1

Introduction to Quantum Measurement Theory

1.1 Pure and Mixed States

Let $\mathcal{H}$ be a complex Hilbert space; for simplicity we restrict consideration to finite-dimensional spaces.[1] We recall that a *pure quantum state* can be represented by a normalized vector of $\mathcal{H}$, namely, $||\psi|| = 1$. Two vectors that differ only by a phase, that is, $\psi = e^{i\theta}\phi$, represent the same quantum state. Under consideration of a single state, its phase does not play any role, but by manipulating a few states, the relative phases play a crucial role, for example, in the interference effect.

In physical literature the Dirac notation $|\psi\rangle$ is used for vectors of Hilbert space. We shall use both notations, ψ and $|\psi\rangle$.

The space of linear operators acting in $\mathcal{H}$ is denoted as $\mathcal{L}(\mathcal{H})$. (In the infinite-dimensional case this symbol denotes the space of bounded linear operators.) We remark that $\mathcal{L}(\mathcal{H})$ is a linear space by itself. It can be endowed with the Hilbert space structure

$$\langle \hat{A}|\hat{B}\rangle = \mathrm{Tr}\,\hat{A}^{\star}\hat{B}. \tag{1.1}$$

(In the infinite-dimensional case the scalar product is defined on the space of Hilbert–Schmidt operators.)

A linear operator $\hat{A}$ is called Hermitian if $\hat{A} = \hat{A}^{\star}$. The space of Hermitian operators $\mathcal{L}_H$ is the real Hilbert space with scalar product (1.1). A linear operator $\hat{A}$ is called positive semidefinite if $\langle\psi|\hat{A}|\psi\rangle \geq 0$ for any $\psi \in \mathcal{H}$ or symbolically $\hat{A} \geq 0$.

A density operator $\hat{\rho}$ is determined by the following conditions:

- $\hat{\rho} = \hat{\rho}^{\star}$ (Hermitian operator);
- $\hat{\rho} \geq 0$ (positive semidefinite operator);
- $\mathrm{Tr}\,\hat{\rho} = 1$ (trace one operator).

The space of density operators is denoted by the symbol $D \equiv D(\mathcal{H})$.

[1] Sometimes infinite-dimensional spaces will be considered. We make remarks about this case in the corresponding places.

https://doi.org/10.1017/9781009313469.002 Published online by Cambridge University Press

We note that each pure state $|\psi\rangle$ can be represented by a density operator, projection on the state vector. In the Dirac notation, for a pure state $|\psi\rangle$, the corresponding density operator (projector) is symbolically written as $|\psi\rangle\langle\psi|$.

Typically one says that a density operator represents a mixed quantum state – a statistical mixture of pure states. However, such an interpretation is ambiguous, since the same density operator can be represented by a variety of probabilistically weighted sums of projectors corresponding to pure states.

1.2 Observables, Born's Rule, Projection Postulate

By axiomatics of QM due to von Neumann [365] an observable A (with a discrete range of values) is represented by a Hermitian operator,

$$\hat{A} = \sum_{\alpha} \alpha \hat{E}_A(\alpha), \tag{1.2}$$

where $\hat{E}_A(\alpha)$ is a projection onto the space $\mathcal{H}_A(\alpha)$ of eigenvectors for the eigenvalue α.

Observables are denoted by the symbols $A, B, \ldots$ and the corresponding operators by the symbols $\hat{A}, \hat{B}, \ldots$. To simplify notation, the unit operator is denoted as I. If operators are used as indexes, the hat symbol is omitted, for example, $\hat{\rho}$, but P_ρ.

For a pure state $|\psi\rangle$, the probability of getting the outcome $A = \alpha$ is given by *Born's rule*:

$$P_\psi(A = \alpha) = \|\hat{E}_A(\alpha)\psi\|^2 = \langle \hat{E}_A(\alpha)\psi|\psi\rangle. \tag{1.3}$$

(The latter equality is due to idempotence of projectors, $E_A(\alpha)E_A(\alpha) = E_A(\alpha)$.)

For each state ψ and quantum observable $\hat{A}$, formula (1.3) determines the probability distribution and classical probability theory (Section 1.5) can be applied. So, the mean value (average) of a quantum observable is given by the integral formula (1.22). In quantum terms it can be expressed as

$$\langle A\rangle_\psi = \langle \hat{A}\psi|\psi\rangle. \tag{1.4}$$

A measurement with the outcome $A = \alpha$ generates back-action onto the system's state:

$$|\psi\rangle \to |\psi\rangle_{A=\alpha} = \hat{E}_A(\alpha)\psi / \|\hat{E}_A(\alpha)\psi\|. \tag{1.5}$$

This is the projection postulate, one of the axioms of QM. In this general form it was formulated by Lüders [274]. Often one refers to (1.5) as the von Neumann projection postulate. However, von Neumann explored (1.5) only for observables represented by Hermitian operators with nondegenerate spectra [365]. In the case

https://doi.org/10.1017/9781009313469.002 Published online by Cambridge University Press

of degenerate spectra he considered more general state updates, which in the future would lead to quantum instrument theory.

Measurements of another observable, say B, conditioned on the outcome $A = \alpha$ lead to probability:

$$P_\psi(B = \beta | A = \alpha) = \|\hat{E}_B(\beta)\psi_{A=\alpha}\|^2 = \frac{\|\hat{E}_B(\beta)\hat{E}_A(\alpha)\psi\|^2}{\|\hat{E}_A(\alpha)\psi\|^2}. \tag{1.6}$$

This is quantum *conditional probability* (see Chapters 10 and 13 for the mathematical details, interpretation, and applications).

The preceding formulas can be easily generalized to states given by density operators, but we do this directly for quantum instruments; see the following discussion. Observables represented mathematically by Hermitian operators and generating the projection state update are often called von Neumann observables and referred to as performing accurate measurements. Generally quantum instruments describe noisy instruments.

1.3 Quantum Instruments

We consider also linear operators acting in $\mathcal{L}(\mathcal{H})$, *superoperators*. A superoperator is called positive if it maps the set of positive semidefinite operators into itself: for $\hat{\rho} \geq 0$, $\hat{T}(\hat{\rho}) \geq 0$.

The detailed presentation of quantum instruments theory can be found in Section 11.3 (for technical details see [291]-[300]). Now we say just a few words about them. In this book they will be used only in Chapter 10 as an example of a measurement procedure of the nonprojection type and in Chapter 18 devoted to applications of the quantum formalism and methodology outside of physics (quantum-like modeling). The rest of the book is based solely on the von Neumann measurement theory with observables given by Hermitian operators.

Consider an observable A with the following discrete range of values:

$$X = \{x_1, \ldots, x_m, \ldots\}.$$

Any map $x \to \mathcal{I}_A(x)$, where, for each $x \in X$, the map $\mathcal{I}_A(x)$ is a positive superoperator and

$$\sum_x \mathcal{I}_A(x) \colon D \to D, \tag{1.7}$$

is called a *quantum instrument.* It represents one of the measurement procedures of an observable A (see Section 11.3).

The probability for the output $A = x$ is given by the generalized Born rule in the form

$$P_\rho(A = x) = \mathrm{Tr}\ [\mathcal{I}_A(x)\hat{\rho}]. \tag{1.8}$$

https://doi.org/10.1017/9781009313469.002 Published online by Cambridge University Press

We note that a measurement with the output $A = x$ generates the state-update by the following transformation:

$$\hat{\rho} \to \hat{\rho}_x = \frac{\mathcal{I}_A(x)\hat{\rho}}{\mathrm{Tr}\,\mathcal{I}_A(x)\hat{\rho}}. \tag{1.9}$$

Thus, the projection postulate is no longer a requirement for the state update. An observable A can be measured by a variety of instruments generating the same probability distribution, but different state updates.

A special class of quantum instruments was introduced by Davis and Levis (see, e.g., [98]) and we call them *the Davis–Levis instruments.* This class is too general for applications in quantum physics (although it might be useful in quantum-like modeling; see Chapter 18). In quantum physics the special class of quantum instruments is used (see, e.g., Ozawa [298]); such instruments are given by *completely positive maps* $\mathcal{I}(x)$.

A superoperator $\hat{\mathcal{T}}\colon \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ is called *completely positive* if its natural extension $\hat{\mathcal{T}} \otimes I$ to the tensor product $\mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(M) = \mathcal{L}(\mathcal{H} \otimes L)$, where L is an arbitrary finite-dimensional complex Hilbert space, namely, $\hat{\mathcal{T}} \otimes I(|\psi\rangle|\phi\rangle) = \hat{\mathcal{T}}|\psi\rangle|\phi\rangle$, is again a positive superoperator on $\mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(L)$.

Complete positivity is the mathematical description of the following physical statement. Let S be a system with the state space $\mathcal{H}$ and let $\hat{\mathcal{T}}$ be a superoperator mapping states into states. We can always add to S another system, say S' with state space M. Assume that systems do not interact. The unit map $I\colon M \to M$ does not change states of S'. Consider now compound system $\mathbf{S} = (S, S')$. Its state space is given by the tensor product $\mathcal{H} \otimes L$. It is natural (from the physical viewpoint) that the superoperator map $\hat{\mathcal{T}} \otimes I$ would map the states of $\mathbf{S}$ into its states, that is, it should preserve positivity as $\hat{\mathcal{T}}$ does. Adding to consideration another system that does not interact with the original one cannot change the physical situation and it would be surprising if a state transformation corresponding to measurement on S would be impossible to extend identically to the state space of the compound system. In terms of observables, violation of complete positivity means that observation on S cannot be treated as observation on the compound system $\mathbf{S}$.

According to the Kraus theorem, for any instrument $\mathcal{I}$, there exists a family $\{\hat{M}_j(x)\}$ of operators, called the *measurement operators* for $\mathcal{I}$, in $\mathcal{H}$ such that

$$\mathcal{I}(x)\rho = \sum_j \hat{M}_j(x)\hat{\rho}\hat{M}_j^\star(x) \tag{1.10}$$

for any state ρ. In this case, we have

$$\sum_{xj} \hat{M}_j^\star(x)\hat{M}_j(x) = I. \tag{1.11}$$

https://doi.org/10.1017/9781009313469.002 Published online by Cambridge University Press

Conversely, any family $\{M_j(x)\}$ of operators in $\mathcal{H}$ satisfying (1.11) defines an instrument $\mathcal{I}$.

Each quantum instrument determines a generalized quantum observable, which is mathematically represented as

$$\hat{\Pi}(x) = \mathcal{I}^{\star}(x)I. \tag{1.12}$$

Operators $\hat{\Pi}(x), x \in X$, are called *effects*; they are positive semidefinite Hermitian and sum up to the unit operator:

$$\sum_{x\in X} \hat{\Pi}(x) = I.$$

The family of operators $\hat{\Pi} = (\hat{\Pi}(x), x \in X)$ is called a *positive operator valued measure* (POVM); see Section 11.3 for the details. For any $O \subset X$, we set

$$\hat{\Pi}(O) = \sum_{x\in O} \hat{\Pi}(x). \tag{1.13}$$

The map $O \to \hat{\Pi}(O)$ is additive on the collection of all subsets of X, that is, it is really an operator-valued measure. (We remark that we consider a finite of countable set X. Generally it should be endowed with some σ-algebra (Section 1.5) of subsets of X.)

Consider a quantum instrument such that

$$\mathcal{I}_A(x)\hat{\rho} = \hat{E}(x)\hat{\rho}\hat{E}(x), \tag{1.14}$$

where, for any $x \in X, \hat{E}(x)$ is projection. Such an instrument is called the projection instrument. Any Hermitian operator $\hat{A}$ induces the projection instrument based on its spectral projections, $\hat{E}_A = (\hat{E}_A(x) \colon x \in X)$, where $X \subset \mathbb{R}$ is its spectrum. Generally the set of the observable's outcomes X need not be a subset of the real line. A system of mutually orthogonal projections $\hat{E} = (\hat{E}(x) \colon x \in X)$ that is normalized to the unit operator, namely

$$\sum_{x} \hat{E}(x) \mathrel{\hat{=}} I, \tag{1.15}$$

is called a projection valued measure (PVM). As earlier, for any $O \subset X$, we set

$$\hat{E}(O) = \sum_{x\in O} \hat{E}(x). \tag{1.16}$$

The map $O \to \hat{E}(O)$ is additive of the collection of all subsets of X, that is, it is PVM.

https://doi.org/10.1017/9781009313469.002 Published online by Cambridge University Press

1.4 Indirect Measurement Model

Construction of quantum instruments is based on the indirect measurement model. This scheme formalizes the following situation. As was pointed out by Bohr (Chapter 3), the measurement's outcomes are created through interaction of a system S with a measurement apparatus M. This apparatus consists of a complex physical device interacting with S and a pointer that shows the result of measurement, say spin up or spin down. An observer can see only outputs of the pointer and they associate these outputs with the values of the observable A for the system S. So, the observer approaches only the pointer, not the system by itself. This is a good place to note that the same observable A can be measured by a variety of apparatuses, so the map $A \to M_A$ is multivalued.

Thus, the indirect measurement scheme involves the following:

- the states of the systems S and the apparatus M_A for measurement of some (physical or mental) observable s;
- the unitary operator $\hat{U}$ representing the interaction dynamics for the system $S + M_A$;
- the meter observable $\hat{M}_A$ giving outputs of the pointer of the apparatus M designed for measurements of observable A.

Formally, an *indirect measurement model*, introduced in [291] as a "(general) measuring process," is a quadruple

$$(\mathcal{K}, \hat{\sigma}, \hat{U}, \hat{M}_A)$$

consisting of a Hilbert space $\mathcal{K}$, a density operator $\hat{\sigma} \in D(\mathcal{K})$, a unitary operator $\hat{U}$ on the tensor product of the state spaces of S and M, $U: \mathcal{H} \otimes \mathcal{K} \to \mathcal{H} \otimes \mathcal{K}$, and a self-adjoint operator $\hat{M}_A$ on $\mathcal{K}$. By this measurement model, the Hilbert space $\mathcal{K}$ describes the states of the apparatus M_A, the unitary operator $\hat{U}$ describes the time-evolution of the composite system $S + M_A$, the density operator $\hat{\sigma}$ describes the initial state of the apparatus M_A, and the self-adjoint operator $\hat{M}_A$ describes the meter observable of the apparatus M_A. Then, the output probability distribution $P_\rho(A = x)$ in the system state $\hat{\rho}$ is given by

$$P_\rho(A = x) = \mathrm{Tr}\,[(I \otimes \hat{E}_{M_A}(x))\hat{U}(\hat{\rho} \otimes \hat{\sigma})\hat{U}^\star], \tag{1.17}$$

where $\hat{E}_{M_A}(x)$ is the spectral projection of operator $\hat{M}_A$ for the eigenvalue x.

The change of the state $\hat{\rho}$ of the system S caused by the measurement for the outcome $A = x$ is represented with the aid of the map $\mathcal{I}_A(x)$ in the space of density operators defined as

$$\mathcal{I}_A(x)\hat{\rho} = \mathrm{Tr}_{\mathcal{K}}[(I \otimes \hat{E}_{M_A}(x))\hat{U}(\hat{\rho} \otimes \hat{\sigma})\hat{U}^\star], \tag{1.18}$$

https://doi.org/10.1017/9781009313469.002 Published online by Cambridge University Press

where $\mathrm{Tr}_{\mathcal{K}}$ is the partial trace over $\mathcal{K}$. Then, the map $x \mapsto \mathcal{I}_A(x)$ turns out to be a quantum instrument. Thus, the statistical properties of measurements realized by indirect measurement model $(\mathcal{K}, \hat{\sigma}, \hat{U}, \hat{M}_A)$ are described by a quantum instrument. We remark that conversely any quantum instrument can be represented via the indirect measurement model [291]. Thus, quantum instruments mathematically characterize the statistical properties of all the physically realizable quantum measurements.

Now we make the following remark on terminology. In the textbook quantum theory one speaks about observation performed on systems, where an observable A is mathematically given by Hermitian operator $\hat{A} \in \mathcal{L}(\mathcal{H})$, where $\mathcal{H}$ is the state space of S. This scheme is straightforward and it does not involve the apparatus' state space

Now, we point to a few details which were omitted in the preceding considerations. The measuring interaction between the system S and the apparatus M turns on at time t_0, the time of measurement, and turns off at time $t = t_0 + \Delta t$. We assume that the system S and the apparatus M_A do not interact with each other before t_0 nor after $t = t_0 + \Delta t$ and that the compound system $S + M_A$ is isolated in the time interval (t_0, t).

Now we describe the structure of the apparatus M_A in more detail. The *probe system* P is defined to be the minimal part of apparatus M_A such that the compound system $S + P$ is isolated in the time interval (t_0, t).

Then the preceding scheme is applied to the probe system P instead of the whole apparatus M. The rest of the apparatus M performs the pointer measurement on the probe P. In particular, the unitary evolution operator U describing the state-evolution of the system $S + P$ has the form $\hat{H} = e^{-i\Delta t\hat{H}}$, where

$$\hat{H} = \hat{H}_S + \hat{H}_P + \hat{H}_{SP}$$

is a Hamiltonian of $S + P$ with the terms $\hat{H}_S$ and $\hat{H}_P$ representing the internal dynamics in the subsystems S and P of the compound system and $\hat{H}_{SP}$ describing the interaction between the subsystems.

Introduction of probe systems may be seen as unnecessary complication of the scheme of indirect measurements. However, it is useful if the apparatus M_A is a very complex system that interacts (often in parallel) with many systems $S_j, j = 1, 2, \ldots, m$. Its different probes are involved solely in interaction with the concrete systems, P_j with S_j. And the system $S_j + P_j$ can be considered as an isolated system; in particular, from interactions with other systems S_i and probes P_i.

The indirect measurement scheme is part of the theory of open quantum systems. Instead of a measurement apparatus M_A, we can consider the surrounding environment $\mathcal{E}$ of the system S.

https://doi.org/10.1017/9781009313469.002 Published online by Cambridge University Press

1.5 Classical Probability Space

Classical probability theory was mathematically formalized in a set and measure theoretical framework by Kolmogorov in 1933 [250, 251]. Here we briefly introduce its main notions.

Let Ω be a set of any origin; its points are called elementary events. Consider a collection of subsets $\mathcal{F}$ of Ω forming a σ-algebra, that is, it is closed w.r.t. countable unions and intersections and the operation of complement. (In American literature the term σ-field is used.) If Ω is finite, then $\mathcal{F}$ is the collection of all its subsets. The sets belonging to $\mathcal{F}$ represent events. The set operations correspond to the logical operations as follows:

- union – disjunction,
- intersection – conjunction,
- complement – negation.

So, the σ-algebra of events $\mathcal{F}$ is the set representation of classical Boolean logic.

Let P be a probability measure on $\mathcal{F}$. It assigns to each event $A \in \mathcal{F}$ nonnegative weight $P(A)$; these weights are normalized as $P(\Omega) = 1$. The main property of P is its additivity: if two events A_1, A_2 are disjoint, $A_1 \cap A_2 = \emptyset$, then the probability of disjunction of these events equals the sum of probabilities

$$P(A_1 \cup A_2) = P(A_1) + P(A_2). \tag{1.19}$$

To make probability theory consistent with Lebesgue integration theory, it is supposed that probability is not only additive, but even countably additive. Consider a sequence of pairwise disjoint events, (A_k), that is, $A_i \cap A_j = \emptyset, i \neq j$; then the probability of disjunction of these events equals the sum of probabilities

$$P(\cup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k). \tag{1.20}$$

The triple $K = (\Omega, \mathcal{F}, P)$ is called *(Kolmogorov) probability space*.

A random variable is map $A\colon \Omega \to \mathbb{R}$ having a special property – measurability. Measurability is equivalent to the following condition. For any $x \in \mathbb{R}$, the set

$$\Omega_{A<x} = \{\omega \in \Omega\colon A(\omega) < x\} \tag{1.21}$$

belongs to the σ-algebra $\mathcal{F}$. The probability distribution of A, a measure p_A on $\mathbb{R}$, is determined by its values for intervals, $p^A((a, b]), \ldots, p^A((a, b))$, for example,

$$p^A([a, b]) = P(\omega \in \Omega\colon A(\omega) \in [a, b]).$$

https://doi.org/10.1017/9781009313469.002 Published online by Cambridge University Press

The mean value (mathematical expectation, average) of a random variable $A = A(\omega)$ is defined as its integral:

$$E[A] \equiv \langle A \rangle = \int_{\Omega} A(\omega) dP(\omega) = \int_{\mathbb{R}} x dp^A(x). \tag{1.22}$$

For a discrete random variable with the range of values $X = (x_n)$, its probability distribution is determined by the following set of probabilities:

$$p^A(x) = P(\omega \in \Omega : A(\omega) = x), \ x \in X.$$

The mathematical expectation is given by the sum

$$E[A] \equiv \langle A \rangle = \sum_{x \in X} x p_A(x). \tag{1.23}$$

Let A and B be random variables. We recall *the Bayes formula* for conditional probability,

$$P(B = y | A = x) = P(B = y, A = x)/P(A = x), \tag{1.24}$$

for $P(A = x) > 0$. It should be emphasized that the Bayes formula is the *definition* of conditional probability. It is not a theorem; it cannot be derived from other "natural postulates." Here

$$P(B = y, A = x) = P(\omega \in \Omega : A(\omega) = x, B(\omega) = y)$$

is the probability of the event that the variables take the values x and y. In quantum physics such events are not always defined, for example, for position and momentum, or spin projections on different axes. The Bayes formula can't be used. The main difference between classical and quantum probability theories is in conditional probability.

One of the basic classical laws of probability is given by the *formula of total probability* (FTP):

$$P(B = y) = \sum_{x \in X_A} P(B = y | A = x) P(A = x). \tag{1.25}$$

It is one of the key elements of the Bayesian probability inference. FTP is derived on the basis of the Bayes formula for conditional probability and additivity of a probability measure P. As was noted, the Bayes definition of conditional probability is not applicable in quantum probability. Hence, one can expect that FTP can also be violated for quantum observables.

We remark that Kolmogorov considered the condition of countable additivity (1.20) as purely mathematical [250]; it is needed to have the Lebesque integral representation of the mean value of a random variable. He pointed out that this condition is not experimentally testable: one can't realize experimentally an infinite

https://doi.org/10.1017/9781009313469.002 Published online by Cambridge University Press

sequence of events. Only the condition of finite additivity (1.19) has the practical meaning. It is also interesting that he did not consider the axioms of his theory as given by God. In particular, in the process of formulation of the axiomatics of probability theory he seriously considered the possibility of proceeding with nonadditive probability measures. But he gave up: such theory would be too complicated. It is interesting that Feynman [120] pointed to the violation of additivity as the main property of quantum probability distinguishing it from classical probability. (He did not know about the Kolmogorov axiomatics of probability theory, so he referred to the Laplace probability theory.)

In my monograph [179] reader can find the references on Kolmogorov's works written in 1920s in which he discussed possible variations of the axiomatics of probability theory. And I was lucky to meet Kolmogorov and discuss the axiomatics and its possible modifications, when I was a graduate student. My supervisor Smolyanov introduced me to Kolmogorov in the process of the discussion on recommendation of our article [342] to Doklady USSR. We developed a generalization of probability theory with complex-valued distributions and proved an analog of the central limit theorem. In this theorem the role of the limiting Gaussian distribution was played by Feynman distribution (see [342] for its mathematical definition). We have some problems with publication of this note due to opposition from the probability community which was rigidly structured within the Kolmogorov theory. So, Smolyanov decided to speak directly with Kolmogorov. And Kolmogorov was open to our theory with "complex-valued probability distributions."

Smolyanov asked once Kolmogorov why he had never tried to work with quantum generalization of probability. Kolmogorov's answer was surprising. He said something as "it was to difficult to proceed deeply at the mathematical level of rigorousness."

https://doi.org/10.1017/9781009313469.002 Published online by Cambridge University Press