the basis of the probabilities for the results α_i for the a -observable and conditional probabilities. Of course, the main nonclassical feature of this probability update rule is the presence of phase angles. In the case of dichotomous observables of the von Neumann–Lüders type the phase angles ϕ_j can be expressed in terms of probabilities.

10 Quantum Logic

von Neumann and Birkhoff [11, 61] suggested to represent *events* (propositions) by orthogonal projectors in complex Hilbert space H .

For an orthogonal projector P , we set $H_P = P(H)$, its image, and vice versa, for subspace L of H , the corresponding orthogonal projector is denoted by the symbol P_L .

The set of orthogonal projectors is a *lattice* with the order structure: $P \leq Q$ iff $H_P \subset H_Q$ or equivalently, for any $\psi \in H$, $\langle \psi | P \psi \rangle \leq \langle \psi | Q \psi \rangle$.

We recall that the lattice of projectors is endowed with operations “and” (\wedge) and “or” (\vee). For two projectors P_1, P_2 , the projector $R = P_1 \wedge P_2$ is defined as the projector onto the subspace $H_R = H_{P_1} \cap H_{P_2}$ and the projector $S = P_1 \vee P_2$ is defined as the projector onto the subspace H_S defined as the minimal linear subspace containing the set-theoretic union $H_{P_1} \cup H_{P_2}$ of subspaces H_{P_1}, H_{P_2} : this is the space of all linear combinations of vectors belonging to these subspaces. The operation of negation is defined as the orthogonal complement: $P^\perp = \{y \in H : \langle y | x \rangle = 0 \text{ for all } x \in H_P\}$.

In the language of subspaces the operation “and” coincides with the usual set-theoretic intersection, but the operations “or” and “not” are nontrivial deformations of the corresponding set-theoretic operations. It is natural to expect that such deformations can induce deviations from classical Boolean logic.

Consider the following simple example. Let H be two dimensional Hilbert space with the orthonormal basis (e_1, e_2) and let $v = (e_1 + e_2)/\sqrt{2}$. Then $P_v \wedge P_{e_1} = 0$ and $P_v \wedge P_{e_2} = 0$, but $P_v \wedge (P_{e_1} \vee P_{e_2}) = P_v$. Hence, for quantum events, in general the distributivity law is violated:

$$P \wedge (P_1 \vee P_2) \neq (P \wedge P_1) \vee (P \wedge P_2). \quad (44)$$

The lattice of orthogonal projectors is called *quantum logic*. It is considered as a (very special) generalization of classical Boolean logic. Any sub-lattice consisting of commuting projectors can be treated as classical Boolean logic.

At the first sight the representation of events by projectors/linear subspaces might look exotic. However, this is simply a prejudice which springs from too common usage of the set-theoretic representation of events (Boolean logic) in the modern classical probability theory. The tradition to represent events by subsets was firmly established by A. N. Kolmogorov in 1933. We remark that before him the basic

classical probabilistic models were not of the set-theoretic nature. For example, the main competitor of the Kolmogorov model, the von Mises frequency model, was based on the notion of a collective.

As we have seen, quantum logic relaxes some constraints posed on the operations of classical Boolean logic, in particular, the distributivity constraint. This provides novel possibilities for logically consistent reasoning.

Since human decision makers violate FTP [32, 38]—the basic law of classical probability, it seems that they process information by using nonclassical logic. Quantum logic is one of the possible candidates for logic of human reasoning. However, one has to remember that *in principle other types of nonclassical logic may be useful for mathematical modeling of human decision making.*

11 Space of Square Integrable Functions as a State Space

Although we generally proceed with finite-dimensional Hilbert spaces, it is useful to mention the most important example of infinite-dimensional Hilbert space used in QM. Consider the space of complex valued functions, $\psi : \mathbb{R}^m \rightarrow \mathbb{C}$, which are square integrable with respect to the Lebesgue measure on \mathbb{R}^m :

$$\|\psi\|^2 = \int_{\mathbb{R}^m} |\psi(x)|^2 dx < \infty. \quad (45)$$

It is denoted by the symbol $L^2(\mathbb{R}^m)$. Here the scalar product is given by

$$\langle \psi_1 | \psi_2 \rangle = \int_{\mathbb{R}^m} \bar{\psi}_1(x) \psi_2(x) dx.$$

A delicate point is that, for some measurable functions, $\psi : \mathbb{R}^m \rightarrow \mathbb{C}$, which are not identically zero, the integral

$$\int_{\mathbb{R}^m} |\psi(x)|^2 dx = 0. \quad (46)$$

We remark that the latter equality implies that $\psi(x) = 0$ a.e. (almost everywhere). Thus the quantity defined by (45) is, in fact, not norm: $\|\psi\| = 0$ does not imply that $\psi = 0$. To define a proper Hilbert space, one has to consider as its elements not simply functions, but classes of equivalent functions, where the equivalence relation is defined as $\psi \sim \phi$ if and only if $\psi(x) = \phi(x)$ a.e. In particular, all functions satisfying (46) are equivalent to the zero-function.

12 Operation of Tensor Product

Let both state spaces be L^2 -spaces, the spaces of complex valued square integrable functions: $H_1 = L^2(\mathbb{R}^k)$ and $L^2(\mathbb{R}^m)$.

Take two functions: $\psi \equiv \psi(x)$ belongs to H_1 and $\phi \equiv \phi(y)$ belongs to H_2 . By multiplying these functions we obtain the function of two variables $\Psi(x, y) = \psi(x) \times \phi(y)$, where \times denotes the usual point wise product.⁵ It is easy to check that this function belongs to the space $H = L^2(\mathbb{R}^{k+m})$. Take now n functions, $\psi_1(x), \dots, \psi_n(x)$, from H_1 and n functions, $\phi_1(y), \dots, \phi_n(y)$, from H_2 and consider the sum of their pairwise products:

$$\Psi(x, y) = \sum_i \psi_i(x) \times \phi_i(y). \quad (47)$$

This function also belongs to H .

It is possible to show that any function belonging to H can be represented as (47), where the sum is in general infinite. Multiplication of functions is the basic example of the operation of the tensor product. The latter is denoted by the symbol \otimes . Thus in the example under consideration $\psi \otimes \phi(x, y) = \psi(x) \times \phi(y)$. The tensor product structure on $H = L^2(\mathbb{R}^{k+m})$ is symbolically denoted as $H = H_1 \otimes H_2$.

Consider now two arbitrary orthonormal bases in spaces $H_k, (e_j^{(k)}), k = 1, 2$. Then functions $(e_{ij} = e_i^{(1)} \otimes e_j^{(2)})$ form an orthonormal basis in H : any $\Psi \in H$ can be represented as

$$\Psi = \sum c_{ij} e_{ij} \equiv \sum c_{ij} e_i^{(1)} \otimes e_j^{(2)}, \quad (48)$$

where

$$\sum |c_{ij}|^2 < \infty. \quad (49)$$

Consider now two arbitrary finite-dimensional Hilbert spaces, H_1, H_2 . For each pair of vectors $\psi \in H_1, \phi \in H_2$, we form a new formal entity denoted by $\psi \otimes \phi$. Then we consider the sums $\Psi = \sum_i \psi_i \otimes \phi_i$. On the set of such formal sums we can introduce the linear space structure. (To be mathematically rigorous, we have to constraint this set by some algebraic relations to make the operations of addition and multiplication by complex numbers well defined.) This construction gives us the tensor product $H = H_1 \otimes H_2$. In particular, if we take orthonormal bases in $H_k, (e_j^{(k)}), k = 1, 2$, then $(e_{ij} = e_i^{(1)} \otimes e_j^{(2)})$ form an orthonormal basis in H , any $\Psi \in H$ can be represented as (48) with (49).

⁵Here it is convenient to use this symbol, not just write as $\Psi(x, y) = \psi(x)\phi(y)$.

The latter representation gives the simplest possibility to define the tensor product of two arbitrary (i.e., may be infinite-dimensional) Hilbert spaces as the space of formal series (48) satisfying the condition (49).

Besides the notion of the tensor product of states, we shall also use the notion of the tensor product of operators. Consider two linear operators $\widehat{a}_i : H_i \rightarrow H_i, i = 1, 2$. Their tensor product $\widehat{a} \equiv \widehat{a}_1 \otimes \widehat{a}_2 : H \rightarrow H$ is defined starting with the tensor products of two vectors:

$$\widehat{a} \psi \otimes \phi = (\widehat{a}_1 \psi) \otimes (\widehat{a}_2 \phi).$$

Then it is extended by linearity. By using the coordinate representation (48) the tensor product of operators can be represented as

$$\widehat{a} \Psi = \sum c_{ij} \widehat{a} e_{ij} \equiv \sum c_{ij} \widehat{a}_1 e_i^{(1)} \otimes \widehat{a}_2 e_j^{(2)}, \quad (50)$$

If operators $\widehat{a}_i, i = 1, 2$, are represented by matrices $(a_{kl}^{(i)})$, with respect to the fixed bases, then the matrix $(a_{kl, nm})$ of the operator \widehat{a} with respect to the tensor product of these bases can be easily calculated.

In the same way one defines the tensor product of Hilbert spaces, H_1, \dots, H_n , denoted by the symbol $H = H_1 \otimes \dots \otimes H_n$. We start with forming the formal entities $\psi_1 \otimes \dots \otimes \psi_n$, where $\psi_j \in H_j, j = 1, \dots, n$. Tensor product space is defined as the set of all sums $\sum_j \psi_{1j} \otimes \dots \otimes \psi_{nj}$ (which has to be constrained by some algebraic relations, but we omit such details). Take orthonormal bases in $H_k, (e_j^{(k)}), k = 1, \dots, n$. Then any $\Psi \in H$ can be represented as

$$\Psi = \sum_{\alpha} c_{\alpha} e_{\alpha} \equiv \sum_{\alpha=(j_1 \dots j_n)} c_{j_1 \dots j_n} e_{j_1}^{(1)} \otimes \dots \otimes e_{j_n}^{(n)}, \quad (51)$$

where $\sum_{\alpha} |c_{\alpha}|^2 < \infty$.

13 Ket- and Bra-Vectors: Dirac's Symbolism

Dirac's notations [21] are widely used in quantum information theory. Vectors of H are called *ket-vectors*, they are denoted as $|\psi\rangle$. The elements of the dual space H' of H , the space of linear continuous functionals on H , are called *bra-vectors*, they are denoted as $\langle\psi|$.

Originally the expression $\langle\psi|\phi\rangle$ was used by Dirac for the duality form between H' and H , i.e., $\langle\psi|\phi\rangle$ is the result of application of the linear functional $\langle\psi|$ to the vector $|\phi\rangle$. In mathematical notation it can be written as follows. Denote the functional $\langle\psi|$ by f and the vector $|\phi\rangle$ by simply ϕ . Then $\langle\psi|\phi\rangle \equiv f(\phi)$. To simplify the model, later Dirac took the assumption that H is Hilbert space, i.e., the H' can be identified with H . We remark that this assumption is an axiom simplifying

the mathematical model of QM. However, in principle Dirac's formalism [21] is applicable for any topological linear space H and its dual space H' ; so it is more general than von Neumann's formalism [60] rigidly based on Hilbert space.

Consider an observable a given by the Hermitian operator \hat{a} with nondegenerate spectrum and restrict our consideration to the case of finite dimensional H . Thus the normalized eigenvectors e_i of A form the orthonormal basis in H . Let $\hat{a}e_i = \alpha_i e_i$. In Dirac's notation e_i is written as $|\alpha_i\rangle$ and, hence, any pure state can be written as

$$|\psi\rangle = \sum_i c_i |\alpha_i\rangle, \quad \sum_i |c_i|^2 = 1. \quad (52)$$

Since the projector onto $|\alpha_i\rangle$ is denoted as $P_{\alpha_i} = |\alpha_i\rangle\langle\alpha_i|$, the operator \hat{a} can be written as

$$\hat{a} = \sum_i \alpha_i |\alpha_i\rangle\langle\alpha_i|. \quad (53)$$

Now consider two Hilbert spaces H_1 and H_2 and their tensor product $H = H_1 \otimes H_2$. Let $(|\alpha_i\rangle)$ and $(|\beta_j\rangle)$ be orthonormal bases in H_1 and H_2 corresponding to the eigenvalues of two observables A and B . Then vectors $|\alpha_i\rangle \otimes |\beta_j\rangle$ form the orthonormal basis in H . Typically in physics the sign of the tensor product is omitted and these vectors are written as $|\alpha_i\rangle|\beta_j\rangle$ or even as $|\alpha_i\beta_j\rangle$. Thus any vector $\psi \in H = H_1 \otimes H_2$ can be represented as

$$\psi = \sum_{ij} c_{ij} |\alpha_i\beta_j\rangle, \quad (54)$$

where $c_{ij} \in \mathbf{C}$ (in the infinite-dimensional case these coefficients are constrained by the condition $\sum_{ij} |c_{ij}|^2 < \infty$).

14 Qubit

In particular, in quantum information theory typically qubit states are represented with the aid of observables having the eigenvalues 0, 1. Each qubit space is two dimensional:

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle, \quad |c_0|^2 + |c_1|^2 = 1. \quad (55)$$

A pair of qubits is represented in the tensor product of single qubit spaces, here pure states can be represented as superpositions:

$$|\psi\rangle = c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle, \quad (56)$$

where $\sum_{ij} |c_{ij}|^2 = 1$. In the same way the n -qubit state is represented in the tensor product of n one-qubit state spaces (it has the dimension 2^n) :

$$|\psi\rangle = \sum_{x_j=0,1} c_{x_1\dots x_n} |x_1 \dots x_n\rangle, \quad (57)$$

where $\sum_{x_j=0,1} |c_{x_1\dots x_n}|^2 = 1$. We remark that the dimension of the n qubit state space grows exponentially with the growth of n . The natural question about possible physical realizations of such multi-dimensional state spaces arises. The answer to it is not completely clear; it depends very much on the used interpretation of the wave function.

15 Entanglement

Consider the tensor product $H = H_1 \otimes H_2 \otimes \dots \otimes H_n$ of Hilbert spaces $H_k, k = 1, 2, \dots, n$. The states of the space H can be *separable and non-separable* (entangled). We start by considering pure states. The states from the first class, separable pure states, can be represented in the form:

$$|\psi\rangle = \otimes_{k=1}^n |\psi_k\rangle = |\psi_1 \dots \psi_n\rangle, \quad (58)$$

where $|\psi_k\rangle \in H_k$. The states which cannot be represented in this way are called non-separable or *entangled*. Thus mathematically the notion of entanglement is very simple, it means impossibility of tensor factorization.

For example, let us consider the tensor product of two one-qubit spaces. Select in each of them an orthonormal basis denoted as $|0\rangle, |1\rangle$. The corresponding orthonormal basis in the tensor product has the form $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. Here we used Dirac's notations, see Sect. 13, near the end. Then the so-called Bell's states

$$|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}, \quad |\Phi^-\rangle = (|00\rangle - |11\rangle)/\sqrt{2}; \quad (59)$$

$$|\Psi^+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}, \quad |\Psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2} \quad (60)$$

are entangled.

Although the notion of entanglement is mathematically simple, its physical interpretation is one of the main problems of modern quantum foundations. The common interpretation is that entanglement encodes quantum nonlocality, the possibility of action at the distance (between parts of a system in an entangled state). Such an interpretation implies the drastic change of all classical physical presentations about nature, at least about the microworld. In the probabilistic terms entanglement induces correlations which are too strong to be described by classical probability theory. (At least this is the common opinion of experts in quantum

information theory and quantum foundations.) Such correlations violate the famous Bell inequality which can be derived only in classical probability framework. The latter is based on the use of a single probability space covering probabilistic data collected in a few incompatible measurement contexts.

Now consider a quantum state given by density operator ρ in H . This state is called separable if it can be factorized in the product of density operators in spaces H_k :

$$\rho = \otimes_{k=1}^n \rho_k, \quad (61)$$

otherwise the state ρ is called entangled. We remark that an interpretation of entanglement for mixed states is even more complicated than for pure states.

16 Violation of Formula of Total Probability in Two-Slit Experiment

Consider the famous two-slit experiment with the symmetric setting: the source of photons is located symmetrically with respect to two slits, Fig. 1.

Consider the following pair of observables a and b . We select a as the “slit passing observable,” i.e., $a = 0, 1$, see Fig. 1 (we use indexes 0, 1 to be close to qubit notation) and observable b as the position on the photo-sensitive plate, see Fig. 2. We remark that the b -observable has the continuous range of values, the position x on the photo-sensitive plate. We denote $\mathbf{P}(a = i)$ by $\mathbf{P}(i)$ ($i = 0, 1$), and

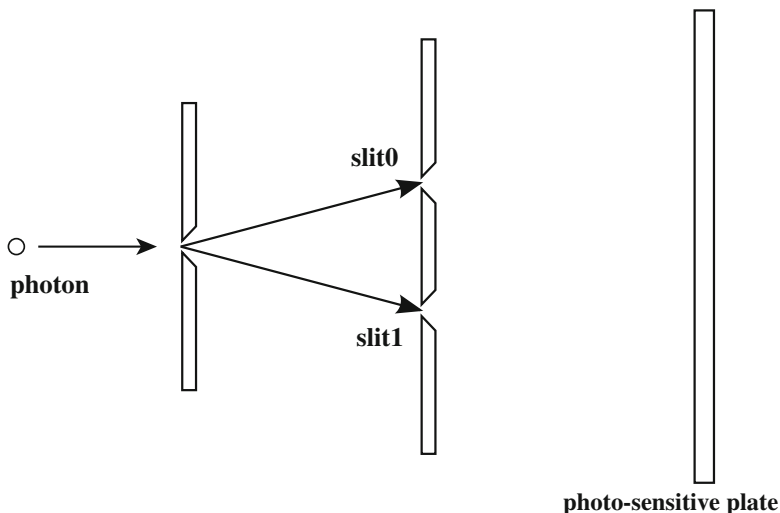


Fig. 1 Experimental setup

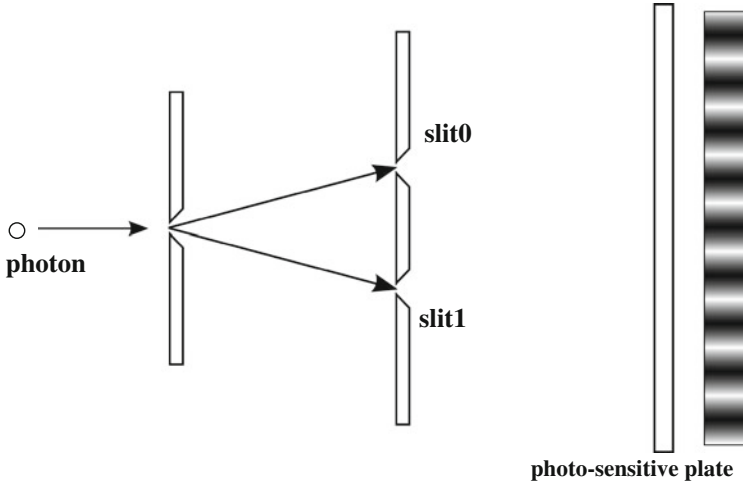


Fig. 2 Context with both slits is open

$\mathbf{P}(b = x)$ by $\mathbf{P}(x)$. Physically the a -observable corresponds to measurement of position (coarse grained to “which slit?”) and the b -observable represents measurement of momentum.

In quantum foundational studies, various versions of the two-slit experiment have been successfully performed, not only with photons, but also with electrons and even with macroscopic molecules (by Zeilinger’s group). All those experiment demonstrated matching with predictions of QM. Experimenters reproduce the interference patterns predicted by QM and calculated by using the wave functions.

The probability that a photon is detected at position x on the photo-sensitive plate is represented as

$$\begin{aligned} \mathbf{P}(x) &= \left| \frac{1}{\sqrt{2}}\psi_0(x) + \frac{1}{\sqrt{2}}\psi_1(x) \right|^2 \\ &= \frac{1}{2} |\psi_0(x)|^2 + \frac{1}{2} |\psi_1(x)|^2 + |\psi_0(x)| |\psi_1(x)| \cos \theta, \end{aligned} \tag{62}$$

where ψ_0 and ψ_1 are two wave functions, whose squared absolute values $|\psi_i(x)|^2$ give the distributions of photons passing through the slit $i = 0, 1$, see Figs. 3 and 4. Here we explored the rule of addition of complex probability amplitudes, a quantum analog of the rule of addition of probabilities. This rule is the direct consequence of the linear space structure of quantum state spaces.

The term

$$|\psi_0(x)| |\psi_1(x)| \cos \theta$$

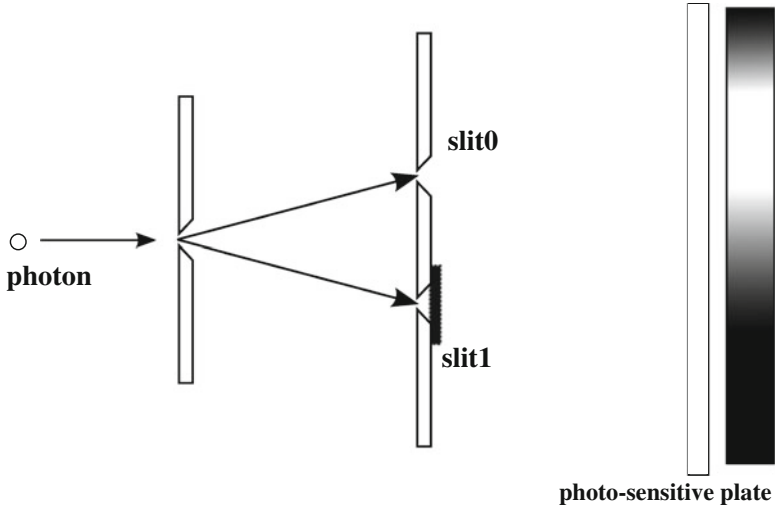


Fig. 3 Context with one slit is closed-I

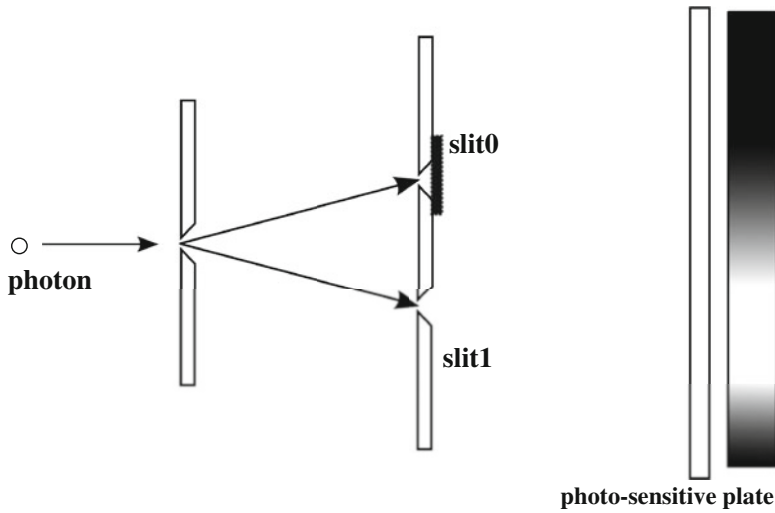


Fig. 4 Context with one slit is closed-II

implies the interference effect of two wave functions. Let us denote $|\psi_i(x)|^2$ by $\mathbf{P}(x|i)$, then Eq. (62) is represented as

$$\mathbf{P}(x) = \mathbf{P}(0)\mathbf{P}(x|0) + \mathbf{P}(1)\mathbf{P}(x|1) + 2\sqrt{\mathbf{P}(0)\mathbf{P}(x|0)\mathbf{P}(1)\mathbf{P}(x|1)} \cos \theta. \quad (63)$$

Here the values of probabilities $\mathbf{P}(0)$ and $\mathbf{P}(1)$ are equal to $1/2$ since we consider the symmetric settings. For general experimental settings, $\mathbf{P}(0)$ and $\mathbf{P}(1)$ can be taken

as the arbitrary nonnegative values satisfying $\mathbf{P}(0) + \mathbf{P}(1) = 1$. In the above form, the classical probability law (FTP)

$$\mathbf{P}(x) = \mathbf{P}(0)\mathbf{P}(x|0) + \mathbf{P}(1)\mathbf{P}(x|1) \quad (64)$$

is violated, and the term of interference $2\sqrt{\mathbf{P}(x|0)\mathbf{P}(0)\mathbf{P}(x|1)\mathbf{P}(1)} \cos \theta$ specifies the violation.

The crucial point is that the two-slit experiment has the multi-contextual structure: $C_i, i = 0, 1$, only the i th slit is open, and C_{01} , both slits are open, see Figs. 3, 4, and 2. Comparison of possibilities is represented as comparison of the corresponding probability distributions $\mathbf{P}(x|i)$, $\mathbf{P}(x)$. In the contextual notations they can be written as

$$p_{C_i}^b(x) \equiv \mathbf{P}(b = x|C_i), p_{C_{01}}^b(x) \equiv \mathbf{P}(b = x|C_{01}).$$

Here conditioning is not classical probabilistic event conditioning, but context conditioning: different contexts are mathematically represented by different Kolmogorov probability spaces. The general contextual probability theory including its representation in complex Hilbert space is presented in very detail in my monograph [37].

References

1. Aerts, D., Sozzo, S., & Tapia, J. (2012). A quantum model for the Ellsberg and Machina Paradoxes. *Quantum Interaction Lecture Notes in Computer Science*, 7620, 48–59.
2. Asano, M., Basieva, I., Khrennikov, A., Ohya, M., Tanaka, Y., & Yamato, I. (2015). Quantum information biology: From information interpretation of quantum mechanics to applications in molecular biology and cognitive psychology. *Foundations of Physics*, 45(10), 1362–1378.
3. Asano, M., Basieva, I., Khrennikov, A., & Yamato, I. (2017). A model of differentiation in quantum bioinformatics. *Progress in Biophysics and Molecular Biology*, 130(Part A), 88–98.
4. Asano, M., Basieva, I., Khrennikov, A., Ohya, M., & Tanaka, Y. (2017). A quantum-like model of selection behavior. *Journal of Mathematical Psychology*, 78, 2–12.
5. Asano, M., Khrennikov, A., Ohya, M., Tanaka, Y., & Yamato, I. (2014). Violation of contextual generalization of the Leggett-Garg inequality for recognition of ambiguous figures. *Physica Scripta*, T163, 014006.
6. Asano, M., Khrennikov, A., Ohya, M., Tanaka, Y., & Yamato, I. (2015). *Quantum adaptivity in biology: From genetics to cognition*. Heidelberg: Springer.
7. Ballentine, L. E. (1989). The statistical interpretation of quantum mechanics. *Reviews of Modern Physics*, 42, 358–381.
8. Ballentine, L. E. (1998). *Quantum mechanics: A modern development*. Singapore: WSP.
9. Ballentine, L. E. (2001). Interpretations of probability and quantum theory. In A. Y. Khrennikov (Ed.), *Foundations of probability and physics. Quantum probability and white noise analysis* (Vol. 13, pp. 71–84). Singapore: WSP.
10. Basieva, I., Khrennikova, P., Pothos, E. M., Asano, M., & Khrennikov, A. (2018). Quantum-like model of subjective expected utility. *Journal of Mathematical Economics*, 78, 150–162.
11. Birkhoff, J., & von Neumann, J. (1936). The logic of quantum mechanics. *Annals of Mathematics*, 37(4), 823–843.

12. Bohr, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Physics Review*, 48, 696–702.
13. Bohr, N. (1938). The causality problem in atomic physics. In J. Faye & H. J. Folse (Eds.) (1987). *The philosophical writings of Niels Bohr*. Causality and Complementarity, Supplementary Papers (Vol. 4, pp. 94–121). Woodbridge: Ox Bow Press.
14. Bohr, N. (1987). *The philosophical writings of Niels Bohr* (Vol. 3). Woodbridge: Ox Bow Press.
15. Broekaert, J., Basieva, I., Blasiak, P., & Pothos, E. M. (2017). Quantum dynamics applied to cognition: A consideration of available options. *Philosophical Transactions of the Royal Society A*, 375(2016), 375–405.
16. Brukner, C., & Zeilinger, A. (1999). Malus' law and quantum information. *Acta Physica Slovaca*, 49(4), 647–652.
17. Brukner, C., & Zeilinger, A. (2009). Information invariance and quantum probabilities. *Foundations of Physics*, 39, 677–689.
18. Busemeyer, J. R., & Bruza, P. D. (2012). *Quantum models of cognition and decision*. Cambridge: Cambridge University Press.
19. Chiribella, G., D'Ariano, G. M. & Perinotti, P. (2010). Probabilistic theories with purification. *Physical Review A*, 81, 062348.
20. D'Ariano, G. M. (2007). Operational axioms for quantum mechanics. In: Adenier et al. (Eds.), *Foundations of probability and physics-3*. AIP conference proceedings (Vol. 889, pp. 79–105).
21. Dirac, P. A. M. (1995). *The principles of quantum mechanics*. Oxford: Clarendon Press.
22. Feynman, R., & Hibbs, A. (1965). *Quantum mechanics and path integrals*. New York: McGraw-Hill.
23. Feynman, R. P. (1951). The concept of probability in quantum mechanics. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability* (pp. 533–541). California: University of California Press.
24. Frommholz, I., Larsen, B., Piowowski, B., Lalmas, M., Ingwersen, P., & van Rijsbergen, K. (2010). Supporting polyrepresentation in a quantum-inspired geometrical retrieval framework. In *IliX'10* (pp. 115–124).
25. Fuchs, C. A. (2002). Quantum mechanics as quantum information (and only a little more). In A. Khrennikov (Ed.), *Quantum theory: Reconsideration of foundations. Series mathematical modeling* (Vol. 2, pp. 463–543). Växjö: Växjö University Press.
26. Fuchs, C. A., Mermin, N. D., & Schack, R. (2014). An Introduction to QBism with an application to the locality of quantum mechanics. *American Journal of Physics*, 82, 749.
27. Haven, E., & Khrennikov, A. (2009). Quantum mechanics and violation of the sure-thing principle: The use of probability interference and other concepts. *Journal of Mathematical Psychology*, 53, 378–388.
28. Haven, E., & Khrennikov, A. (2013). *Quantum social science*. Cambridge: Cambridge University Press.
29. Khrennikov, A. (2009). *Interpretations of probability* (2nd ed.). Tokyo/Berlin: VSP International SC Publisher/De Gruyter.
30. Khrennikov, A. Y. (2001). Linear representations of probabilistic transformations induced by context transitions. *Journal of Physics A: Mathematical and General*, 34, 9965–9981.
31. Khrennikov, A. Y. (2001). Origin of quantum probabilities. In A. Khrennikov (Ed.), *Foundations of probability and physics* (pp. 180–200). Singapore: Växjö-2000/WSP.
32. Khrennikov, A. (2004). *Information dynamics in cognitive, psychological, social, and anomalous phenomena. Fundamental theories of physics*. Dordrecht: Kluwer.
33. Khrennikov, A. (2004). Växjö interpretation-2003: Realism of contexts. In *Proceeding International Conference Quantum Theory: Reconsideration of Foundations. Mathematical Modelling in Physics and Engineering, and Cognitive Science* (Vol. 10, pp. 323–338). Växjö: Växjö University Press.
34. Khrennikov, A. (2005). Reconstruction of quantum theory on the basis of the formula of total probability. In *AIP Conference Proceeding of the Foundations of Probability and Physics—3* (Vol. 750, pp. 187–218). Melville: American Institute of Physics.

35. Khrennikov, A. (2005). The principle of supplementarity: A contextual probabilistic viewpoint to complementarity, the interference of probabilities, and the incompatibility of variables in quantum mechanics. *Foundations of Physics*, 35(10), 1655–1693.
36. Khrennikov, A. (2008). Algorithm for quantum-like representation: Transformation of probabilistic data into vectors on Bloch's sphere. *Open Systems and Information Dynamics*, 15, 223–230.
37. Khrennikov, A. (2009). *Contextual approach to quantum formalism*. Berlin: Springer.
38. Khrennikov, A. (2010). *Ubiquitous quantum structure: From psychology to finances*. Berlin: Springer.
39. Khrennikov, A. (2015). Quantum-like model of unconscious-conscious dynamics. *Frontiers in Psychology*, 6, Art. 997.
40. Khrennikov, A. (2015). Towards information lasers. *Entropy*, 17(10), 6969–6994.
41. Khrennikov, A. (2015). Quantum version of Aumann's approach to common knowledge: Sufficient conditions of impossibility to agree on disagree. *Journal of Mathematical Economics*, 60, 89–104.
42. Khrennikov, A. (2016). Analog of formula of total probability for quantum observables represented by positive operator valued measures. *International Journal of Theoretical Physics*, 55, 859–3874.
43. Khrennikov A. Yu. (2016). *Probability and randomness: Quantum versus classical*. Singapore: World Scientific.
44. Khrennikov, A. (2016). Quantum Bayesianism as the basis of general theory of decision-making. *Philosophical Transactions of the Royal Society A*, 374, 20150245.
45. Khrennikova, P. (2016, April). Quantum dynamical modeling of competition and cooperation between political parties: The coalition and non-coalition equilibrium model. *Journal of Mathematical Psychology*, 71, 39–50.
46. Khrennikov, A., & Basieva, I. (2014). Quantum model for psychological measurements: From the projection postulate to interference of mental observables represented as positive operator valued measures. *NeuroQuantology*, 12, 324–336.
47. Khrennikova, P., & Haven, E. (2016). Instability of political preferences and the role of mass-media: a dynamical representation in a quantum framework. *Philosophical Transactions of the Royal Society A*, 374, 20150106.
48. Khrennikova, P., Haven, E., & Khrennikov, A. (2014) An application of the theory of open quantum systems to model the dynamics of party governance in the US political System. *International Journal of Theoretical Physics*, 53(4), 1346–1360.
49. Kolmogoroff, A. N. (1933). *Grundbegriffe der wahrscheinlichkeitsrechnung*. Berlin: Springer.
50. Kolmogorov, A. N. (1956). *Foundations of the probability theory*. New York: Chelsea Publishing Company.
51. Melucci, M. (2013). Deriving a quantum information retrieval basis. *The Computer Journal*, 56(11), 1279–1291.
52. Melucci, M. (2015). *Introduction to information retrieval and quantum mechanics*. Berlin: Springer.
53. Plotnitsky, A. (2012). *Niels Bohr and complementarity: An introduction*. Berlin: Springer.
54. Plotnitsky, A., & Khrennikov, A. (2015). Reality without realism: On the ontological and epistemological architecture of quantum mechanics. *Foundations of Physics*, 45(10), 269–1300.
55. Pothos, E. M., & Busemeyer, J. R. (2009). A quantum probability explanation for violation of rational decision theory. *Proceedings of the Royal Society B*, 276, 2171–2178.
56. Song, D., Lalmas, M., van Rijsbergen, C. J., Frommholz, I., Piwowski, B., Wang, J., et al. (2010). How quantum theory is developing the field of information retrieval? In *Quantum informatics for cognitive, social, and semantic processes: Papers from the AAAI fall symposium (FS-10-08)* (pp. 105–108).
57. Sozzo, S. (2014). A Quantum probability explanation in Fock space for borderline contradictions. *Journal of Mathematical Psychology*, 58(1), 1–12.

58. van Rijsbergen, C. J. (2004). *The geometry of information retrieval*. Cambridge: Cambridge University Press.
59. van Rijsbergen, C. J. (2011). What is Quantum information retrieval? In G. Amati, F. Crestani, (Eds.), *Advances in information retrieval theory. ICTIR 2011. Lecture notes in computer science* (Vol. 6931). Berlin: Springer.
60. von Neuman, J. (1955). *Mathematical foundations of quantum mechanics*. Princeton: Princeton University Press.
61. von Neumann, J., & Morgenstern, O. (1953). *Theory of games and economic behavior*. Princeton: Princeton University Press.