

Chapter 2

Fundamentals of Classical Probability and Quantum Probability Theory

Abstract In this chapter we present briefly the basic notions of classical and quantum theories of probability and information. This chapter is especially important for biologists, psychologists, experts in cognition, and sociologists who were not trained in quantum theory (but even classical theory is presented in a simple manner). We start with the presentation of the standard measure-theoretic formulation of the modern classical probability theory (Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin [1]). Then we turn to fundamentals of quantum formalism, including theory of open quantum systems and its generalizations.

Keywords Foundations of classical and quantum probability · Quantum formalism · Quantum states and observables · Quantum information · Born rule for probabilities · Quantum master equation

2.1 Short Introduction to Classical Probability Theory

2.1.1 Probability Space

The modern axiomatics of probability theory was invented by Andrei Nikolaevich Kolmogorov in 1933 [1]. A crucial point is the *representation of events by subsets* of some basic set Ω . The collection of subsets representing events should be sufficiently rich—to be able to perform set-theoretic operations such as the intersection, the union, and the difference of sets. However, at the same time it should be reasonably rich. If a too extended system of subsets is selected to represent events, then it may contain “events” which cannot be interpreted in a reasonable way. After selection of a proper system of sets to represent events, one assigns weights to these subsets:

$$A \mapsto P(A). \quad (2.1)$$

The probabilistic weights are chosen to be nonnegative real numbers and normalized by 1: $P(\Omega) = 1$, the probability that something happens equals one. An event with large weight is more probable than an event with small weight. The weight of an

event A that can be represented as the disjoint union of events A_1 and A_2 is equal to the sum of weights of these events. The latter property is called *additivity*. (There is the evident similarity with mass, area, volume.)

It is useful to impose some restrictions on the system of sets representing events: (a) set Ω containing all possible events and the empty set \emptyset are events (something happens and nothing happens); (b) the union of two sets representing events represents an event; (c) the intersection of two sets representing events represents an event; (d) the complement of a set representing an event, i.e., the collection of all points that do not belong to this set, again represents an event. These set-theoretic operations correspond to the basic operations of (Boolean) logic: “or”, “and”, “no”. And the modern set-theoretic representation of events is a mapping of propositions describing events onto sets with preservation of the logical structure. At the beginning of the mathematical formalization of probability theory the map (2.1) was defined on an algebraic structure corresponding to the logical structure, the *Boolean algebra* (invented by Boole, the creator of “Boolean logic” [2]). The set-system with properties (a)–(d) is called the *algebra of sets* (in the American literature, the field of sets).

In the case of finite Ω the map given by (2.1) with the above-mentioned properties is (measure-theoretic) probability. (Since Ω can contain billions of points, this model is useful in a huge class of applications.) Here $\Omega = \{\omega_1, \dots, \omega_N\}$. To determine any map (2.1), it is enough to assign to each elementary event its weight

$$0 \leq P(\omega_j) \leq 1, \quad \sum_j P(\omega_j) = 1.$$

Then by additivity this map is extended to the set-algebra consisting of all subsets of Ω :

$$P(A) = \sum_{\{\omega_j \in A\}} P(\omega_j).$$

However, if Ω is countable, i.e., it is infinite and its points can be enumerated, or “continuous”—e.g., a segment of the real line \mathbf{R} , then simple additivity is not sufficient to create a fruitful mathematical model. The map (2.1) has to be additive with respect to countable unions of disjoint events:

$$P(A_1 \cup \dots \cup A_n \cup \dots) = P(A_1) + \dots + P(A_n) + \dots, \quad (2.2)$$

and to work fruitfully with such maps (e.g., to integrate), one has to impose special restrictions on the system of sets representing events. It has to be not simply a set-algebra, but a σ -algebra of sets (in the American literature, a σ -field), i.e., (b) and (c) must be valid for countable unions and intersections of sets. In logical terms, it means that the operations “or” and “and” can be applied infinitely many times to form new events. Of course, this is a mathematical idealization of the real situation. One of the most important “continuous probability model” is based on $\Omega = \mathbf{R}$, i.e.,

elementary events are represented by real numbers. Typically a σ -algebra is selected as the *Borel σ -algebra*: it is generated by all half-open intervals, $[\alpha, \beta)$, $\alpha < \beta$, with the aid of the operations of the union, intersection, and complement.

We remark that although probability models with infinite spaces of elementary events play an important role in many applications, one can, in principle, read practically the whole book using finite Ω , the collection of events represented by all its subsets, and additive probability given by assigning weights to points of Ω .

Let Ω be a set and let \mathcal{F} be a σ -algebra of its subsets. A *probability measure* \mathbf{P} is a map from \mathcal{F} to the segment $[0, 1]$ normalized $P(\Omega) = 1$ and σ -additive, i.e., the equality (2.2) holds for disjoint sets belonging to \mathcal{F} .

By the Kolmogorov axiomatics [1] the *probability space* is a triple

$$\mathcal{P} = (\Omega, \mathcal{F}, P).$$

Points ω of Ω are said to be *elementary events*, elements of \mathcal{F} are *events*, P is *probability*.

Random observations are represented by random variables. We start with mathematically simplest random variables, the discrete ones. In fact, the majority of observables (classical and quantum) considered in this book are discrete-valued.

Discrete random variables on the Kolmogorov space \mathcal{P} are by definition functions $a : \Omega \rightarrow X_a$, where $X_a = \{\alpha_1, \dots, \alpha_n, \dots\}$ is a countable set (the *range of values*) such that the sets

$$C_\alpha^a = \{\omega \in \Omega : a(\omega) = \alpha\}, \alpha \in X_a, \quad (2.3)$$

belong to \mathcal{F} .

It is typically assumed that the range of values X_a is a subset of the real line. We will proceed under this assumption practically everywhere, but sometimes, e.g., in cognitive and psychological modeling, it will be more convenient to consider Boolean labels, e.g., $\alpha = \text{yes, no}$.

The probability distribution of a (discrete) random variable a is defined as

$$p^a(\alpha) \equiv P(\omega \in \Omega : a(\omega) = \alpha).$$

We remark that

$$p^a(\alpha_1) + \dots + p^a(\alpha_n) + \dots = 1, \quad p^a(\alpha_n) \geq 0. \quad (2.4)$$

The *average* (mathematical expectation) of a random variable a is defined as

$$\bar{a} \equiv Ea = \alpha_1 p^a(\alpha_1) + \dots + \alpha_n p^a(\alpha_n) + \dots. \quad (2.5)$$

If the set of values of ξ is infinite, then the average is well defined if the series in the right-hand side of (2.5) converges absolutely.

For a family of random variables a_1, \dots, a_m , their *joint probability distribution* is defined as

$$p^{a_1 \dots a_m}(\alpha_{j_1}^1, \dots, \alpha_{j_m}^m) = P(\omega \in \Omega : a_1(\omega) = \alpha_{j_1}^1, \dots, a_m(\omega) = \alpha_{j_m}^m). \quad (2.6)$$

We remark that the joint probability is symmetric with respect to permutations; e.g., for two random variables a and b , we have

$$p^{ab}(\alpha, \beta) = P(\omega \in \Omega : a(\omega) = \alpha, b(\omega) = \beta) = p^{ba}(\beta, \alpha). \quad (2.7)$$

For two (discrete) random variables a and b , their *covariance* is defined as

$$\text{cov}(a, b) = E(a - \bar{a})(b - \bar{b}) = \sum_{\alpha\beta} (\alpha - \bar{a})(\beta - \bar{b}) p^{ab}(\alpha, \beta). \quad (2.8)$$

It is easy to see that

$$\text{cov}(a, b) = \overline{ab} - \bar{a}\bar{b}. \quad (2.9)$$

We remark that covariance is symmetric

$$\text{cov}(a, b) = \text{cov}(b, a). \quad (2.10)$$

Suppose now that the results of random measurement cannot be represented by a finite or countable set. Thus such an observable cannot be represented as a discrete random variable.

A *random variable* on the Kolmogorov space \mathcal{P} is by definition any function $\xi : \Omega \rightarrow \mathbf{R}$ such that for any set Γ belonging to the Borel σ -algebra, its pre-image belongs to the σ -algebra of events $\mathcal{F} : \xi^{-1}(\Gamma) \in \mathcal{F}$, where $\xi^{-1}(\Gamma) = \{\omega \in \Omega : \xi(\omega) \in \Gamma\}$.

In this case the mathematical formalism is essentially more complicated. The main mathematical difficulty is to define the integral with respect to a probability measure on \mathcal{F} , the *Lebesgue integral* [1]. In fact, classical probability theory, being based on measure theory, is mathematically more complicated than quantum probability theory. In the latter integration is replaced by the trace operation for linear operators and the trace is always the discrete sum. Here we are not able to present theory of Lebesgue integration. We formally invent the symbol of integration.¹

¹ We remark that, for a discrete random variable, the integral coincides with the sum for the mathematical expectation, see (2.5). And a discrete random variable is integrable if its mathematical expectation is well defined. In general any integrable random variable can be approximated by integrable discrete random variables and its integral is defined as the limit of the integrals for the approximating sequence.

The *average* (mathematical expectation) of a random variable a is defined as

$$\bar{a} \equiv Ea = \int_{\Omega} \xi(\omega) dP(\omega). \quad (2.11)$$

The probability distribution of a random variable a is defined (for Borel subsets of the real line) as

$$p^a(\Gamma) \equiv P(\omega \in \Omega : a(\omega) \in \Gamma).$$

This is a probability measure on the Borel σ -algebra. And the calculation of the average can be reduced to integration with respect to p^a :

$$\bar{a} \equiv Ea = \int_{\mathbf{R}} x dp^a(x). \quad (2.12)$$

For simplicity, further we shall proceed only with discrete random variables. The general case can be treated by using everywhere integrals instead of sums.

2.1.2 Conditional Probability

Kolmogorov's probability model is based on a probability space equipped with the operation of conditioning. In this model *conditional probability* is defined by the well known *Bayes' formula*

$$P(B|C) = P(B \cap C)/P(C), \quad P(C) > 0. \quad (2.13)$$

By Kolmogorov's interpretation it is the *probability of an event B to occur under the condition that an event C has occurred*.

We remark that $P_C(B) \equiv P(B|C)$ is again a probability measure on \mathcal{F} . For a set $C \in \mathcal{F}$, $P(C) > 0$, and a (discrete) random variable a , the conditional probability distribution is defined as

$$p_C^a(\alpha) \equiv P(a = \alpha|C).$$

We naturally have

$$p_C^a(\alpha_1) + \cdots + p_C^a(\alpha_n) + \cdots = 1, \quad p_C^a(\alpha_n) \geq 0. \quad (2.14)$$

The conditional expectation of a random variable a is defined by

$$E(a|C) = \alpha_1 p_C^a(\alpha_1) + \cdots + \alpha_n p_C^a(\alpha_n) + \cdots \quad (2.15)$$

For two random variables a and b , consider conditional probabilities

$$p_{\alpha\beta} \equiv P(b = \beta|a = \alpha), \quad p_{\beta\alpha} \equiv P(a = \alpha|b = \beta).$$

These conditional probabilities are called *transition probabilities*. (This terminology is also used in quantum probability, Sect. 4.1.1.)

These conditional probabilities can also be written in the form

$$p_{\alpha\beta} = P(b = \beta | C_\alpha^a), \quad p_{\beta\alpha} = P(a = \alpha | C_\beta^b), \quad (2.16)$$

where, e.g., C_α^a is defined by (2.3). It is, of course, assumed that in the first case $p^a(\alpha) > 0$ and in the second case $p^b(\beta) > 0$.

It is useful to consider the matrices of transition probabilities

$$\mathbf{P}^{b|a} = (p_{\alpha\beta}), \quad \mathbf{P}^{a|b} = (p_{\beta\alpha}). \quad (2.17)$$

We remark that these matrices are always *left stochastic*. The left stochastic matrix is a square matrix whose columns consist of nonnegative real numbers whose sum is 1. For example, for the matrix $\mathbf{P}^{b|a}$, we have

$$\sum_{\beta} p_{\alpha\beta} = \sum_{\beta} P(b = \beta | a = \alpha) = 1. \quad (2.18)$$

for any fixed $a = \alpha$. This is a consequence of the fact that, for any set C of strictly positive probability, P_C is also a probability measure.

We point to the following equality connecting the joint probability distribution of two random variables a and b with their transition probabilities

$$p^{ab}(\alpha, \beta) = p^a(\alpha) p_{\alpha\beta} = p^b(\beta) p_{\beta\alpha} = p^{ba}(\beta, \alpha). \quad (2.19)$$

In our further considerations one special class of matrices of transition probabilities will play an important role. These are so called doubly stochastic matrices. We recall that in a *doubly stochastic matrix* all entries are nonnegative and in all rows and all columns they sum to 1. Of course, in general, $\mathbf{P}^{b|a}$ is only left stochastic, but not doubly stochastic. In Sect. 4.1.1 we shall see that for quantum probability, the matrices of transition probabilities are always doubly stochastic for observables of the most important for applications class, observables with non-degenerate spectra.

Consider now a pair of dichotomous random variables $a = \alpha_1, \alpha_2$ and $b = \beta_1, \beta_2$. The matrix of transition probabilities $\mathbf{P}^{b|a}$ has the form

$$\mathbf{P}^{b|a} = \begin{pmatrix} p_{\alpha_1\beta_1} & p_{\alpha_2\beta_1} \\ p_{\alpha_1\beta_2} & p_{\alpha_2\beta_2} \end{pmatrix} \quad (2.20)$$

It is doubly stochastic iff $p_{11} = p_{22}$ and $p_{12} = p_{21}$, i.e.,

$$\mathbf{P}^{b|a} = \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix} \quad (2.21)$$

In particular, it is automatically symmetric.

2.1.3 Formula of Total Probability

In our further considerations the important role will be played by the *formula of total probability* (FTP). This is a theorem of the Kolmogorov model. Let us consider a countable family of disjoint sets A_k belonging to \mathcal{F} such that their union is equal to Ω and $P(A_k) > 0, k = 1, \dots$. Such a family is called a *partition* of the space Ω .

Theorem 2.1 *Let $\{A_k\}$ be a partition. Then, for every set $B \in \mathcal{F}$, the following formula of total probability holds*

$$P(B) = P(A_1)P(B|A_1) + \dots + P(A_k)P(B|A_k) + \dots \quad (2.22)$$

Proof We have

$$P(B) = P\left(B \cap \left(\bigcup_{k=1}^{\infty} A_k\right)\right) = \sum_{k=1}^{\infty} P(B \cap A_k) = \sum_{k=1}^{\infty} P(A_k) \frac{P(B \cap A_k)}{P(A_k)}.$$

Especially interesting for us is the case where a partition is induced by a discrete random variable a taking values $\{\alpha_k\}$. Here,

$$A_k = C_{\alpha_k}^a = \{\omega \in \Omega : a(\omega) = \alpha_k\}. \quad (2.23)$$

Let b be another random variable. It takes values $\{\beta_j\}$. For any $\beta \in X_b$, we have

$$P(b = \beta) = P(a = \alpha_1)P(b = \beta|a = \alpha_1) + \dots + P(a = \alpha_k)P(b = \beta|a = \alpha_k) + \dots \quad (2.24)$$

or in compact notation,

$$p^b(\beta) = p^a(\alpha_1)p_{\alpha_1\beta} + \dots + p^a(\alpha_k)p_{\alpha_k\beta} + \dots \quad (2.25)$$

2.2 Short Introduction of Quantum Probability Theory

We shall present the formalism of quantum mechanics (and quantum information) by small portions, from section to section. Advanced notions such as, e.g., channels and lifting, will be introduced in Chap. 4. However, the later notions will be used to construct a generalization of the conventional quantum formalism, which will serve for biological purposes. We start with the presentation of the basic structure of the quantum formalism.

2.2.1 States and Observables

Following Heisenberg, observables are represented by Hermitian matrices or, in the abstract framework, by Hermitian operators. These operators act in the complex Hilbert space \mathcal{H} , i.e., a complex linear space endowed with a scalar product denoted as $\langle \psi_1, \psi_2 \rangle$.

Here, let us recall the mathematical definition and properties of scalar product.

The scalar product is a function from the Cartesian product $\mathcal{H} \times \mathcal{H}$ 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: $\langle \psi_1, \psi_2 \rangle = \overline{\langle \psi_2, \psi_1 \rangle}$ ²
3. Linearity with respect to the first argument:

$$\langle k_1 \psi_1 + k_2 \psi_2, \phi \rangle = k_1 \langle \psi_1, \phi \rangle + k_2 \langle \psi_2, \phi \rangle,$$

where k_1, k_2 are complex numbers.

From the second and third properties it is easy to obtain that for the second argument, $\langle \phi, k_1 \psi_1 + k_2 \psi_2 \rangle = \bar{k}_1 \langle \phi, \psi_1 \rangle + \bar{k}_2 \langle \phi, \psi_2 \rangle$. By fixing in \mathcal{H} an orthonormal basis (e_j) , i.e., $\langle e_i, e_j \rangle = \delta_{ij}$, we represent vectors by their coordinates $\psi_1 = (z_1, \dots, z_n, \dots)$, $\psi_2 = (w_1, \dots, w_n, \dots)$. In the coordinate representation the scalar product has the form $\langle \psi_1, \psi_2 \rangle = \sum_j z_j \bar{w}_j$. By using this representation the reader can easily verify the aforementioned properties of the scalar product.

We remark that the usage of *complex numbers* plays the crucial role. One cannot proceed with real Hilbert spaces. There are experimental statistical data, which cannot be embedded in the real model.

The norm (an abstract analog of the Euclidean length) of a vector is defined as $\|\psi\| = \sqrt{\langle \psi, \psi \rangle}$. In the fixed system of coordinates $\|\psi\| = \sqrt{\sum_j |z_j|^2}$. Normalized vectors of \mathcal{H} , i.e., ψ such that $\|\psi\| = 1$, represent a special (and the most important) class of states of quantum systems, *pure states*. Each pure state ψ can be represented as an operator acting in \mathcal{H} , which is the orthogonal projections π onto the vector ψ . In terms of the scalar product, the orthogonal projections can be written as $\pi\phi = \langle \phi, \psi \rangle \psi$. By fixing an orthonormal basis in \mathcal{H} , the pure state is expressed by a matrix $\rho = (\rho_{ij})$ satisfying

- (a) Hermitian: $\rho_{ij} = \bar{\rho}_{ji}$, in particular, the diagonal elements are real,
- (b) Positive definiteness: $\langle \rho\phi, \phi \rangle \geq 0$ for any vector ϕ ,
- (c) Its trace equals 1: $\text{tr}\rho = \sum_j \rho_{jj} = 1$.

The eigenvalues of Hermitian matrix are real numbers. In addition, a Hermitian matrix with the positive definiteness has non-negative eigenvalues. Finally, from the condition c), the sum of all the eigenvalues equals 1.

As an example in the two dimensional space $\mathcal{H} = \mathbb{C}^2$, we introduce the operator given by the following matrix

² Here bar denotes complex conjugation; for $z = x + iy$, $\bar{z} = x - iy$.

$$\rho_{\text{qubit}} = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} = \begin{pmatrix} |\alpha|^2 & \alpha^* \beta \\ \alpha \beta^* & |\beta|^2 \end{pmatrix} \quad (2.26)$$

for the corresponding pure state vector

$$\psi = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

Here α and β are complex numbers satisfying $|\alpha|^2 + |\beta|^2 = 1$. The above ρ_{qubit} expresses the state of *quantum bit* (shortly called *qubit*), which is often seen in quantum information theory (see Sect. 2.2.4). One can easily see that ρ_{qubit} satisfies the following condition:

$$\rho_{\text{qubit}}^2 = \rho_{\text{qubit}}. \quad (2.27)$$

This condition guarantees that the operator ρ_{qubit} is a projection on a vector ψ . In general, there are operators that are not projectors but satisfy the conditions (a)–(c).

The next step in the development of the quantum formalism (due to Landau and von Neumann) was proceeding without the Eq. 2.27 constraint, i.e., considering all possible matrices satisfying conditions (a)–(c). They are called *density matrices* and they represent the most general states of quantum systems. In the abstract framework one considers operators satisfying conditions (a)–(c), density operators. Each density operator can be written as a weighted sum of projection operators corresponding to pure states. If such a sum contains more than one element, then the state represented by this density operator is called a *mixed state*: the mixture of pure states with some weights. Although this terminology is widely used, it is ambiguous. The representation of a density operator as a weighted sum of projectors corresponding to pure states is not unique. Thus by using the terminology mixed state one has to take into account this non-uniqueness.

For simplicity let us restrict our consideration to the case of a finite dimensional Hilbert space.

Take an arbitrary quantum observable represented by a Hermitian operator A . For any Hermitian operator, there exists an orthogonal basis of \mathcal{H} consisting of its eigenvectors. The values of the observable operationally represented by A are encoded in the eigenvalues of this operator, a_1, \dots, a_n . (We shall use the same symbol for an observable and its operator representation.)

Consider now the case of nondegenerate spectrum: $a_j \neq a_i$ if $j \neq i$. For an ensemble of quantum systems prepared in the same state represented by the density operator ρ , the probability to obtain a fixed value $P(a_j)$ is encoded in the corresponding matrix element of the operator ρ (in the basis of eigenvectors):

$$P(a_j) = \rho_{jj}. \quad (2.28)$$

This is one of the basic postulates of quantum mechanics. It connects experimental probabilities with the operator representation of observables. (In the simplest form,

for the pure state, this postulate was proposed by Born, see (2.32); the form (2.28) is due to von Neumann [3]).

Thus all information about possible results of measurements is encoded in two Hermitian operators, the observable A and the state ρ . This is very compact and convenient representation.

In fact, mathematically, quantum formalism is a linear algebra (with elements of functional analysis for the infinite-dimensional case). Thus it is very simple; there is nothing simpler than a *linear representation*. The corresponding dynamical equations (e.g., the Schrödinger's equation) are linear.

The simplicity of the quantum linear representation of measurable quantities is one of the reasons to use this formalism, in particular, in biology.

Surprisingly, practically any theory of statistical measurements can be operationally represented in a linear space. Here we have no possibility to discuss this problem that was studied in detail already in 1950th by Mackey; recently one of the authors of this book has proposed a *quantum-like representation algorithm* [4]. This algorithm constructs the QL representation of statistical data collected for two observables (random variables) under some restrictions on the matrix of transition probabilities for these observables. Hence, all “natural operational representations” are reduced to linear representations (in some situations one has to use number systems different from complex numbers, e.g., the so called hyperbolic numbers [4]).

2.2.2 Superposition

Let the state space of some system (physical or biological) be represented as a finite-dimensional Hilbert space \mathcal{H} . Consider a pure state ψ and an observable A , denote its eigenvalues by a_1, \dots, a_m and the corresponding eigenvectors by e_1, \dots, e_m . This is an orthonormal basis in \mathcal{H} . (We again proceed under the assumption that all eigenvalues are different.) We expand the vector ψ with respect to this basis:

$$\psi = c_1 e_1 + \dots + c_m e_m, \quad (2.29)$$

where (c_j) are complex numbers such that the sum of their squared absolute values are equal to one (this is the coordinate expression of the normalization by one of a pure state vector):

$$|c_1|^2 + \dots + |c_m|^2 = 1. \quad (2.30)$$

By using the terminology of linear algebra we say that the pure state ψ is a *superposition* of pure states e_j .

The density matrix corresponding to ψ has the elements

$$\rho_{ij} = c_i \bar{c}_j. \quad (2.31)$$

Hence, for the pure state ψ , the basic probabilistic postulate of quantum mechanics, (2.28), has the form

$$P(a_j) = \rho_{jj} = c_j \bar{c}_j = |c_j|^2. \quad (2.32)$$

This postulate can be written without using the coordinates of the state vector ψ with respect to the basis of eigenvectors of a quantum observable. We remark that, since the basis of eigenvectors of a Hermitian operator can always be selected as orthonormal, the coordinates c_j can be expressed in the form: $c_j = \langle \psi, e_j \rangle$. Hence, the Born's rule takes the form:

$$P(a_j) = |\langle \psi, e_j \rangle|^2. \quad (2.33)$$

Projection Postulate; Resolution of Uncertainty

The next natural question is about the post-measurement state. What will happen with the state ψ after a measurement? By the von Neumann *projection postulate* the superposition (2.29) is reduced to just one term, the state e_j corresponding to the eigenvalue a_j obtained in the measurement. We remark that we consider the simplest case: all eigenvalues are different from each other; the case of nondegenerate spectrum.

This procedure can be interpreted in the following way:

This superposition encodes the uncertainty in the results of measurements for the observable A. Roughly speaking, before a measurement a quantum system “does not know how it will answer to the question A.” The mathematical expressions (2.29) and (2.32) encode potentialities for different answers. Thus a quantum system in the superposition state ψ does not have any value of A as its objective property.³ After a measurement the superposition is reduced to just one term in the expansion (2.29) corresponding the value of A obtained in the process of measurement.

We remark that the state reduction is often called *state collapse*. Some experts in quantum foundations treat superposition physically and not simply operationally; for them, the collapse is also a physical event.⁴ We state again that in this book we proceed with the operational interpretation of the quantum formalism. By this interpretation superposition (2.29) expresses the uncertainty in the expected results of the A-measurement. And nothing more! When the result a_j is detected, this uncertainty is resolved. Hence, “collapse” takes place not in the physical space, but in the information space.

Encoding of the uncertainty in bio-systems by superpositions is one of the cornerstones of quantum bio-informatics. A bio-system as well as a quantum physical system can be in a state of uncertainty on possible reactions to measurements. Such states are *mathematically encoded* as linear superpositions. Measurement resolves such superpositions.

³ See Sect. 1.3 for a further discussion.

⁴ In particular, by the orthodox Copenhagen interpretation ψ is interpreted as the physical state of a system. As a consequence, its collapse is a physical event.

Degenerate Eigenvalues

Consider now the general case: the eigenvalues can be degenerate and eigensubspaces⁵ need not be one dimensional. Suppose that by measuring an observable represented by the Hermitian operator A its eigenvalue a was obtained; denote the projector to this eigensubspace by π_a . Then by the projection postulate the input pure state ψ is transformed into (again pure) state

$$\psi_{\text{out};a} = \frac{\pi_a \psi}{\|\pi_a \psi\|}. \quad (2.34)$$

We remark that although this is the standard definition used in modern quantum theory, von Neumann distinguished sharply the cases of non-degenerate spectrum (i.e., in the finite dimensional case, all eigenvalues are different) and degenerate spectrum. In the latter case the definition (2.34) is due to Lüders. Originally von Neumann assumed that a measurement can transfer a pure state into a mixed state given by a density operator, see [3]. However, we would not disturb biologists by such foundational problems. Hence, we proceed with the Lüders form of the projection postulate even in the case of degenerate spectrum (as the simplest form of the projection postulate).⁶ If the input state is given by the density operator ρ , then

$$\rho_{\text{out};a} = \frac{\pi_a \rho \pi_a}{\text{tr} \pi_a \rho \pi_a}. \quad (2.35)$$

Suppose now that by measuring A we are interested in the state (mixed) representing the ensemble of states of all systems after measurement, so this ensemble is the mixture of the ensembles corresponding to the results $A = a_j$ for different a_j . Then this state is mathematically represented by the density operator

$$\rho_{\text{out}} = \sum_a \pi_a \rho \pi_a. \quad (2.36)$$

It can be represented as a mixture of the states $\rho_{\text{out};a}$:

$$\rho_{\text{out}} = \sum_a P_a \rho_{\text{out};a}, \text{ where } P_a = \text{tr} \pi_a \rho \pi_a. \quad (2.37)$$

Self-measurements

We stress that in quantum bio-informatics we interpret the notion of measurement in a more general way than in quantum physics. In particular, measurements can be *self-measurements*, see Sect. 1.3; we also treat *decoherence* as a form of measurement.

⁵ The set of all the eigenvectors is a linear subspace of \mathcal{H} .

⁶ Even the majority of physicists have never read the von Neumann's book [3] and they have no idea that von Neumann distinguished degenerate and non-degenerate cases. We are aware that this distinction may play an important role in biology.

Decoherence is the process of transformation of superposition (2.29) into the diagonal density matrix, $\rho = \text{diag}(\rho_{11}, \dots, \rho_{mm})$ corresponding to the total resolution of uncertainty.⁷

Self-measurements play an important role in the brain science, cognitive science, psychology. Decoherence will be applied to describe processes of decision making (in cognitive psychology and game theory, Chap. 4) and in modeling of biological evolution, including evolution on the cellular level, Chap. 8.

2.2.3 Dirac Notation

Dirac was one of the inventors of the quantum formalism [5]. He introduced his own symbolic notation for the Hilbert space linear algebra which was not used in mathematics. This notation became very common in quantum physics and especially in quantum information theory. We shall now present shortly Dirac's symbolic notation. In this books we shall use it very often (but not always).

Consider an observable A , denote its eigenvalues by a_1, \dots, a_m and the corresponding eigenvectors by e_1, \dots, e_m . Suppose again that all eigenvalues are different. Then by Dirac e_j is denoted as $|a_j\rangle$. Even an arbitrary pure state ψ is often written as $|\psi\rangle$ (just for convenience). In Dirac's notation superposition (2.29) is written as

$$|\psi\rangle = c_1|a_1\rangle + \dots + c_m|a_m\rangle, \quad (2.38)$$

or simply $|\psi\rangle = \sum_a c_a|a\rangle$. (We repeat that this is just symbolic expression for the operator with the set of eigenvectors $|a\rangle$.) The sum “can be continuous”, i.e., it will be an integral representation (in the case of observables with continuous spectra such as e.g. position and momentum of a quantum particle): $|\psi\rangle = \int c(a)|a\rangle da$, where $\int |c(a)|^2 da = 1$.

The Born's rule (2.33) is now written as

$$P(a_j) = |\langle\psi|a_j\rangle|^2. \quad (2.39)$$

Consider some pure state $|\psi\rangle$, then the symbol $|\psi\rangle\langle\psi|$ denotes the operator of orthogonal projection onto the vector $|\psi\rangle$. Thus the operator A with the system of distinct eigenvalues a_1, \dots, a_m can be written as

$$A = \sum_{i=1}^m a_i |a_i\rangle\langle a_i|,$$

⁷ Decoherence is a complicated interpretational issue of quantum mechanics. Some (but not all) researchers treat decoherence as a form of measurement.

or simply $A = \sum_a a|a\rangle\langle a|$. We shall come back to the problem of encoding of quantum information in Sect. 2.5.3, after the introduction of state spaces of compound systems (given by tensor products of Hilbert state spaces of subsystems).

2.2.4 Qubit

The basic notion of quantum information theory is quantum bit (qubit), the quantum analog of classical bit of information. Consider any dichotomous observable $A = a_1, a_2$ with $a_1 < a_2$, in particular, it acts in the two dimensional Hilbert space. Let us encode 0, 1 by its eigenvectors. We can always calibrate the pointer of the corresponding measurement device in such a way that $a_1 = 0$ and $a_2 = 1$, the encoding rule has the form $0 \rightarrow |0\rangle, 1 \rightarrow |1\rangle$. The crucial point is that, besides the states $|\alpha\rangle, \alpha = 0, 1$, a quantum system can be in superposition of these states,

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle, \quad |c_0|^2 + |c_1|^2 = 1. \quad (2.40)$$

Thus a single quantum system can carry with its state not just either 0 or 1, i.e., one bit of information, but both 0 and 1, with some probabilistic weights. However, by measurement of A it is possible to extract just one bit of information by getting either the result $A = 0$ or $A = 1$. The crucial point is that (2.40) represents not simply the classical probability distribution with two weights $p_0 = |c_0|^2, p_1 = |c_1|^2$. Qubit can generate interference effects, but the classical probability distribution not. We shall discuss this problem in more detail later.

2.2.5 Wave Function

Up to now we have restricted the presentation to the finite dimensional case. In real physics the Hilbert space of quantum states is *infinite dimensional*, the basic example is given by the space of square integrable complex valued functions, $x \rightarrow \psi(x)$, where x is the spatial variable, i.e., functions such that $\int |\psi(x)|^2 dx < \infty$. We denote this space by the symbol $L^2(\mathbb{R}^n)$, where $x \in \mathbb{R}^n$. A pure state ψ , i.e., a function such that

$$\int |\psi(x)|^2 dx = 1, \quad (2.41)$$

is known as a *wave function*.

For the position observable, the spatial variable x plays the role of the discrete index j in (2.32) (as well as in (2.29), (2.30)). Since the dimension of state space is infinite, eigenfunctions are in general generalized, i.e., they do not belong to the state space. A larger space has to be in use. In our case such a larger space containing eigenfunctions is some space of distributions (generalized functions). For the position operator, the (generalized) eigenfunction corresponding to the (generalized) eigenvalue x_0 is given by the Dirac δ -function whose argument is shifted by x_0 , i.e.,

$e_{x_0}(x) = \delta(x - x_0)$. Thus eigenfunctions $e_{x_0}(x)$ are “enumerated” with the aid of the continuous index x_0 . And formally any square integrable function can be represented as the superposition of eigenfunctions of the position operator:

$$\psi(x) = \int \psi(x_0) e_{x_0}(x) dx_0. \quad (2.42)$$

This is a continuous analog of superposition (2.29). A continuous analog of the normalization condition (2.30) is given by the equality (2.41). For a pure state, the probabilistic postulate of quantum mechanics has the form:

$$P(x) = |\psi(x)|^2, \quad (2.43)$$

“the probability to find a quantum system with the wave function ψ at the point x is equal to the squared absolute value of this function at this point.” This is the original formulation of the basic probabilistic postulate of quantum mechanics, *Born’s rule*. This is a postulate; it cannot be derived in the conventional quantum theory; however, cf. [6].

In the real quantum physics state spaces are always of the L^2 -type, i.e., they are infinite-dimensional. However, in quantum information one typically selects just a few degrees of freedom, which are then used for the representation of information and proceed in finite dimensional state space. In this book we shall apply the same strategy. However, physical state space and wave function will be used in Sect. 4.1.1, which is devoted to the presentation of interference of photons.

2.3 Schrödinger Dynamics and Its Role in Quantum Bio-informatics

We can assume that, like in a quantum physical system, the evolution of the QL-state of a bio-system isolated from the environment is described by the *Schrödinger equation* (with a minor modification related to the usage of the Planck constant in physics):

$$i\gamma \frac{d\psi}{dt} = H\psi(t), \quad \psi(0) = \psi^0. \quad (2.44)$$

Instead of the Planck constant \hbar , we use a constant γ , which has the dimension of time. We recall that the *Planck constant* has the dimension of action, i.e., energy \times time. Here H is a Hermitian operator acting in the space of QL-states of bio-systems. In quantum physics \mathcal{H} is the energy operator; it is called *Hamiltonian*. The energy dimension of this operator matches with the action dimension of the Planck constant. In quantum bio-informatics this operator does not represent the physical energy of a bio-system. (We state again that quantum bio-informatics is not at all quantum biology.) It is the generator of purely information dynamics. It is convenient to have

it dimensionless and this motivates the usage of the time scale constant γ , instead of the action constant \hbar . In coming QL-models γ gives the time scale of information processing in biological systems. We do not claim that there exists some universal constant scaling dynamics in all bio-systems. Each population of bio-systems can have its own temporal scale.

We shall also call the generator of the Schrödinger dynamics Hamiltonian (the operator H) and we hope that it will not cause misunderstanding.

The Schrödinger equation is, in fact, a system of linear differential equations with complex coefficients; in the one dimensional case \mathcal{H} is just a real number and the general solution has the form of the imaginary exponent: $\psi(t) = e^{\frac{-it\mathcal{H}}{\gamma}} \psi^0$. In the general case \mathcal{H} is an operator and the solution is represented in the form of imaginary operator-exponent (for the fixed basis it is simply the exponent of the matrix):

$$\psi(t) = U_t \psi^0, \quad U_t = e^{\frac{-it\mathcal{H}}{\gamma}}. \quad (2.45)$$

As well as the one dimensional imaginary exponent, the operator-exponent describes *oscillating dynamics*. It is more complicated than in the one dimensional case; it is a mixture of many oscillating imaginary exponents. For us, the most important point is that such a dynamics will never stabilize to any definite solution.

We point to the following fundamental property of the Schrödinger dynamics. The evolution operator U_t , see (2.45), is a *unitary operator*, i.e., it preserves the scalar product:

$$\langle U_t \psi | U_t \psi \rangle = \langle \psi | \psi \rangle. \quad (2.46)$$

Thus this dynamics transfers a pure quantum state to another pure quantum state.

In quantum physics, the Schrödinger equation plays a crucial role. It seems not to be the case in biology. Here it is very difficult and practically impossible to provide the complete isolation from environment. Take a single cell and isolate it from other cells. If we still provide the supply of chemical stuffs, this cell will survive, but its behavior in the absence of signaling from other cells will be very artificial. For example, such an important process as cell differentiation will be impossible. If we try to isolate a cell from the supply of all chemical stuffs, it will die. The same can be said about the isolated (mentally or/and physically) brain.

Since any general quantum state, a density operator, can be represented as a mixture of density operators corresponding to pure states, the Schrödinger dynamics for pure states implies the following dynamics for density operators:

$$\gamma \frac{d\rho}{dt}(t) = -i[H, \rho(t)], \quad \rho(0) = \rho^0. \quad (2.47)$$

(In quantum physics the Planck constant \hbar is used instead of the time scaling factor γ). This equation is known as the *Von Neumann equation* [3]. By using representation (2.45) of the Schrödinger evolution for the pure state we represent the evolution of the density operator (“mixed state”) in the form

$$\rho(t) = U_t^* \rho^0 U_t, \quad (2.48)$$

where, for an operator W , the symbol W^* denotes its *adjoint operator*. The latter is defined by the equality

$$\langle W\psi_1 | \psi_2 \rangle = \langle \psi_1 | W^* \psi_2 \rangle; \quad (2.49)$$

by denoting the matrix elements of these operators as w_{ij} and w_{ij}^* we have $w_{ij}^* = \bar{w}_{ji}$.

2.4 Theory of Open Quantum Systems in Biology

What kind of the mathematical apparatus can be used to model adaptive biological phenomena operationally? An important class of adaptive dynamical systems can be described by the apparatus of the theory of *open quantum systems*, see, e.g., Ingarden et al. [7], Ohya and Volovich [8]. In complete accordance with the operational viewpoint to the formalism of quantum physics, we can apply the theory of open quantum systems in biology, from genetics and cellular biology to the brain science, cognitive science and psychology. By the theory of open quantum systems, the dynamics of the QL-state of a bio-system interacting with an environment is approximately described by *quantum master equation*, the Gorini-Kossakowski-Sudarshan-Lindblad (GKSL) equation, see, e.g., [7].⁸

In cellular QL-models at the beginning of interaction with an environment the QL-state of a population of cells is characterized by a *high degree of uncertainty about possible changes* (cf. Sect. 2.2.2), which can be generated via the coupling with an environment. The quantum master equation describes the process of resolution of this state of uncertainty and approaching the complete matching with the environment, Sect. 2.4. This process can be considered as *decoherence of the QL-state* through interaction with an environment. As a result, the state loses its fundamentally quantum(-like) feature, *superposition* of a few alternatives, and the final situation can be described by classical probability theory. Mathematically, this is the process of approaching a *steady state* solution of the quantum master equation.

The first application of the theory of open quantum systems to biology was in the domain of cognitive psychology and game theory, see [9, 10]. The classical probability distribution corresponding to the steady state was used to model decision making by cognitive systems; in particular, by gamblers in games of the Prisoner's Dilemma type; cf. with QL-modeling performed on the basis of Schrödinger equation [11, 12].

One of the novel proposals realized in this book is *to use the machinery of the theory of open quantum systems and to describe dynamics of the QL-state of a population of bio-systems* by using the quantum master equation (the GKSL-equation). We state again that this equation can be used to describe transitions from states

⁸ In quantum information theory this equation is often referred as Lindblad equation.

of uncertainty given by superpositions to classical probability distributions. Hence, such an equation cannot be an equation with respect to the pure state represented as a *vector* belonging to complex Hilbert space and normalized by one (We remark that the Schrödinger equation is an equation with respect to the pure states!). Pure states represent superpositions of possibilities which have to disappear at the end: when a bio-system will make its decision.⁹ We have to use general quantum states represented by *density operators*. As was stressed in Sect. 2.2, initial pure quantum states (superpositions describing uncertainty) and the final classical probability distribution can be represented by density operators. (Although the picture of a population of bio-systems making decisions is very illustrative in the brain science, cognitive science and psychology, in general we use the operational interpretation: the output (classical) probabilities give us the probability distribution of decisions.)

In the quantum Markovian approximation, the dynamics of the state of a system interacting with an environment is described by the GKSL-equation (again with a minor modification related to the usage of the Planck constant in quantum physics):

$$\gamma \frac{d\rho}{dt} = -i[H, \rho(t)] + \mathscr{W} \rho(t), \rho(0) = \rho^0, \quad (2.50)$$

where H is a Hermitian operator determining the internal dynamics of the QL-state of a population of bio-systems isolated from the environmental pressure (“cell’s Hamiltonian”) and the linear operator \mathscr{W} describes the environmental pressure.

Opposite to H (which is simply a Hermitian operator; for a fixed basis it is represented by a Hermitian matrix), in general the operator \mathscr{W} has a complex mathematical structure. It has such a form that starting with a density operator ρ^0 we shall get density operators at all instances of time. For a moment, the specific structure of \mathscr{W} is not important for us; see, e.g., [7, 8], and Sect. 8.2.3 for mathematical details. Biologically this operator is determined by the properties of the environment, including the initial state of the environment.

Here γ is the time scale constant, it determines the temporal scale of the biological dynamics. By using such a scaling factor of the dimension of time, we are able to proceed with dimensionless Hamiltonian H and the environmental operator \mathscr{W} .

We state once again that for our QL-modeling it is crucial that, for a very general class of GKSL-equations, the environmental operator \mathscr{W} drives (in the limit $t \rightarrow \infty$) the QL-state of a biological population, $\rho(t)$, to the *steady solution*: $\rho(t) \rightarrow \rho_{\text{st}}$. Typically the uncertainty (in the form of superposition) is eliminated from the asymptotic state ρ_{st} .

In our QL-model such a steady state is considered as the result of the biological dynamics in the environment (mathematically represented by the operator \mathscr{W}). For example, we can consider *epigenetic evolution* (Chap. 8). Here the limiting probability distribution $\rho_{\text{epi;st}}$ describes the probability distribution of epimutations which took place in a cell population as a consequence of interaction with the environment.

⁹ We treat the notion of decision very generally: from decisions made by people to “cell’s decisions”, e.g., to undergo epimutation.

Internal uncertainty, to epimutate or not epimutate, was resolved and a stable *phenotype* was created.

Finally, we remark that under natural restrictions a selection operator produces the *same steady state for all possible initial states*, Sect. 8.3.1. Such an open system dynamics simulates, for example, the spreading of mutations or epimutations in a population. The pressure of the environment can be so strong that the same genotype or phenotype is created independently of the initial states of populations. Such a dynamics can also have applications in social science and political technologies (See, e.g., [13, 14] for modeling of dynamics of *party governance* in the USA political system. Here a QL-model of decision-making of American voters was elaborated. In the framework of open quantum systems a possibility of driving populations of voters to the fixed state, e.g., to vote for republicans independently of their initial preferences was modeled.)

We shall continue our discussion on the theory of open quantum systems and the theory of adaptive quantum dynamics in Sect. 2.6 after the introduction to the theory of compound quantum systems.

2.5 Compound Systems

The quantum description of a compound system S consisting of two subsystems S_1 and S_2 with state (Hilbert) spaces \mathcal{H}_1 and \mathcal{H}_2 , respectively, is based on the representation of states of such a system in the tensor product space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. Since the notion *tensor product* is not used so much in biology,¹⁰ in the following sections we present briefly the construction of Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$. Although in quantum information theory and in this book we shall use the formal algebraic definition, which is especially useful for finite dimensional Hilbert spaces, we prefer to start with the construction originally used by Von Neumann [3], namely, the tensor product of two spaces of square integrable functions.

Tensor product of functional spaces. Let now both state spaces be L^2 -spaces, $\mathcal{H}_1 = L^2(\mathbb{R}^n)$ and $\mathcal{H}_2 = L^2(\mathbb{R}^m)$, see Sect. 2.2.5. Take two functions; $\psi \equiv \psi(x)$ belongs to \mathcal{H}_1 and $\phi \equiv \phi(y)$ belongs to \mathcal{H}_2 . By multiplying these functions we obtain the function of two variables $\Psi(x, y) = \psi(x) \times \phi(y)$. It is easy to check that this function belongs to the space $\mathcal{H} = L^2(\mathbb{R}^{n+m})$. Take now n functions, $\psi_1(x), \dots, \psi_n(x)$, from \mathcal{H}_1 and n functions, $\phi_1(y), \dots, \phi_n(y)$, from \mathcal{H}_2 and consider the sum of their pairwise products:

$$\Psi(x, y) = \sum_i \psi_i(x) \times \phi_i(y). \quad (2.51)$$

¹⁰ In the coordinate form tensor products of vectors and matrices are also known under the name *Kronecker product*. This structure is widely used in various computational algorithms including computational biology.

This function also belongs to \mathcal{H} .

It is possible to show that any function belonging to \mathcal{H} can be represented as (2.51), 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 $\mathcal{H}_2 = L^2(\mathbb{R}^m)$ is symbolically denoted as $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$.

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

$$\Psi = \sum c_{ij} e_{ij} \equiv \sum c_{ij} e_i^{(1)} \otimes e_j^{(2)}, \quad (2.52)$$

where

$$\sum |c_{ij}|^2 < \infty. \quad (2.53)$$

Those who work with electromagnetic signals in biology, e.g., in the brain research, have experience in expanding electromagnetic signals with respect to various bases, e.g., using the Fourier expansion or the wavelet expansion. Some bases are indexed by continuous parameters; integrals take place of sums. Thus the notion of basis in the L^2 -space is widely known. However, there is a crucial difference between the classical field and quantum mechanical representations of compound systems. The state of a classical bi-signal consisting of two components is represented in the *Cartesian product* of the corresponding L^2 -spaces. And the state of a quantum bi-system, e.g., bi-photon, is represented in the tensor product space (Sect. 2.5.1). One may state that the crucial difference between the classical and quantum physical models is in the representation of states of compound systems. (Although, as we have already seen in previous sections, the descriptions of non-compound systems differ essentially.)

Tensor product, the algebraic definition. Consider now two finite dimensional Hilbert spaces, $\mathcal{H}_1, \mathcal{H}_2$. For each pair of vectors $\psi \in \mathcal{H}_1, \phi \in \mathcal{H}_2$, 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 $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. In particular, if we take orthonormal bases in \mathcal{H}_k , $(e_j^{(k)})$, $k = 1, 2$, then $(e_{ij} = e_i^{(1)} \otimes e_j^{(2)})$ form an orthonormal basis in \mathcal{H} , any $\Psi \in \mathcal{H}$, can be represented as (2.52) with (2.53).

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 (2.52) satisfying the condition (2.53).

Besides the notion of tensor product of states, we shall also use the notion of tensor product of operators. Consider two linear operators $A_i : \mathcal{H}_i \rightarrow \mathcal{H}_i$, $i = 1, 2$. Their tensor product $A \equiv A_1 \otimes A_2 : \mathcal{H} \rightarrow \mathcal{H}$ is defined starting with the tensor products of two vectors: $A\psi \otimes \phi = (A_1\psi) \otimes (A_2\phi)$. Then it is extended by linearity.

By using the coordinate representation (2.52) the tensor product of operators can be represented as

$$A\Psi = \sum c_{ij} A e_{ij} \equiv \sum c_{ij} A_1 e_i^{(1)} \otimes A_2 e_j^{(2)}, \quad (2.54)$$

If the operators A_i , $i = 1, 2$, are represented by matrices (with respect to the fixed bases), i.e., $A_i = (A_{kl}^{(i)})$, then the matrix $A = (A_{kl, nm})$ 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, $\mathcal{H}_1, \dots, \mathcal{H}_n$, denoted by the symbol $\mathcal{H} = \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_n$. We start with forming the formal entities $\psi_1 \otimes \dots \otimes \psi_n$, where $\psi_j \in \mathcal{H}_j$, $j = 1, \dots, n$. 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 \mathcal{H}_k , $(e_j^{(k)})$, $k = 1, \dots, n$. Then any $\Psi \in \mathcal{H}$ can be represented as

$$\Psi = \sum_{\alpha} c_{\alpha} e_{\alpha} \equiv \sum_{\alpha=(j_1 \dots j_n)} c_{j_1 \dots j_n} e_{j_1}^{(1)} \otimes \dots \otimes e_{j_n}^{(n)}, \quad (2.55)$$

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

2.5.1 Entanglement

Consider a compound quantum system $S = (S_1, S_2)$, where the subsystems S_i , $i = 1, 2$, have state spaces \mathcal{H}_i . Then the state space of S is $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. States and observables are defined as it was presented in Sect. 2.2. However, the presence of the tensor product structure generates a new important notion, namely, *entanglement of states*. Consider the case of two uncorrelated systems S_1 and S_2 in states ψ_1 and ψ_2 . Then the state of the compound system S is given by the tensor product

$$\Psi = \psi_1 \otimes \psi_2. \quad (2.56)$$

(In particular, for wave functions we have $\Psi(x, y) = \psi_1(x) \times \psi_2(y)$.) However, if there are some correlations between the degrees of freedom in S_1 and S_2 , then the state of S cannot be represented in the form (2.56). Such a state is called an *entangled state*. The notion of entanglement can be generalized to nonpure states (which are represented by density operators acting in \mathcal{H} .) The interpretation of this notion is still the topic of hot debates [4]. There can be found variety of interpretations: from the original interpretation of Einstein et al. [15], who considered entanglement as the Hilbert space representation of classical correlations between systems to modern interpretations such as the subjective probability interpretation – entanglement as the correlation of probabilistic knowledge about subsystems in the mind of an observer, e.g., Fuchs and coauthors [16–18]. In this book we shall not try to select

one of possible interpretations of entanglement. For the operational approach, we are fine even with Fuchsian interpretation. However, we cannot reject a possibility of classical probabilistic interpretation of entanglement, see [6, 19–24]. In quantum bio-informatics we shall simply use the tensor product representation (and, hence, entanglement) of states for compound bio-systems. For example, this representation plays an important role in modeling correlations between the expressions of different genes in genome as well as correlations between mutations (as well as epimutations) in different genes, Chap. 8. In the theory of decision making we shall consider entanglement between states of different parties (typically named Alice and Bob), Chap. 4.

Is the usage of tensor product representation and entanglement biologically justified? It seems that this is not a proper question. In quantum physics we cannot find “real physical reasons” for the usage of the aforementioned mathematical constructions. They are fruitful, because they work well for the representation of statistical experimental data. In quantum bio-informatics we shall proceed in the same way. We remark that in the abstract algebraic framework elements of tensor products are defined in a formal way; there is no heuristic coupling between Ψ belonging to $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ and its components, e.g., (2.52), belonging to \mathcal{H}_i . In the operational approach this situation is natural. We just need a symbolic representation of states of compound quantum systems.¹¹

2.5.2 Tensor Products and Contextuality of Observables for a Single System

Tensor product structure can appear not only in the description of a state space of a compound system, but even in the case of a single system. Sometimes it is possible to factorize the state space \mathcal{H} of a single quantum system, say a neutron, into the tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ in such a way that two compatible observables (Hermitian operators A_1, A_2 acting in \mathcal{H}) can be represented in the form $A_1 = a_1 \otimes I$ and $A_2 = I \otimes a_2$, where $a_i, i = 1, 2$, are Hermitian operators in \mathcal{H}_i and I denotes the unit operator, $Ix = x$ (For neutrons, A_1 and A_2 can be selected as the position and spin observables.) Here we also can consider entangled states. However, this is *entanglement of observables* for a single system and not states of a few systems. This approach is very useful for quantum bio-informatics. For example, we can consider the entanglement of various genetic or epigenetic markers of a single cell

¹¹ “What is beyond this symbolism?”—this is a separate question (Sect. 9.2; see also [19]). In a series of papers [6, 20, 21, 23, 24] quantum systems were represented by classical random fields. In this approach elements of the tensor product can be visualized via the functional representation.

(Sect. 8.5). Using of the entanglement (-like)¹² model can explain the high speed of the epigenetic evolution, in which inheritable epigenetic markers can be created during only one generation of a cellular population.

2.5.3 A Few Words About Quantum Information

The notion of qubit was introduced in Sect. 2.2.4. Now we consider n -qubit states. One qubit space is two dimensional with the basis encoding 0 and 1: $|0\rangle, |1\rangle$. It is isomorphic to \mathbb{C}^2 . Now we consider the tensor product of n qubit spaces. The basis in this space has the form $|x\rangle \equiv |x_1\rangle \otimes \cdots \otimes |x_n\rangle$, where $x_j = 0, 1$. Thus each basis vector $|x\rangle$ encodes the string of zeros and ones of the length n . As in the one dimensional case, it is possible to form weighted superpositions of these basis vectors. These are quantum n -bit states. The dimension of the n -qubit state space is equal to 2^n . This space is isomorphic to \mathbb{C}^{2^n} . Typically one omits the signs of the tensor product and writes $|x\rangle \equiv |x_1 \cdots x_n\rangle$. Thus a general n -qubit state has the form

$$|\psi\rangle = \sum_x c_x |x\rangle, \text{ where } \sum_x |c_x|^2 = 1. \quad (2.57)$$

The main distinguishing feature of quantum information theory is that the dimension of a state space increases exponentially with the linear increasing of the number n of subsystems of a compound system or more generally entangled degrees of freedom (may be even of a single system, see Sect. 2.5.2). This is the main computational resource. And it is huge. A relatively small physical system (say $n = 100$) can process a gigantic information state (in our example, the superposition of 2^n bits of information). However, as in the single qubit case, this huge information resource is unapproachable: a measurement can give us only one result.

Quantum computations are based on the physical realization of the Schrödinger dynamics (unitary transformations) in the n -qubit spaces and the final measurement. The last step has to be designed in an intelligent way, since the solution of a problem has to be found in a single basis state obtained via measurement (Lüders projection). This is one of the reasons why only a special class of problems can be solved by using quantum computations. This class is very restricted. Ohya et al. proposed to combine quantum algorithms with classical chaotic dynamics [25, 26]. This approach

¹² We remark that information features of entanglement can be modeled (mimic) by using coarse graining procedure for classical stochastic processes, even for classical Brownian motion [19]. In the latter paper the entanglement is exhibited at the level of observables corresponding to coarse graining.

provides a possibility to solve new NP-problems¹³, e.g., the SAT-problem, see Ohya and Volovich [8, 27, 28].

However, in general the usage of unitary transformations makes physical realization of quantum algorithms very difficult, since unitary dynamics is possible only in isolation from the environment. The latter is practically impossible. The entangled quantum n -qubit states are subjected to quick decoherence (destruction of superposition (2.57) and spontaneous reduction to one of the basis states). There is simply not enough time to perform quantum calculations. Therefore we do not think that bio-systems can perform real quantum calculations based on unitary transformations. Our conjecture is that bio-systems can process entangled states, but by using non-unitary dynamics of the GKSL-type (Sect. 2.4), see Sect. 8.4.2 for more detail.

2.6 From Open Quantum System Dynamics to State-Observable Adaptive Dynamics

In Sect. 2.4 we considered the dynamics of the state of a system, say S , interacting with an environment. To derive this dynamics, the Eq. (2.50), one proceeds in the following framework. The states of the system under consideration are represented in a Hilbert space \mathcal{H} and the states of an environment (bath) are represented in another Hilbert space \mathcal{K} . Since an environment is a huge system, its state space has very high dimension, in a mathematical model it can be considered as infinite-dimensional. The states of the compound system, the system S and the environment, are represented in the Hilbert space $\mathcal{H} \otimes \mathcal{K}$. In the theory of open quantum systems the evolution of a pure state of such a compound system is described (in the complete accordance with the postulates of quantum mechanics) by the Schrödinger equation. (In quantum bio-informatics we use its slight modification by using the dimensionless Hamiltonian and the time scaling constant γ , instead of the Planck constant, see Sect. 2.3. In this section we also proceed with such γ . In physics one can, as usual, proceed with Hamiltonians having the dimension of physical energy and the Planck constant instead of the scaling factor γ .) The Hamiltonian of a compound system is in general very complicated. Therefore in real calculations its various approximations are in use. Set $U_t = e^{-itH/\gamma}$, the evolutionary operator of S interacting with the environment. Then unitary dynamics of the pure state is given as

$$\Psi(t) = U_t \Psi_0, \quad (2.58)$$

¹³ In computability theory, a decision problem, which has two possible answers, “yes” or “no”, for an input of question is studied well, and its complexity is classified into several complexity classes. NP (Nondeterministic Polynomial time) problem is a decision problem, whose solution is not given in polynomial time on a non-deterministic Turing machine, and NP-complete problem is a NP problem reduced in polynomial time from any other NP problems. Whether NP-complete problem can be reduced to polynomial problem is one of the millennium prize problems, it has been discussed for thirty years.

where Ψ_0 is the initial state of S combined with the environment. For general states given by density operators, the dynamics is given by the von Neumann equation and, hence, it can be represented in the form, see (2.48),

$$R(t) = U_t R_0 U_t^*, \quad (2.59)$$

where R_0 is the initial state of the compound system. Typically one is interested only in the dynamics of S (and not interested in what happens with an environment around S). This dynamics is obtained by taking the trace of the density operator of the compound system with respect to all environmental degrees of freedom:

$$\rho(t) = \text{tr}_{\mathcal{H}} R(t) = \text{tr}_{\mathcal{H}} U_t R_0 U_t^* \quad (2.60)$$

where $\text{tr}_{\mathcal{H}}$ denotes the partial trace with respect to the subspace \mathcal{H} of the complete state space $\mathcal{H} \otimes \mathcal{H}$.

This dynamics is very complex because of the complexity of Hamiltonian of the total system and, in particular, the complexity of interaction between S and environment. Then under a number of assumptions it is possible to prove that the dynamics given by (2.60) can be reduced to the quantum master Eq. (2.50). One of the main assumptions is Markovness of dynamics. Another important assumption is that interaction of S with the environment cannot change the state of the environment (or if it changes then this change can be considered as negligible). (In Sect. 8.2.2 we shall discuss conditions of derivation of (2.50) from (2.60) in more detail and in the biological framework; in application to biological evolution.) We emphasize that in general dynamics (2.60) is non-Markovian. In the latter case the system S “remembers” the states of the compound system, S and the environment for sufficiently long period and it makes its state update on the basis of this information.

We think that to match completely with biology, the theory of open quantum systems has to be generalized. First of all we are critical to the approach assuming a possibility to construct interaction Hamiltonian. This is very difficult in physics. And in biology this problem is even more complicated. It is difficult or may be even impossible to describe formally the character of the interaction of a bio-system S with its environment. Such an environment has not only physical and chemical counterparts, but also a complex information component. And the latter plays the crucial role in many biological processes (e.g., signaling of surrounding cells is crucial in the process of cell differentiation, Chap. 8. Starting with the Hamiltonian, one postulates the Schrödinger dynamics for pure states and then the von Neumann dynamics is its trivial consequence. However, the assumption that such a complex system as an environment interacting with a system S can be in a pure state has no justification, especially in biology. It is more natural to assume the presence of mixture of states. Of course, one can easily avoid the last problem by postulating directly the von Neumann equation. However, even for a density operator, the usage of the *unitary dynamics* (2.59) is not justified.)

We think that, especially in quantum bio-informatics, the direct usage of the quantum Hamiltonian approach to describe the dynamics of a system in some

environment is not justified. Moreover, a bio-system is able to perform measurements, both on surrounding environment and on itself (self-measurements, Sect. 2.2.2). Such measurements are not described by unitary operators. Since environment information component is typically based on the activity of other biological systems, e.g., cells surrounding the cell under consideration, we cannot assume that by testing the environment a bio-system does not change the state of this environment. Therefore, instead of tracing of the unitary dynamics of the compound system, a system S interacting with an environment, see (2.60), and then approximating it in some way (Markovian or even non-Markovian), we develop a new theory—the theory of *state-observable adaptive dynamical systems*, Chap. 4.

By this theory *a system S (e.g., a cell) permanently updates its state ρ by performing measurements on the state of the surrounding environment (hence, the system state update depends on the state σ of the environment and measurement performed by S) and successive self-measurement*. This state update cannot be done without changing the state of the environment. Therefore formally we can describe both measurement on the environment and self-measurement of S as a measurement, say Q , on the compound system, S and environment, see Chap. 4 for the formal mathematical presentation based on the apparatus of lifting maps (quantum information theory). Since we use the Hilbert space representation, the adaptive dynamics model under consideration can be called quantum(-like).

References

1. Kolmogoroff, A.N.: Grundbegriffe der Wahrscheinlichkeitsrechnung. Springer, Berlin (1933); English translation: Kolmogorov, A.N.: Foundations of Theory of Probability. Chelsea Publishing Company, New York (1956)
2. Boole, G.: An Investigation of the Laws of Thought. Dover Edition, New York (1958)
3. Von Neuman, J.: Mathematical Foundations of Quantum Mechanics. Princeton University Press, Princeton (1955)
4. Khrennikov, A.: Contextual Approach to Quantum Formalism. Springer, Berlin-Heidelberg-New York (2009)
5. Dirac, P.A.M.: The Principles of Quantum Mechanics. Clarendon Press, Oxford (1995)
6. Khrennikov, A.: Detection model based on representation of quantum particles by classical random fields: Born's rule and beyond. Found. Phys. **39**, 997–1022 (2009)
7. Ingarden, R.S., Kossakowski, A., Ohya, M.: Information Dynamics and Open Systems: Classical and Quantum Approach. Kluwer, Dordrecht (1997)
8. Ohya, M., Volovich, I.V.: Mathematical Foundations of Quantum Information and Computation and Its Applications to Nano- and Bio-systems. Springer, Berlin-New York (2011)
9. Asano, M., Ohya, M., Khrennikov, A.: Quantum-like model for decision making process in two players game. Found. Phys. **41**(3), 538–548 (2010)
10. Basieva, I., Khrennikov, A., Ohya, M., Yamato, I.: Quantum-like interference effect in gene expression: glucose-lactose destructive interference. Syst. Synth. Biol. **5**, 59–68 (2011)
11. Pothos, E.M., Busemeyer, J.R.: A quantum probability explanation for violation of rational decision theory. Proc. R. Soc. B **276**(1165), 2171–2178 (2009)
12. Wang, Z., Busemeyer, J. R.: A quantum question order model supported by empirical tests of an a priori and precise prediction. Top. Cogn. Sci. (2013), to be published

13. Zorn, C.H., Smith, Ch.: Pseudo-classical nonseparability and mass politics in two-party systems. In: Song, D., Melucci, M., Frommholz, I., Zhang, P., Wang, L., Arafat, S. (eds.) *Quantum Interaction-3. Lectures Notes in Computer Science*, vol. 7052, pp. 83–94. Springer, Heidelberg (2011)
14. Khrennikova, P., Khrennikov, A., Haven, E.: The quantum-like description of the dynamics of party governance in the US political system (2012). [arXiv:1206.2888](https://arxiv.org/abs/1206.2888) [physics.gen-ph]
15. Einstein, A., Podolsky, B., Rosen, N.: Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47**(10), 777–780 (1933)
16. Fuchs, C.A.: Quantum mechanics as quantum information (and only a little more). In: Khrennikov, A. (ed.) *Quantum Theory: Reconsideration of Foundations. Series Mathematical Modeling 2*, pp. 463–543. Växjö University Press, Växjö (2002)
17. Caves, C.M., Fuchs, Ch.A., Schack, R.: Quantum probabilities as Bayesian probabilities. *Phys. Rev. A* **65**, 022305 (2002)
18. Fuchs, Ch.A., Schack, R.: A quantum-Bayesian route to quantum-state space. *Found. Phys.* **41**, 345–356 (2011)
19. Allahverdyan, A.E., Balian, R., Nieuwenhuizen, Th.M.: The quantum measurement process in an exactly solvable model. In: *Foundations of Probability and Physics-3. American Institute of Physics. Series Conference Proceedings*, vol. 750, pp. 16–24. Melville (2005)
20. Khrennikov, A.: Description of composite quantum systems by means of classical random fields. *Found. Phys.* **40**, 1051–1064 (2010)
21. Khrennikov, A., Ohya, M., Watanabe, N.: Classical signal model for quantum channels. *J. Russ. Laser Res.* **31**, 462–468 (2010)
22. Khrennikov, A.: *Ubiquitous Quantum Structure: From Psychology to Finance*. Springer, Heidelberg-Berlin-New York (2010)
23. Khrennikov, A.: Prequantum classical statistical field theory: Schrödinger dynamics of entangled systems as a classical stochastic process. *Found. Phys.* **41**, 317–329 (2011)
24. Khrennikov, A., Ohya, M., Watanabe, N.: Quantum probability from classical signal theory. *Int. J. Quantum Inf.* **9**, 281–292 (2011)
25. Ohya, M.: Complexities and their applications to characterization of chaos. *Int. J. Theor. Phys.* **37**, 495–505 (1998)
26. Ohya, M.: Some aspects of quantum information theory and their applications to irreversible processes. *Rep. Math. Phys.* **27**, 19–47 (1989)
27. Ohya, M., Volovich, I.V.: New quantum algorithm for studying NP-complete problems. *Rep. Math. Phys.* **52**, 25–33 (2003)
28. Ohya, M.: Adaptive dynamics and its applications to chaos and NPC problem. In: *QP-PQ: Quantum Probability and White Noise Analysis. Quantum Bio-Informatics*, vol. 21, pp. 181–216 (2007)

Quantum Adaptivity in Biology: From Genetics to
Cognition

Asano, M.; Khrennikov, A.; Ohya, M.; Tanaka, Y.; Yamato,
I.

2015, XX, 173 p. 27 illus., Hardcover

ISBN: 978-94-017-9818-1