Questionable and Unquestionable in Quantum Mechanics

László E. Szabó[(1)], Márton Gömöri[(1)(2)] and Zalán Gyenis[(3)]

[(1)] *Department of Logic, Institute of Philosophy, Eötvös University, Budapest*
[(2)] *Institute of Philosophy, Research Centre for the Humanities, Budapest*
[(3)] *Department of Logic, Jagiellonian University, Krakow*

Abstract

According to the Kolmogorovian Censorship Hypothesis, everything that quantum theory says about the world in the language of the quantum mechanical Hilbert space formalism is actually about relationships between ordinary relative frequencies expressible in operational terms using classical Kolmogorovian probability theory. In other words, a quantum theoretical description of a system should in principle be translatable into a purely operational–probabilistic description. However, our goal in this paper is different; we do not want to deal with the problem how to translate the known theory of quantum mechanics into operational terms, or to reconstruct the theory from postulates which can be interpreted in operational terms. Our aim is somewhat broader and points in the opposite direction. We start with a general scheme for the operational description of an arbitrary physical system. The description is based solely on the notion of observable events (measurement operations and measurement results) and on general, empirically established simple laws concerning their relative frequency. These laws are so simple and fundamental that they apply equally to any physical system—no plausibly conceivable physical system is known that would violate our basic assumptions. In the first part of the paper, we outline the basic elements of such an operational–probabilistic theory; such as the notion of state, the mathematical description of state space, and the basic notions of dynamics. All these notions are expressed in classical terms, within the framework of Kolmogorovian probability theory, and, since our goal is not necessarily to reproduce standard quantum mechanics, we try to avoid making assumptions that are restrictive and would not hold in the most general case. In the second part of the paper, we discuss how this operational–probabilistic description compares to the quantum mechanical description and to what extent the standard Hilbert space quantum mechanics can be regarded as a reformulation of the general operational–probabilistic theory.

1 Introduction

The main motivation of this paper is the so called Kolmogorovian Censorship Hypothesis, which was first formulated in (Szabó 1995). For long decades it had been the widely accepted view that quantum mechanics not only teaches us that the world is essentially of probabilistic nature, but also that the notion of probability itself, as it is used in quantum mechanics, is a new one, essentially different from the traditional Kolmogorovian notion of probability. The main difference, and the source of many other differences, is that the event algebra in quantum probability theory is not a Boolean algebra, but a more general algebraic structure, an orthomodular lattice, isomorphic with the subspace lattice of a Hilbert space. On the one hand, this insight has helped us to understand more deeply the curiosities of quantum theory, but on the other hand, it has raised a number of additional foundational problems. Nevertheless, it was agreed that classical Kolmogorovian probability theory was unable to accommodate quantum phenomena and that some form of quantum probability theory was inevitable.

The Kolmogorovian Censorship Hypothesis challenged this picture. The hypothesis is a generalization of the observation that whenever quantum probabilities appear in the description of real physical phenomena, they appear in combination with classical Kolmogorovian probabilities, such that, on the surface, the probabilities of all observable, or at least ontologically occurring, physical events are classical Kolmogorovian probabilities, which can be interpreted as ordinary relative frequencies.

Take the simplest example. Let A be a quantum observable, a self-adjoint operator with spectral decomposition $A = \sum_i \alpha_i P_i$ where α_i denotes an eigenvalue and P_i denotes the corresponding spectral projector. Let the system be in quantum state W. The most fundamental probabilistic claim of quantum theory is that the quantum probability of getting value α_i in an A-measurement is

$$q(\alpha_i) = \mathrm{tr}(WP_i)$$

However, comparing this prediction of the theory with the laboratory observations, what we actually verify is that

$$p([\alpha_i]) = \mathrm{tr}(WP_i)\, p(a) \tag{1}$$

where $p([\alpha_i])$ is the relative frequency of the outcome event $[\alpha_i]$, say, that the pointer's position is "α_i," and $p(a)$ is the relative frequency of the event a consisting in that an A-measurement is performed. That is, quantum probability $\mathrm{tr}(WP_i)$ never stands naked, "facing the tribunal of sense experience," but is surrounded by classical Kolmogorovian probabilities.

While this problem does not arise if we restrict ourselves to the spectral projectors of a single observable, it is easy to see that an arbitrary set of quantum probabilities, collectively, cannot constitute the relative frequencies of events. (Due to violation of Bell–type inequalities, Pitowsky 1989, Ch. 2. See Szabó 1998; 2001; 2008.) Don't be misled by the fact that in expressions like (1), the quantum probability can be interpreted as a classical conditional probability,

$$\operatorname{tr}(WP_i) = \frac{p([\alpha_i])}{p(a)} = p([\alpha_i] | a)$$

assuming that $p(a) \neq 0$ and $p([\alpha_i] \wedge a) = p([\alpha_i])$, and that the conditional probability can be interpreted as another probability function on the same event algebra. The different quantum probabilities can be interpreted as classical conditional probabilities, but they may belong to different conditioning events, and therefore do not together form a classical probability function over the event algebra in question; or over an arbitrary event algebra in general.

This means that there can be no events or any states of affairs in the ontology of the physical world whose relative frequencies, counted on the Humean mosaic, are equal to quantum probabilities. According to the Kolmogorovian Censorship Hypothesis, however, quantum probabilities can always be expressed in terms of the relative frequencies of such real events. Typically, as suggested by (1), in terms of the relative frequencies of the outcomes of measurements and the relative frequencies of executions of measurements.

It should be noted that the possibility of interpreting quantum probability as formulated in the Hypothesis can be formally proved under different special conditions (Bana and Durt 1997; Szabó 2001; Rédei 2010; Hofer-Szabó *et al.* 2013, Ch. 9).

Thus, according to the Kolmogorovian Censorship Hypothesis, everything that quantum theory says about the world in the language of the quantum mechanical Hilbert space formalism is actually about relationships between ordinary relative frequencies expressible in operational terms using classical Kolmogorovian probability theory. In other words, a quantum theoretical description of a system should in principle be translatable into a purely operational–probabilistic description.

To translate quantum mechanics into operational terms is not a new idea, of course (e.g. Jauch and Piron 1963; Ludwig 1970; Foulis and Randall 1974; Davies 1976; Busch, Grabowski, and Lahti 1995; Spekkens 2005; Barum *et al.* 2007; 2008; Hardy 2008; Aerts 2009; Abramsky and Heunen 2016; Schmid, Spekkens, and Wolfe 2018). Nevertheless, what such a translation of standard quantum mechanics looks like is not a self-evident question. However, we do not wish to discuss this question here. Our goal is different; we do not want to translate the known theory of quantum mechanics into operational terms, or to reconstruct the theory from postulates which can be interpreted in operational terms. The aim of this paper is somewhat broader and points in the opposite direction.

We start with a general scheme for the operational description of an arbitrary physical system. The description is based solely on the notion of observable events (measurement operations and measurement results) and on general, empirically established simple laws concerning their relative frequency. These laws are so simple and fundamental that they apply equally to any physical system, whether it is traditionally considered as classical or quantum, or even "more general than quantum" (Müller 2021, Sec. 2). In other words, our goal is not necessarily to reproduce standard quantum mechanics, and therefore we try to avoid making assumptions that are restrictive and would not hold in the most general case—no plausibly conceivable physical system is known that would violate our basic assumptions. In this sense, our operational–probabilistic model significantly differs from the other similar approaches in the above mentioned literature; including the most recent GPT approaches (Hardy 2008; Holevo 2011; Müller 2021). The main differences are briefly highlighted in several places throughout the paper; on two key aspects we reflect in more detail in Appendices 2 and 3.

Although our main—technically non-trivial—result concerns how this operational–probabilistic description compares to the quantum mechanical description and to what extent the standard Hilbert space quantum mechanics can be regarded as a reformulation of the general operational–probabilistic theory, we would like to draw the reader's attention to the first part of the paper in which we outline the basic elements of such an operational–probabilistic theory; such as the notion of state, the mathematical description of state space, and the basic notions of dynamics. All these notions are expressed in classical terms, within the framework of Kolmogorovian probability theory. It remains an intriguing question how to continue this project, how to apply it, for example, for the description of well-known quantum mechanical systems with the known symmetries and dynamics, etc.; to which the authors plan to return somewhere else.

The paper is structured as follows. We describe a typical empirical scenario in the following way: One can perform different measurement operations on a physical system, each of which may have different possible outcomes. The performance of a measuring operation is regarded as a physical event on par with the measurement outcomes. Empirical data are, exclusively, the observed relative frequencies of how many times different measurement operations are performed and how many times different outcome events occur, including the joint performances of two or more measurements and the conjunctions of their outcomes. In terms of the observed relative frequencies we stipulate two empirical conditions, **(E1)** and **(E2)**, which are simple, plausible, and empirically testable.

Of course, the observed relative frequencies essentially depend on the frequencies with which the measurement operations are performed; that is, on circumstances external to the physical system under consideration; for example, on the free choice of a human. Under a further empirically testable assumption about the observed frequencies, **(E3)**, we can isolate a notion which is independent of the relative frequencies of the measurement operations and can be identified with the system's own state; in the sense that it characterizes the system's probabilistic behavior against all possible measurement operations. The largest part of our further investigation is at the level of generality defined by assumptions **(E1)**–**(E3)**.

In Section 3, we derive important theorems, solely from conditions **(E1)**–**(E3)**, concerning the possible states of the system. In Section 4, we characterize the time evolution of these states, first in its most general form under conditions **(E1)**–**(E3)** alone, then on the basis of a further, empirically testable assumption **(E4)**. Section 5 considers various possible ontological pictures consistent with our probabilistic notion of state.

All these investigations are expressed in terms of relative frequencies, which by definition satisfy the Kolmogorovian axioms of classical probability theory. This means that any physical system—traditionally categorized as classical or quantum, or "more general than quantum"—that can be described in operational terms can be described within classical Kolmogorovian probability theory; including the system's state, time evolution or ontology. In the second part of the paper, at the same time, we will show that anything that can be described in these operational terms can, if we wish, be represented in the Hilbert space quantum mechanical formalism. It will be proved that there always exists:

- a suitable Hilbert space, such that
- the outcomes of each measurement can be represented by a system of pairwise orthogonal closed subspaces, spanning the whole Hilbert space,
- the states of the system can be represented by pure state operators with suitable state vectors, and
- the probabilities of the measurement outcomes can be reproduced by the usual trace formula of quantum mechanics.

Moreover, if appropriate, one can label the possible outcomes of a measurement with numbers, and talk about them as the measured values of a physical quantity. Each such quantity

- can be associated with a suitable self-adjoint operator, such that
- the expectation value of the quantity, in all states of the system, can be reproduced by the usual trace formula applied to the associated self-adjoint operator,
- the possible measurement results are exactly the eigenvalues of the operator, and
- the corresponding outcome events are represented by the eigenspaces pertaining to the eigenvalues respectively, according to the spectral decomposition of the operator in question.

This suggests that the basic postulates of quantum theory are in fact analytic statements: they do not tell us anything about a physical system beyond the fact that the system can be described in operational terms. This is almost true. Nevertheless, it must be mentioned that the quantum-mechanics-like representation we will obtain is not completely identical with standard quantum mechanics. The interesting fact is that most of the deviations from the quantum mechanical folklore, discussed in Section 8, are related to exactly those issues in the foundations of quantum mechanics that have been hotly debated for long decades.

2 The General Operational Schema

Consider a general experimental scenario: we can perform different measurement operations denoted by $a_1, a_2, \ldots a_r, \ldots a_m$ on a physical system. We shall use the same notation a_r for the physical event that the measurement operation a_r happened. Each measurement a_r may have different outcomes denoted by $X_1^r, X_2^r, \ldots X_{n_r}^r$. Let $M = \sum_{r=1}^m n_r$, and let I^M denote the following set of indices:

$$I^M \quad = \quad \left\{ {}_i^r \mid 1 \leq r \leq m, 1 \leq i \leq n_r \right\} \tag{2}$$

Sometimes we perform two or more measurement operations simultaneously—that is, in the same run of the experiment. So we also consider the double, triple, and higher conjunctions of measurement operations and the possible outcome events. In general, we consider the free Boolean algebra $\mathcal{A}$ generated by the set of all measurement operation and measurement outcome events

$$G \quad = \quad \{a_r\}_{r=1,2,\ldots m} \cup \{X_i^r\}_{{}_i^r \in I^M} \tag{3}$$

with the usual Boolean operations, denoted by $\wedge$, $\vee$ and $\neg$. Introduce the following concise notation: let S_{max}^M denote the set of the indices of all double, triple, and higher conjunctions of the outcome events in G. That is, for example, ${}_{i_1 i_2 \ldots i_L}^{r_1 r_2 \ldots r_L} \in S_{max}^M$ will stand for the conjunction $X_{i_1}^{r_1} \wedge X_{i_2}^{r_2} \ldots \wedge X_{i_L}^{r_L}$, etc.

The event algebra $\mathcal{A}$ has 2^{M+m} atoms, each having the form of

$$\Delta_{\vec{\varepsilon},\vec{\eta}} \quad = \quad \left(\bigwedge_{{}_i^r \in I^M} [X_i^r]^{\varepsilon_i^r} \right) \wedge \left(\bigwedge_{s=1}^m [a_s]^{\eta_s} \right) \tag{4}$$

where $\vec{\varepsilon} = (\varepsilon_i^r) \in \{0,1\}^M$, $\vec{\eta} = (\eta_s) \in \{0,1\}^m$, and

$$[X_i^r]^{\varepsilon_i^r} \quad = \quad \begin{cases} X_i^r & \text{if } \varepsilon_i^r = 1 \\ \neg X_i^r & \text{if } \varepsilon_i^r = 0 \end{cases}$$

$$[a_s]^{\eta_s} \quad = \quad \begin{cases} a_s & \text{if } \eta_s = 1 \\ \neg a_s & \text{if } \eta_s = 0 \end{cases}$$

And, of course, all events in algebra $\mathcal{A}$ can be uniquely expressed as a disjunction of atoms.

Assume that we can repeat the same experimental situation as many times as needed; that is, we can prepare the same (or identical) physical system in the same way and we can repeat the same measuring operations with the same (or identical) measuring devices, etc. In every run of the experiment we observe which measurement operations are performed and which outcome events occur, including the joint performances of two or more measurements and the conjunctions of their outcomes. In this way, we observe the *relative frequencies* of all elements of the event algebra $\mathcal{A}$. Let p denote this relative frequency function on $\mathcal{A}$. Obviously, $(\mathcal{A}, p)$ constitutes a classical probability model satisfying the

Kolmogorovian axioms. Since the relative frequencies on the whole event algebra are uniquely determined by the relative frequencies of the atoms, p can be uniquely given by

$$p(\Delta_{\vec{\varepsilon},\vec{\eta}}) \qquad \vec{\varepsilon} \in \{0,1\}^M \; ; \; \vec{\eta} \in \{0,1\}^m \tag{5}$$

The observed relative frequencies on $\mathcal{A}$ are considered *the* empirical data, exclusively.

We do not make *a priori* assumptions about these relative frequencies. Any truth about them will be regarded as empirical fact observed in the experiment. For example, we do not assume that the stipulated set of measurements $a_1, a_2, \ldots a_r, \ldots a_m$, or a subset of them, is sufficient to "fully characterize the system's state" (in the GPT terminology: "fiducial" measurements, Müller 2021, p. 23). The reason is we have no operationally meaningful notion of "state" prior to setting up a collection of measurements (such a notion will be defined in the next section). If there is any redundancy in $a_1, a_2, \ldots a_r, \ldots a_m$, this will be reflected in the observed relative frequencies on $\mathcal{A}$.

Another example is the fact that two or more measurements $a_{r_1}, a_{r_2}, \ldots a_{r_L}$ cannot be performed simultaneously; which reveals in the observed fact that $p\left(a_{r_1} \wedge a_{r_2} \ldots \wedge a_{r_L}\right)$ *always* equals 0. Though, this "always" needs some further explanation. For, it is obviously true that the frequencies $p(a_r)$ sensitively depend on the will of the experimenter. Therefore, it can be the case that $p\left(a_{r_1} \wedge a_{r_2} \ldots \wedge a_{r_L}\right) = 0$ simply because the experimenter never chooses to perform the measurements $a_{r_1}, a_{r_2}, \ldots a_{r_L}$ simultaneously. At least at first sight this seems to significantly differ from the situation when a certain combination of experiments are never performed due to objective reasons; because the simultaneous performance of the measurement operations is—as we usually express—impossible. Without entering into the metaphysical disputes about possibility–impossibility, we only say that the impossibility of a combination of measurements is a contingent fact of the world; the measuring devices and the measuring operations are such that the joint measurement $a_{r_1} \wedge a_{r_2} \ldots \wedge a_{r_L}$ never occurs. Let us denote by $\mathfrak{I} \subset \mathcal{P}\left(\{1,2,\ldots m\}\right)$ (where $\mathcal{P}(A)$ is the power set of set A) the set of indices of such "impossible" conjunctions. That is, for all $2 \leq L \leq m$,

$$p\left(a_{r_1} \wedge a_{r_2} \ldots \wedge a_{r_L}\right) \quad = \quad 0 \qquad \text{if } \{r_1, r_2, \ldots r_L\} \in \mathfrak{I} \tag{6}$$

In contrast, let $\mathfrak{P} \subset \mathcal{P}\left(\{1,2,\ldots m\}\right)$ denote the set of indices of the "possible" conjunctions:

$$\mathfrak{P} \quad = \quad \left\{ \{r_1, r_2, \ldots r_L\} \in \mathcal{P}\left(\{1,2,\ldots m\}\right) \middle| \, 2 \leq L \leq m; \{r_1, r_2, \ldots r_L\} \notin \mathfrak{I} \right\}$$

(E1) We assume, as empirically observed fact, that every conjunction of measurements that is possible does occur with some non-zero frequency:

$$p\left(a_{r_1} \wedge a_{r_2} \ldots \wedge a_{r_L}\right) \quad > \quad 0 \qquad \text{if } \{r_1, r_2, \ldots r_L\} \in \mathfrak{P} \tag{7}$$

We also assume that for all $1 \leq r \leq m$,

$$p\left(a_r\right) \quad > \quad 0 \tag{8}$$

Similarly to (2), we introduce the following sets of indices:

$$\begin{aligned} S &= \left\{ {}^{r_1 r_2 \ldots r_L}_{i_1 i_2 \ldots i_L} \in S^M_{max} \mid \{r_1, r_2, \ldots r_L\} \in \mathfrak{P} \right\} \\ S_{\mathfrak{I}} &= \left\{ {}^{r_1 r_2 \ldots r_L}_{i_1 i_2 \ldots i_L} \in S^M_{max} \mid \{r_1, r_2, \ldots r_L\} \in \mathfrak{I} \right\} \end{aligned}$$

(E2) The following assumptions are also regarded as empirically observed regularities: for all ${}^{r}_{i}, {}^{r'}_{i'} \in I^M$ and $\{r_1, r_2, \ldots r_L\} \in \mathfrak{P}$,

$$p\left(a_r \wedge X^r_i\right) = p\left(X^r_i\right) \tag{9}$$

$$\text{if } r = r' \text{ and } i \neq i' \text{ then } p\left(X^r_i \wedge X^{r'}_{i'}\right) = 0 \tag{10}$$

$$\sum_{\substack{k \\ \left({}^{r}_{k} \in I^M\right)}} p\left(X^r_k | a_r\right) = 1 \tag{11}$$

$$\sum_{\substack{k_1 \ldots k_L \\ \left({}^{r_1 \ldots r_L}_{k_1 \ldots k_L} \in S\right)}} p\left(X^{r_1}_{k_1} \wedge \ldots \wedge X^{r_L}_{k_L} | a_{r_1} \wedge \ldots \wedge a_{r_L}\right) = 1 \tag{12}$$

where $p\left(\,|\,\right)$ denotes the usual conditional relative frequency defined by the Bayes rule—$p\left(a_r\right) \neq 0$ and $p\left(a_{r_1} \wedge a_{r_2} \ldots \wedge a_{r_L}\right) \neq 0$, due to (7)–(8). That is to say, an outcome event does not occur without the performance of the corresponding measurement operation; it is never the case that two different outcomes of the same measurement occur simultaneously; whenever a measurement operation is performed, one of the possible outcomes occurs; whenever a conjunction of measurement operations is performed, one of the possible outcome combinations occurs.

In the picture we suggest, an outcome of a measurement is, primarily, a physical event, an occurrence of a certain state of affairs at the end of the measuring process; rather than obtaining a numeric value of a quantity. To give an example, the state of affairs when the rotated coil of a voltmeter takes a new position of equilibrium with the distorted spring is ontologically prior to the number on the scale to which its pointer points at that moment. Nevertheless, in some cases the measurement outcomes are labeled by real numbers that are interpreted as the "measured value" of a real-valued physical quantity:

$$\alpha_r : X^r_i \mapsto \alpha^r_i \in \mathbb{R} \tag{13}$$

In this case, at least formally, it may make sense to talk about conditional expectation value, that is the average of the measured values, given that the measurement is performed:

$$\langle \alpha_r \rangle = \sum_{i=1}^{n_r} \alpha^r_i p\left(X^r_i | a_r\right)$$

About all labelings α_r we will assume that $\alpha^r_i \neq \alpha^r_j$ for $i \neq j$.

3 The State of the System

Of course, the relative frequency p in $(\mathcal{A}, p)$ depends not only on the behavior of the physical system after a certain physical preparation but also on the autonomous decisions of the experimenter to perform this or that measurement operation. One can hope a scientific description of the system only if the two things can be separated. Whether this is possible is a contingent fact of the empirically observed reality, reflected in the observed relative frequencies.

Let $|S|$ denote the number of elements of S. Consider the following vector:

$$\vec{Z} \quad = \quad \left(Z_i^r, Z_{i_1\ldots i_L}^{r_1\ldots r_L}\right) \in \mathbb{R}^{M+|S|} \tag{14}$$

where

$$Z_i^r = p\left(X_i^r | a_r\right) \qquad {}_i^r \in I^M \tag{15}$$

and

$$Z_{i_1\ldots i_L}^{r_1\ldots r_L} = p\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L} | a_{r_1} \wedge \ldots \wedge a_{r_L}\right) \quad {}_{i_1\ldots i_L}^{r_1\ldots r_L} \in S \tag{16}$$

In general, even if the physical preparation of the system is identical in every run of the experiment, the conditional relative frequencies on the right hand sides of (15)–(16), hence the values of Z_i^r and $Z_{i_1\ldots i_L}^{r_1\ldots r_L}$, may vary if the actual frequencies $\{p(a_r)\}_{1\leq r\leq m}$ and $\{p(a_{r_1} \wedge \ldots \wedge a_{r_L})\}_{\{r_1,\ldots r_L\}\in\mathfrak{P}}$ vary, for example, upon the experimenter's decisions.

However, we make the following stipulation as observed empirical fact:

(E3) For all physical preparations, keeping the preparation fixed, $\vec{Z}$ is independent of the actual nonzero values of $\{p(a_r)\}_{1\leq r\leq m}$ and $\{p(a_{r_1} \wedge \ldots \wedge a_{r_L})\}_{\{r_1,\ldots r_L\}\in\mathfrak{P}}$.

In other words, what **(E3)** says is that for all fixed physical preparations, the observed relative frequencies are such that

$$p\left(X_i^r\right) = Z_i^r p\left(a_r\right) \qquad {}_i^r \in I^M \tag{17}$$

$$p\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L}\right) = \begin{cases} Z_{i_1\ldots i_L}^{r_1\ldots r_L} p\left(a_{r_1} \wedge \ldots \wedge a_{r_L}\right) & {}_{i_1\ldots i_L}^{r_1\ldots r_L} \in S \\ 0 & {}_{i_1\ldots i_L}^{r_1\ldots r_L} \in S_{\mathfrak{I}} \end{cases} \tag{18}$$

with *one and the same* $\vec{Z} = \left(Z_i^r, Z_{i_1\ldots i_L}^{r_1\ldots r_L}\right)$. $\vec{Z}$ is therefore determined only by the physical preparation. Notice that **(E3)** does not exclude that different physical preparations lead to the same $\vec{Z}$.

$\vec{Z}$ can be regarded as a characterization of the system's *state* right after the given physical preparation, in the sense that it characterizes the system's probabilistic behavior against all possible measurement operations. This characterization is complete in the following sense:

Theorem 1. *State $\vec{Z}$ together with arbitrary relative frequencies of measurements, $\{p(a_r)\}_{1\leq r\leq m}$ and $\{p(a_{r_1} \wedge \ldots \wedge a_{r_L})\}_{\{r_1,\ldots r_L\}\in\mathfrak{P}}$, uniquely determine the relative frequency function p on the whole event algebra $\mathcal{A}$.*

Proof. Using the same notations we introduced in (4), each atom has the form of

$$\Delta_{\vec{\varepsilon},\vec{\eta}} = \underbrace{\left(\bigwedge_{\substack{r\\i}\in I^M} [X_i^r]^{\varepsilon_i^r}\right)}_{\Gamma_{\vec{\varepsilon}}} \wedge \left(\bigwedge_{s=1}^{m} [a_s]^{\eta_s}\right) \tag{19}$$

Notice that the part $\Gamma_{\vec{\varepsilon}} = \bigwedge_{\substack{r\\i}\in I^M} [X_i^r]^{\varepsilon_i^r}$ in (19) uniquely determines the whole $\Delta_{\vec{\varepsilon},\vec{\eta}}$, whenever $p\left(\Delta_{\vec{\varepsilon},\vec{\eta}}\right) \neq 0$. Namely, due to (9) and (11),

$$p\left(\Delta_{\vec{\varepsilon},\vec{\eta}}\right) \neq 0 \text{ implies that for all } 1 \leq r \leq m,\ \sum_{i=1}^{n_r} \varepsilon_i^r = 0 \text{ iff } \eta_r = 0 \tag{20}$$

In other words, for each $\vec{\varepsilon} \in \{0,1\}^M$ there is exactly one $\vec{\eta} \in \{0,1\}^m$ for which (9) and (11) do not imply that $p\left(\Delta_{\vec{\varepsilon},\vec{\eta}}\right) = 0$. Let us denote it by $\vec{\eta}\left(\vec{\varepsilon}\right)$; and, for the sake of brevity, introduce the following notation: $\delta_{\vec{\varepsilon}} = p\left(\Delta_{\vec{\varepsilon},\vec{\eta}(\vec{\varepsilon})}\right)$. (It is not necessarily the case that $\delta_{\vec{\varepsilon}} \neq 0$. For example, the empirical fact (10) will be accounted for in terms of the values on the right hand side of (24) below.)

It must be also noticed that $\{\Gamma_{\vec{\varepsilon}}\}_{\vec{\varepsilon}\in\{0,1\}^M}$ constitute the atoms of the free Boolean algebra $\mathcal{A}^M$ generated by the set $\{X_i^r\}_{\substack{r\\i}\in I^M}$. Events X_i^r and $X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L}$ on the right hand sides of (15)–(16) are elements of $\mathcal{A}^M$, and have therefore a unique decomposition into disjunction of atoms of $\mathcal{A}^M$. Accordingly, taking into account (20), we have

$$\sum_{\vec{\varepsilon}\in\{0,1\}^M} \delta_{\vec{\varepsilon}} = 1 \tag{21}$$

$$\sum_{\vec{\varepsilon}\in\{0,1\}^M} {}^{r}_{i}R_{\vec{\varepsilon}}\, \delta_{\vec{\varepsilon}} = p\left(X_i^r\right) = Z_i^r p\left(a_r\right) \qquad {}^{r}_{i} \in I^M \tag{22}$$

$$\begin{aligned}\sum_{\vec{\varepsilon}\in\{0,1\}^M} {}^{r_1\ldots r_L}_{i_1\ldots i_L}R_{\vec{\varepsilon}}\, \delta_{\vec{\varepsilon}} &= p\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L}\right) \\ &= Z_{i_1\ldots i_L}^{r_1\ldots r_L} p\left(a_{r_1} \wedge \ldots \wedge a_{r_L}\right) \qquad {}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S\end{aligned} \tag{23}$$

$$\sum_{\vec{\varepsilon}\in\{0,1\}^M} {}^{r_1\ldots r_L}_{i_1\ldots i_L}R_{\vec{\varepsilon}}\, \delta_{\vec{\varepsilon}} = p\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L}\right) = 0 \qquad {}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S_{\mathfrak{I}} \tag{24}$$

with

$$\begin{aligned}{}^{r}_{i}R_{\vec{\varepsilon}} &= \begin{cases} 1 & \text{if } \Gamma_{\vec{\varepsilon}} \subseteq X_i^r \\ 0 & \text{if } \Gamma_{\vec{\varepsilon}} \nsubseteq X_i^r \end{cases} \\ {}^{r_1\ldots r_L}_{i_1\ldots i_L}R_{\vec{\varepsilon}} &= \begin{cases} 1 & \text{if } \Gamma_{\vec{\varepsilon}} \subseteq X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L} \\ 0 & \text{if } \Gamma_{\vec{\varepsilon}} \nsubseteq X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L} \end{cases}\end{aligned}$$

where $\subseteq$ is meant in the sense of the partial ordering in $\mathcal{A}^M$.

Now, (21)–(24) constitute a system of $1 + M + \left|S_{max}^M\right| = 2^M$ linear equations with 2^M unknowns $\delta_{\vec{\varepsilon}}, \vec{\varepsilon} \in \{0,1\}^M$. The equations are linearly independent due to the uniqueness of decomposition into disjunction of atoms of $\mathcal{A}^M$, and due to the fact that there are only conjunctions on the right hand side. (A similar equation for, say, $X_{i_1}^{r_1} \vee X_{i_2}^{r_2}$ could be expressed as the sum of equations for $X_{i_1}^{r_1}$ and $X_{i_2}^{r_2}$ minus the one for $X_{i_1}^{r_1} \wedge X_{i_2}^{r_2}$.) Therefore, the system has a unique solution for all $\delta_{\vec{\varepsilon}}$, that is, for the relative frequencies of $\left\{\Delta_{\vec{\varepsilon},\vec{\eta}(\vec{\varepsilon})}\right\}_{\vec{\varepsilon}\in\{0,1\}^M}$. The rest of the atoms of $\mathcal{A}$ have zero relative frequency. □

Thus the notion of state we introduced aligns with the traditional notion of state in a probabilistic and operational context. In fact, it corresponds precisely to Lucien Hardy's formulation, which has been widely used in recent GPT-like approaches:

> The state associated with a particular preparation is defined to be (that thing represented by) any mathematical object that can be used to determine the probability associated with the outcomes of any measurement that may be performed on a system prepared by the given preparation. (2008, p. 2)

To avoid misunderstandings, however, it is worthwhile pointing out two important differences. On the one hand, in our definition the state determines not only the probability associated with the outcomes of any measurement but also the probability associated with the outcomes of any *conjunction* of measurements that can be jointly performed on the system. By contrast, in the GPT approach measurement conjunctions are only introduced in connection with "composite systems," in a way that one measurement is performed on one "subsystem" and simultaneously another measurement on a different "subsystem" (Müller 2021, pp. 24-28). In this paper we do not want to talk about "composite systems" and "subsystems," and just note that any operationally meaningful conception of those will be based on the notion of an overall state which comprises the statistics of all measurements performable on the total system, including all possible measurement *conjunctions*.

On the other hand, it must be clear that the state, in itself, does not determine the probabilities of the measurement outcome events; only the state of the system $\vec{Z}$ and the relative frequencies of the measurements $\{p\,(a_r)\}_{1\leq r\leq m}$ and $\{p\,(a_{r_1}\wedge\ldots\wedge a_{r_L})\}_{\{r_1,\ldots r_L\}\in\mathfrak{P}}$ together. And the fact that the frequencies of the measurements in (17)–(18) can be arbitrary does not imply that the components of $\vec{Z}$

$$\left\{Z_i^r, Z_{i_1\ldots i_L}^{r_1\ldots r_L}\right\}_{\substack{r\\i}\in I^M;\ \substack{r_1\ldots r_L\\i_1\ldots i_L}\in S}$$

constitute relative frequencies of the corresponding outcome events

$$\left\{X_i^r, X_{i_1}^{r_1}\wedge\ldots\wedge X_{i_L}^{r_L}\right\}_{\substack{r\\i}\in I^M;\ \substack{r_1\ldots r_L\\i_1\ldots i_L}\in S}$$

(or events whatsoever), as will be shown in Section 5.

It is essential in our present analysis that the measurement operations are treated on par with the outcome events; they belong to the ontology. However, as it is clearly seen from (21)–(24), the notion of $\vec{Z}$ detaches the "system's contribution" to the totality of statistical facts observed in the measurements from the "experimenter's contribution".

Still, the state of the system depends not only on the features intrinsic to the system in itself, but also on the content of $\mathfrak{I}$, i.e., which combinations of measuring operations cannot be performed simultaneously. This means that the measuring devices and measuring operations, by means of which we establish the empirically meaningful semantics of our physical description of the system, play a *constitutive* role in the notion of state *attributed to the system.* This kind of constitutive role of the semantic conventions is however completely natural in all empirically meaningful physical theories (Szabó 2020).

The following lemma will be important for our further investigations:

Lemma 2. *For all states,*

$$Z_{i_1 \ldots i_L}^{r_1 \ldots r_L} \quad \leq \quad \min \left\{ Z_{i_{\gamma_1} \ldots i_{\gamma_{L-1}}}^{r_{\gamma_1} \ldots r_{\gamma_{L-1}}} \right\}_{\{\gamma_1, \ldots \gamma_{L-1}\} \subset \{1, \ldots L\}} \tag{25}$$

where ${}_{i_1 \ldots i_L}^{r_1 \ldots r_L} \in S$.

Proof. It is known that similar inequality holds for arbitrary relative frequencies. Therefore,

$$p\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L}\right) \leq \min \left\{ p\left(X_{i_{\gamma_1}}^{r_{\gamma_1}} \wedge \ldots \wedge X_{i_{\gamma_{L-1}}}^{r_{\gamma_{L-1}}}\right) \right\}_{\{\gamma_1, \ldots \gamma_{L-1}\} \subset \{1, \ldots L\}} \tag{26}$$

for all ${}_{i_1 \ldots i_L}^{r_1 \ldots r_L} \in S_{max}^M$, and

$$p\left(a_{r_1} \wedge \ldots \wedge a_{r_L}\right) \leq \min \left\{ p\left(a_{r_{\gamma_1}} \wedge \ldots \wedge a_{r_{\gamma_{L-1}}}\right) \right\}_{\{\gamma_1, \ldots \gamma_{L-1}\} \subset \{1, \ldots L\}} \tag{27}$$

for all $2 \leq L \leq m$, $1 \leq r_1, \ldots r_L \leq m$. It follows from the definition of state that

$$p\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L}\right) \quad = \quad Z_{i_1 \ldots i_L}^{r_1 \ldots r_L} p\left(a_{r_1} \wedge \ldots \wedge a_{r_L}\right) \tag{28}$$

$$p\left(X_{i_{\gamma_1}}^{r_{\gamma_1}} \wedge \ldots \wedge X_{i_{\gamma_{L-1}}}^{r_{\gamma_{L-1}}}\right) \quad = \quad Z_{i_{\gamma_1} \ldots i_{\gamma_{L-1}}}^{r_{\gamma_1} \ldots r_{\gamma_{L-1}}} p\left(a_{r_{\gamma_1}} \wedge \ldots \wedge a_{r_{\gamma_{L-1}}}\right) \tag{29}$$

for all ${}_{i_1 \ldots i_L}^{r_1 \ldots r_L} \in S$ and $\{\gamma_1, \ldots \gamma_{L-1}\} \subset \{1, \ldots L\}$. Consequently, from (26) we have

$$\frac{Z_{i_1 \ldots i_L}^{r_1 \ldots r_L}}{Z_{i_{\gamma_1} \ldots i_{\gamma_{L-1}}}^{r_{\gamma_1} \ldots r_{\gamma_{L-1}}}} \quad \leq \quad \frac{p\left(a_{r_{\gamma_1}} \wedge \ldots \wedge a_{r_{\gamma_{L-1}}}\right)}{p\left(a_{r_1} \wedge \ldots \wedge a_{r_L}\right)} \tag{30}$$

Since, according to the definition of state, (28)–(29) hold for all possible relative frequencies $\{p(a_r)\}_{1 \leq r \leq m}$ and $\{p(a_{r_1} \wedge \ldots \wedge a_{r_L})\}_{\{r_1, \ldots r_L\} \in \mathfrak{P}}$, inequality (30) must hold for the minimum value of the right hand side, which is equal to 1, due to (27). And this is the case for all ${}_{i_1 \ldots i_L}^{r_1 \ldots r_L} \in S$ and $\{\gamma_1, \ldots \gamma_{L-1}\} \subset \{1, \ldots L\}$. $\square$

It is of course an empirical question what states a system has after different physical preparations. In what follows, we will answer the question: what can we say about the "space" of theoretically possible states of a system? Where by "theoretically possible states" we mean all vectors constructed by means of definition (14)–(16) from arbitrary relative frequencies satisfying (7)–(12) and (17)–(18). Here we should note that the general probabilistic description includes the possibility—again, as an eventual empirical fact observed from the frequencies (5)—that the system is deterministic, meaning that $\vec{Z} \in \{0,1\}^{M+|S|}$, or at least it behaves deterministically in some states.

We will show that the possible state vectors constitute a closed convex polytope in $\mathbb{R}^{M+|S|}$, which we will denote by $\varphi(M,S)$. First we will prove an important lemma.

Lemma 3. *If $\vec{Z}_1$ and $\vec{Z}_2$ are possible states then their convex linear combination $\vec{Z}_3 = \lambda_1 \vec{Z}_1 + \lambda_2 \vec{Z}_2$ ($\lambda_1, \lambda_2 \geq 0 \;\; \lambda_1 + \lambda_2 = 1$) also constitutes a possible state.*

Proof. According to the definition of state, the observed relative frequencies of the measurement outcomes in the two states are

$$p_1\left(X_i^r\right) = Z_1{}_i^r p_1\left(a_r\right) \tag{31}$$

$$p_1\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L}\right) = Z_1{}_{i_1 \ldots i_L}^{r_1 \ldots r_L} p_1\left(a_{r_1} \wedge \ldots \wedge a_{r_L}\right) \tag{32}$$

and

$$p_2\left(X_i^r\right) = Z_2{}_i^r p_2\left(a_r\right) \tag{33}$$

$$p_2\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L}\right) = Z_2{}_{i_1 \ldots i_L}^{r_1 \ldots r_L} p_2\left(a_{r_1} \wedge \ldots \wedge a_{r_L}\right) \tag{34}$$

for all ${}_i^r \in I^M$ and ${}_{i_1 \ldots i_L}^{r_1 \ldots r_L} \in S$. Due to **(E3)**, $p_1(a_r)$ and $p_1(a_{r_1} \wedge \ldots \wedge a_{r_L})$ as well as $p_2(a_r)$ and $p_2(a_{r_1} \wedge \ldots \wedge a_{r_L})$ can be arbitrary relative frequencies satisfying (6)–(8). Therefore, without loss of generality, we can take the case of

$$p_1(a_r) = p_2(a_r) = p_0(a_r)$$
$$p_1(a_{r_1} \wedge \ldots \wedge a_{r_L}) = p_2(a_{r_1} \wedge \ldots \wedge a_{r_L}) = p_0(a_{r_1} \wedge \ldots \wedge a_{r_L})$$

Now, consider the convex linear combination $p_3 = \lambda_1 p_1 + \lambda_2 p_2$. Obviously, p_3 satisfies (7)–(11), and

$$\begin{aligned} p_3(a_r) &= p_0(a_r) \\ p_3(a_{r_1} \wedge \ldots \wedge a_{r_L}) &= p_0(a_{r_1} \wedge \ldots \wedge a_{r_L}) \end{aligned}$$

Accordingly, we have

$$\begin{aligned} p_3\left(X_i^r\right) &= \lambda_1 p_1\left(X_i^r\right) + \lambda_2 p_2\left(X_i^r\right) = \left(\lambda_1 Z_1{}_i^r + \lambda_2 Z_2{}_i^r\right) p_3(a_r) \\ p_3\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L}\right) &= \lambda_1 p_1\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L}\right) + \lambda_2 p_2\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L}\right) \\ &= \left(\lambda_1 Z_1{}_{i_1 \ldots i_L}^{r_1 \ldots r_L} + \lambda_2 Z_2{}_{i_1 \ldots i_L}^{r_1 \ldots r_L}\right) p_3\left(a_{r_1} \wedge \ldots \wedge a_{r_L}\right) \end{aligned}$$

This means that $\vec{Z}_3 = \lambda_1 \vec{Z}_1 + \lambda_2 \vec{Z}_2$ satisfies condition (17)–(18), as $p_3(a_r)$ and $p_3(a_{r_1} \wedge \ldots \wedge a_{r_L})$ can be arbitrary frequencies satisfying (6)–(8). That is, $\vec{Z}_3$ complies with the definition of state, meaning that $\vec{Z}_3$ is a possible state of the system. □

3.1 The State Space – Polytope View

Now we turn to the question of the "space" of possible states. Let $\vec{e}^{\,r}_{\,i}, \vec{e}^{\,r_1\ldots r_L}_{\,i_1\ldots i_L} \in \mathbb{R}^{M+|S|}$ denote the ${}^{r}_{i}$-th and ${}^{r_1\ldots r_L}_{i_1\ldots i_L}$-th coordinate base vector of $\mathbb{R}^{M+|S|}$, where ${}^{r}_{i} \in I^M$ and ${}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S$, and let $\vec{f} = \left(f^r_i, f^{r_1\ldots r_L}_{i_1\ldots i_L}\right) \in \mathbb{R}^{M+|S|}$ denote an arbitrary vector.

The empirical facts **(E1)**–**(E3)**, partly through Lemmas 2–3, imply that the possible state vectors constitute a closed convex polytope $\varphi(M,S) \subset \mathbb{R}^{M+|S|}$ defined by the following system of linear inequalities:

$$f^r_i \geq 0 \tag{35}$$

$$f^r_i \leq 1 \tag{36}$$

$$f^{r_1\ldots r_L}_{i_1\ldots i_L} \geq 0 \tag{37}$$

$$f^{r_1\ldots r_L}_{i_1\ldots i_L} - f^{r_{\gamma_1}\ldots r_{\gamma_{L-1}}}_{i_{\gamma_1}\ldots i_{\gamma_{L-1}}} \leq 0 \qquad \{\gamma_1,\ldots \gamma_{L-1}\} \subset \{1,\ldots L\} \tag{38}$$

$$\sum_{\substack{k \\ \left({}^{r}_{k} \in I^M\right)}} f^r_k = 1 \tag{39}$$

$$\sum_{\substack{k_1, k_2 \ldots k_L \\ \left({}^{r_1\ldots r_L}_{k_1\ldots k_L} \in S\right)}} f^{r_1\ldots r_L}_{k_1\ldots k_L} = 1 \tag{40}$$

$$f^{r'_1\ldots r'_L}_{i'_1\ldots i'_L} = 0 \qquad {}^{r'_1\ldots r'_L}_{i'_1\ldots i'_L} \in S_0 \tag{41}$$

for all ${}^{r}_{i} \in I^M$, ${}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S$, and

$$S_0 = \left\{ {}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S \middle| r_{\gamma_1} = r_{\gamma_2}, i_{\gamma_1} \neq i_{\gamma_2}, \{\gamma_1, \gamma_2\} \subset \{1,\ldots L\} \right\}$$

Denote by $l(M,S) \subset \mathbb{R}^{M+|S|}$ the closed convex polytope defined by the first group of inequalities (35)–(38). As is well known (Pitowsky 1989, pp. 51 and 65), the vertices of $l(M,S)$ are all the vectors $\vec{v} \in \mathbb{R}^{M+|S|}$ such that

(a) $v^r_i, v^{r_1\ldots r_L}_{i_1\ldots i_L} \in \{0,1\}$ for all ${}^{r}_{i} \in I^M$ and ${}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S$.

(b) $v^{r_1\ldots r_L}_{i_1\ldots i_L} \leq \prod_{\{\gamma_1,\gamma_2,\ldots\gamma_{L-1}\} \subset \{1,2,\ldots L\}} v^{r_{\gamma_1}\ldots r_{\gamma_{L-1}}}_{i_{\gamma_1}\ldots i_{\gamma_{L-1}}}$ for all ${}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S$.

A vertex is called classical if the equality holds everywhere in (b), and non-classical otherwise.

Obviously, $\varphi(M,S) \subseteq l(M,S)$. What can be said about the vertices of $\varphi(M,S)$?

Lemma 4. *The vertices of $\varphi(M,S)$ are all the vectors $\vec{f} \in \varphi(M,S)$ such that $f^r_i, f^{r_1\ldots r_L}_{i_1\ldots i_L} \in \{0,1\}$ for all ${}^{r}_{i} \in I^M$ and ${}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S$.*

Proof. One direction is trivial: if $\vec{f} \in \varphi(M,S)$ and $f_i^r, f_{i_1\ldots i_L}^{r_1\ldots r_L} \in \{0,1\}$ for all ${}^r_i \in I^M$ and ${}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S$, then $\vec{f}$ is a vertex. For, if there exist $\vec{f'}, \vec{f''} \in \varphi(M,S)$ such that $\vec{f} = \lambda\vec{f'} + (1-\lambda)\vec{f''}$ with some $0 < \lambda < 1$, then obviously $\vec{f'} = \vec{f''} = \vec{f}$.

The proof of the other direction is quite involved. For a more concise notation, introduce the following sets of indices:

$$\begin{aligned}
I &= \Big\{1|^r_i,\, 2|^r_i,\, 3|^{r_1\ldots r_L}_{i_1\ldots i_L},\, 4|^{r_1\ldots r_L}_{i_1\ldots i_L}|^{r_{\gamma_1}\ldots r_{\gamma_{L-1}}}_{i_{\gamma_1}\ldots i_{\gamma_{L-1}}},\, 5|r,\, 6|r_1\ldots r_L,\, 7|^{r'_1\ldots r'_L}_{i'_1\ldots i'_L} \,\Big|\, \text{for all} \\
&\qquad {}^r_i \in I^M,\, {}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S, {}^{r'_1\ldots r'_L}_{i'_1\ldots i'_L} \in S_0, \text{and } \{\gamma_1,\ldots\gamma_{L-1}\} \subset \{1,\ldots L\}\Big\} \\
I^0 &= \Big\{1|^r_i,\, 2|^r_i,\, 3|^{r_1\ldots r_L}_{i_1\ldots i_L},\, 4|^{r_1\ldots r_L}_{i_1\ldots i_L}|^{r_{\gamma_1}\ldots r_{\gamma_{L-1}}}_{i_{\gamma_1}\ldots i_{\gamma_{L-1}}} \,\Big|\, \text{for all } {}^r_i \in I^M, {}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S, \\
&\qquad \text{and } \{\gamma_1,\ldots\gamma_{L-1}\} \subset \{1,\ldots L\}\Big\} \\
I^+ &= \Big\{5|r,\, 6|r_1\ldots r_L,\, 7|^{r'_1\ldots r'_L}_{i'_1\ldots i'_L} \,|\, \text{for all } 1 \le r \le m,\, \{r_1\ldots r_L\} \in \mathfrak{P}, \\
&\qquad \text{and } {}^{r'_1\ldots r'_L}_{i'_1\ldots i'_L} \in S_0\Big\}
\end{aligned}$$

Obviously, $I = I^0 \cup I^+$ and $I^0 \cap I^+ = \emptyset$. Rewrite (35)–(41) in the following standard form:

$$\left\langle \vec{\omega}_\mu, \vec{f} \right\rangle - b_\mu \le 0 \quad \text{for all } \mu \in I^0 \tag{42}$$

$$\left\langle \vec{\omega}_\mu, \vec{f} \right\rangle - b_\mu = 0 \quad \text{for all } \mu \in I^+ \tag{43}$$

with the following $\vec{\omega}_\mu \in \mathbb{R}^{M+|S|}$ and $b_\mu \in \mathbb{R}$:

$$\vec{\omega}_{1|^r_i} = (0\ldots0\ \underbrace{-1}^{\substack{r\\i}}\ 0\ldots0) \tag{44}$$

$$b_{1|^r_i} = 0 \tag{45}$$

$$\vec{\omega}_{2|^r_i} = (0\ldots0\,\underbrace{1}^{\substack{r\\i}}\,0\ldots0) \tag{46}$$

$$b_{2|^r_i} = 1 \tag{47}$$

$$\vec{\omega}_{3|^{r_1\ldots r_L}_{i_1\ldots i_L}} = (0\ldots0\ \underbrace{-1}^{\substack{r_1\ldots r_L\\i_1\ldots i_L}}\ 0\ldots0) \tag{48}$$

$$b_{3|^{r_1\ldots r_L}_{i_1\ldots i_L}} = 0 \tag{49}$$

$$\vec{\omega}_{4|^{r_1\ldots r_L}_{i_1\ldots i_L}|^{r_{\gamma_1}\ldots r_{\gamma_{L-1}}}_{i_{\gamma_1}\ldots i_{\gamma_{L-1}}}} = (0\ldots0\ \underbrace{-1}^{\substack{r_{\gamma_1}\ldots r_{\gamma_{L-1}}\\i_{\gamma_1}\ldots i_{\gamma_{L-1}}}}\ 0\ldots0\ \underbrace{1}^{\substack{r_1\ldots r_L\\i_1\ldots i_L}}\ 0\ldots0) \tag{50}$$

$$b_{4|^{r_1\ldots r_L}_{i_1\ldots i_L}|^{r_{\gamma_1}\ldots r_{\gamma_{L-1}}}_{i_{\gamma_1}\ldots i_{\gamma_{L-1}}}} = 0 \tag{51}$$

$$\vec{\omega}_{5|r} = (0\ldots0\,\underbrace{1}^{\substack{r\\1}}\underbrace{1}^{\substack{r\\2}}\underbrace{1}^{\substack{r\\3}}\ldots\underbrace{1}^{\substack{r\\n_r}}\,0\ldots0) \tag{52}$$

$$b_{5|r} = 1 \tag{53}$$

$$\vec{\omega}_{6|r_1\ldots r_L} = (0\ldots0 \quad \overset{\substack{r_1 r_2 \ldots r_L \\ 1\ 1\ \ldots 1}}{\breve{1}} \quad \overset{\substack{r_1 r_2 \ldots r_L \\ 2\ 1\ \ldots 1}}{\breve{1}} \quad \ldots \quad \overset{\substack{r_1\ \ r_2\ \ \ldots r_L \\ n_{r_1} n_{r_2} \ldots n_{r_L}}}{\breve{1}} \quad 0\ldots0) \tag{54}$$

$$b_{6|r_1\ldots r_L} = 1 \tag{55}$$

$$\vec{\omega}_{7|\substack{r'_1\ldots r'_L \\ i'_1\ldots i'_L}} = (0\ldots0 \quad \overset{\substack{r'_1\ldots r'_L \\ i'_1\ldots i'_L}}{\breve{1}} \quad 0\ldots0) \tag{56}$$

$$b_{7|\substack{r'_1\ldots r'_L \\ i'_1\ldots i'_L}} = 0 \tag{57}$$

where ${}^{r}_{i} \in I^M$, ${}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S$, $\{\gamma_1, \ldots \gamma_{L-1}\} \subset \{1, \ldots L\}$, and ${}^{r'_1\ldots r'_L}_{i'_1\ldots i'_L} \in S_0$. Notice that $l\,(M, S)$ is defined by (42).

For an arbitrary $\vec{f} \in l\,(M, S)$ we define the following sets:

$$I_{\vec{f}} = \left\{ \mu \in I \,\middle|\, \left\langle \vec{\omega}_\mu, \vec{f} \right\rangle - b_\mu = 0 \right\}$$
$$I^0_{\vec{f}} = \left\{ \mu \in I^0 \,\middle|\, \left\langle \vec{\omega}_\mu, \vec{f} \right\rangle - b_\mu = 0 \right\}$$

Notice that if $\vec{f} \in \varphi\,(M, S)$, then $I_{\vec{f}} = I^0_{\vec{f}} \cup I^+$, due to the fact that (39)–(41) can be satisfied only with equality.

$I_{\vec{f}}$ constitutes the so called 'active index set' for $\vec{f} \in \varphi\,(M, S)$; and according to a known theorem (see Appendix 1) $\vec{f}$ is a vertex of $\varphi\,(M, S)$ if and only if

$$\text{span}\,\{\vec{\omega}_\mu\}_{\mu \in I_{\vec{f}}} = \mathbb{R}^{M+|S|} \tag{58}$$

Similarly, a vector $\vec{f} \in l\,(M, S)$ is a vertex of $l\,(M, S)$ if and only if

$$\text{span}\,\{\vec{\omega}_\mu\}_{\mu \in I^0_{\vec{f}}} = \mathbb{R}^{M+|S|} \tag{59}$$

For all $\vec{f} \in l\,(M, S)$ define

$$J_{\vec{f}} = \left\{ {}^{r}_{i} \,\middle|\, {}^{r}_{i} \in I^M \text{ and } 0 < f^r_i < 1 \right\}$$
$$J'_{\vec{f}} = \left\{ {}^{r_1\ldots r_L}_{i_1\ldots i_L} \,\middle|\, {}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S \text{ and } 0 < f^{r_1\ldots r_L}_{i_1\ldots i_L} < 1 \right\}$$

Notice that for all ${}^{r}_{i} \in I^M$ and ${}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S$,

$$\vec{e}^{\,r}_{i} \in \text{span}\,\{\vec{\omega}_\mu\}_{\mu \in I^0_{\vec{f}}} \quad \text{if} \quad f^r_i \in \{0, 1\} \tag{60}$$

$$\vec{e}^{\,r_1\ldots r_L}_{i_1\ldots i_L} \in \text{span}\,\{\vec{\omega}_\mu\}_{\mu \in I^0_{\vec{f}}} \quad \text{if} \quad f^{r_1\ldots r_L}_{i_1\ldots i_L} \in \{0, 1\} \tag{61}$$

since the corresponding inequalities (35)–(38) must hold with equality. The only case that requires a bit of reflection is when $f^{r_1\ldots r_L}_{i_1\ldots i_L} = 1$. For example, if

$f^{r_1 r_2}_{i_1 i_2} = 1$ then (38) is satisfied with equality, so that $f^{r_1}_{i_1} = 1$, therefore (36) is also satisfied with equality. Consequently,

$$\begin{aligned} \vec{\omega}_{2|^{r_1}_{i_1}} &\in \operatorname{span}\{\vec{\omega}_\mu\}_{\mu\in I^0_{\vec{f}}} \\ \vec{\omega}_{4|^{r_1 r_2}_{i_1 i_2}|^{r_1}_{i_1}} &\in \operatorname{span}\{\vec{\omega}_\mu\}_{\mu\in I^0_{\vec{f}}} \end{aligned}$$

At the same time, as it can be seen from (46) and (50),

$$\vec{e}^{\,r_1 r_2}_{i_1 i_2} = \vec{\omega}_{2|^{r_1}_{i_1}} + \vec{\omega}_{4|^{r_1 r_2}_{i_1 i_2}|^{r_1}_{i_1}} \in \operatorname{span}\{\vec{\omega}_\mu\}_{\mu\in I^0_{\vec{f}}}$$

This can be recursively continued for the triple and higher conjunction indices.

Assume now that $\vec{f} \in \varphi(M,S)$ is such that $J_{\vec{f}} \cup J'_{\vec{f}} \neq \emptyset$, and at the same time it is a vertex of $\varphi(M,S)$, that is, (58) is satisfied. We are going to show that this leads to contradiction.

Due to (60)–(61), the assumption that $\vec{f}$ is a vertex implies that all base vectors of $\mathbb{R}^{M+|S|}$ must belong to $\operatorname{span}\{\vec{\omega}_\mu\}_{\mu\in I^0_{\vec{f}}}$ save for some $\vec{e}^{\,r}_{i}$'s with ${}^{r}_{i} \in J_{\vec{f}}$ and/or some $\vec{e}^{\,r_1\ldots r_L}_{i_1\ldots i_L}$'s with ${}^{r_1\ldots r_L}_{i_1\ldots i_L} \in J'_{\vec{f}}$. On the other hand, $\vec{f}$ being a vertex implies that (58) holds, therefore

$$\begin{aligned} \vec{e}^{\,r}_{i} &\in \operatorname{span}\{\vec{\omega}_\mu\}_{\mu\in I_{\vec{f}}} &\quad& \text{for all } {}^{r}_{i} \in J_{\vec{f}} \\ \vec{e}^{\,r_1\ldots r_L}_{i_1\ldots i_L} &\in \operatorname{span}\{\vec{\omega}_\mu\}_{\mu\in I_{\vec{f}}} && \text{for all } {}^{r_1\ldots r_L}_{i_1\ldots i_L} \in J'_{\vec{f}} \end{aligned}$$

Taking into account that $I_{\vec{f}} = I^0_{\vec{f}} \cup I^+$, it means that for all ${}^{r}_{i} \in J_{\vec{f}}$ and arbitrary $\tau^r_i \neq 0$ there exist vectors

$${}^{r}_{i}\vec{v} \quad \in \quad \operatorname{span}\{\vec{\omega}_\mu\}_{\mu\in I^0_{\vec{f}}} \tag{62}$$

such that,

$$\begin{aligned} \tau^r_i \vec{e}^{\,r}_{i} \quad = \quad & {}^{r}_{i}\vec{v} + \sum_{s=1}^{m} {}^{r}_{i}\kappa_s \vec{\omega}_{5|s} + \sum_{\{r_1,\ldots r_{L'}\}\in\mathfrak{P}} {}^{r}_{i}\kappa'_{r_1\ldots r_{L'}} \vec{\omega}_{6|r_1\ldots r_{L'}} \\ & + \sum_{{}^{r'_1\ldots r'_L}_{i'_1\ldots i'_L}\in S_0} {}^{r}_{i}\lambda_{{}^{r'_1\ldots r'_L}_{i'_1\ldots i'_L}} \vec{\omega}_{7|^{r'_1\ldots r'_L}_{i'_1\ldots i'_L}} \end{aligned} \tag{63}$$

with some real numbers ${}^{r}_{i}\kappa_s$, ${}^{r}_{i}\kappa'_{r_1\ldots r_{L'}}$, and ${}^{r}_{i}\lambda_{{}^{r'_1\ldots r'_L}_{i'_1\ldots i'_L}}$. From the definitions of $\vec{\omega}_{5|r}$, $\vec{\omega}_{6|r_1\ldots r_L}$, and $\vec{\omega}_{7|^{r'_1\ldots r'_L}_{i'_1\ldots i'_L}}$ in (44)–(57) we can write:

$$\tau^r_i \vec{e}^{\,r}_{i} = {}^{r}_{i}\,\vec{v} + \sum_{{}^{s}_{j}\in I^M} {}^{r}_{i}\kappa_s \vec{e}^{\,s}_{j} + \sum_{{}^{s_1\ldots s_{L'}}_{j_1\ldots j_{L'}}\in S} {}^{r}_{i}\kappa'_{s_1\ldots s_{L'}} \vec{e}^{\,s_1\ldots s_{L'}}_{j_1\ldots j_{L'}}$$

$$+ \sum_{\substack{r'_1 \ldots r'_L \\ i'_1 \ldots i'_L} \in S_0} {}^{r}_{i}\lambda_{\substack{r'_1 \ldots r'_L \\ i'_1 \ldots i'_L}} \vec{e}^{\,r'_1 \ldots r'_L}_{i'_1 \ldots i'_L} \tag{64}$$

Mutatis mutandis, we have the same equation for $\vec{e}^{\,r_1 \ldots r_L}_{i_1 \ldots i_L}$ for all ${}^{r_1 \ldots r_L}_{i_1 \ldots i_L} \in J'_{\vec{f}}$ with arbitrary $\tau^{r_1 \ldots r_L}_{i_1 \ldots i_L} \neq 0$ and with some numbers ${}^{r_1 \ldots r_L}_{i_1 \ldots i_L}\kappa_s$, ${}^{r_1 \ldots r_L}_{i_1 \ldots i_L}\kappa'_{r_1 \ldots r_{L'}}$, and ${}^{r_1 \ldots r_L}_{i_1 \ldots i_L}\lambda_{\substack{r'_1 \ldots r'_L \\ i'_1 \ldots i'_L}}$:

$$\tau^{r_1 \ldots r_L}_{i_1 \ldots i_L} \vec{e}^{\,r_1 \ldots r_L}_{i_1 \ldots i_L} = {}^{r_1 \ldots r_L}_{i_1 \ldots i_L}\vec{v} + \sum_{{}^s_j \in I^M} {}^{r_1 \ldots r_L}_{i_1 \ldots i_L}\kappa_s \vec{e}^{\,s}_j + \sum_{\substack{s_1 \ldots s_{L'} \\ j_1 \ldots j_{L'}} \in S} {}^{r_1 \ldots r_L}_{i_1 \ldots i_L}\kappa'_{s_1 \ldots s_{L'}} \vec{e}^{\,s_1 \ldots s_{L'}}_{j_1 \ldots j_{L'}}$$

$$+ \sum_{\substack{r'_1 \ldots r'_L \\ i'_1 \ldots i'_L} \in S_0} {}^{r_1 \ldots r_L}_{i_1 \ldots i_L}\lambda_{\substack{r'_1 \ldots r'_L \\ i'_1 \ldots i'_L}} \vec{e}^{\,r'_1 \ldots r'_L}_{i'_1 \ldots i'_L} \tag{65}$$

With some rearrangement, from (64) we have

$$\sum_{\substack{{}^s_j \in J_{\vec{f}} \\ {}^s_j \neq {}^r_i}} {}^{r}_{i}\kappa_s \vec{e}^{\,s}_j + \left({}^{r}_{i}\kappa_r - \tau^r_i\right) \vec{e}^{\,r}_i + \sum_{\substack{s_1 \ldots s_{L'} \\ j_1 \ldots j_{L'}} \in J'_{\vec{f}}} {}^{r}_{i}\kappa'_{s_1 \ldots s_{L'}} \vec{e}^{\,s_1 \ldots s_{L'}}_{j_1 \ldots j_{L'}}$$

$$= - \sum_{\substack{r'_1 \ldots r'_L \\ i'_1 \ldots i'_L} \in S_0} {}^{r}_{i}\lambda_{\substack{r'_1 \ldots r'_L \\ i'_1 \ldots i'_L}} \vec{e}^{\,r'_1 \ldots r'_L}_{i'_1 \ldots i'_L} - \sum_{\substack{{}^{s_1 \ldots s_{L'}}_{j_1 \ldots j_{L'}} \in S \\ {}^{s_1 \ldots s_{L'}}_{j_1 \ldots j_{L'}} \notin J'_{\vec{f}}}} {}^{r}_{i}\kappa'_{s_1 \ldots s_{L'}} \vec{e}^{\,s_1 \ldots s_{L'}}_{j_1 \ldots j_{L'}}$$

$$- \sum_{\substack{{}^s_j \in I^M \\ {}^s_j \notin J_{\vec{f}}}} {}^{r}_{i}\kappa_s \vec{e}^{\,s}_j - {}^{r}_{i}\vec{v} \tag{66}$$

for all ${}^r_i \in J_{\vec{f}}$.

Similarly, from (65) we have

$$\sum_{{}^s_j \in J_{\vec{f}}} {}^{r_1 \ldots r_L}_{i_1 \ldots i_L}\kappa_s \vec{e}^{\,s}_j + \sum_{\substack{{}^{s_1 \ldots s_{L'}}_{j_1 \ldots j_{L'}} \in J'_{\vec{f}} \\ {}^{s_1 \ldots s_{L'}}_{j_1 \ldots j_{L'}} \neq {}^{r_1 \ldots r_L}_{i_1 \ldots i_L}}} {}^{r_1 \ldots r_L}_{i_1 \ldots i_L}\kappa'_{s_1 \ldots s_{L'}} \vec{e}^{\,s_1 \ldots s_{L'}}_{j_1 \ldots j_{L'}}$$

$$+ \left({}^{r_1 \ldots r_L}_{i_1 \ldots i_L}\kappa'_{r_1 \ldots r_{L'}} - \tau^{r_1 \ldots r_L}_{i_1 \ldots i_L}\right) \vec{e}^{\,r_1 \ldots r_L}_{i_1 \ldots i_L}$$

$$= - \sum_{\substack{r'_1 \ldots r'_L \\ i'_1 \ldots i'_L} \in S_0} {}^{r_1 \ldots r_L}_{i_1 \ldots i_L}\lambda_{\substack{r'_1 \ldots r'_L \\ i'_1 \ldots i'_L}} \vec{e}^{\,r'_1 \ldots r'_L}_{i'_1 \ldots i'_L}$$

$$- \sum_{\substack{{}^{s_1 \ldots s_{L'}}_{j_1 \ldots j_{L'}} \in S \\ {}^{s_1 \ldots s_{L'}}_{j_1 \ldots j_{L'}} \notin J'_{\vec{f}}}} {}^{r_1 \ldots r_L}_{i_1 \ldots i_L}\kappa'_{s_1 \ldots s_{L'}} \vec{e}^{\,s_1 \ldots s_{L'}}_{j_1 \ldots j_{L'}}$$

$$-\sum_{\substack{{}^{s}_{j}\in I^{M}\\ {}^{s}_{j}\notin J_{\vec{f}}}} {}^{r_1\ldots r_L}_{i_1\ldots i_L}\kappa_s \vec{e}^{\,s}_{j} - {}^{r_1\ldots r_L}_{i_1\ldots i_L}\vec{v} \tag{67}$$

for all ${}^{r_1\ldots r_L}_{i_1\ldots i_L}\in J'_{\vec{f}}$.

Denote the right hand side of (66) by $\vec{B}_{{}^{r}_{i}}$ and the right hand side of (67) by $\vec{B}_{{}^{r_1\cdots r_L}_{i_1\cdots i_L}}$. Notice that the vectors $\vec{B}_{{}^{r}_{i}}$ and $\vec{B}_{{}^{r_1\cdots r_L}_{i_1\cdots i_L}}$ are contained in span $\{\vec{\omega}_\mu\}_{\mu\in I^0_{\vec{f}}}$, due to (62), and (60)–(61). So, in (66)–(67), together, we have a system of linear equations with vector-variables $\left\{\vec{e}^{\,s}_{j}\right\}_{{}^{s}_{j}\in J_{\vec{f}}}$ and $\left\{\vec{e}^{\,s_1\ldots s_{L'}}_{j_1\ldots j_{L'}}\right\}_{{}^{s_1\cdots s_{L'}}_{j_1\cdots j_{L'}}\in J'_{\vec{f}}}$, which can be written in the following form:

$$\sum_{{}^{s}_{j}\in J_{\vec{f}},\,{}^{s_1\cdots s_{L'}}_{j_1\cdots j_{L'}}\in J'_{\vec{f}}} \beta_{\left({}^{r}_{i},\,{}^{r_1\cdots r_L}_{i_1\cdots i_L}\right)\left({}^{s}_{j},\,{}^{s_1\cdots s_{L'}}_{j_1\cdots j_{L'}}\right)} \left(\vec{e}^{\,s}_{j},\,\vec{e}^{\,s_1\ldots s_{L'}}_{j_1\ldots j_{L'}}\right) = \left(\vec{B}_{{}^{r}_{i}},\vec{B}_{{}^{r_1\cdots r_L}_{i_1\cdots i_L}}\right) \tag{68}$$

where $\beta_{\left({}^{r}_{i},\,{}^{r_1\cdots r_L}_{i_1\cdots i_L}\right)\left({}^{s}_{j},\,{}^{s_1\cdots s_{L'}}_{j_1\cdots j_{L'}}\right)}$ is a $\left(|J_{\vec{f}}|+|J'_{\vec{f}}|\right)\times\left(|J_{\vec{f}}|+|J'_{\vec{f}}|\right)$ matrix with diagonal elements

$$\begin{aligned}\beta_{{}^{r}_{i}{}^{r}_{i}} &= {}^{r}_{i}\kappa_r - \tau^{r}_{i}\\ \beta_{{}^{r_1\cdots r_L}_{i_1\cdots i_L}{}^{r_1\cdots r_L}_{i_1\cdots i_L}} &= {}^{r_1\ldots r_L}_{i_1\ldots i_L}\kappa'_{r_1\ldots r_L} - \tau^{r_1\ldots r_L}_{i_1\ldots i_L}\end{aligned}$$

The off diagonal elements depend only on ${}^{r}_{i}\kappa_s$'s and ${}^{r_1\ldots r_L}_{i_1\ldots i_L}\kappa'_{s_1\ldots s_{L'}}$'s. Since the numbers $\tau^{r}_{i}\neq 0$ and $\tau^{r_1\ldots r_L}_{i_1\ldots i_L}\neq 0$ in the diagonal can be chosen arbitrarily, we may assume that $\det\beta_{\left({}^{r}_{i},\,{}^{r_1\cdots r_L}_{i_1\cdots i_L}\right)\left({}^{s}_{j},\,{}^{s_1\cdots s_{L'}}_{j_1\cdots j_{L'}}\right)}\neq 0$. Therefore, the system of linear equations (68) has a unique solution for all vector-variables $\vec{e}^{\,r}_{i}$ and $\vec{e}^{\,r_1\ldots r_L}_{i_1\ldots i_L}$, namely,

$$\left(\vec{e}^{\,r}_{i},\,\vec{e}^{\,r_1\ldots r_L}_{i_1\ldots i_L}\right) = \sum_{{}^{s}_{j}\in J_{\vec{f}},\,{}^{s_1\cdots s_{L'}}_{j_1\cdots j_{L'}}\in J'_{\vec{f}}} \beta^{-1}_{\left({}^{r}_{i},\,{}^{r_1\cdots r_L}_{i_1\cdots i_L}\right)\left({}^{s}_{j},\,{}^{s_1\cdots s_{L'}}_{j_1\cdots j_{L'}}\right)} \left(\vec{B}_{{}^{s}_{j}},\,\vec{B}_{{}^{s_1\cdots s_{L'}}_{j_1\cdots j_{L'}}}\right)$$

Taking into account that $\vec{B}_{{}^{r}_{i}},\vec{B}_{{}^{r_1\cdots r_L}_{i_1\cdots i_L}}\in$ span $\{\vec{\omega}_\mu\}_{\mu\in I^0_{\vec{f}}}$, this all means that for all ${}^{r}_{i}\in J_{\vec{f}}$, the base vectors $\vec{e}^{\,r}_{i}$, and for all ${}^{r_1\ldots r_L}_{i_1\ldots i_L}\in J'_{\vec{f}}$, the base vectors $\vec{e}^{\,r_1\ldots r_L}_{i_1\ldots i_L}$ can be expressed as linear combinations of vectors contained in span $\{\vec{\omega}_\mu\}_{\mu\in I^0_{\vec{f}}}$. As all the rest of base vectors belong to span $\{\vec{\omega}_\mu\}_{\mu\in I^0_{\vec{f}}}$ (as we have already mentioned above), we have

$$\text{span}\,\{\vec{\omega}_\mu\}_{\mu\in I^0_{\vec{f}}} = \mathbb{R}^{M+|S|}$$

meaning that $\vec{f}$ must be a vertex of $l\,(M,S)$. Due to the fact that all components of a vertex of $l\,(M,S)$ are necessarily 0 or 1, there cannot exists a vertex $\vec{f}\in\varphi\,(M,S)$ with $0<f^{r}_{i}<1$ and/or $0<f^{r_1\ldots r_L}_{i_1\ldots i_L}<1$. $\square$

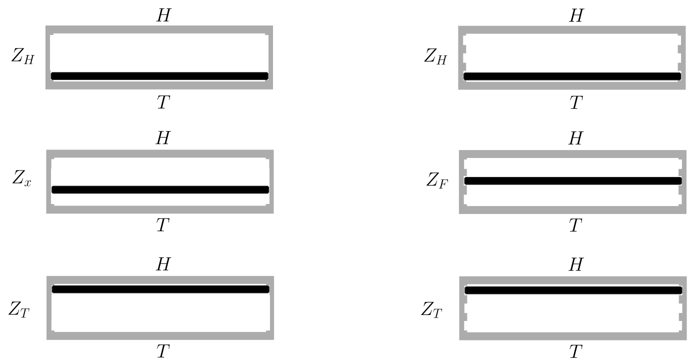

Figure 1: Baised coin

All this means that the vertices of $\varphi(M,S)$ are those vertices of $l(M,S)$ that satisfy the further restrictions (39)–(41).

To sum up, the "space" of possible states is a closed convex polytope $\varphi(M,S) \subset \mathbb{R}^{M+|S|}$ whose vertices are the vectors $\vec{w} \in \mathbb{R}^{M+|S|}$ such that

(a) $w_i^r, w_{i_1\ldots i_L}^{r_1\ldots r_L} \in \{0,1\}$ for all ${}_i^r \in I^M$ and ${}_{i_1\ldots i_L}^{r_1\ldots r_L} \in S$.

(b) $w_{i_1\ldots i_L}^{r_1\ldots r_L} \leq \prod_{\{\gamma_1,\ldots\gamma_{L-1}\}\subset\{1,\ldots L\}} w_{i_{\gamma_1}\ldots i_{\gamma_{L-1}}}^{r_{\gamma_1}\ldots r_{\gamma_{L-1}}}$ for all ${}_{i_1\ldots i_L}^{r_1\ldots r_L} \in S$.

(c) $w_{i_1\ldots i_L}^{r_1\ldots r_L} = 0$ for all ${}_{i_1\ldots i_L}^{r_1\ldots r_L} \in S_0$.

(d) For all $1 \leq r \leq m$ there is exactly one $1 \leq i_*^r \leq n_r$ such that $w_{i_*^r}^r = 1$.

(e) For all $\{r_1, \ldots r_L\} \in \mathfrak{P}$ there is exactly one ${}_{i_1^*\ldots i_L^*}^{r_1\ldots r_L} \in S$ such that $w_{i_1^*\ldots i_L^*}^{r_1\ldots r_L} = 1$.

(We note in advance that property (d) will be crucial in the proof of Theorem 9.) Let $\mathcal{W} = \{\vec{w}_\vartheta\}_{\vartheta\in\Theta}$ denote the set of vertices of $\varphi(M,S)$.

Again, we emphasize that $\varphi(M,S)$ is the total space of theoretically possible states, determined by the totality of possible relative frequency functions over $\mathcal{A}$ that satisfy conditions **(E1)**–**(E3)**. It may be that the empirically determined possible states of the system for different physical preparations constitute only a subset $\Phi \subseteq \varphi(M,S)$. Note that such a subset is not necessarily a convex subset in $\varphi(M,S)$—notice that Lemma 3 is about the whole $\varphi(M,S)$.

As a simple illustration, consider a coin made of a very light material, which is empty inside. Inside we install a heavy metal disk, such that the disk can be fixed in different positions. (Fig. 1) In this way, the system can be prepared in different states. There is only one measurement we may perform, tossing the coin, with two possible outcomes, Heads and Tails, and one conjunction. Keeping the preparation fixed, that is, keeping the position of the disk fixed, we can take the statistics of Heads and Tails. The observed data will perfectly satisfy **(E1)**–**(E3)**, and the state space $\varphi(M,S)$ (where $M=2$ and $S=\left\{\begin{smallmatrix}1&1\\1&2\end{smallmatrix}\right\}$) is a one-dimensional polytope in $\mathbb{R}^3$, determined by two vertices $(1,0,0)$ and $(0,1,0)$. Denote an arbitrary state vector by $\vec{Z}=(z_H,z_T,0)$, meaning that whenever we perform the coin-toss and the coin is in state $\vec{Z}$, we get Heads with relative frequency z_H, Tails with z_T, but never the Heads and Tails in conjunction.

Consider now the particular physical situation on the left hand side of Fig. 1. The position of the disk can be continuously varied between two extreme positions; one in which the coin is maximally biased for Heads,

$$\vec{Z}_H=(0.8,0.2,0)$$

the other in which the coin is maximally biased for Tails,

$$\vec{Z}_T=(0.2,0.8,0)$$

Each intermediate position is possible and the corresponding state vector $\vec{Z}_x$ falls between vectors $\vec{Z}_H$ and $\vec{Z}_T$. So, the physically possible sates are restricted to a proper subset $\Phi\subset\varphi(M,S)$, the line segment between $\vec{Z}_H$ and $\vec{Z}_T$; but, Φ is closed under convex combination.

Consider now another physical situation shown on the right hand side of Fig. 1. There are three separate slots where the disk can be clicked. Accordingly, there are only three possible positions, the two extremes and a third in the middle. The corresponding state vectors are $\vec{Z}_H$, $\vec{Z}_T$, and the "fair" state

$$\vec{Z}_F=(0.5,0.5,0)$$

Now,

$$\Phi=\left\{\vec{Z}_H,\vec{Z}_F,\vec{Z}_T\right\}\subset\varphi(M,S) \tag{69}$$

and it is obviously not closed under convex combination. While for example $\frac{1}{2}\vec{Z}_H+\frac{1}{2}\vec{Z}_T=\vec{Z}_F$ is indeed contained in Φ, $\frac{2}{3}\vec{Z}_H+\frac{1}{3}\vec{Z}_T=(0.6,0.4,0)$ is not.

Thus, the set of physically possible states Φ can be a strongly restricted subset of the total space of theoretically possible states $\varphi(M,S)$; and it is not necessary closed under convex combination. (We do not see any reason to adopt for example the a priori argumentation in the GPT literature that Φ must be convex; see Appendix 2.) The actual content of $\Phi\subseteq\varphi(M,S)$ is determined by further physical facts beyond the stipulated **(E1)**–**(E3)**. We do not wish to impose such an additional restriction, even if it could be empirically justified for certain types of physical systems. The point of assumptions **(E1)**-**(E3)** is precisely that they are general; we know of no physical system whose description in operational terms does not satisfy them.

Keeping all this in mind, for the sake of generality, we will consider $\varphi(M,S)$ as if it were the space of states without any restrictions. For the purposes of our analysis below, this is of no particular relevance, and all of our results below can be easily modified for a particular subset $\Phi\subset\varphi(M,S)$.

3.2 The State Space – Manifold View

A closed convex polytope like $\varphi(M,S) \subset \mathbb{R}^{M+|S|}$ is a $\dim(\varphi(M,S))$-dimensional manifold with boundary. Any coordinate system in the affine hull of $\varphi(M,S)$ can be a natural coordination of $\varphi(M,S)$.

Thus, $\varphi(M,S)$ as a manifold with boundary is a perfect mathematical representation of the states of the system; in fact, it is the most straightforward one, expressible directly in empirical terms. This is however not the only one. Any mathematical object can represent the state of the system that determines the system's probabilistic behavior against all possible measurement operations. For example, for our later purposes the convex decomposition

$$\vec{Z} = \sum_{\vartheta\in\Theta} \lambda_\vartheta \vec{w}_\vartheta \qquad \lambda_\vartheta \geq 0, \sum_{\vartheta\in\Theta} \lambda_\vartheta = 1 \tag{70}$$

will be a more suitable characterization of a point of the state space. However, in general, this decomposition is not unique. In fact there are continuum many ways of such decomposition for all $\vec{Z} \in \operatorname{Int}\varphi(M,S)$; and a unique one if $\vec{Z}$ is on the boundary. As we will show, there are various good solutions for obtaining a unique representation of states in terms of their vertex decomposition (70).

Introduce the following notation: $\vec{\lambda} = (\lambda_\vartheta)_{\vartheta\in\Theta} \in \mathbb{R}^{|\Theta|}$. Let

$$\Lambda = \left\{ \vec{\lambda} \in \mathbb{R}^{|\Theta|} \,\middle|\, \lambda_\vartheta \geq 0, \sum_{\vartheta\in\Theta} \lambda_\vartheta = 1 \right\}$$

Λ is the $(|\Theta|-1)$-dimensional standard simplex in $\mathbb{R}^{|\Theta|}$. Obviously,

$$D : \Lambda \to \varphi(M,S)\,; \quad D\left(\vec{\lambda}\right) = \sum_{\vartheta\in\Theta} \lambda_\vartheta \vec{w}_\vartheta \tag{71}$$

is a continuous projection, and it preserves convex combination.

Lemma 5. *For all $\vec{Z} \in \varphi(M,S)$, $D^{-1}\left(\vec{Z}\right)$ is a polytope contained in Λ.*

Proof. To satisfy (70), beyond being contained in Λ, $\vec{\lambda}$ has to satisfy the following system of linear equations:

$$\sum_{\vartheta\in\Theta} \lambda_\vartheta {w_\vartheta}_i^r = Z_i^r \qquad {}_i^r \in I^M \tag{72}$$

$$\sum_{\vartheta\in\Theta} \lambda_\vartheta {w_\vartheta}_{i_1\ldots i_L}^{r_1\ldots r_L} = Z_{i_1\ldots i_L}^{r_1\ldots r_L} \qquad {}_{i_1\ldots i_L}^{r_1\ldots r_L} \in S \tag{73}$$

For a given $\vec{Z}$, the set of solutions constitute an affine subspace $\mathfrak{a}_{\vec{Z}} \subset \mathbb{R}^{|\Theta|}$ with difference space $\mathcal{B} \subset \mathbb{R}^{|\Theta|}$ constituted by the solutions of the homogeneous equations

$$\sum_{\vartheta\in\Theta} \lambda_\vartheta {w_\vartheta}_i^r = 0 \qquad {}_i^r \in I^M$$

$$\sum_{\vartheta\in\Theta} \lambda_\vartheta w_{\vartheta}{}^{r_1\ldots r_L}_{i_1\ldots i_L} \quad = \quad 0 \qquad {}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S$$

Notice that $D^{-1}\left(\vec{Z}\right) = \Lambda \cap \mathfrak{a}_{\vec{Z}}$. Due to the fact that an intersection of a polytope with an affine subspace is a polytope (Henk *et al.* 2004), each $D^{-1}\left(\vec{Z}\right)$ is a polytope contained in Λ. $\square$

Lemma 6. *$D^{-1}\left(\vec{Z}\right)$, as a subset of $\mathbb{R}^{|\Theta|}$, continuously depends on $\vec{Z}$ in the following sense:*

$$\lim_{\vec{Z}'\to\vec{Z}} \max_{\vec{\lambda}\in D^{-1}\left(\vec{Z}\right)} d\left(\vec{\lambda}, D^{-1}\left(\vec{Z}'\right)\right) \quad = \quad 0 \tag{74}$$

$$\lim_{\vec{Z}'\to\vec{Z}} \max_{\vec{\lambda}\in D^{-1}\left(\vec{Z}'\right)} d\left(\vec{\lambda}, D^{-1}\left(\vec{Z}\right)\right) \quad = \quad 0 \tag{75}$$

where $d\,(\, , \,)$ denotes the usual distance of a point from a set.

Proof. We have to show that (74)–(75) hold approaching from all possible directions to $\vec{Z}$. In other words, if $t \in [0,1]$ and $\Delta\vec{Z} \in \mathbb{R}^{M+|S|}$ is an arbitrary non-zero vector such that $\vec{Z} - \Delta\vec{Z} \in \varphi\,(M,S)$, then

$$\lim_{t\to 0} \max_{\vec{\lambda}\in D^{-1}\left(\vec{Z}\right)} d\left(\vec{\lambda}, D^{-1}\left(\vec{Z} - t\Delta\vec{Z}\right)\right) \quad = \quad 0 \tag{76}$$

$$\lim_{t\to 0} \max_{\vec{\lambda}\in D^{-1}\left(\vec{Z}-t\Delta\vec{Z}\right)} d\left(\vec{\lambda}, D^{-1}\left(\vec{Z}\right)\right) \quad = \quad 0 \tag{77}$$

Let $\Delta\vec{\lambda}$ be a solution of equations (72)–(73) with $\Delta\vec{Z}$:

$$\sum_{\vartheta\in\Theta} \Delta\lambda_\vartheta w_{\vartheta}{}^{r}_{i} \quad = \quad \Delta Z^{r}_{i} \qquad {}^{r}_{i} \in I^M \tag{78}$$

$$\sum_{\vartheta\in\Theta} \Delta\lambda_\vartheta w_{\vartheta}{}^{r_1\ldots r_L}_{i_1\ldots i_L} \quad = \quad \Delta Z^{r_1\ldots r_L}_{i_1\ldots i_L} \qquad {}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S \tag{79}$$

$\Delta\vec{\lambda}$ can be orthogonally decomposed as follows:

$$\Delta\vec{\lambda} = \Delta\vec{\lambda}^{\perp} + \Delta\vec{\lambda}^{\parallel} \qquad \Delta\vec{\lambda}^{\perp} \in \mathcal{B}^{\perp} \text{ and } \Delta\vec{\lambda}^{\parallel} \in \mathcal{B}$$

Obviously, $\Delta\vec{\lambda}^{\perp}$ is uniquely determined by $\Delta\vec{Z}$; accordingly, replacing $\Delta\vec{Z}$ with $t\Delta\vec{Z}$ on the right hand side of (78)–(79) we get $t\Delta\vec{\lambda}^{\perp}$ in place of $\Delta\vec{\lambda}^{\perp}$. Notice that $\left|t\Delta\vec{\lambda}^{\perp}\right|$ is the distance between the affine subspaces of solutions $\mathfrak{a}_{\vec{Z}}$ and $\mathfrak{a}_{\vec{Z}-t\Delta\vec{Z}}$; tending to zero if $t \to 0$.

Let $\vec{\lambda}$ be an arbitrary point in $D^{-1}\left(\vec{Z}\right)$ and let $\vec{\lambda}'$ be the point in $D^{-1}\left(\vec{Z} - \Delta\vec{Z}\right)$ closest to $\vec{\lambda}$, that is,

$$d\left(\vec{\lambda}, D^{-1}\left(\vec{Z}-\Delta\vec{Z}\right)\right) \quad = \quad \left|\vec{\lambda}'-\vec{\lambda}\right|$$

Consider the point

$$\vec{\lambda}_t=\vec{\lambda}+t\left(\vec{\lambda}'-\vec{\lambda}\right)$$

Obviously, $\vec{\lambda}_t \in \Lambda$ and $\vec{\lambda}_t \in \mathfrak{a}_{\vec{Z}-t\Delta\vec{Z}}$ for all $t \in [0,1]$, that is,

$$\vec{\lambda}_t \in D^{-1}\left(\vec{Z}-t\Delta\vec{Z}\right)$$

Therefore,

$$d\left(\vec{\lambda}, D^{-1}\left(\vec{Z}-t\Delta\vec{Z}\right)\right) \leq t\left|\vec{\lambda}'-\vec{\lambda}\right|$$

which implies (76).

Also, notice that

$$\lim_{t\to 0} \max_{\vec{\lambda}\in D^{-1}\left(\vec{Z}-t\Delta\vec{Z}\right)} d\left(\vec{\lambda}, \mathfrak{a}_{\vec{Z}}\right)=0$$

which implies (77), otherwise there would exist a convergent sequence of points from different $D^{-1}\left(\vec{Z}-t\Delta\vec{Z}\right)$ sets such that the limiting point is not contained in $D^{-1}\left(\vec{Z}\right)$, contradicting to the facts that Λ is closed and $D^{-1}\left(\vec{Z}\right)=\Lambda\cap \mathfrak{a}_{\vec{Z}}$. □

Lemma 5 and 6 mean that the states of the system can be represented in a continuous way by a disjoint family of polytopes contained in Λ. This is of course a very unusual and inconvenient way of representation. However, we can easily make it more convenient by assigning a point in each $D^{-1}\left(\vec{Z}\right)$ representing the entire polytope. There are several possibilities: for example, the center of mass, or any other notion of the center of a polytope. Here we will use the notion of the point of maximal entropy, which is perhaps physically also meaningful (Pitowsky 1989, p. 47).

The point of maximal entropy of an arbitrary polytope $\mathcal{S}\subset\Lambda$:

$$\vec{c}(\mathcal{S})=\begin{cases}\text{maximize} & H\left(\vec{\lambda}\right)=-\sum_{\vartheta\in\Theta}\lambda_\vartheta \log\lambda_\vartheta\\ \text{subject to} & \vec{\lambda}\in\mathcal{S}\end{cases}$$

Since $\mathcal{S}$ is contained in Λ, this maximization problem always has a solution. Meaning that $\vec{c}(\mathcal{S})$ is uniquely determined and always contained in $\mathcal{S}$.

Lemma 7. *Let us define the following section of the bundle projection (71):*

$$\begin{aligned}\sigma &: \varphi(M,S)\to\Lambda\\ \sigma\left(\vec{Z}\right) &= \vec{c}\left(D^{-1}\left(\vec{Z}\right)\right) \ \in D^{-1}\left(\vec{Z}\right)\end{aligned} \tag{80}$$

Then, $\sigma\left(\vec{Z}\right)$ is continuous in $\vec{Z}$, that is, for all $\vec{Z},\vec{Z}'\in\varphi(M,S)$,

$$\lim_{\vec{Z}'\to\vec{Z}}\sigma\left(\vec{Z}'\right)=\sigma\left(\vec{Z}\right)$$

Proof. Consider a sufficiently fine division of the unit cube $C^{|\Theta|}\subset\mathbb{R}^{|\Theta|}$ into equally sized small cubes of volume ΔV. Denote the i-th such elementary cube by C_i. The point of maximal entropy of a polytope $\mathcal{S}\subset\Lambda\subset C^{|\Theta|}$ can be approximated with arbitrary precision in the following way:

$$\vec{c}(\mathcal{S})\simeq\begin{cases}\text{maximize} & H\left({}^{i}\vec{\lambda}\right)=-\sum_{\vartheta\in\Theta}{}^{i}\lambda_{\vartheta}\log{}^{i}\lambda_{\vartheta}\\ \text{subject to} & i\in\{j\,|\,\mathcal{S}\cap C_j\neq\emptyset\}\end{cases}\tag{81}$$

where ${}^{i}\vec{\lambda}$ is, say, the center of C_i. Due to Lemma 6, for all $\Delta V>0$ there is an $\varepsilon>0$ such that, for all elementary cube C_i,

$$D^{-1}\left(\vec{Z}'\right)\cap C_i\neq\emptyset\;\Leftrightarrow\;D^{-1}\left(\vec{Z}\right)\cap C_i\neq\emptyset\qquad\text{if}\left|\vec{Z}'-\vec{Z}\right|<\varepsilon$$

Meaning that, for a sufficiently small ε, approximation (81) leads to the same result for $D^{-1}\left(\vec{Z}'\right)$ and $D^{-1}\left(\vec{Z}\right)$. Therefore,

$$\lim_{\vec{Z}'\to\vec{Z}}\vec{c}\left(D^{-1}\left(\vec{Z}'\right)\right)=\vec{c}\left(D^{-1}\left(\vec{Z}\right)\right)$$

□

By means of σ (or any similar continuous section) the whole state space $\varphi(M,S)$ can be lifted into a $\dim(\varphi(M,S))$-dimensional submanifold with boundary:

$$\Lambda_{\sigma}=\sigma(\varphi(M,S))\subset\Lambda\tag{82}$$

4 Dynamics

So far, nothing has been said about the dynamics of the system, that is, about the time evolution of the state $\vec{Z}$. First we have to introduce the concept of time evolution in general operational terms. Let us start with the most general case.

Imagine that the system is in state $\vec{Z}(t_0)$ after a certain physical preparation at time t_0. According to the definition of state, this means that the system responds to the various measurement operations right after time t_0 in a way described in (17)–(18). Let then the system evolve under a given set of circumstances until time t. Let $\vec{Z}(t)$ be the system's state at moment t. Again, this means that the system responds to the various measurement operations right after time t in a way described in (17)–(18) with $\vec{Z}(t)$. Thus, we have a temporal path of the system in the space of states $\varphi(M,S)$.

By means of a continuous cross section like (80), of course, $\vec{Z}(t)$ can be lifted into Λ_σ and expressed as a curve $\sigma\left(\vec{Z}(t)\right)$ on Λ_σ.

At this level of generality, we say nothing about the temporal path $\vec{Z}(t)$. Whether it has some specific feature or the time evolution of the system shows any regularity whatsoever, is a matter of empirical facts reflected in the observed relative frequencies under various circumstances. As a possible empirically observed such regularity, we formulate a typical situation when the time evolution $\vec{Z}(t)$ can be generated by a one-parameter group of transformations of $\varphi(M,S)$.

(E4) The time evolutions of states are such that there exists a one-parameter group of transformations of $\varphi(M,S)$, F_t, satisfying the following conditions:

$$F_t : \varphi(M,S) \to \varphi(M,S) \text{ is one-to-one}$$
$$F : \mathbb{R} \times \varphi(M,S) \to \varphi(M,S)\,;\ \left(t,\vec{Z}\right) \mapsto F_t\left(\vec{Z}\right) \text{ is continuous}$$
$$F_{t+s} = F_s \circ F_t$$
$$F_{-t} = F_t^{-1}\text{; consequently, } F_0 = id_{\varphi(M,S)}$$

and the time evolution of an arbitrary initial state $\vec{Z}(t_0) \in \varphi(M,S)$ is $\vec{Z}(t) = F_{t-t_0}\left(\vec{Z}(t_0)\right)$.

It is worth mentioning that although the state space $\varphi(M,S)$ is closed under convex combination, and, in some cases, the subset $\Phi \subseteq \varphi(M,S)$ of the actually observable states may be closed under convex combination, the stipulated empirical facts **(E1)**–**(E3)** do not imply that the time evolution should preserve convex combinations. That is,

$$\vec{Z}_3(t_0) = \lambda_1 \vec{Z}_1(t_0) + \lambda_2 \vec{Z}_2(t_0) \qquad \lambda_1, \lambda_2 \geq 0; \lambda_1 + \lambda_2 = 1$$

at time t_0 generally does not imply that

$$\vec{Z}_3(t) = \lambda_1 \vec{Z}_1(t) + \lambda_2 \vec{Z}_2(t)$$

at any other moment of time t. This does not follow even if the time evolution additionally satisfies condition **(E4)**.

As an illustration, consider our previously discussed example with the coin that can be prepared into different biased states, shown on the left hand side of Fig. 1, with the following modification. Imagine that the disk is mounted on a threaded rod that rotates in a given direction at a constant angular velocity. (Fig. 2) The threading of the rod is not uniform, but becomes denser and denser towards the two ends, and the density of threads tends to infinity as the two extreme disk positions are approached. Due to the rotation of the rod, the disk, when placed in an arbitrary initial position, moves upwards with velocity determined by the threading. Meaning, that the coin continuously evolves from

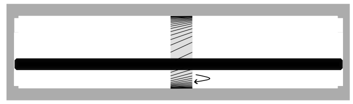

Figure 2: Moving disk

being maximally biased for Heads towards being maximally biased for Tails. One can imagine continuously many ways of such threading. For example, let us assume that the resulted change of probabilities of Heads and Tails, $F_t : (z_H, z_T, 0) \mapsto (F_t(z_H), F_t(z_T), 0)$, is something like this:

$$F_t(z_H) = 0.8 - 0.6 \frac{\frac{(0.8-z_H)}{0.6}}{\frac{(0.8-z_H)}{0.6} + \left(1 - \frac{(0.8-z_H)}{0.6}\right)\exp(-t)} \tag{83}$$

$$F_t(z_T) = 1 - F_t(z_H) \tag{84}$$

This is a physically entirely possible dynamics, and it satisfies **(E4)**. But, as it must be obvious from the formula (83) itself, it does not preserve convex combination. (For a concrete numeric example, see the end of Appendix 3.)

Thus, in our general operational-probabilistic framework based on the assumptions **(E1)**–**(E3)**, even if **(E4)** is satisfied, the time-evolution of states does not necessarily preserve convex combination; and, as the above simple example suggests, there is no reason to assume that experience shows such a restriction on the possible dynamics. (This is in stark contrast to stipulations in the GPT framework; see Appendix 3.)

By means of the continuous cross section (80), F_t generates a one-parameter group of transformations on Λ_σ, $K_t = \sigma \circ F_t \circ D$, with exactly the same properties:

$K_t : \Lambda_\sigma \to \Lambda_\sigma$ is one-to-one

$K : \mathbb{R} \times \Lambda_\sigma \to \Lambda_\sigma$; $\left(t, \vec{Z}\right) \mapsto K_t\left(\vec{Z}\right)$ is continuous

$K_{t+s} = K_s \circ K_t$

$K_{-t} = K_t^{-1}$; consequently, $K_0 = id_{\Lambda_\sigma}$

Despite all the mathematical attractiveness of **(E4)**, and despite the fact that it is satisfied for many physical systems, **(E4)** is not taken to be satisfied by the time-evolution of a system in general, and so will not be assumed in the remainder of this paper—with some rare exceptions where we will indicate this. The main reason is that **(E4)** is too strong a requirement, the violation of which doesn't mean that the system cannot have a meaningful time-evolution. For instance, consider the example depicted in Fig. 2 with the modification that the density of threads does not go to infinity at the ends, therefore the disk reaches its extreme positions in finite time, but there is a mechanism which changes the direction of the rod's rotation whenever the disk reaches an endpoint. In this way, the state $(z_H, z_T, 0)$ will continuously oscillate between the two extremes $(0.8, 0.2, 0)$ and $(0.2, 0.8, 0)$. Consequently, the state—in the operational–probabilistic sense—does not in itself determine its subsequent value, since that will depend on whether it is in the downward or upward period. So, **(E4)** is not satisfied because there does not exist such an $F_t : \varphi(M, S) \rightarrow \varphi(M, S)$ function. Yet, it is worth emphasizing that there is a definite regularity according to which temporal development takes place in the underlying ontology. This underlying temporal development is, in the above example, deterministic and Markovian (cf. Barandes 2023). The only lesson is that the operational–probabilistic notion of state is conceptually different from the notion of Cauchy data for the underlying dynamics.

5 Ontology

So, at a given time instant, the operational–probabilistic state fully characterizes the probabilistic behavior of the system with respect to all possible measurements at that time instant—according to Theorem 1. In general, however, such a probabilistic state admits different underlying ontological pictures even at the given time instant. Though, as we will see, some of those underlying ontologies imply further conditions on the observed relative frequencies. We will mention three important cases, but various combinations are conceivable.

Case 1 In the most general case, without any further restriction on the observed relative frequencies, the outcomes of the measurements are random events produced in the measurement process itself. The state $\vec{Z}$ characterizes the system in a dispositional sense: the system has a propensity to behave in a certain way, that is, to produce a certain statistics of outcomes, if a given combination of measurements is performed. In general, the produced statistics is such that, for example,

$$p\left(X_i^r | a_r \wedge a_{r'}\right) \neq p\left(X_i^r | a_r\right) \qquad \{r, r'\} \in \mathfrak{P} \tag{85}$$

meaning that the underlying process is "contextual" in the sense that the system's statistical behavior against measurement a_r can be influenced by the performance of another measurement $a_{r'}$.

Case 2 In the second case we assume that there is no such cross-influence in the underlying ontology. That is, the observed relative frequencies satisfy the following general condition:

$$p\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L} | a_{r_1} \wedge \ldots \wedge a_{r_L} \wedge a_{r'_1} \wedge \ldots \wedge a_{r'_{L'}}\right)$$

$$= p\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L} | a_{r_1} \wedge \ldots \wedge a_{r_L}\right) \tag{86}$$

for all L, L', $2 \leq L + L' \leq m$, ${}^{r_1 \ldots r_L}_{i_1 \ldots i_L} \in S$, and $\{r_1, \ldots r_L, r'_1, \ldots r'_{L'}\} \in \mathfrak{P}$. This does not mean that there cannot be correlation between the outcomes $X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L}$ and the performance of measurement $a_{r'_1} \wedge \ldots \wedge a_{r'_{L'}}$ It only means that the correlation must be the consequence of the fact that the measurement operations $a_{r'_1} \wedge \ldots \wedge a_{r'_{L'}}$ and $a_{r_1} \wedge \ldots \wedge a_{r_L}$ are correlated; $a_{r_1} \wedge \ldots \wedge a_{r_L}$ must be the common cause responsible for the correlation. Indeed, (86) is equivalent with the following "screening off" condition:

$$\begin{aligned} p\left(a_{r'_1} \wedge \ldots \wedge a_{r'_{L'}} \wedge X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L} | a_{r_1} \wedge \ldots \wedge a_{r_L}\right) \\ = p\left(a_{r'_1} \wedge \ldots \wedge a_{r'_{L'}} | a_{r_1} \wedge \ldots \wedge a_{r_L}\right) \\ \times p\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L} | a_{r_1} \wedge \ldots \wedge a_{r_L}\right) \end{aligned} \tag{87}$$

for all L, L', $2 \leq L + L' \leq m$, ${}^{r_1 \ldots r_L}_{i_1 \ldots i_L} \in S$, and $\{r_1, \ldots r_L, r'_1, \ldots r'_{L'}\} \in \mathfrak{P}$.

All this means that the state of the system $\vec{Z}$ reflects the propensities of the system to produce a certain statistics of outcomes against each possible measurement/measurement combination, separately. The observed statistics reveals the propensity in question, but, in general, we are not entitled to say that a single outcome (of a measurement/measurement combination) reveals an element of reality existing independently of the measurement(s). As we will see below, that would require a stronger restriction on the observed frequencies.

Case 3 Assume that the underlying ontology contains such elements of reality. Let us denote them by $\#X_i^r$ (${}^r_i \in I^M$). More precisely, let $\#X_i^r$ denote the event that the element of reality revealed in the outcome X_i^r is present in the given run of the experiment. Certainly, every such event $\#X_i^r$, even if hidden to us, must have some relative frequency. That is to say, there must exists a relative frequency function p' on the extended free Boolean algebra $\mathcal{A}'$ generated by the set

$$G' = \{a_r\}_{r=1,2,\ldots m} \cup \{X_i^r\}_{{}^r_i \in I^M} \cup \left\{\#X_j^s\right\}_{{}^s_j \in I^M} \tag{88}$$

such that

$$p'\big|_{\mathcal{A} \subset \mathcal{A}'} = p \tag{89}$$

The ontological assumption that $\#X_i^r$ is revealed by the measurement outcome X_i^r means that

$$p'\left(X_i^r | a_r \wedge \#X_i^r\right) = 1 \tag{90}$$

$$p'\left(X_i^r | a_r \wedge \neg\#X_i^r\right) = 0 \tag{91}$$

$$p'\left(a_r \wedge \#X_i^r\right) = p'\left(a_r\right) p'\left(\#X_i^r\right) \tag{92}$$

Similarly,

$$p'\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L} | a_{r_1} \wedge \ldots \wedge a_{r_L} \wedge \#X_{i_1}^{r_1} \wedge \ldots \wedge \#X_{i_L}^{r_L}\right) = 1 \tag{93}$$

$$p'\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L} | a_{r_1} \wedge \ldots \wedge a_{r_L} \wedge \neg \left(\# X_{i_1}^{r_1} \wedge \ldots \wedge \# X_{i_L}^{r_L}\right)\right) = 0 \tag{94}$$

$$\begin{aligned} p'\left(a_{r_1} \wedge \ldots \wedge a_{r_L} \wedge \# X_{i_1}^{r_1} \wedge \ldots \wedge \# X_{i_L}^{r_L}\right) = p'\left(a_{r_1} \wedge \ldots \wedge a_{r_L}\right) \\ \times p'\left(\# X_{i_1}^{r_1} \wedge \ldots \wedge \# X_{i_L}^{r_L}\right) \end{aligned} \tag{95}$$

for all ${}^{r_1 \ldots r_L}_{i_1 \ldots i_L} \in S$.

Now, (90)–(95) and (89) imply that

$$p'\left(\# X_i^r\right) = p\left(X_i^r | a_r\right) = Z_i^r \tag{96}$$

$$\begin{aligned} p'\left(\# X_{i_1}^{r_1} \wedge \ldots \wedge \# X_{i_L}^{r_L}\right) &= p\left(X_{i_1}^{r_1} \wedge \ldots \wedge X_{i_L}^{r_L} | a_{r_1} \wedge \ldots \wedge a_{r_L}\right) \\ &= Z_{i_1 \ldots i_L}^{r_1 \ldots r_L} \end{aligned} \tag{97}$$

for all ${}^{r}_{i} \in I^M$ and ${}^{r_1 \ldots r_L}_{i_1 \ldots i_L} \in S$.

Notice that on the right hand side of (96)–(97) we have the components of $\vec{Z}$. At the same time, on the left hand side of (96)–(97) we have numbers that are values of relative frequencies. Therefore the components of $\vec{Z}$ must constitute values of relative frequencies (of the occurrences of elements of reality $\# X_i^r$ and $\# X_{i_1}^{r_1} \ldots \wedge \# X_{i_L}^{r_L}$). Since values of relative frequencies satisfy the Kolmogorovian laws of classical probability, $\vec{Z}$ must be in the so-called classical correlation polytope (Pitowsky 1989, Ch. 2):

$$\vec{Z} \in c(M, S) \tag{98}$$

(Equivalently, the components of $\vec{Z}$ must satisfy the corresponding Bell-type inequalities.) In this case the physical state of the system admits a more fine-grained characterization than the probabilistic description provided by $\vec{Z}$: in each run of the experiment the system can be thought of as being in an underlying physical state (fixing whether the elements of reality $\# X_i^r$ and $\# X_{i_1}^{r_1} \ldots \wedge \# X_{i_L}^{r_L}$ are present or not) that predetermines the outcome of every possible measurement, given that the measurement in question is performed.

Thus, as we have seen from the above examples, the probabilistic–operational notion of state admits different underlying ontologies, depending on whether some further conditions are met or not. Note that condition (86) in Case 2 is sometimes called "no-signaling condition"; and Case 3 is usually interpreted as "admitting deterministic non-contextual hidden variables". In what follows, we do not assume anything more about the observed relative frequencies than we stipulated in **(E1)**–**(E3)**. Meaning that we remain within the most general framework of Case 1.

6 Quantum Representation

So far in the previous sections, we have stayed within the framework of classical Kolmogorovian probability theory; including the notion of state, which is a simple vector constructed from classical conditional probabilities. Meaning that any physical system—traditionally categorized as classical or quantum, or "more general than quantum"—that can be described in operational terms can be described within classical Kolmogorovian probability theory. It is worth pointing out that this is also the case when the system is traditionally described in terms of the Hilbert space quantum mechanical formalism. That is, all the empirically expressible content of the quantum mechanical description can be described in the language of classical Kolmogorovian probabilities; including what we refer to as "quantum probability", given by the usual trace formula, which can be expressed simply as classical conditional probability. All this is in perfect alignment with the Kolmogorovian Censorship Hypothesis.

In the remainder of the paper we will show that the opposite is also true: anything that can be described in operational terms can be represented in the Hilbert space quantum mechanical formalism, if we wish. From assumptions **(E1)-(E3)** alone, we will show that there always exists:

(Q1) a suitable Hilbert space, such that

(Q2) the outcomes of each measurement can be represented by a system of pairwise orthogonal closed subspaces, spanning the whole Hilbert space,

(Q3) the states of the system can be represented by pure state operators with suitable state vectors, and

(Q4) the probabilities of the measurement outcomes, with arbitrarily high precision, can be reproduced by the usual trace formula of quantum mechanics.

Moreover, in the case of real-valued quantities,

(Q5) each quantity, if we wish, can be associated with a suitable self-adjoint operator, such that

(Q6) in all states of the system, the expectation value of the quantity can be reproduced, with arbitrarily high precision, by the usual trace formula applied to the associated self-adjoint operator,

(Q7) the possible measurement results are the eigenvalues of the operator,

(Q8) and the corresponding outcome events are represented by the eigenspaces pertaining to the eigenvalues respectively, according to the spectral decomposition of the operator in question.

In preparation for our quantum representation theorem, first we prove a lemma, which is a straightforward consequence of previous results in Pitowsky's *Quantum Probability – Quantum Logic.*

Lemma 8. *For each vector* $\vec{f} \in l(M,S)$ *there exists a Hilbert space* $^{(\vec{f})}H$ *and closed subspaces* $^{(\vec{f})}E_i^r$ *in the subspace lattice* $L\left(^{(\vec{f})}H\right)$ *and a pure state* $P_{\Psi_{\vec{f}}}$

with a suitable unit vector $\Psi_{\vec{f}} \in {}^{(\vec{f})}H$*, such that*

$$f_i^r \simeq tr\left(P_{\Psi_{\vec{f}}}{}^{(\vec{f})}E_i^r\right) \tag{99}$$

$$f_{i_1\ldots i_L}^{r_1\ldots r_L} \simeq tr\left(P_{\Psi_{\vec{f}}}\left({}^{(\vec{f})}E_{i_1}^{r_1} \wedge \ldots \wedge {}^{(\vec{f})}E_{i_L}^{r_L}\right)\right) \tag{100}$$

for all ${}_i^r \in I^M$ *and* ${}_{i_1\ldots i_L}^{r_1\ldots r_L} \in S$*.*

Proof. It follows from a straightforward generalization of a theorem in (Pitowsky 1989, p. 65) that the so called quantum polytope $q(M,S)$, constituted by the vectors satisfying (99)–(100) with exact equality, is a dense convex subset of $l(M,S)$; it is essentially $l(M,S)$ save for some points on the boundary of $l(M,S)$, namely the finite number of non-classical vertices. $q(M,S)$ contains the interior of $l(M,S)$. This means that arbitrary vector $\vec{f} \in l(M,S)$ can be regarded as "an element of" $q(M,S)$ *with arbitrary precision.* That is, there exists a Hilbert space ${}^{(\vec{f})}H$ and for each ${}_i^r \in I^M$ a closed subspace/projector ${}^{(\vec{f})}E_i^r$ in the subspace/projector lattice $L\left({}^{(\vec{f})}H\right)$ and a suitable unit vector $\Psi_{\vec{f}} \in {}^{(\vec{f})}H$, such that the approximate equalities (99)–(100) hold. □

Theorem 9. *There exists a Hilbert space* H *and for each outcome event* X_i^r *a closed subspace/projector* E_i^r *in the subspace/projector lattice* $L(H)$*, such that for each state* $\vec{Z}$ *of the system there exists a pure state* $P_{\Psi_{\vec{Z}}}$ *with a suitable unit vector* $\Psi_{\vec{Z}} \in H$*, such that*

$$Z_i^r \simeq tr\left(P_{\Psi_{\vec{Z}}}E_i^r\right) \tag{101}$$

$$Z_{i_1\ldots i_L}^{r_1\ldots r_L} \simeq tr\left(P_{\Psi_{\vec{Z}}}\left(E_{i_1}^{r_1} \wedge \ldots \wedge E_{i_L}^{r_L}\right)\right) \tag{102}$$

and

$$E_i^r \perp E_j^r \qquad i \neq j \tag{103}$$

$$\bigvee_{k=1}^{n_r} E_k^r = H \tag{104}$$

for all ${}_i^r, {}_j^r \in I^M$ *and* ${}_{i_1\ldots i_L}^{r_1\ldots r_L} \in S$*.*

Proof. The proof is essentially based on Lemma 4 and proceeds in two major steps.

Step I

Consider the vertices of $\varphi(M,S)$, $\{\vec{w}_\vartheta\}_{\vartheta\in\Theta}$. Each $\vec{w}_\vartheta$ is a vector in $l(M,S)$. Therefore, due to Lemma 8, for each $\vec{w}_\vartheta$ there exists a Hilbert space ${}^\vartheta\tilde{H}$ and closed subspaces ${}^\vartheta\tilde{E}_i^r$ in the subspace lattice $L\left({}^\vartheta\tilde{H}\right)$ and a pure state $P_{\tilde{\Psi}_\vartheta}$ with a suitable unit vector $\tilde{\Psi}_\vartheta \in {}^\vartheta\tilde{H}$, such that

$$w_{\vartheta}{}_i^r \simeq tr\left(P_{\tilde{\Psi}_\vartheta}{}^\vartheta\tilde{E}_i^r\right) \tag{105}$$

$$w_\vartheta{}^{r_1\ldots r_L}_{i_1\ldots i_L} \simeq tr\left(P_{\tilde{\Psi}_\vartheta}\left({}^\vartheta\tilde{E}^{r_1}_{i_1}\wedge\ldots\wedge{}^\vartheta\tilde{E}^{r_L}_{i_L}\right)\right) \tag{106}$$

for all ${}^r_i \in I^M$ and ${}^{r_1\ldots r_L}_{i_1\ldots i_L} \in S$.

Now, let

$${}^\vartheta H = H^{n_1}\otimes H^{n_2}\otimes\ldots\otimes H^{n_m}\otimes{}^\vartheta\tilde{H} \tag{107}$$

where $H^{n_1}, H^{n_2}, \ldots H^{n_m}$ are Hilbert spaces of dimension $n_1, n_2, \ldots n_m$. Let $e_1^r, e_2^r, \ldots e_{n_r}^r$ be an orthonormal basis in H^{n_r}. Define the corresponding subspace for each event X_i^r as follows:

$${}^\vartheta E_i^r = H^{n_1}\otimes\ldots H^{n_{r-1}}\otimes[e_i^r]\otimes H^{n_{r+1}}\ldots\otimes H^{n_m}\otimes{}^\vartheta\tilde{E}_i^r \tag{108}$$

where $[e_i^r]$ stands for the one-dimensional subspace spanned by e_i^r in H^{n_r}. Notice that, for all r,

$${}^\vartheta E_i^r \perp {}^\vartheta E_j^r \qquad \text{if} \quad i\neq j \tag{109}$$

due to the fact that $e_1^r, e_2^r, \ldots e_{n_r}^r$ is an orthonormal basis in H^{n_r}.

Due to Lemma 4, for all $1\leq r\leq m$ there is exactly one $1\leq {}^\vartheta i_*^r \leq n_r$ such that $w_\vartheta{}^r_{{}^\vartheta i_*^r} = 1$. This makes it possible to define the state vector in ${}^\vartheta H$ as the following unit vector:

$$\Psi_\vartheta = e^1_{{}^\vartheta i_*^1}\otimes e^2_{{}^\vartheta i_*^2}\otimes\ldots\otimes e^r_{{}^\vartheta i_*^r}\otimes\ldots\otimes e^m_{{}^\vartheta i_*^m}\otimes\tilde{\Psi}_\vartheta$$

Now, it is easily verifiable that

$$w_\vartheta{}^r_i \simeq tr\left(P_{\Psi_\vartheta}{}^\vartheta E_i^r\right) \tag{110}$$

$$w_\vartheta{}^{r_1\ldots r_L}_{i_1\ldots i_L} \simeq tr\left(P_{\Psi_\vartheta}\left({}^\vartheta E^{r_1}_{i_1}\wedge\ldots\wedge{}^\vartheta E^{r_L}_{i_L}\right)\right) \tag{111}$$

for all ${}^r_i\in I^M$, ${}^{r_1\ldots r_L}_{i_1\ldots i_L}\in S$, and for all $\vartheta\in\Theta$. For example:
If $w_\vartheta{}^r_i = 1$, and so $i ={}^\vartheta i_*^r$, then

$$tr\left(P_{\Psi_\vartheta}{}^\vartheta E_i^r\right) = \underbrace{tr\left(P_{e^1_{{}^\vartheta i_*^1}}H^{n_1}\right)}_{1} tr\left(P_{e^2_{{}^\vartheta i_*^2}}H^{n_2}\right)\ldots\underbrace{tr\left(P_{e^r_{{}^\vartheta i_*^r}}\left[e^r_{{}^\vartheta i_*^r}\right]\right)}_{1}\ldots$$
$$tr\left(P_{e^m_{{}^\vartheta i_*^m}}H^{n_m}\right)\underbrace{tr\left(P_{\tilde{\Psi}_\vartheta}{}^\vartheta\tilde{E}_i^r\right)}_{\simeq w_\vartheta{}^r_i=1}\simeq 1$$

If $w_\vartheta{}^r_i = 0$, and so $i\neq{}^\vartheta i_*^r$, then

$$tr\left(P_{\Psi_\vartheta}{}^\vartheta E_i^r\right) = \underbrace{tr\left(P_{e^1_{{}^\vartheta i_*^1}}H^{n_1}\right)}_{1} tr\left(P_{e^2_{{}^\vartheta i_*^2}}H^{n_2}\right)\ldots\underbrace{tr\left(P_{e^r_{{}^\vartheta i_*^r}}\left[e^r_{i\neq{}^\vartheta i_*^r}\right]\right)}_{0}\ldots$$
$$tr\left(P_{e^m_{{}^\vartheta i_*^m}}H^{n_m}\right)\underbrace{tr\left(P_{\tilde{\Psi}_\vartheta}{}^\vartheta\tilde{E}_i^r\right)}_{\simeq w_\vartheta{}^r_i=0} = 0$$

Similarly, if ${w_\vartheta}_{i_1}^{r_1}=0$, ${w_\vartheta}_{i_2}^{r_2}=1$ then

$$tr\left(P_{\Psi_\vartheta}\left({}^\vartheta E_{i_1}^{r_1}\wedge{}^\vartheta E_{i_2}^{r_2}\right)\right)=\underbrace{tr\left(P_{e^1_{\vartheta\, i_*^1}}\left(H^{n_1}\wedge H^{n_1}\right)\right)}_{1}\ldots$$

$$\underbrace{tr\left(P_{e^{r_1}_{\vartheta\, i_*^{r_1}}}\left(\left[e^{r_1}_{i_1\neq^\vartheta i_*^{r_1}}\right]\wedge H^{n_{r_1}}\right)\right)}_{0}\ldots\underbrace{tr\left(P_{e^{r_2}_{\vartheta\, i_*^{r_2}}}\left(H^{n_{r_2}}\wedge\left[e^{r_2}_{i_2=^\vartheta i_*^{r_2}}\right]\right)\right)}_{1}\ldots$$

$$tr\left(P_{e^m_{\vartheta\, i_*^m}}\left(H^{n_m}\wedge H^{n_m}\right)\right)\underbrace{tr\left(P_{\tilde\Psi_\vartheta}\left({}^\vartheta\tilde E_{i_1}^{r_1}\wedge{}^\vartheta\tilde E_{i_2}^{r_2}\right)\right)}_{\simeq {w_\vartheta}_{i_1i_2}^{r_1r_2}=0}=0$$

in accordance with that ${}^\vartheta w_{i_1i_2}^{r_1r_2}$ must be equal to 0, due to (38).
If ${w_\vartheta}_{i_1}^{r_1}=1$, ${w_\vartheta}_{i_2}^{r_2}=1$ then

$$tr\left(P_{\Psi_\vartheta}\left({}^\vartheta E_{i_1}^{r_1}\wedge{}^\vartheta E_{i_2}^{r_2}\right)\right)=\underbrace{tr\left(P_{e^1_{\vartheta\, i_*^1}}\left(H^{n_1}\wedge H^{n_1}\right)\right)}_{1}\ldots$$

$$\underbrace{tr\left(P_{e^{r_1}_{\vartheta\, i_*^{r_1}}}\left(\left[e^{r_1}_{i_1=^\vartheta i_*^{r_1}}\right]\wedge H^{n_{r_1}}\right)\right)}_{1}\ldots\underbrace{tr\left(P_{e^{r_2}_{\vartheta\, i_*^{r_2}}}\left(H^{n_{r_2}}\wedge\left[e^{r_2}_{i_2=^\vartheta i_*^{r_2}}\right]\right)\right)}_{1}\ldots$$

$$tr\left(P_{e^m_{\vartheta\, i_*^m}}\left(H^{n_m}\wedge H^{n_m}\right)\right)\underbrace{tr\left(P_{\tilde\Psi_\vartheta}\left({}^\vartheta\tilde E_{i_1}^{r_1}\wedge{}^\vartheta\tilde E_{i_2}^{r_2}\right)\right)}_{\simeq {w_\vartheta}_{i_1i_2}^{r_1r_2}}\simeq {w_\vartheta}_{i_1i_2}^{r_1r_2}$$

Step II

Consider an arbitrary state $\vec{Z}$. Since $\vec{Z}\in\varphi(M,S)$, it can be decomposed in terms of the vertices $\{\vec{w}_\vartheta\}_{\vartheta\in\Theta}$ in the fashion of (70) with some coefficients $\{\lambda_\vartheta\}_{\vartheta\in\Theta}$.

Now we construct the Hilbert space H and the state vector $\Psi_{\vec{Z}}$:

$$H = \bigoplus_{\vartheta\in\Theta}{}^\vartheta H \tag{112}$$

$$\Psi_{\vec{Z}} = \bigoplus_{\vartheta\in\Theta}\sqrt{\lambda_\vartheta}\Psi_\vartheta \tag{113}$$

Obviously,

$$\left\langle\Psi_{\vec{Z}},\Psi_{\vec{Z}}\right\rangle=\sum_{\vartheta\in\Theta}\lambda_\vartheta\left\langle\Psi_\vartheta,\Psi_\vartheta\right\rangle=1$$

The subspaces E_i^r representing the outcome events will be defined further below. First we consider the following subspaces of H:

$${}^*E_i^r = \bigoplus_{\vartheta\in\Theta}{}^\vartheta E_i^r$$

Since

$$\begin{aligned} tr\left(P_{\Psi_{\vec{Z}}} {}^*E_i^r\right) &= \left\langle \Psi_{\vec{Z}}, {}^*E_i^r \Psi_{\vec{Z}} \right\rangle = \sum_{\vartheta\in\Theta} \left\langle \sqrt{\lambda_\vartheta}\Psi_\vartheta, {}^\vartheta E_i^r \sqrt{\lambda_\vartheta}\Psi_\vartheta \right\rangle \\ &= \sum_{\vartheta\in\Theta} \lambda_\vartheta tr\left(P_{\Psi_\vartheta} {}^\vartheta E_i^r\right) \end{aligned}$$

$$\begin{aligned} tr\left(P_{\Psi_{\vec{Z}}}\left({}^*E_{i_1}^{r_1} \wedge \ldots \wedge {}^*E_{i_L}^{r_L}\right)\right) &= \left\langle \Psi_{\vec{Z}}, \left({}^*E_{i_1}^{r_1} \wedge \ldots \wedge {}^*E_{i_L}^{r_L}\right) \Psi_{\vec{Z}} \right\rangle \\ &= \sum_{\vartheta\in\Theta} \left\langle \sqrt{\lambda_\vartheta}\Psi_\vartheta, \left({}^\vartheta E_{i_1}^{r_1} \wedge \ldots \wedge {}^\vartheta E_{i_L}^{r_L}\right) \sqrt{\lambda_\vartheta}\Psi_\vartheta \right\rangle \\ &= \sum_{\vartheta\in\Theta} \lambda_\vartheta tr\left(P_{\Psi_\vartheta}\left({}^\vartheta E_{i_1}^{r_1} \wedge \ldots \wedge {}^\vartheta E_{i_L}^{r_L}\right)\right) \end{aligned}$$

from (110)–(111) and (70) we have

$$Z_i^r \simeq tr\left(P_{\Psi_{\vec{Z}}} {}^*E_i^r\right) \tag{114}$$

$$Z_{i_1 \ldots i_L}^{r_1 \ldots r_L} \simeq tr\left(P_{\Psi_{\vec{Z}}}\left({}^*E_{i_1}^{r_1} \wedge \ldots \wedge {}^*E_{i_L}^{r_L}\right)\right) \tag{115}$$

Also, as direct sum preserves orthogonality, from (109) we have

$${}^*E_i^r \perp {}^*E_j^r \quad \text{if} \quad i \neq j \tag{116}$$

For all $1 \leq r \leq m$, let ${}^*E_{i_0}^r \in \left\{{}^*E_1^r, {}^*E_2^r, \ldots {}^*E_{n_r}^r\right\}$ be arbitrarily chosen, and let ${}^*E_\perp^r = \left(\bigvee_{i=1}^{n_r} {}^*E_i^r\right)^\perp = \bigwedge_{i=1}^{n_r} \left({}^*E_i^r\right)^\perp$. We define the subspaces representing the outcome events as follows:

$$E_i^r = \begin{cases} {}^*E_i^r & i \neq i_0 \\ {}^*E_{i_0}^r \vee {}^*E_\perp^r & i = i_0 \end{cases} \tag{117}$$

Obviously, (116) implies ${}^*E_{i_0}^r \leq \bigwedge_{i\neq i_0} \left({}^*E_i^r\right)^\perp$. Due to the orthomodularity of the subspace lattice $L(H)$, we have

$${}^*E_{i_0}^r \vee \left(\underbrace{\left({}^*E_{i_0}^r\right)^\perp \wedge \left(\bigwedge_{i\neq i_0} \left({}^*E_i^r\right)^\perp\right)}_{{}^*E_\perp^r}\right) = \bigwedge_{i\neq i_0} \left({}^*E_i^r\right)^\perp$$

meaning that

$$E_{i_0}^r = \bigwedge_{i\neq i_0} \left({}^*E_i^r\right)^\perp$$

Therefore, taking into account (116) and (117),

$$E_i^r \perp E_j^r \quad \text{if} \quad i \neq j \tag{118}$$

Also, it is obviously true that

$$\bigvee_{i=1}^{n_r} E_i^r = H \tag{119}$$

Both (118) and (119) hold for all $1 \leq r \leq m$. There remains to show (101)–(102).

It follows from (117) that

$$E_i^r \geq {}^*E_i^r$$

for all ${}^r_i \in I^M$. Similarly,

$$E_{i_1}^{r_1} \wedge \ldots \wedge E_{i_L}^{r_L} \geq {}^*E_{i_1}^{r_1} \wedge \ldots \wedge {}^*E_{i_L}^{r_L}$$

for all ${}^{r_1 \ldots r_L}_{i_1 \ldots i_L} \in S$. Therefore, for all $\vec{Z} \in \varphi(M, S)$,

$$\left\langle \Psi_{\vec{Z}}, E_i^r \Psi_{\vec{Z}} \right\rangle \geq \left\langle \Psi_{\vec{Z}}, {}^*E_i^r \Psi_{\vec{Z}} \right\rangle \tag{120}$$

and

$$\left\langle \Psi_{\vec{Z}}, E_{i_1}^{r_1} \wedge \ldots \wedge E_{i_L}^{r_L} \Psi_{\vec{Z}} \right\rangle \geq \left\langle \Psi_{\vec{Z}}, {}^*E_{i_1}^{r_1} \wedge \ldots \wedge {}^*E_{i_L}^{r_L} \Psi_{\vec{Z}} \right\rangle \tag{121}$$

Now, (11) and (114) imply that

$$\sum_{\substack{i \\ \left({}^r_i \in I^M\right)}} \left\langle \Psi_{\vec{Z}}, {}^*E_i^r \Psi_{\vec{Z}} \right\rangle \simeq 1$$

At the same time, taking into account (118)–(119), we have

$$1 = \sum_{\substack{i \\ \left({}^r_i \in I^M\right)}} \left\langle \Psi_{\vec{Z}}, E_i^r \Psi_{\vec{Z}} \right\rangle \geq \sum_{\substack{i \\ \left({}^r_i \in I^M\right)}} \left\langle \Psi_{\vec{Z}}, {}^*E_i^r \Psi_{\vec{Z}} \right\rangle \simeq 1 \tag{122}$$

From (120) and (122), therefore,

$$tr\left(P_{\Psi_{\vec{Z}}} E_i^r\right) \simeq tr\left(P_{\Psi_{\vec{Z}}} {}^*E_i^r\right) \tag{123}$$

Similarly, on the one hand, (12) and (115) imply that

$$\sum_{\substack{i_1, i_2 \ldots i_L \\ \left({}^{r_1 \ldots r_L}_{i_1 \ldots i_L} \in S\right)}} \left\langle \Psi_{\vec{Z}}, {}^*E_{i_1}^{r_1} \wedge \ldots \wedge {}^*E_{i_L}^{r_L} \Psi_{\vec{Z}} \right\rangle \simeq 1 \tag{124}$$

On the other hand, $\left\{E_{i_1}^{r_1} \wedge \ldots \wedge E_{i_L}^{r_L}\right\}_{\substack{i_1, i_2 \ldots i_L \\ \left({}^{r_1 \ldots r_L}_{i_1 \ldots i_L} \in S\right)}}$ is an orthogonal system of subspaces. Therefore,

$$1 \geq \sum_{\substack{i_1, i_2 \ldots i_L \\ \left({}^{r_1 \ldots r_L}_{i_1 \ldots i_L} \in S\right)}} \left\langle \Psi_{\vec{Z}}, E_{i_1}^{r_1} \wedge \ldots \wedge E_{i_L}^{r_L} \Psi_{\vec{Z}} \right\rangle$$

$$\geq \sum_{\substack{i_1, i_2 \ldots i_L \\ \left(\substack{r_1 \ldots r_L \\ i_1 \ldots i_L} \in S\right)}} \left\langle \Psi_{\vec{Z}}, {}^*E_{i_1}^{r_1} \wedge \ldots \wedge {}^*E_{i_L}^{r_L} \Psi_{\vec{Z}} \right\rangle \simeq 1 \tag{125}$$

From (121) and (125), for all $\substack{r_1 \ldots r_L \\ i_1 \ldots i_L} \in S$, we have

$$tr\left(P_{\Psi_{\vec{Z}}} E_{i_1}^{r_1} \wedge \ldots \wedge E_{i_L}^{r_L}\right) \simeq tr\left(P_{\Psi_{\vec{Z}}} {}^*E_{i_1}^{r_1} \wedge \ldots \wedge {}^*E_{i_L}^{r_L}\right) \tag{126}$$

Thus, (114)–(115) together with (123) and (126) imply (101)–(102). □

With Theorem 9 we have accomplished (Q1)–(Q4). The next two theorems cover statements (Q5)–(Q8).

Theorem 10. *Let a_r be the measurement of a real valued quantity with labeling (13). On the Hilbert space H, there exists a self-adjoint operator A_r such that for every state of the system $\vec{Z}$,*

$$\langle \alpha_r \rangle_{\vec{Z}} \simeq tr\left(P_{\Psi_{\vec{Z}}} A_r\right) \tag{127}$$

Proof. Let

$$A_r = \sum_{i=1}^{n_r} \alpha_i^r E_i^r \tag{128}$$

A_r is obviously a self-adjoint operator, and

$$\begin{aligned} \langle \alpha_r \rangle_{\vec{Z}} &= \sum_{i=1}^{n_r} \alpha_i^r p\left(X_i^r | a_r\right) = \sum_{i=1}^{n_r} \alpha_i^r Z_i^r \simeq \sum_{i=1}^{n_r} \alpha_i^r tr\left(P_{\Psi_{\vec{Z}}} E_i^r\right) \\ &= tr\left(P_{\Psi_{\vec{Z}}} \sum_{i=1}^{n_r} \alpha_i^r E_i^r\right) = tr\left(P_{\Psi_{\vec{Z}}} A_r\right) \end{aligned}$$

□

Theorem 11. *The possible measurement results of the α_r-measurement are exactly the eigenvalues of the associated operator A_r. The subspace E_i^r representing the outcome event labeled by α_i^r is the eigenspace pertaining to eigenvalue α_i^r. Accordingly, (128) constitutes the spectral decomposition of A_r.*

Proof. First, let $\psi \in E_i^r$. Then, due to (103), $A_r\psi = \left(\sum_{i=1}^{n_r} \alpha_i^r E_i^r\right)\psi = \alpha_i^r \psi$. Meaning that every α_i^r is an eigenvalue of A_r. Now consider an arbitrary eigenvector of A_r, that is, a vector $\psi \in H$ such that

$$A_r \psi = x\psi \tag{129}$$

with some $x \in \mathbb{R}$. Due to (103)–(104), $\left\{E_1^r, E_2^r, \ldots E_{n_r}^r\right\}$ constitutes an orthogonal decomposition of H, meaning that arbitrary $\psi \in H$ can be decomposed as

$$\psi = \sum_{i=1}^{n_r} \psi_i \quad \psi_i \in E_i^r$$

From (128) we have

$$\sum_{i=1}^{n_r} \alpha_i^r \psi_i = \sum_{i=1}^{n_r} x\psi_i \tag{130}$$

About the labeling we have assumed that $\alpha_i^r \neq \alpha_j^r$ for $i \neq j$, therefore, (130) implies that

$$\begin{aligned} x &= \alpha_i^r \quad \text{for one } \alpha_i^r \\ \psi_j &= 0 \quad \text{for all } j \neq i \end{aligned}$$

that is, $\psi \in E_i^r$. Meaning that (128) is the spectral decomposition of A_r. □

A consequence of Theorems 10 and 11 is that if $f : \mathbb{R} \to \mathbb{R}$ is an arbitrary injective function "re-labeling" the outcomes, then

$$\begin{aligned} \langle f(\alpha_r)\rangle_{\vec{Z}} &= \sum_{i=1}^{n_r} f(\alpha_i^r) p\left(X_i^r | a_r\right) = \sum_{i=1}^{n_r} f(\alpha_i^r) Z_i^r \simeq \sum_{i=1}^{n_r} f(\alpha_i^r)\, tr\left(P_{\Psi_{\vec{Z}}} E_i^r\right) \\ &= tr\left(P_{\Psi_{\vec{Z}}} \sum_{i=1}^{n_r} f(\alpha_i^r) E_i^r\right) = tr\left(P_{\Psi_{\vec{Z}}} f(A_r)\right) \end{aligned}$$

7 Representation of Dynamics

Notice that not all unit vectors of H are involved in the representation of states. In order to specify the ones being involved, consider the following subspace $\mathcal{H} \subset H$:

$$\mathcal{H} = \mathsf{span}\left\{\Psi_\vartheta\right\}_{\vartheta\in\Theta}$$

where $\left\{\Psi_\vartheta\right\}_{\vartheta\in\Theta}$ is the set of vectors in the direct sum (113), understood as being pairwise orthogonal, unit-length elements of H. Denote by $\mathcal{O}$ the closed first hyperoctant (orthant) of the $(|\Theta| - 1)$-dimensional sphere of unit vectors in $\mathcal{H}$:

$$\mathcal{O} = \left\{ \sum_{\vartheta\in\Theta} o_\vartheta \Psi_\vartheta \,\middle|\, o_\vartheta \geq 0 \quad \sum_{\vartheta\in\Theta} o_\vartheta^2 = 1 \right\}$$

Obviously, there is a continuous one-to-one map between the Λ and $\mathcal{O}$:

$$O : \Lambda \to \mathcal{O};\ O\left(\vec{\lambda}\right) = \sum_{\vartheta\in\Theta} \sqrt{\lambda_\vartheta}\Psi_\vartheta$$

As we have shown, however, the states of the system actually are represented on the $\dim\left(\varphi\left(M, S\right)\right)$-dimensional slice $\Lambda_\sigma \subset \Lambda$ (see (82)). Accordingly, the quantum mechanical representation of states constitutes a $\dim\left(\varphi\left(M, S\right)\right)$-dimensional submanifold with boundary: $\mathcal{O}_\sigma = O\left(\Lambda_\sigma\right) \subset \mathcal{O}$.

Consequently, the time evolution of state $\vec{Z}\left(t\right)$ will be represented by a path in $\mathcal{O}_\sigma$:

$$\Psi(t) = O \circ \sigma\left(\vec{Z}\left(t\right)\right)$$

The representation is of course not unique, as it depends on the choice of cross section σ. This is however inessential; just like a choice of a coordinate system.

As emphasized at the beginning of Section 6, the quantum representation was derived exclusively from the assumptions **(E1)-(E3)**. If in addition **(E4)** holds, that is the time evolution $\vec{Z}(t)$ can be generated by a one-parameter group of transformations on $\varphi(M,S)$, $\vec{Z}(t) = F_{t-t_0}\left(\vec{Z}(t_0)\right)$, then the same is true for $\mathcal{O}_\sigma$. Let $G_t = O \circ \sigma \circ F_t \circ D \circ O^{-1}$. Obviously, G_t is a map $\mathcal{O}_\sigma \to \mathcal{O}_\sigma$, such that

$G_t : \mathcal{O}_\sigma \to \mathcal{O}_\sigma$ is one-to-one

$G : \mathbb{R} \times \mathcal{O}_\sigma \to \mathcal{O}_\sigma$; $(t,\Psi) \mapsto G_t(\Psi)$ is continuous

$G_{t+s} = G_s \circ G_t$

$G_{-t} = G_t^{-1}$; consequently, $G_0 = id_{\mathcal{O}_\sigma}$

and the time evolution of an arbitrary initial state $\Psi(t_0) \in \mathcal{O}_\sigma$ is $\Psi(t) = G_{t-t_0}(\Psi(t_0))$.

8 Questionable and Unquestionable in Quantum Mechanics

What we have *proved* in the above theorems, that is, statements (Q1)–(Q8), are nothing but the basic postulates of quantum theory. This means that the basic postulates of quantum theory are in fact analytic statements: they do not tell us anything about a physical system beyond the fact that the system can be described in empirical/operational terms—even if this logical relationship is not so evident. In this sense, of course, these postulates of quantum theory are unquestionable. Though, as we have seen, the Hilbert space quantum mechanical formalism is only an optional mathematical representation of the probabilistic behavior of a system—empirical facts do not necessitate this choice.

Nevertheless, it must be mentioned that the quantum-mechanics-like representation, characterized by (Q1)–(Q8), is not completely identical with standard quantum mechanics. There are several subtle deviations:

(D1) There is no one-to-one correspondence between operationally meaningful physical quantities and self-adjoint operators. First of all, it is not necessarily true that every self-adjoint operator represents some operationally meaningful quantity.

(D2) There is no necessary connection between commutation of the associated self-adjoint operators and joint measurability of the corresponding physical quantities. In general, there is no obvious role of the mathematically definable algebraic structures over the self-adjoint operators in the operational context. First of all because those mathematically "natural" structures are mostly meaningless in an operational sense. As we have already mentioned, the outcome *events* are ontologically prior to the labeling of the outcomes by means of numbers; and the events themselves are well represented in the subspace/projector lattice, prior to any self-adjoint operator associated with a numerical coordination.

For example, consider three real-valued physical quantities with labelings $\alpha_{r_1}, \alpha_{r_2}, \alpha_{r_3}$. The three physical quantities reflect three different features of the system defined by three different measurement operations. A functional relationship $\alpha_{r_1} = f\left(\alpha_{r_2}, \alpha_{r_3}\right)$ means that whenever we perform the measurements $a_{r_1}, a_{r_2}, a_{r_3}$ in conjunction (meaning that $\{r_1, r_2, r_3\} \in \mathfrak{P}$) the outcomes $X_{i_1}^{r_1}, X_{i_2}^{r_2}, X_{i_3}^{r_3}$ are strongly correlated: if $X_{i_2}^{r_2}$ and $X_{i_3}^{r_3}$ are the outcomes of a_{r_2} and a_{r_3}, labeled by $\alpha_{i_2}^{r_2}$ and $\alpha_{i_3}^{r_3}$, then the outcome of measurement a_{r_1}, $X_{i_1}^{r_1}$, is the one labeled by $\alpha_{i_1}^{r_1} = f\left(\alpha_{i_2}^{r_2}, \alpha_{i_3}^{r_3}\right)$. That is, in probabilistic terms:

$$\begin{aligned} p\left(\alpha_{r_1}^{-1}\left(f\left(\alpha_{i_2}^{r_2}, \alpha_{i_3}^{r_3}\right)\right) \wedge \alpha_{r_2}^{-1}\left(\alpha_{i_2}^{r_2}\right) \wedge \alpha_{r_3}^{-1}\left(\alpha_{i_3}^{r_3}\right) | a_{r_1} \wedge a_{r_2} \wedge a_{r_3}\right) \\ = p\left(\alpha_{r_2}^{-1}\left(\alpha_{i_2}^{r_2}\right) \wedge \alpha_{r_3}^{-1}\left(\alpha_{i_3}^{r_3}\right) | a_{r_1} \wedge a_{r_2} \wedge a_{r_3}\right) \end{aligned} \tag{131}$$

This contingent fact of regularity in the observed relative frequencies of physical events is what is a part of the ontology. And it is well reflected in our quantum mechanical representation, in spite of the fact that the relationship (131) is generally not reflected in some algebraic or other functional relation of the associated self-adjoint operators A_{r_1}, A_{r_2} and A_{r_3}.

The fact that in our quantum-mechanics-like representation there is no correspondence between commutation and co-measurability explains why there is no need to require satisfaction of the Cirel'son inequalities (cf. Cirenl'son 1980; Popescu and Rohrlich 1994; Müller 2021), which would mean further restriction on the observed relative frequencies beyond **(E1)**–**(E3)**.

(D3) It is worthwhile emphasizing that the Hilbert space of representation is finite dimensional and real. It is of course no problem to embed the whole representation into a complex Hilbert space of the same dimension. As it follows from (103) and (107), the required minimal dimension increases with increasing the number of possible measurements m, and/or increasing the number of possible outcomes n_r. In any event, it is finite until we have a finite operational setup. Employing complex Hilbert spaces is only necessary if, in addition to the stipulated operational setup, we have some further algebraic requirements, for example, in the form of commutation relations, and the likes. How those further requirements are justified in operational terms, of course, can be a question.

(D4) There is no problem with the empirical meaning of the lattice-theoretic meet of subspaces/projectors representing outcome events: the meet represents the empirically meaningful conjunction of the outcome events, regardless whether the corresponding projectors commute or not. Of course, by definition (18), the conjunctions that do not belong to S have zero probability in all states of the system.

In contrast, the lattice-theoretic joins and orthocomplements, in general, have nothing to do with the disjunctions and negations of the outcome events. Nevertheless, as we have seen, the quantum state uniquely determines the probabilities on the whole event algebra, including the conjunctions, disjunctions and negations of all events—in the sense of Theorem 1.

(D5) All possible states of the system, $\vec{Z} \in \varphi(M,S)$, are represented by *pure* states. That is to say, the quantum mechanical notion of mixed state is not needed. The reason is very simple. $\varphi(M,S)$ is a convex polytope being closed under convex linear combinations. The state of the system intended to be represented by a mixed state, say,

$$W = \mu_1 P_{\Psi_{\vec{Z}_1}} + \mu_2 P_{\Psi_{\vec{Z}_2}} \qquad \mu_1, \mu_2 \geq 0; \;\; \mu_1 + \mu_2 = 1$$

is nothing but another element of $\varphi(M,S)$,

$$\vec{Z}_3 = \mu_1 \vec{Z}_1 + \mu_2 \vec{Z}_2 \in \varphi(M,S)$$

However, in our representation theorem (Theorem 9) the Hilbert space and the representations of the outcome events are constructed in a way that all states $\vec{Z} \in \varphi(M,S)$ are represented by a suitable state vector in one and the same Hilbert space. Therefore, $\vec{Z}_3$ is also represented by a pure state $P_{\Psi_{\vec{Z}_3}}$ with a suitably constructed state vector $\Psi_{\vec{Z}_3}$. Namely, given that

$$\begin{aligned} \vec{Z}_1 &= \sum_{\vartheta\in\Theta} \lambda^1_\vartheta \vec{w}_\vartheta \quad \lambda^1_\vartheta \geq 0, \sum_{\vartheta\in\Theta} \lambda^1_\vartheta = 1 \\ \vec{Z}_2 &= \sum_{\vartheta\in\Theta} \lambda^2_\vartheta \vec{w}_\vartheta \quad \lambda^2_\vartheta \geq 0, \sum_{\vartheta\in\Theta} \lambda^2_\vartheta = 1 \end{aligned}$$

we have

$$\vec{Z}_3 = \sum_{\vartheta\in\Theta} \left(\mu_1 \lambda^1_\vartheta + \mu_2 \lambda^2_\vartheta\right) \vec{w}_\vartheta$$

therefore, from (113),

$$\Psi_{\vec{Z}_3} = \bigoplus_{\vartheta\in\Theta} \sqrt{\mu_1 \lambda^1_\vartheta + \mu_2 \lambda^2_\vartheta}\, \Psi_\vartheta$$

To avoid a possible misunderstanding, it is worthwhile mentioning that all we said above is not in contradiction with the mathematical fact that the density operators W and $P_{\Psi_{\vec{Z}_3}}$ generate different "quantum probability" measures over the *whole* subspace lattice $L(H)$. The two measures will coincide on those elements of $L(H)$ that represent operationally meaningful events—$E^r_i, E^{r_1}_{i_1} \wedge \ldots \wedge E^{r_L}_{i_L}$ for ${}^{r}_{i} \in I^M, {}^{r_1 \ldots r_L}_{i_1 \ldots i_L} \in S$. This reinforces the point in (D4) that there is no one-to-one correspondence between the operationally meaningful events and the elements of $L(H)$.

(D6) We don't need to invoke the entire Hilbert space for representing the totality of operationally meaningful possible states of the system; subspace $\mathcal{H}$ is sufficient. Even in this restricted sense, there is no one-to-one correspondence between the rays of the subspace $\mathcal{H} \subset H$ and the states of the system. The unit vectors involved in the representation are the ones pointing to $\mathcal{O}_\sigma$, a $\dim(\varphi(M,S))$-dimensional submanifold with boundary on the unit sphere of $\mathcal{H}$.

(D7) The so called "superposition principle" does not hold. The ray determined by the linear combination of two different vectors pointing to $\mathcal{O}_\sigma$ does not necessarily intersect $\mathcal{O}_\sigma$; meaning that such a linear combination, in general, has nothing to do with a third state of the system. Neither has it anything to do with the logical/probability theoretic notion of "disjunction" of events, of course. Nevertheless, as we have already emphasized in (D4) and (D5), all possible states of the system are well represented in $\mathcal{O}_\sigma$; and these states uniquely determine the probabilities on the whole event algebra of operationally meaningful events, including their disjunctions too.

(D8) The dynamics of the system can be well represented in the usual way, by means of temporal evolution on the state manifold $\mathcal{O}_\sigma$. In the case where (E4) is also satisfied, this temporal evolution can be generated by a one-parameter group of transformations of $\mathcal{O}_\sigma$. However, these transformations are in no way related to the unitary transformations of H (or $\mathcal{H}$); because they do not respect the linear structure of the Hilbert space or orthogonality; but they do respect that the state space $\mathcal{O}_\sigma$ is a manifold with boundary.

It is remarkable that most of the above mentioned deviations from the quantum mechanical folklore are related with exactly those issues in the foundations of quantum mechanics that have been hotly debated for long decades (e.g. Strauss 1936; Reichenbach 1944; Popper 1967; Park and Margenau 1968; 1971; Ross 1974; Bell 1987; Gudder 1988; Malament 1992; Leggett 1998; Griffiths 2013; Cassinelli and Lahti 2017; Fröhlich and Pizzo 2022). The fact that so much of the core of quantum theory can be unquestionably deduced from three elementary empirical conditions, equally true about all physical systems whether classical or quantum, or beyond, may shed new light on these old problems in the foundations.

Appendices

Appendix 1

The following theorem is mentioned as an exercise in most texts. We formulate it using the same notation we used in the proof of Lemma 4.

Theorem 12. *Let P be a polytope in $\mathbb{R}^d$, defined by the following set of linear inequalities:*

$$\left\langle \vec{\omega}_\mu, \vec{f} \right\rangle - b_\mu \quad \leq \quad 0 \quad \textit{for all } \mu \in I \tag{132}$$

For each $\vec{f} \in P$, define the active index set:

$$I_{\vec{f}} := \left\{ \mu \in I \,\middle|\, \langle \vec{\omega}_\mu, \vec{f} \rangle - b_\mu = 0 \right\}$$

$\vec{f} \in P$ is a vertex of P if and only if

$$\text{span}\left\{\vec{\omega}_\mu\right\}_{\mu \in I_{\vec{f}}} = \mathbb{R}^d \tag{133}$$

Proof. First, suppose $\vec{f}$ is vertex of P, but $\text{span}\left\{\vec{\omega}_\mu\right\}_{\mu \in I_{\vec{f}}} \neq \mathbb{R}^d$. Then choose a non-zero $\vec{g} \in \left(\text{span}\left\{\vec{\omega}_\mu\right\}_{\mu \in I_{\vec{f}}}\right)^\perp$. Obviously, if $\mu \notin I_{\vec{f}}$ then there exists a neighborhood U of $\vec{f}$ such that $\mu \notin I_{\vec{f}_*}$ for all $\vec{f}_* \in U$. Consider the points $\vec{f} + \lambda\vec{g}$. If λ is small enough, both $\vec{f} + \lambda\vec{g}$ and $\vec{f} - \lambda\vec{g}$ are in P, since (132) are satisfied. Now, we can write

$$\vec{f} = \frac{1}{2}\left(\left(\vec{f} + \lambda\vec{g}\right) + \left(\vec{f} - \lambda\vec{g}\right)\right)$$

which contradicts the fact that $\vec{f}$ is vertex of P.

Second, now suppose that $\vec{f} \in P$ and $\text{span}\left\{\vec{\omega}_\mu\right\}_{\mu \in I_{\vec{f}}} = \mathbb{R}^d$. Suppose $\vec{f} = \lambda\vec{f}_* + (1-\lambda)\vec{f}_{**}$ with some $\vec{f}_*, \vec{f}_{**} \in P$ and $0 < \lambda < 1$. We know that $\mu \in I_{\vec{f}}$ implies

$$\left\langle \vec{\omega}_\mu, \vec{f} \right\rangle = \lambda \left\langle \vec{\omega}_\mu, \vec{f}_* \right\rangle + (1-\lambda)\left\langle \vec{\omega}_\mu, \vec{f}_{**} \right\rangle = b_\mu$$

On the other hand, from (132) we have

$$\begin{aligned} \left\langle \vec{\omega}_\mu, \vec{f}_* \right\rangle &\leq b_\mu \\ \left\langle \vec{\omega}_\mu, \vec{f}_{**} \right\rangle &\leq b_\mu \end{aligned}$$

which implies that $\left\langle \vec{\omega}_\mu, \vec{f} \right\rangle = \left\langle \vec{\omega}_\mu, \vec{f}_* \right\rangle = \left\langle \vec{\omega}_\mu, \vec{f}_{**} \right\rangle$ (for all $\mu \in I_{\vec{f}}$). Therefore,

$$\left(\vec{f} - \vec{f}_*\right), \left(\vec{f} - \vec{f}_{**}\right) \in \left(\text{span}\left\{\vec{\omega}_\mu\right\}_{\mu \in I_{\vec{f}}}\right)^\perp = \emptyset$$

meaning that $\vec{f} = \vec{f}_* = \vec{f}_{**}$. Therefore, $\vec{f}$ is a vertex. □

Appendix 2

The standard argumentation in the GPT literature (Hardy 2008; Holevo 2011, pp. 4–5; Müller 2021, p. 14) that the space of physically possible states Φ *must* be convex is based on a problematic notion of "statistical mixture of preparation procedures". A "mixture" of two preparation procedures, resulting in states $\vec{Z}_1$ and $\vec{Z}_2$, would be a procedure in which we alternate between two preparations in some ratio of λ_1 and λ_2 ($\lambda_1+\lambda_2 = 1$), say randomly, with probabilities λ_1 and λ_2. Under some circumstances—excluding any tricky correlations, for example, between the choice of preparations and the choice of measurements, or between the system's behavior in one run and its preparation in the previous run, etc.—the surface observed statistics would indeed be as if the system were in state $\lambda_1\vec{Z}_1 + \lambda_2\vec{Z}_2$. "The situation described above can be considered as a special way of state preparation" (Holevo 2011, p. 5), the argument says, therefore $\lambda_1\vec{Z}_1 + \lambda_2\vec{Z}_2 \in \mathit{\Phi}$.

We believe that this argument is conceptually flawed. After all, the notion of relative frequency itself is based on a series of measurements that are performed in the same probabilistic setup, and not in a setup changing from one run to the next. Taking into account what the probabilistic setup is that is fixed in the "mixing" procedure, the "mixture" $\lambda_1\vec{Z}_1 + \lambda_2\vec{Z}_2$ is not a state of the original system, but a state of the composed system consisting of the preparation device (denote it by $\mathscr{D}$) and the original physical system (denote it by $\mathscr{S}$). We should not be mislead by the fact that different physical things can be described by the same mathematical object. In order to avoid the confusion it is better to use different notation for the physically different things. When we are talking about the mixture, we are talking about states of the composed ($\mathscr{D}$+$\mathscr{S}$) system:

${}^{(\mathscr{D}+\mathscr{S})}\vec{Z}_1 \in {}^{(\mathscr{D}+\mathscr{S})}\mathit{\Phi}$: the state of ($\mathscr{D}$+$\mathscr{S}$) in which the preparation device $\mathscr{D}$ has propensity 1 to set system $\mathscr{S}$ into state $\vec{Z}_1 \in \mathit{\Phi}$

${}^{(\mathscr{D}+\mathscr{S})}\vec{Z}_2 \in {}^{(\mathscr{D}+\mathscr{S})}\mathit{\Phi}$: the state of ($\mathscr{D}$+$\mathscr{S}$) in which the preparation device $\mathscr{D}$ has propensity 1 to set system $\mathscr{S}$ into state $\vec{Z}_2 \in \mathit{\Phi}$

${}^{(\mathscr{D}+\mathscr{S})}\vec{Z}_3 \in {}^{(\mathscr{D}+\mathscr{S})}\mathit{\Phi}$: the state of ($\mathscr{D}$+$\mathscr{S}$) in which the preparation device $\mathscr{D}$ has propensity λ_1 to set system $\mathscr{S}$ into state $\vec{Z}_1 \in \mathit{\Phi}$ and propensity λ_2 to set system $\mathscr{S}$ into state $\vec{Z}_2 \in \mathit{\Phi}$

Indeed,

$${}^{(\mathscr{D}+\mathscr{S})}\vec{Z}_3 = \lambda_1{}^{(\mathscr{D}+\mathscr{S})}\vec{Z}_1 + \lambda_2{}^{(\mathscr{D}+\mathscr{S})}\vec{Z}_2$$

and, if the preparation device is such that propensities λ_1 and λ_2 ($\lambda_1 + \lambda_2 = 1$) can be arbitrary, then ${}^{(\mathscr{D}+\mathscr{S})}\mathit{\Phi}$ is closed under convex combination.

Just take our Biased Coin example we discussed at the end of section 3.1. Consider a composed ($\mathscr{D}$+$\mathscr{S}$) system, where $\mathscr{S}$ is the coin on the right hand side of Fig. 1, and $\mathscr{D}$ is a device setting the disk inside the coin into one of the possible positions Z_H, Z_F, or Z_T with some probabilities λ_H, λ_F, and λ_T. Consider only one possible measurement, tossing the coin, with two possible outcomes, Heads and Tails, and one conjunction. Now, keeping the preparation of the ($\mathscr{D}$+$\mathscr{S}$) system fixed, that is, the coin always has the same three slots where the disk can be clicked and the preparation device always has the same propensities λ_H, λ_F, and λ_T, we can take the statistics of Heads and Tails. The observed data will satisfy **(E1)**–**(E3)**, and the state space $\varphi(M,S)$ (where $M=2$ and $S=\left\{\begin{smallmatrix}1&1\\1&2\end{smallmatrix}\right\}$) is the one-dimensional polytope in $\mathbb{R}^3$, determined by vertices $(1,0,0)$ and $(0,1,0)$—just as it was the case in our coin example. Denote an arbitrary state vector of the $(\mathscr{D}+\mathscr{S})$ system by ${}^{(\mathscr{D}+\mathscr{S})}\vec{Z}=\left({}^{(\mathscr{D}+\mathscr{S})}z_H,{}^{(\mathscr{D}+\mathscr{S})}z_T,0\right)\in\varphi(M,S)$, meaning that whenever we perform the coin-toss and the $(\mathscr{D}+\mathscr{S})$ system is in state ${}^{(\mathscr{D}+\mathscr{S})}\vec{Z}$, we get Heads with relative frequency ${}^{(\mathscr{D}+\mathscr{S})}z_H$, Tails with ${}^{(\mathscr{D}+\mathscr{S})}z_T$, but never the Heads and Tails in conjunction. Assuming that the device $\mathscr{D}$ is such that its propensities λ_H,λ_F, and λ_T can be arbitrary ($\lambda_H+\lambda_F+\lambda_T=1$), the set of the physically possible states of the $(\mathscr{D}+\mathscr{S})$ system, ${}^{(\mathscr{D}+\mathscr{S})}\Phi\subset\varphi(M,S)$, is the line segment between $(0.8,0.2,0)$ and $(0.2,0.8,0)$. Hence ${}^{(\mathscr{D}+\mathscr{S})}\Phi$ is closed under convex combination. While, recall (69), the set of the physically possible states of the coin, the original system $\mathscr{S}$ in itself, is

$$\Phi=\{(0.8,0.2,0),(0.5,0.5,0),(0.2,0.8,0)\}$$

which is not closed under convex combination.

In contrast, consider an $(\mathscr{D}+\mathscr{S})$ system such that $\mathscr{S}$ is a coin like the one on the left hand side of Fig. 1, but the preparation device $\mathscr{D}$ has only two possible propensity-states: placing the disk in position Z_H with probability 1, or placing the disk in position Z_T with probability 1. In this case, the set of the physically possible states of the $(\mathscr{D}+\mathscr{S})$ system consists of only two points,

$${}^{(\mathscr{D}+\mathscr{S})}\Phi=\{(0.8,0.2,0),(0.2,0.8,0)\}$$

hence it is not closed under convex combination. While, the set of the physically possible states of the coin in itself, Φ, is the whole line segment between $(0.8,0.2,0)$ and $(0.2,0.8,0)$, which is closed under convex combination.

To sum up, the "statistical mixture of preparation procedures" is a misleading conception. It actually changes the notion of the system in question from the original $\mathscr{S}$ to a composed system $(\mathscr{D}+\mathscr{S})$. And even in this sense it involves unjustified a priori assumptions, without regard to the actual physical properties of the preparation device.

Appendix 3

In the GPT literature, in order to achieve the dynamics be linear, the following three assumptions are made (Hardy 2008, Appendix 1). In our own notations:

(a) The time-evolution map F_t preserves convex combination.

(b) The set of states contains the null vector.

(c) The time evolution maps the null vector into itself.

Notice that assumptions (b) and (c) directly contradict our assumed empirical facts (11)–(12) in **(E2)**; $\varphi(M,S)$ does not contain the null vector. It is essential to clarify this contradiction. First of all, it should be noted that these assumptions are based on conceptually untenable claims, for example that the null vector is that state of the system "when the system is not present" (Hardy 2008,

p. 5). But how do we imagine a measurement on a physical system that is not present? Fortunately, however, assumptions (b) and (c) are not necessary for the time-evolution F_t be linear. Indeed, assumption (a) implies in itself that F_t is an $\mathbb{R}^{M+|S|} \to \mathbb{R}^{M+|S|}$ affine transformation restricted to the convex set of states (see Meyer and Kay 1973, Theorem 4). F_t is linear if and only if it satisfies (a) and

(d) its affine extension on $\mathbb{R}^{M+|S|}$ preserves the null vector.

And condition (d) does not require the null vector to be in the set of states.

In our view, of course, the fulfillment of condition (d) is a matter of further empirical information, beyond the stipulated empirical facts **(E1)**–**(E4)**. In any case, let us come to hypothesis (a) itself. How is it that assumption (a) does not hold even in an example as simple and physically completely plausible as the one with the coin discussed in section 4? In our view, the reason is that the usual argumentation (see Müller 2021, pp. 20-21) in favor of assumption (a) is conceptually flawed, as it operates with the same problematic notion of "statistical mixture of preparation procedures" that we have already criticized in Appendix 2. The argument goes as follows. Consider the following two procedures:

(i) The preparation device prepares the state $\vec{Z}_i$ with probability λ_i. Take the "statistical mixture of these states $\sum_i \lambda_i \vec{Z}_i$"; then we let this "mixed state" evolve in time, resulting in some final state $\vec{Z}'$.

(ii) The preparation device prepares the state $\vec{Z}_i$ with probability λ_i. We let each $\vec{Z}_i$ evolve in time into $F_t\vec{Z}_i$. Finally, we take the "statistical mixture of states $F_t\vec{Z}_i$ with the same λ_i-s, $\vec{Z}'' = \sum_i \lambda_i F_t \vec{Z}_i$

"Clearly, (i) and (ii) are different descriptions of one and the same laboratory procedure; they must hence result in the exact same statistics of any measurements that we may decide to perform in the end, and therefore lead to the same final state[.]" (Müller 2021, p. 21) That is,

$$\vec{Z}' = \vec{Z}'' = \sum_i \lambda_i F_t \vec{Z}_i \tag{134}$$

So far, with some clarification below, we agree with the argument. As, indeed, if the "time-evolution" of the "mixed state $\sum_i \lambda_i \vec{Z}_i$" makes sense at all, it probably means the same process as described in (ii). Where we disagree is the next claim: "But this implies that" (*ibid.)*

$$F_t\left(\sum_i \lambda_i \vec{Z}_i\right) = \sum_i \lambda_i F_t \vec{Z}_i \tag{135}$$

Because F_t has nothing to do with a "mixed state." As we pointed out in Appendix 2, a "mixed state $\sum_i \lambda_i \vec{Z}_i$" is not the state of the system, conceptually, but the state of the composed system consisting of the preparation device

($\mathscr{D}$) and the original system ($\mathscr{S}$)—independently of whether or not the vector $\sum_i \lambda_i \vec{Z}_i$ is among the possible states of the original system. Applying the more precise notations we introduced in Appendix 2, (134) reads as follows:

$$^{(\mathscr{D}+\mathscr{S})}\vec{Z}' = {}^{(\mathscr{D}+\mathscr{S})}\vec{Z}'' = \sum_i \lambda_i \, {}^{(\mathscr{D}+\mathscr{S})}\left(F_t \vec{Z}_i\right) \tag{136}$$

$F_t : \Phi \to \Phi$ is not the time-evolution map of "mixed states," states of the joint system ($\mathscr{D}$+$\mathscr{S}$). The time-evolution of the ($\mathscr{D}$+$\mathscr{S}$) system is something else; a map $T_t : {}^{(\mathscr{D}+\mathscr{S})}\Phi \to {}^{(\mathscr{D}+\mathscr{S})}\Phi$, which is perhaps correctly claimed to be equivalent to the procedure described in point (ii). So, (134), more precisely (136), means, so to say, by the definition of T_t, that

$$T_t\left(\sum_i \lambda_i \, {}^{(\mathscr{D}+\mathscr{S})}\vec{Z}_i\right) = \sum_i \lambda_i \, {}^{(\mathscr{D}+\mathscr{S})}\left(F_t \vec{Z}_i\right) \tag{137}$$

instead of (135).

For instance, in our Biased Coin example, consider the following states:

$^{(\mathscr{D}+\mathscr{S})}\vec{Z}_1 = (0.7, 0.3, 0) \in {}^{(\mathscr{D}+\mathscr{S})}\Phi$: the state of ($\mathscr{D}$+$\mathscr{S}$) in which the preparation device $\mathscr{D}$ has propensity 1 to set the coin into state $\vec{Z}_1 = (0.7, 0.3, 0) \in \Phi$

$^{(\mathscr{D}+\mathscr{S})}\vec{Z}_2 = (0.3, 0.7, 0) \in {}^{(\mathscr{D}+\mathscr{S})}\Phi$: the state of ($\mathscr{D}$+$\mathscr{S}$) in which the preparation device $\mathscr{D}$ has propensity 1 to set the coin into state $\vec{Z}_2 = (0.3, 0.7, 0) \in \Phi$

$^{(\mathscr{D}+\mathscr{S})}\vec{Z}_{mix} = (0.5, 0.5, 0) \in {}^{(\mathscr{D}+\mathscr{S})}\Phi$: the state of ($\mathscr{D}$+$\mathscr{S}$) in which the preparation device $\mathscr{D}$ has propensity $\frac{1}{2}$ to to set the coin into state $\vec{Z}_1 = (0.7, 0.3, 0) \in \Phi$ and propensity $\frac{1}{2}$ to set the coin into state $\vec{Z}_2 = (0.3, 0.7, 0) \in \Phi$

Obviously, $^{(\mathscr{D}+\mathscr{S})}\vec{Z}_{mix} = \frac{1}{2}{}^{(\mathscr{D}+\mathscr{S})}\vec{Z}_1 + \frac{1}{2}{}^{(\mathscr{D}+\mathscr{S})}\vec{Z}_2$. Now, adopting the assumption that (137) holds, and applying (83)–(84) with, say, $t = 2$,

$$\begin{aligned}
T_2\left({}^{(\mathscr{D}+\mathscr{S})}\vec{Z}_{mix}\right) &= T_2\left(\frac{1}{2}{}^{(\mathscr{D}+\mathscr{S})}\vec{Z}_1 + \frac{1}{2}{}^{(\mathscr{D}+\mathscr{S})}\vec{Z}_2\right) \\
&= \frac{1}{2}F_2\vec{Z}_1 + \frac{1}{2}F_2\vec{Z}_2 = \frac{1}{2}F_2\left(0.7, 0.3, 0\right) + \frac{1}{2}F_2\left(0.3, 0.7, 0\right) \\
&\approx \frac{1}{2}(0.44, 0.56, 0) + \frac{1}{2}(0.22, 0.78, 0) = (0.33, 0.67, 0)
\end{aligned}$$

At the same time,

$$F_2\left(\frac{1}{2}\vec{Z}_1 + \frac{1}{2}\vec{Z}_2\right) = F_2\left(\frac{1}{2}\left(0.7, 0.3, 0\right) + \frac{1}{2}\left(0.3, 0.7, 0\right)\right) = F_2\left(0.5, 0.5, 0\right) \approx (0.27, 0.73, 0)$$

Thus, as expected from the formulas (83)–(84), F_t does not preserve convex combination, whether or not (137) is satisfied.

Funding

This work was supported by Hungarian National Research, Development and Innovation Office (Grant No. K134275).

References

Abramsky, S. and Heunen, C. (2016): Operational Theories and Categorical Quantum Mechanics. In J. Chubb, A. Eskandarian, and V. Harizanov (eds.), *Logic and Algebraic Structures in Quantum Computing,* Cambridge University Press, Cambridge.

Aerts, D. (2009): Operational Quantum Mechanics, Quantum Axiomatics and Quantum Structures. In D. Greenberger, K. Hentschel and F. Wienert (Eds.), *Compendium of Quantum Physics Concepts. Experiments, History and Philosophy,* Springer-Verlag, Berlin, Heidelberg.

Bana, G. and Durt, T. (1997): Proof of Kolmogorovian Censorship, *Found. Phys.* **27**, 1355.

Barandes, J. (2023): The Stochastic-Quantum Correspondence, https://philsci-archive.pitt.edu/22501.

Barnum, H., Barrett, J., Leifer, M., and Wilce, A. (2007): Generalized No-Broadcasting Theorem, *Phys. Rev. Let.* **99**, 240501.

Barnum, H., Barrett, J., Leifer, M., and Wilce, A. (2008): Teleportation in General Probabilistic Theories, arXiv:0805.3553 [quant-ph]

Bell, J. S. (1987): *Speakable and unspeakable in quantum mechanics*, Cambridge University Press, Cambridge.

Busch, P., Grabowski, M., and Lahti, P. J. (1995): *Operational Quantum Physics*, Springer-Verlag, Berlin, Heidelberg.

Cassinelli G. and Lahti, P. (2017): Quantum mechanics: why complex Hilbert space? *Phil. Trans. R. Soc. A* **375**, 20160393.

Cirel'son, B. S. (1980): Quantum Generalizations Of Bell'S Inequality, *Letters in Mathematical Physics* **4**, 93–100.

Davies, E. B. (1976): *Quantum Theory of Open Systems*, Academic Press, London.

Foulis, D. J. and Randall, C. H. (1974): Empirical Logic and Quantum Mechanics, *Synthese* **29**, 81–111.

Fröhlich J. and Pizzo A. (2022): The Time-Evolution of States in Quantum Mechanics according to the ETH-Approach, *Commun. Math. Phys.* **389**, 1673–1715

Gudder, S. (1988): Two-Hole Experiment (Ch. 2. Sec. 5), *Quantum probability*, Academic Press, Boston.

Griffiths, R. B. (2013): Hilbert space quantum mechanics is noncontextual, *Studies in History and Philosophy of Modern Physics* **44**, 174–181.

Hardy, L. (2008): Quantum Theory From Five Reasonable Axioms, http://arxiv.org/abs/quant-ph/0101012v4.

Henk M., Richter-Gebert J., and Ziegler G. M. (2004): Basic Properties Of Convex Polytopes, in *Handbook of Discrete and Computational Geometry*, *Second Edition*, J. E. Goodman and J. O'Rourke (eds.), Chapman & Hall/CRC, Boca Raton, Fla.

Hofer-Szabó, G., Rédei, M., and Szabó, L.E. (2013): *The Principle of the Common Cause*, Cambridge University Press, Cambridge.

Holevo, A. (2011): *Probabilistic and statistical aspects of quantum theory*, Edizioni della Normale, Springer, Pisa.

Jauch, I. M. and Piron, C. (1963): Can Hidden Variables be Excluded in Quantum Mechanics?, *Helv. Phys. Acta.* **36**, 827–837.

Leggett, A. J. (1998): Macroscopic Realism: What Is It, and What Do We Know about It from Experiment?, in: *Quantum Measurement: Beyond Paradox*, University of Minnesota Press, Minneapolis.

Ludwig, G. (1970): *Deutung des Begriffs "physikalische Theorie" und axiomatische Grundlagung der Hilbertraumstruktur der Quantenmechanik durch Hauptsätze des Messens*, Springer-Verlag, Berlin, Heidelberg.

Malament, D. B. (1992): Critical notice: Itamar Pitowsky's 'Quantum Probability – Quantum Logic', *Philosophy of Science* **59**, 300–320.

Meyer, W. and Kay, D. C. (1973): A convexity structure admits but one real linearization of dimension greater than one, *J. London Math. Soc.* **2**, 124–130.

Müller, M. P. (2021): Probabilistic theories and reconstructions of quantum theory, *SciPost Phys. Lect. Notes* 28 (doi: 10.21468/SciPostPhysLectNotes.28).

Pitowsky, I. (1989): *Quantum Probability – Quantum Logic*, Springer-Verlag, Berlin, Heidelberg.

Park, J. L. and Margenau, H. (1968): Simultaneous Measurability in Quantum Theory, *Int. J. Theoretical Physics* **1**, 211–283.

Park, J. L. (1970): The Concept of Transition in Quantum Mechanics, *Foundations of Physics* **1**, 23–33.

Park, J. L. and Margenau, H. (1971): The Logic of Noncommutability of Quantum-Mechanical Operators–and Its Empirical Consequences, in: *Perspectives in Quantum Theory – Essays in Honor of Alfred Landé*, W. Yourgrau and A. van der Merwe (eds.), The MIT Press, Cambridge, Massachusetts.

Popescu S. and Rohrlich, D. (1994): Quantum Nonlocality as an Axiom, *Foundations of Physics* **24**, 379–385.

Popper, K. R. (1967): Quantum mechanics without 'the observer', in: *Quantum theory and reality,* M. Bunge (Ed.), Springer, Berlin, Heidelberg.

Rédei, M. (2010): Kolmogorovian Censorship Hypothesis For General Quantum Probability Theories, *Manuscrito* **33,** 365-380.

Reichenbach, H. (1944): Analysis of an interference experiment (Ch. 1 § 7), in: *Philosophical foundations of quantum mechanics*, University of California Press, Los Angeles.

Ross, D. J. (1974): Operator–observable correspondence, *Synthese* **29**, 373–403.

Schmid, D., Spekkens, R.W., and Wolfe, E. (2018): All the noncontextuality inequalities for arbitrary prepare-and-measure experiments with respect to any fixed set of operational equivalences, *Phys. Rev. A* **97**, 062103.

Schroeck, F. E., Jr. and Foulis, D. J. (1990): Stochastic Quantum Mechanics Viewed from the Language of Manuals, *Foundations of Physics* **20**, 823–858.

Spekkens, R.W. (2005): Contextuality for preparations, transformations, and unsharp measurements, *Phys. Rev. A* **71**, 052108.

Strauss, M. (1936): The logic of complementarity and the foundation of quantum theory, in: *The Logico-Algebraic Approach To Quantum Mechanics, Volume I, Historical Evolution*, C. A. Hooker (ed.), D. Reidel Publishing Company, Dordrecht, 1975.

Szabó, L. E. (1995): Is quantum mechanics compatible with a deterministic universe? Two interpretations of quantum probabilities, *Foundations of Physics Letters* **8**, 421.

Szabó, L. E. (1998): Quantum structures do not exist in reality, *Int. J. of Theor. Phys.* **37**, 449–456.

Szabó, L. E. (2001): Critical reflections on quantum probability theory, in: *John von Neumann and the Foundations of Quantum Physics*, M. Rédei, M. Stoeltzner (eds.), Kluwer Academic Publishers, Dordrecht.

Szabó, L. E. (2008): The Einstein--Podolsky--Rosen Argument and the Bell Inequalities, *Internet Encyclopedia of Philosophy.* (https://iep.utm.edu/einstein-podolsky-rosen-argument-bell-inequalities).

Szabó, L. E. (2020): Intrinsic, extrinsic, and the constitutive a priori, *Foundations of Physics* **50**, 555–567.